Biosynthesis and regulation of terpene production in accessions of chamomile (*Matricaria recutita* L.)

Dissertation

zur Erlangung des
Doktorgrades der Naturwissenschaften (Dr. rer. nat.)
der
Naturwissenschaftlichen Fakultät I – Biowissenschaften –

der Martin-Luther-Universität
Halle-Wittenberg

vorgelegt

von Prodyut Mondal

geboren am 21.11.1983 in Kolkata, India

Reviewers (Gutachter)

1. Prof. Dr. Jörg Degenhardt
Institute of Pharmacy, Martin Luther University, Halle-Wittenberg
Germany

2. Prof. Dr. Timo Niedermeyer
Institute of Pharmacy, Martin Luther University, Halle-Wittenberg
Germany

3. Prof. Dr. Timothy F Sharbel
Global Institute for Food Security (GIFS), University of Saskatchewan, Saskatoon, Canada

Tag der öffentlichen Verteidigung: 27.02.2020

Table of Contents

List of Figures

List of Figures

List of Tables

Abbreviations

aa	amino acid
ATP	adenosine tri-phosphate
C. recutita	*Chamomilla recutita*
CBDA	cannabidiolic acid
CBGA	cannabigerolic acid
cDNA	complementary Deoxyribonucleic acid
Conc	concenration
DMAPP	dimethylallyl pyrophosphate
DNA	deoxyribonucleic acid
dNTP	deoxynucleoside triphosphate
DTT	dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
eV	electron volt
FID	flame ionization detector
FPP	farnesyl diphosphate
GC-MS	gas chromatography mass spectrometry
GGPP	geranylgeranyl diphosphate
GPP	geranyl pyrophosphate
HPLC	high pressure liquid chromatography
HS-SPME	head space-solid phase microextraction
IPP	isopentenyl pyrophosphate
kb	kilobase pairs
LB	Luria-Bertani (medium)
LC-MS	liquid chromatography with mass spectrometry
LMAO	low molecular antioxidants
LPP	Linally pyrophosphate
MEP	methylerythritol phosphate
m-RNA	messenger-Ribonucleic acid
MrTPS	*Matricaria recutita* α-bisabolol synthase
MS	mass spectrometry
MVA	mevalonic acid
NADPH	Nicotinamide adenine dinucleotide phosphate hydrogen
NCBI	National Center for Biotechnology Information

Abbreviations

NMR	Nuclear Magnetic Resonance
OD	optical density
OPP	Pyrophosphate
ORF	open reading frame
PAL	phenylalanine ammonia-lysae
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulfonyl fluoride
qRT-PCR	quantitative real-time polymerase chain reaction
RACE	rapid amplification of cDNA ends
ROS	reactive oxygen species
rpm	revolutions per minute
SA	salicylic acid
SOC	super optimal broth with catabolite repression
SPME	solid phase microextraction
TAE	Tris-acetate-ethylenediaminetetraacetic acid
TIC	total ion chromatogram
TLC	thin-layer chromatography
TPSs	terpene synthases
USA	United State of America
VOC	volatile organic component

1 Introduction

1.1 Essential oils in pharmaceutically used plants

Essential oils are composed of lipophilic compounds with a molecular weight below 500 Da. They can contain several hundred different compounds of isoprenoids, phenylpropanoids or derivatives of fatty acids (Atta-ur-Rahman et al., 2000; Hüsnü et al., 2007). Essential oils are produced in plants in separate compartments, in oil cells, in secretion ducts or in glands. Structurally, two types of oil glands exist, capitate cells with apical secretory cells, basal cell and stalk cell, and peltate cells with an eight-cell secretory apparatus, stalk and basal cell (Hanlidou et al., 1991; Tiwari, 2016; Wang et al., 2008). Trichome studies suggest that secretory cells are the main site of monoterpene and sesquiterpene biosynthesis (Huchelmann et al., 2017). Because of their shorter C-chains, monoterpenoids (consisting of ten carbon atoms), are often more volatile and more hydrophilic than sesquiterpenoids (fifteen carbon atoms). Trichome research gained importance because of the biosynthesis of pharmaceutically interesting essential oils (Wang, 2014).

Fig. 1.1 Grandular trichomes in plant (image acknowledgement: Andreucci et. al, 2008)

The composition of essential oils differs widely between plant species (Table 1.1). In addition, large intraspecific variation has been observed. Often, plant populations contain groups of individuals with distinct essential oil profiles which are termed chemotypes.

In *Thymus vulgaris* L. the chemotypic variation found depends upon the dominance of a single monoterpene, categorized as A-type (α-terpineol), C-type (carvacrol), G-type (geraniol), L-type (linalool), T-type (thymol) and U-type (*(E)*-sabinene hydrate) (Vernet et al., 1986). The Australian *Eucalyptus dives* contains piperitone as the chief constituent in the oil, but other types are known which produce principally phellandrene or cineole, while still others produce oils that are intermediate in composition (Evans, 2009). For chamomile, different accessions are distinguished by the presence of α-bisabolol and its oxides (type A to

D). In type-A, bisabolol oxide B is dominant (α-bisabolol oxide B > α-bisabolol oxide A >α-bisabolol), α-bisabolol oxide A is predominant in type-B (α-bisabolol oxide A > α-bisabolol oxide B > α-bisabolol), in type C the α-bisabolol is higher than oxide B (α-bisabolol >α-bisabolol oxide B >α-bisabolol oxide A) and both α-bisabolol and its oxides are balanced in type D of chamomile plants (Lawrence B.M., 1987; Rolf Franke, 2005).

Table 1.1 Selected plant families and their essential oil composition.

Family	Plant	Compounds	References
Asteraceae	*Chamomilla recutita* L.	It contains 0.24%–1.9% of volatile oils, approximately 120 secondary metabolites have been identified in chamomile, including 28 terpenoids and 36 flavonoids.	(Mann and Staba, 1986)
Labiatae	*Salvia officinalis*	40 components were identified, mainly about 50% of α- and β-thujone together with cineole, borneol and other monoterpenes.	(Raal et al., 2007)
Labiatae	*Rosmarinus officinalis* L	Several phenolic abietane diterpenoids, carnosic acid together with its degradation products carnosol, rosmanol, epirosmanol and 7-methylepirosmanol and flavonoids.	(Mahmoud et al., 2005)
Labiatae	*Origanum Onites* L.	2.5% oil and a minimum of 1.5% carvacol and thymol.	(Kintzios, 2002)
Labiatae	*Lavandula Angustifolia* Miller	66% of the total essential oils, within these α-pinene Camphene, Sabinene β-pinene β-phellandrene, cymene, 1,8-cineole and limonene are major one.	(De Pascual Teresa et al., 2007)
Lamiaceae	*Thumus. vulgaris* L.	Depending on the dominant terpene, they are categorized as: α-terpineol (A-type), carvacrol (C-type), geraniol (G-type), linalool (L-type), thymol (T-type), and (E)-sabinene hydrate (U-type).	(Vernet et al., 1986)
Lamiaceae	*Mentha piperita*	Major compounds are limonene 1.0–5.0%, cineole 3.5–14.0%, menthone 14–32%, menthofuran 1–9%, isomenthone 1.5–	(William Charles Evans, 2009)

		10.0%, menthyl acetate 2.8–10.0% and menthol 30–55%.	
Lamiaceae	*Ocimum basilicum*	Monoterpenoids, sesquiterpenoids and phenylpropanoids. Major components: geranial, neral, linalool, methyl chavicol, eugenol and methyl eugenol.	(Grayer et al., 1996)
Lauraceae	*Cinnamomum zeylanicum*	Cinnamon oil contains about 60–75% w/w of trans-cinnamic aldehyde, 4–10% of phenols, hydrocarbons (pinene, phellandrene and caryophyllene) and small quantities of ketones, alcohols and esters.	(William Charles Evans, 2009)
Myristicaceae	*Myristica fragrans*	Nutmegs yield 5–15% of volatile oil, 30–40% fat, phytosterin, starch, amylodextrin, coloring matter and saponins.	(Mukherjee et al., 2007)
Myrtaceae	*Syzygium aromaticum (Eugenia caryophyllus)*	Clove oil contains 84–95% of phenols (eugenol with about 3% of acetyl-eugenol) sesquiterpenes (α and β-caryophyllene) and small quantities of esters, ketones and alcohols.	(William Charles Evans, 2009)
Myrtaceae	*Melaleuca alternifolia*	Cyclic monoterpenes are the principal components of the oil, terpeien-4-ol (30% minimum), α-terpinene(10–28%), *p*-cymene (0.5–12.0%) and cineole (less than 15.0%).	(An Southwell, 1999)
Myrtaceae	*Eucalyptus globules* Labill.	The chief requirements are a high cineole and pinene (α and β).	(Douglas et al., 2018)
Pinaceae	*Pinus palustris* (long leaf pine) and *P. elliottii* (slash pine)	The principal constituents and official limits are α-pinene (70–85%), β-pinene (11–20%) and limonene (1–7%). Other components in small amounts are camphene, β-myrcene, longifolene, β-caryophyllene and caryophyllene oxide.	(William Charles Evans, 2009)

Rosaceae	*Rosa damascena*	The terpenes mainly consist of geraniol, citronellol, nerol and 2-phenylethanol with smaller quantities of esters and other odorous principles.	(Oka et al., 1998)
Rutaceae	*Citrus aurantium*	The aroma constituents of orange peel are terpenes, mainly monoterpenes, and certain aldehydes. The most abundant component is limonene.	(Sueramania nt, 1988)
Santalaceae	*Santalum album*	The heartwood of mature trees (10 years old) contain essential oils, chiefly the sesquiterpene alcohols cis-α-santalol, cis-β-santalol, α-trans-bergamotol, epi-cis-β-santalol along with small quantities of trans-β-santalol and cis-lanceol.	(Jones et al., 2006)
Umbelli ferae	*Carum carvi*	Caraway contains 3–7% of volatile and 8–20% of fixed oil, proteins, calcium oxalate, colouring matter and resin.	(Hardman, 1998)
Zingiber aceae	*Elettaria cardamomum*	Samples of cardamom seed yield 2.8–6.2% of volatile oil, starch (up to 50.0%), fixed oil (1–10%) and calcium oxalate (according to the British Pharmacopoeia).	(Kenmogne et al., 2005)

1.1.1 Chamomile: An important medicinal herb

The medicinal herb ‘Chamomile flowers’ are the dried flower heads of *Matricaria recutita* L. (family: Asteraceae). It originates from Southeastern Europe and Western Asia but nowadays it is cultivated throughout the world. Chamomile is an annual plant and its height varies from 20 to 60 cm, the root system is short but quite widespread (Applequist, 2002). The flowers have a pleasant sweet-herbaceous odor with a slightly bitter taste (Hänsel, R., Sticher, O., Steinegger, 1999). German chamomile, (*Matricaria chamomilla* L, *Matricaria recutita* L.) is one of the most widely used herbs in the world. Its flowers, bearing white petals around a bright yellow center, are a familiar sight in central Europe. The plant is cultivated for the production of dried tea, pharmaceutical products and essential oils (Andrzejewska and Woropaj-Janczak, 2014). In Germany, the domestic production of medicinal plants increased

during the past few years due to a higher interest in phytopharmaceuticals, better quality and good practice for quality controls of regional products (Brabandt and Ehlert, 2011). Due to growing demand from the pharmaceutical, cosmetic, and food industries for high-quality herbal raw materials, breeding work is undertaken to obtain improved accessions (Andrzejewska and Woropaj-Janczak, 2014). Both, diploid (e.g. Bohemia, Bona, Promyk, Mastar) and tetraploid (e.g. Lutea, Goral, Zloty Lan, Dukat etc.) accessions exist. They often differ from each other in physical appearance with regard to height, branching and flower size, and more importantly in the composition and quantity of essential oils (Azizi et al., 2007; Slovak, 2017).

1.1.2 Pharmacological effects of chamomile essential oil

Chamomile is known to possess different therapeutic activities such as anti-arthritis, antispasmodic and anti-ulcerogenic effects (Srivastava et al., 2010). The curative effect of chamomile has been known from ancient times (Lucknow, 1995). Different classes of bioactive constituents are present and have been isolated and used as medicinal preparations and cosmetics (Ara Der Marderosian, 1989). The α-bisabolol is used in dermatological applications including moisturizing lotions, cleansers, sunscreens, antiperspirants, and for makeup products. The whole plant is used for making herb beers, and creams used externally for toothache, earache, neuralgia and external swellings. Chamomile preparations are commonly used against hay fever, inflammation, muscle spasms, menstrual disorders, insomnia, gastrointestinal disorders, congestive neuralgia and hemorrhoids. The most popular application form is herbal tea. The tea infusion is used as a mouth wash or gargle against inflammation (Fidler et al., 1996). The chamomile flower oil consists mainly of sesquiterpene derivatives (75 -90%) and traces of monoterpenes. The whole plant contains 0.24%–1.9% essential oil. German chamomile is a natural source of the blue oil azulene. Approximately, 120 secondary metabolites have been identified including 28 terpenoids and 36 flavonoids (Mann and Staba, 1986).The principal components are α-bisabolol and its oxide azulene including chamazulene and acetylene derivatives. Among flavonoids, apigenin is the most abundant compound. It is present in very small quantities as free apigenin, but predominantly exists in the form of various glycosides (Avallone et al., 2000; Babenko and Shakhova, 2006; Švehlíková et al., 2004). Most important are α-bisabolol and chamazulene. Chamazulene is an artifact formed from matricin, which is naturally present in the flowers. Wild types of

German chamomile often belong to chemotype A, which is relatively low in *α*-bisabolol and its oxides.

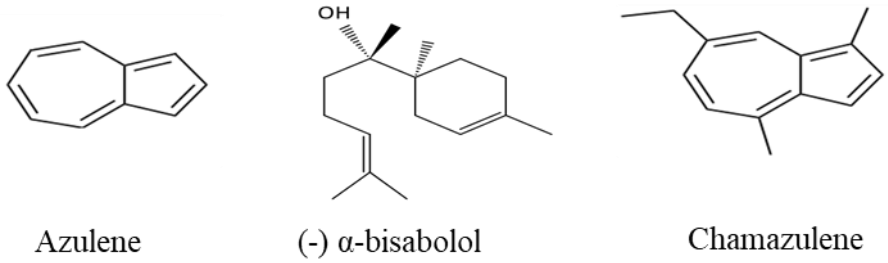


Fig. 1.2 Structures of Azulene, (-) α-bisabolol and Chamazulene.

The anti-inflammatory and skin-soothing properties are used for pain remedies (Russell, Kathryn; Jacob, 2010). It is reported that α-bisabolol induces apoptosis in cancer cell lines. It had an anti-proliferative effect on pancreatic cancer cell lines using *in vitro* and *in vivo* experiments (Uno et al., 2016). These studies demonstrate that α-bisabolol is one of the most significant constituents for commercial perspectives.

1.1.3 The environment affects the composition of essential oil

Several parameters affect the quantity and the quality of essential oils. Mann and Staba, 1986 found that the growth conditions regulate the content of α-bisabolol from 0.4% up to 1.2% in chamomile flowers within α-bisabolol chemotype that contained 15% to 30% of α-(−)-bisabolol in essential oil in the mature stage of the plant (Mann and Staba, 1986). The quantity and quality of essential oils are strongly influenced by environmental factors such as temperature, light, photoperiod, nutrition, moisture and soil salinity (Banchio et al., 2005; Naghdi Badi et al., 2004). Salicylic acid (SA) plays a significant role in plant stress as a signaling molecule. Under normal and heat stress conditions SA was used on three chamomile accessions (Bodegold, Bona and Bushehr). The quantity of chamazulene and α-bisabolol were increased. No significant effect on the quality of bisabolol oxide A was observed in the presence of heat stress (Ghasemi et al., 2017).

1.1.3.1 Genetic efforts to improve the essential oil content

Breeding efforts on chamomile have focused on better plant performance under stress and a higher quality of desired chemical constituents. Wagner et al., 2005 were phenotyping F2-populations. The different chemotypes were named after the dominating bisaboloid. The

Fig. 1.3 Chamomile bud and flowers

breeding populations Akk34, B28, C42 and D46 belong to the α-bisabolol-rich chemotype whereas the accessions Bodegold and Soroksari40 belong to the bisabololoxide A chemotype and Margaritar to the oxide B chemotype (Wagner et al., 2005).

Natural chamomile populations have a large phenotypic variability which has been used for outcrossing. The wild chamomile populations are diploid, but generally tetraploid or diploid lines are chosen for cultivation. Otto et al., 2017 reported that tetraploid lines have larger flower heads and higher seed weight than diploids. Diploid accessions are flowering earlier than tetraploid accessions. Interestingly, tetraploid accessions are not always preferable over diploids accessions in terms of α-bisabolol production (Otto et al., 2017).

Table 1.2 Diploid and tetraploid chamomile accessions

Tetraploids	Diploids
Bodegold, Bohemia, Goral, Lazur, Lutea, Manzana, Margaritar, Pnos, Zloty Lan,	Argenmilla, Bona, Camoflora. Germania, Pop Germany Poherilicky velkokvety.

In the same study, the single bisaboloid was measured individually from different genotypes. Plants with a high level of α-bisabolol often showed a lower content of its oxides including bisabolone-oxide A. An earlier study proposed steps for the oxidative conversion of α-bisabolol to its oxides and bisabolon oxide A (Lanka et al., 2014).

1.1.3.2 Nitrogen and heavy metal affect the essential oil in chamomile

Nitrogen is essential for plant growth and metabolism (Kusano et al., 2011). Plants have to balance their metabolism for carbon-rich metabolites in nitrogen-limited conditions (Matt et al., 2002). Accumulation of phenolic secondary metabolites can be limited by phenylalanine ammonia-lysae (PAL) activity, which is a key enzyme in the biosynthesis leading to flavonoids, benzoic acid as well as coumarin formation. Repcak et al., 2001 reported that during stress conditions, phenolic metabolites such as umbelliferone with anti-oxidative activity increases in chamomile (Repčák et al., 2001). A higher concentration of nitrogen in the field leads to growth and a higher chlorophyll content. However, the opposite effect was observed in response to water loss (Repčák et al., 2001). Heavy metal can be toxic for living organisms (Maksymiec, 2007). Jozef and Martin, 2008 observed a reduction of the oxidative status in *M. recutita* with the introduction of cadmium and copper. Cu and Cd accumulation affected the plant water balance and anti-oxidant capacity (Kováčik and Bačkor, 2008). The coumarin-related compound herniarin was not affected by Cd while its precursors (*Z)*- and (*E*)-2-β-D-glucopyranosyloxy-4-methoxycinnamic acids (GMCAs) increased significantly (Kováčik et al., 2006). *M. recutita* plants selected from areas of medium and high heavy metal levels (Cu, Mn, Ni, Pb, and Zn) showed an increase in the contents of low molecular anti-oxidants (LMAO), including flavonoid (Prokop'ev et al., 2014).

1.1.3.3 Drought effects on chamomile

The effect of soil dehydration and rehydration on two chamomile varieties was studied by Kwinta et al., 2010. Parameters like shoot and leaf growth, vegetative development, protein content, ascorbate peroxidase activity, and gas exchange were measured. Under stress the water content was affected and the effect ceased after rehydration. The growth of long shoots and leaf formation drastically decreased in all plants (Baczek-Kwinta et al., 2010). Several studies focused on morphological, agronomical and phytochemical traits changed due to the effect of saline irrigational water (Baghalian et al., 2008).

1.1.3.4 Biotic stress affects the quality of essential oils

Plant secondary metabolites often function in plant defense and biotic stress responses. The emission or accumulation of terpenoids can be affected by herbivores or pathogens. A classic example is the crop plant maize (*Zea mays L*). Maize leaves emit a complex, volatile blend of mostly (*E*)-ß-farnesene, (*E*)-α-bergamotene, and (*E*)-β-caryophyllene when damaged by lepidopteran larvae. The emitted terpenes attract parasitic wasps acting as an indirect defense

for the plant (Köllner et al., 2008; Richter et al., 2016; T. C. J. Turlings, J. H. Tumlinson, 1990). Plants attacked by herbivorous insects can undergo significant changes in their primary and secondary metabolism, including increased production of toxic, anti-digestive, and anti-nutritive compounds (Kessler and Baldwin, 2002). Table 1.3 summarizes the impact of some common herbivores on chamomile.

Table 1.3 Herbivors of chamomile
According to Andreas Plescher, 2004

Type of herbivory	Scientific name or species	Damage
Chewing herbivory on stem and roots	*Melolontha* spp., *Phyllopertha* spp., *Rhizotrogus* spp.	Herbivores are present in soil, affected to roots and stems.
Gall formation on roots	Nematode Meloidogyne *hapla* (Chitwood).	Nematode causes swelling of roots and gall formation. Plant growth inhibited and infected plants are sensitive to drought.
Chewing herbivory on shoots and leaves	*Cucullia tanaceti* SCHIFF *Phalonia implicata* WCK	Damage to upper parts of the plant and spin webs around them. Plant damaged by skeletal feeding and leaves become dry.
Sap structure on leaves and shoots	Aphid: *Cerosipha gossypii* HB., *Aphis fabe* SCOP *Chlorita viridula* FALL	Aphid causes color patches over tissues. First turns dark green to brown. Leaves and stems deformation occurs.
Stem and leaf mining	*Ceutorhynchus rugulosus* HERBST	The lower part of the stem turns red and then brown, while the leaves turn yellow and wilt. Flowers and stems are easily collapsed.
Chewing on Flowers	Beetels genus Meligethes Larvae: *Pseudostyphlus pilumnus* GYLL *Ceutorhynchus rugulosus* HERBST	Entire flower heads turn brown. Flower harvest getting low.
Sap sucking on Flowers	*Thrips physapus* L. *Thrips tabaci* LIND	Sucking on tubular florets impairs flower head integrity and are dried.

Attacks of viruses, bacteria and fungi (*Fusarium* spp., *Erysiphe cichoracearum* D.C, Peronospora Danica GÄUM, *Puccinia matricariae* SYD and *Puccinia tanaceti*) affect directly the quality of the plant. Chamomile is one of the host plants for the lettuce big vein virus (LBV-V) and the cabbage black ring virus (CBR-V). Roots, flowers, stems and leaves can be damaged.

1.1.4 The content and composition of chamomile essential oil varies with extraction procedures and geographical origin.

Different solvent methods are used for the analysis and extraction of chamomile essential oil. Hydro-distillation, successive solvent extraction, and chromatographic techniques are used. Advanced techniques such as solid phase microextraction (SPME), high performance liquid chromatography (HPLC), high performance liquid chromatography-mass spectrometer (HPLC-MS) with or without Nuclear Magnetic Resonance (NMR), Liquid Chromatography-Diode-Array Detection (LC-DAD), and GC-MS are employed for both qualitative and quantitative detection. Chamomile ligulate flowers and tubular flowers contain predominantly artemisia ketone, (*E*)-β-farnesene and germacrene D as volatile components when analyzed by Head space-solid phase microextraction (HS-SPME). α-bisabolol oxide A, chamazulene and spiroether (*Z*) are extracted by conventional steam distillation-solvent extraction (SDSE) (Rafieiolhossaini et al., 2019). For flavonoid compounds, polar solvents (ethanol or water) are best, but non-polar or organic solvents are best for the extraction of the essential oil. Employing HPLC-MS and HPLC-NMR with the respective solvent extraction mentioned above, Weber et al 2008 found apigenin and its glucosides and α-bisabolol as the main constituents from *M. recutita* (Weber et al., 2008).

Accessions can vary in quantity and quality of essential oils, also depending on the respective author's reports. Rall et al., 2012 found 30% less quantity compared to Weiss and Fintelmannin 2000 (Weiss RF Fintelmann V, 2000). Among the cultivated accessions, the Hungarian *BK-2* line produces more chamazulene (anti-inflammatory agent) in its essential oil than the German Degumil type (Manzana). Within the essential oils, the monoterpenes cineole, limonene, pinene, terpineol etc. as well as the sesquiterpenes chamazulene, caryophyllene, bisabolol and its oxides, farnesene etc. are commercially important. The main constituents in plants from European regions are *(E)*-β-farnesene, α-bisabolol, α-bisabolol oxides A and B, chamazulene, and *(Z)*-en-in-dicycloether (Pintore et al., 2002). The chemical constituents can differ depending on the geographical origin. Commercial samples of *M. recutita* collected from Latvia and Poland have a higher amount of α- bisabolol oxide B than

α-bisabolol oxide A. The opposite was found in samples from Lithuania and Netherlands. Artemisia ketone (7.8%) and *(E)*-β-farnesene (10.9%) are found in a higher quantity in samples from the United States of America (USA). A higher content of chamazulene (3.4 to 4.9% in total) was found in Estonian and Polish chamomile oil (6.7% in total) when compared with samples from the USA (0.2% in total) (Raal et al., 2012).

Orva et al., 2001 reported from Estonian *M. recutita* α-bisabolol, bisabolol oxide A and B, bisabolone oxide A, *(E)*-β-farnesene, chamazulene and en-in-dicycloether as the main constituents derived from solvent distillation (Orav et al., 2001). In *C. recutita* from India, the main constituents are 6-methyl-hept-5-en-2-one (0.5–5.4%), artemisia ketone, artemisia alcohol, *(E)*-β-farnesene (8.6–13.6%), α-bisabolol, α-bisabolol oxide A and B and chamazulene extracted from flower heads by applying hydro distillation techniques (Pintore et al., 2002).

1.2 Biosynthesis of terpenes in plants

Plant terpenes form the largest group of plant secondary compounds and are structurally derived from the C5-hydrocarbon isoprene (Rosenthal, G.A., Berenbaum, 1992). In plants, both the cytosolic mevalonate (MVA) and the plastidic methylerythritol phosphate (MEP) pathways generate the five-carbon compound isopentenyl pyrophosphate (IPP) and its isomer di-methylallyl pyrophosphate (DMAPP) (Fig. 1.4). The condensation of one DMAPP and one IPP molecule results in a ten-carbon geranyl pyrophosphate (GPP). Ultimately, IPP and DMAPP are used for the formation of all terpenes and isoprenoids either in the cytosol or the chloroplasts of plants.

In chloroplasts, prenyltransferases condense DMAPP with three IPP molecules to produce geranyl-geranyl-pyrophosphate (GGPP). In the cytosol, the condensation of one DMAPP molecule with two IPP molecules results in farnesyl pyrophosphate (FPP). All of these precursors are building blocks of many secondary metabolites in plants (Fig. 1.4) (Buchanan et al., 2000; Pichersky et al., 2006)

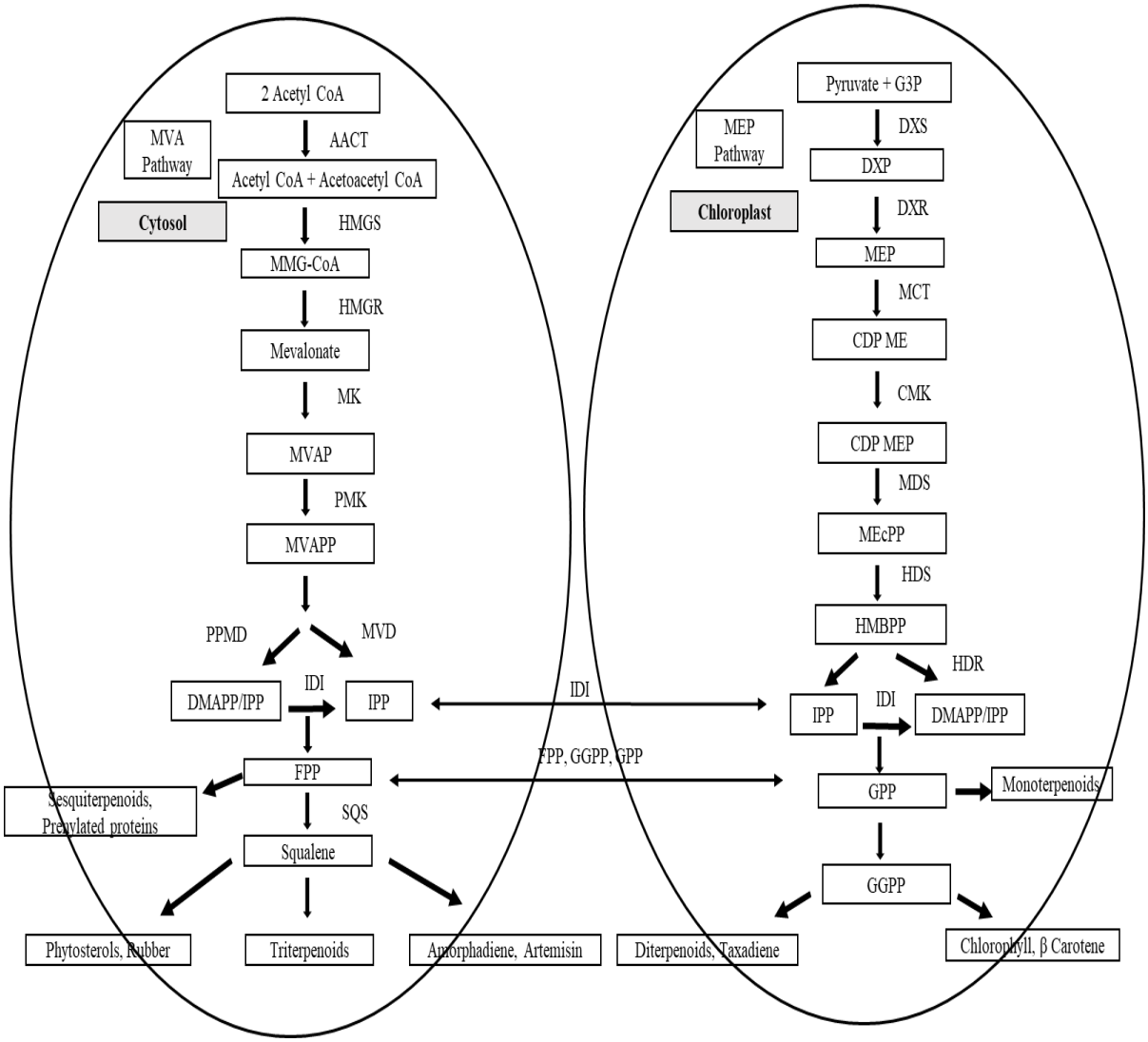


Fig. 1.4 Scheme of terpene biosynthesis in plants (MVA and MEP pathway)
On the left side is the cytosolic MVA pathway, leading to IPP and DMAPP. On the right side is the chloroplastic MEP pathway, also leading to IPP and DMAPP. Abbreviations:
CDP-ME-4-cytidine 5′-diphospho-2- C-methyl-D-erythritol, CDP-MEP - CDP-ME 2-phosphate,DMAPP-dimethylallyl diphosphate, DXP- 1-deoxy-D-xylulose 5-phosphate, FPP- farnesyl diphosphate, GGPP-geranylgeranyl diphosphate, GPP geranyl diphosphate, HMBPP- 1-hydroxy-2-methyl-2-butenyl 4-diphosphate, HMG-CoA- 3-hydroxy-3-methylglutaryl CoA, IPP- isopentenyl diphosphate, MEcPP - 2-C-methyl-D-erythritol 2,4-cyclodiphosphate, MEP-2-C-methyl-D-erythritol 4-phosphate, MVA- mevalonic acid, MVP- 5-phosphomevalonate, MVPP - 5-diphosphomevalonate.
MVA pathways enzymes : AACT - acetoacetyl-CoA thiolase, HMGS- 3-hydroxy-3-methylglutaryl CoA synthase, HMGR: 3-hydroxy-3-methylglutaryl CoA reductase, MVK – mevalonate kinase, PMK - 5-phosphomevalonate kinase, MVD - 5- diphosphomevalonate decarboxylase, IDI- isopentenyl diphosphate isomerase, PPMD - diphospho-mevalonate decarboxylase, SQS - qualene synthase.
MEP pathways enzymes: DXS - 1-deoxy-D-xylulose 5-phosphate synthase, DXR - 1-deoxy-D-xylulose 5-phosphate reductoisomerase, MCT - 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, CMK- 4-(cytidine 5′- diphospho)-2-C-methyl-D-erythritol kinase, MDS - 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, HDS - 1-hydroxy-2-methyl-2-butenyl 4-diphosphate synthase, HDR-1-hydroxy-2-methyl-2-butenyl 4-diphosphate reductase, IPP - isopentenyl diphosphate. Fig. 1.4 modified from Sarada D. Tetali, 2019 and Lio et. al, 2016.

1.3 Terpene synthases create a large structural diversity

The pyrophosphates GPP, FPP, and GGPP are substrates used by terpene synthases to form monoterpenes, sesquiterpenes, and di-terpenes, respectively. The reaction is initiated by the activation of the diphosphate involving a divalent metal ion. Several nucleophilic, dehydrogenation, and cyclization reactions can be catalyzed by terpene synthases (TPS) to generate the large diversity of terpenes (Degenhardt et al., 2009). The structural diversity of sesquiterpenes is higher than that of monoterpenes because an increased number of cyclizations is possible from a precursor with five additional carbon atoms.

1.3.1 Structural and biochemical characteristics of terpene synthases

More than 25,000 terpene structures were identified from plants (J. D. Connolly and R. A. Hill, 1991) and more are to be found. Because many TPSs are multiproduct enzymes, more than one compound can be produced from a single substrate. For example, the enzyme MtTPS5 from *Medicago truncatula* forms 27 products from a single substrate FPP (Garms et al., 2010).

The TPS enzymes have an average length of 550-850 amino acids. Enzymes for the biosynthesis of monoterpenes in the plastids have a length of 600-650 amino acids (Bohlmann and Steele, 1997; Colby et al., 1993; Yuba et al., 1996). Cytosolic sesquiterpene synthases are 50-70 amino acid shorter, but diterpene synthases are approximately 210 amino acids longer than monoterpene synthases (Wildung and Croteau, 1996). Generally, sesquiterpene synthases are located in the cytosol, monoterpene and diterpene synthases are located in plastids. Within the pathway, the enzymes and poly-isoprenoid substrates change bonding, hybridization and configuration via highly reactive carbocation intermediates (McCaskill and Croteau, 1997; Misawa, 2010; Wendt et al., 2000). Crystal structures of monoterpene, sesquiterpene and diterpene synthases reveal α-helices and short connecting loops (Kampranis et al., 2007; Starks et al., 1997; Whittington et al., 2002). Two types of domains are significant, the N-terminal and the C-terminal domain. The N-terminal domain structurally resembles the catalytic core of glycosylhydrolsases (Bohlmann et al., 1998). The C-terminal domain contains the active site. Resolved crystal structures imply that the N-terminal domain facilitates proper folding, stabilizes the active site in the C-terminus and covers the active center after substrate binding (Köllner Schnee, C., Gershenzon, J., Degenhardt, J., 2004; Starks et al., 1997).

1.3.1.1 Specific structural motifs of terpene synthases

The N-terminal signal peptide of monoterpene synthases is responsible for the transport to the plastids. It is rich in threonine and serine but low in acidic amino acids (Bohlmann et al., 1998; Bohlmann and Steele, 1997). Bohlmann et al., 1997 reported that the deletion of this N terminal signal peptide has no effect on the activity of monoterpene synthases. The RRX8W motif is highly conserved for the enzymatic activity and downstream sequences are significant for maturation process of the proteins. Degenhardt et al., 2009 described that the active site of the C-terminal domain is the hydrophobic pocket and formed by six α-helices and covered by two loops on the protein surface (Degenhardt et al., 2009). Further important motifs of terpene synthases are following:

DDxxD aspartate rich region: It is the most abundant structural motif found in almost all plant terpene synthases and microbial synthases (Fig. 1.4). Resolved crystal structures revealed that this aspartate rich motif is responsible for the binding of divalent cations and interact with the diphosphate moiety of the substrate at the entrance of the active site (Lesburg et al., 2007; Tarshis et al., 1994).The mutagenesis of this motif, generally decreases the catalytic activity and produces an abnormal product which can alter the substrate binding (Cane et al., 1996; David E. Cane, Qun Xue, 1996; Rynkiewicz et al., 2002).

NSE/DTE: Another aspartate rich motif is the (L,V)(V,L,A)-(**N**,**D**)D(L,I,V)X(**S**,**T**)XXX**E**, it is designated as NSE/DTE, which is conserved in prenyltransferases at the opposite site of the catalytic center (Christianson, 2006). However, the DDxxD motifs are highly abundant in almost all plant TPSs compared to the NSE/DTE motif. In some sesquiterpene synthases, the NSE/DTE is replaced with another DDxxD which is responsible for catalytic activity (Little and Croteau, 2002). In several terpene synthases, the DDxxD and the NSE/DTE motif together bind three magnesium ions which interact with the diphosphate moiety of the substrate. This causes the initial substrate ionization and protonation of the pyrophosphate group (Christianson, 2006). Additionally, another aspartate rich motif DxDD is present in certain diterpene and triterpene cyclases. The DxDD acts as a proton donor that mobilizes the initial carbocation formation in terpenoid cyclases (Wendt et al., 2000).

Rx8W/RRx8W motif: This motif is located about 60 amino acids from the N-terminus and is predominantly present in monoterpene and sesquiterpene synthases (Fig.1.4). In monoterpene synthases, it is responsible for isomerization of the GPP substrate to form the LPP intermediate (David C. Williams, Douglas J. McGarvey, Eva J. Katahira, 1998). In the absence of this motif, monoterpene synthases do not accept GPP as a substrate, but are able to

metabolize LPP to form an acyclic compound, No cyclic monoterpenes are formed (Bohlmann and Steele, 1997; Krause, 2016).

RxR motif: The RxR motif is abundant in terpene synthases and generally found 35 amino acid upstream of the DDxxD motif. RXR motif is preventing the nucleophilic attack of any carbocationic intermediates. This is responsible for the complexation of the diphosphate function after ionization of the substrate (Starks et al., 1997).

Table 1.4 The conserved regions of sesquiterpene synthase sequences of chamomile.

RX8W
MSTLSVSTPSFSSSPLSSVNKNSTKQHVT**RNSVIFHDSIW**GDQFLEYKEKSNVATEKQLIEEL
KEEVRNELMIRACNEASIDVVERLGLAYHFEKEIEESLQHIYVTYGHKWTNYNNIESLSLWF
RLLRQNGFNVSSDIFENHIDEKGNFQESLCNDPQGMLALYEAAYMRVEGEIILDKALGFTKL
HLGIISNDPSCDSSLRTEIKQALKQPLRRRLPRLEAVRYIAIYQQKASHSEVLLKLAKLD
RXR
FNVLQEMHKDELSQICKWWKDLDIRNKLPYV**RDR**LIEGYFWILGIYFEPQHSRTRMFLMK
DDXXD
TCMWLIVL**DDTFD**NYGTYEELEIFTQAVERWSITCLDELPEYMKLVYHEQFRVHQEMEESL
EKEGKAYQIHYIKEMAKEGTRSLLLLEAKWLKEGYMPTLDEYLSNSLVTCGYALMTARSY
VARDDGIVTEDAFKWVATHPPIVKAACKILRLMDDIATHKEEQERGHIASSIECYRKETGASE
EEACKDFLKQVEDGWKVINQESLMPTDVPFPLLIPAISLARVSNTLYKDNDGYNHADKEVIG

1.3.1.2 The role of metal ion cofactors for TPSs activity

Terpene synthases need different metal cofactors. Manganese and magnesium are commonly found in monoterpene and sesquiterpene synthases. For diterpene synthases, magnesium is a metal cofactor (Frost and West, 1977; Hallahan and Croteau, 1988; Rutering et al., 2016; Schnee et al., 2002). The basic role of the metal ion is to neutralize the negative charge of the pyrophosphate moiety and to assist the ionization of the allylic diphosphate substrate (Lanznaster and Croteau, 1991; Vial et al., 1981). Köllner et al 2008 reported for a (*E*)-β-caryophyllene synthase of maize that K_m values were vastly different with different concentrations of Mg^{2+} (10 mM) and Mn^{2+} (0.25 mM) *in vitro*. Divalent metals are important for the activation of maize TPS1 which catalyzes the formation of (*E*)-β-farnesene, (*E*)-nerolidol, and (*E*,*E*)-farnesol after herbivore damage (Köllner et al., 2008).

The K_m value of TPS1 of maize changes up to 18-fold by using Mg^{2+} and Mn^{2+} as cofactors (Schnee et al., 2002). Mg^{2+} is utilized as a cofactor for the enzyme because the concentration of Mg^{2+} *in vivo* is higher (50 to 100-fold) than the one of Mn^{2+}. The enzyme (*4S*)-limonene synthase is a monoterpene cyclase from the glandular trichomes of peppermint. It showed at

the optimal concentration of Mn^{2+} higher rates of activity with Mg^{2+}. At a higher concentration of divalent metal ions (>10mM $MgCl_2$ and > 2mM $MnCl_2$) an inhibition of the enzymatic reaction occurred, and no synergistic effect was detected between these metal ions. In general, monoterpene synthases have a higher affinity to Mg^{2+} and lower affinity to Mn^{2+}(Biochemistry, 1987; Lewinsohn et al., 1992; Rajaonarivony et al., 1992). With different divalent metal ions like Ca^{2+}, Cd^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Ni^{2+} and Zn^{2+} at 0.5 mM, all as chloride salts and Fe^{2+} as sulfate, the catalytic activity was not greater than 5% (Rajaonarivony et al., 1992).

Multiple studies on terpene synthases revealed that the cyclase reaction proceeds via several nucleophilic and electrophilic reactions as an ionization dependent cyclization-isomerization mechanism. Different, stereo-chemically related monoterpene products are possible from one enzyme at the same catalytic site (Croteau RB, Wheeler CJ, Cane DE, Ebert R, 1987; Savage et al., 1994; Wagschal et al., 1991). Savage et al., 1994 found at a pH optimum of 7.8 for terpenoid cyclases from lodgepole pine (*Pinus contorta*) and Grand fir (*Abies grandis*), that some mono and divalent metal ion cofactors are effective for activation. Terpene synthases without metal ion cofactors are rare. The formation of cannabidiolic acid (CBDA) from cannabigerolic acid (CBGA) in *Cannabis sativa* L. does not require any metal ions as a cofactor (Taura et al., 1996). CBGA has no allylic diphosphate moiety, therefore, CBDA synthase has no requirement of a metal cofactor for CBDA formation.

The ‘Tyr’ residue in the active site is nearly hundred percent conserved across all types (monoterpene, sesquiterpene, diterpene, plant, fungal, and bacterial) of terpene synthases. Alteration or transition of this motif affects the enzymatic activity. Morehouse et al., 2017 mutagenized a limonene synthase from navel orange (*Citrus sinensis*). The ‘Tyr’ at position 565 on the J helix appears to undergo a transition from a position outside of the active site in the apo-(+)-limonene synthase (LS) structure to a position close to the substrate and metal-binding site. Mutation of this ‘Tyr’ to ‘Phe’ decreased the catalytic activity 50-fold (Morehouse et al., 2017). They observed that the protein folding is depending on the presence of hydrophobic amino acids. Increasing numbers of hydrophobic residues slower the folding rate. If water molecules are able to attach to amino acids, the protein folding is accelerated. However, the conformational flexibility of the amino acids has little influence on the folding rates.

1.3.1.3 Terpene synthases in chamomile

The identification and organ-specific gene expression of several TPS in chamomile has been studied recently by Irmisch et al., 2012. The characterization of an α-bisabolol synthase has also been investigated. This bisabolol synthase contains 530–560 amino acids. The signal peptide has 23 amino acid residues. The N-terminus has a direct impact on the activity and may also influence the stability and/or solubility of the enzyme (Son et al., 2014).

1.4 Biosynthesis of bisabolol and its oxides

All sesquiterpenes, several cyclic (β-caryophyllene, germacrene C etc.) and acyclic sesquiterpenes (β-farnesene) are derived from a central FPP molecule. The α-bisabolene is formed via nerolidyl diphosphate (NDP), a sesquiterpene analog of linalyl di-phposphate LDP. Ionization-dependent cyclizations produce several macrocyclic compounds like germacrene C and γ-humulene (Fig.1.5). They are formed by increasing the size of the farnesyl chain and methyl migration along with hydride shifts of initially formed cyclic carbocations. This results in a broad range of structures (Bohlmann et al., 1998).

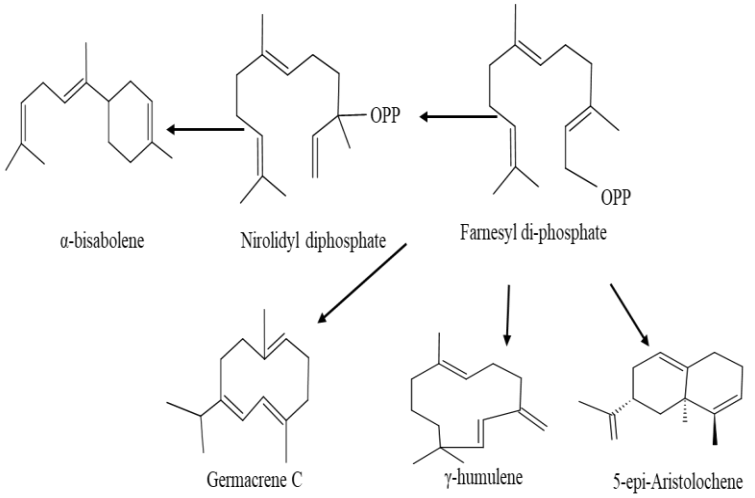


Fig. 1.5 Sesquiterpenes derived from farnesyl diphosphate
Sesquiterpenes are formed through preliminary isomerization of the trans-farnesyl pre-cursor to the sesquiterpene analog of LDP, i.e., nerolidyl diphosphate (NDP), followed by ionization-dependent cylization reactions. OPP: diphosphate molecule

Steam distillation of essential oil extracted from Brazilian candeia tree (*Eremanthus erythropappus*), sage (*Salvia runcinata*) and matricaria herbs serve as the major source for α-bisabolol (de Souza et al., 2008; H. Vuorela, Y. Holm, R. Hiltunen, T. Harvala, 1990).

However, sustainable and bio-controversial issues are arising as a major problem with the candeia tree (Han et al., 2016). α-bisabolol can also be chemically synthesized, but the chemical synthesis requires additional cost-effective purification steps due to the formation of unwanted by-products (Son et al., 2014). The essential oil extracted from flowers of chamomile contents several terpenes including with α-bisabolol and its oxides. During isolation of two oxides (bisabolol oxide A and bisabolol oxide B), a small quantity of crystalline substance was obtained and identified as α-bisabolol oxide C (Scilcher, H., Novotny, L., Ubik, K., Motl, O., Herout, 1976).

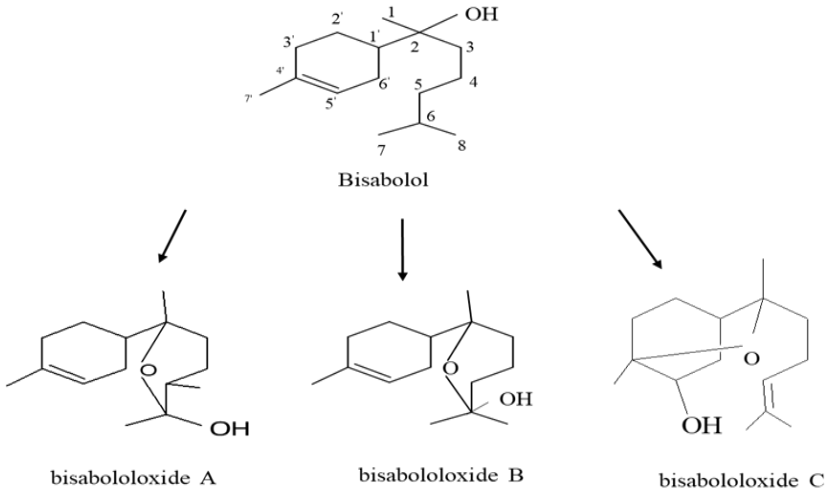


Fig. 1.6 Products from α-bisabolol oxidation.

1.5 Regulation of terpene synthases

1.5.1 Gene transcription can regulate TPS activity

Volatile terpenoids are often synthesized and emitted from specialized plant tissues. Gene expression studies of terpene synthases from different tissues often show a direct correlation of the mRNA level with the terpene products of the enzyme (Nagegowda, 2010). Some variation can occur in different parts of the plant, and sometimes the transcript level is too low to correlate it with the product. Crocoll et.al., 2010 found that ,monoterpene synthase activity of *Origanum vulgare* L. is predominantly regulated on the level of transcription and that the terpene synthase expression is directly correlated with the levels of essential oil (Crocoll et al., 2010a). However, the transcript level of sesquiterpene synthases among the tested plant lines (*Origanum vulgare* L) was too low to detect any correlations with their

products. The linalool synthase (*LIS*) of Basil is highly expressed in the glandular trichomes of an accession producing mostly (*R*)-linalool. The transcription of a geraniol synthase is the highest in accessions which produced geraniol. Similarly, sesquiterpene synthases such as germacrene synthase, γ-cadinene synthase and selinene synthase were highly expressed in leaves of different Basil accessions which showed the highest level of their respective products (Iijima, 2004). The *ZSS1* and *ZSS2*, which encode *Zingiber zerumbet* sesquiterpene synthases showed the highest expression in rhizomes and correlated with a higher content of α-humulene and eudesmaol in the essential oil (Yu et al., 2008). In *Clarkia brewery s*tigma, styles and petals emit linalool. They also showed a higher transcription of linalool synthase genes (Dudareva, 1996). Two monoterpenes, myrcene and ocimene, are emitted from the petals of snapdragon flowers. The rhythmic expression of their terpene synthase genes correlated with the emission (Dudareva, 2003). Furthermore, farnesene (*AdAFS1*) and germacrene synthases (*AdGDS1*) from kiwi fruit *(Actinidia deliciosa*) show a higher expression in flowers than in leaves. Within the floral tissues, m-RNA expression of both genes was the highest in petals and stamens (Nieuwenhuizen et al., 2009).

1.5.1.1 Regulation of TPS activity during development

The *TPS* transcript abundance and product formation can not only vary in different organs, but also during the development (Sangwan et al., 2001). Shimada et al., 2004 found that the expression of four different monoterpene synthases varies within tissues and developmental stages in citrus (*Citrus unshiu*). During the early stages, the expression of all four genes was mainly found in the peel. It disappeared or decreased in fruits in latter stages (Shimada et al., 2004). The *Cstps1* (valencene synthase) transcript abundance was high at fruit maturation and corresponds well with valencene accumulation in *Citrus sinensis* (Sharon-Asa et al., 2003). Higher levels of m-RNA transcription of *HlMTS1* (caryophyllene synthase) and *HlMTS2* (humulene synthase) in hops (*Humulus lupulus*) suggest that the sesquiterpenes caryophyllene and humulene accumulate in trichomes (Wang et al., 2008). Aharoni et al., 2004 described that the enzyme *FaNES1* is responsible for the formation of the monoterpenes nerolidol and linalool in strawberry fruit (*Fragaria ananassa*). *FaNES1* was highly expressed in fruits of different cultivated and wild strawberry lines (Aharoni, 2004). Two TPS genes, *LaLIMS* and *LaLINS* encoding limonene and linalool synthases, are highly expressed during flower development in *Lavandula angustifolia* (Guitton et al., 2010).

1.5.1.2 Regulation of TPS activity in chamomile

In chamomile plants, the transcript abundance of terpene synthases and their products in different plant organs has been described by Irmisch et al., 2012. Five *TPS* genes (*TPS1* to *TPS5*) were cloned from different plant organs. After heterologous gene expression several products, either sesquiterpenes or monoterpenes were formed. The major products were *(E)*-β-caryophyllene, Germacrene D, isocomene (α and β), β-elemene, *(E* or *Z)* and β-ocimene. *TPS3*, *TPS4* and *TPS5* were highly expressed in leaves and in flowers, while *TPS1* and *TPS2* were expressed in flowers and in roots, respectively. This suggests a spatial regulation of *TPS* gene expression qualitatively and quantitatively affecting terpene composition (Irmisch et al., 2012).

Aims of the study

The essential oil of German chamomile (*Matricaria recutita* L.) has significant commercial value for cosmetic, food and pharmaceutical industries. The components α-bisabolol and its oxides are the main interest of this study. The focus is the identification of the regulatory mechanism(s) controlling α-bisabolol production in a group of sixteen chamomile (*Matricaria recutita* L.*)* accessions. The study looks at the variation of terpenes in different plant parts and at different developmental stages of the accessions. We will isolate putative α-bisabolol synthase alleles from chamomile accessions. In addition, we characterize some of their biochemical properties (e.g. K_m), and their transcript levels in flowers. To identify the mechanism(s) responsible for the regulation of α-bisabolol production, we correlate these features with the *in vivo* farnesyl diphosphate (FPP) substrate levels and α-bisabolol product concentrations in flowers. These results will help to understand biosynthesis and regulation of terpene production in chamomile and will assist the design of trait-specific molecular markers for directed breeding strategies to obtain chamomile accessions with high content of α-bisabolol.

2 Materials and methods

2.1 Chemicals

Chemicals were purchased from Carl Roth GmbH (Karlsruhe), unless another source or vendor is given.

2.2 Plant material

Sixteen different accessions of Chamomile were included in the analysis (Table 2.2). For the purpose of this study, we define an origin as either a registered variety, a gene bank accession, a cultivated or wild population of chamomile. The term population applies to a wild natural population as well as to a cultivated one selected for breeding or agriculture. Genotype refers to a genetically individual plant.

Table 2.1 Plant growing conditions

Sl. No.	Parameters	Values
1.	Day temperature (13 h light)	20-22 °C
2.	Night temperature (11 h dark)	18-20 °C
3.	Relative humidity	55%
4.	Light intensity	1mmol m^{-2} s^{-1}

For each experiment three to five plants were used. From each plant disc flowers, ray flowers, leaves and stems were harvested separately. Plant material was frozen in liquid nitrogen and stored at -80 °C until grinding for further analysis.

Table 2.2 The list of accessions of plant chamomile

Sl. No.	Origin	Code	Geographic origin	Source
1.	Argenmilla	ARG	Argentina	Pharmaplant GmbH
2.	Bodegold	BOD	Germany	Pharmasaat GmbH
3.	Bohemia	BOHE	Czech Republic	Semo
4.	Bona	BON	Slovakia	Vet. med. University of Vienna
5.	Camoflora	CAM	Germany	Pharmaplant GmbH
6.	Germania	GER	Egypt	N.L. Chrestensen
7.	Goral	GO	Slovakia	Pharmasaat GmbH
8.	Lazur	LAZ	Bulgaria	Pharmaplant GmbH
9.	Lutea	LEU	Slovakia	Pharmaplant GmbH
10.	*M. discoidea*[*]	MO	Germany	Loki-Schmidt Genbank. Genbank Osnabrück
11.	Manzana	MAN	Germany	Pharmaplant GmbH
12.	Margaritar	MARG	Romania	Pharmaplant GmbH
13.	Pohorelicky Velkokvety[*]	MAT	Czech Republic	Genebank Gatersleben
14.	Pnos	PO	Polish	PONOS Poland
15.	Pop Germany[*]	MAR	Germany	Genebank Gatersleben
16.	Zloty Lan	ZL	Poland	Pharmasaat GmbH

[*]Samples were generated by Apomixis Research Group, Department Plant Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany, see also Otto et al., 2017.

2.3 Nucleic acid techniques

2.3.1 Extraction of RNA

Total RNA was extracted from 100–200 mg homogenized flowers (head/buds), of respective accessions, taken from three to five plants. The RNA was extracted according to the manufacturer's instructions with the "RNeasy Plant Mini Kit" (Qiagen, Hilden, Germany). Genomic DNA was removed with the RQ1 RNAse-free DNAse (Promega GmbH, Mannheim, Germany).

2.3.1.1 DNA (cDNA) synthesis

For use in qualitative and quantitative PCR, cDNA was generated with the "RevertAid First Strand cDNA Synthesis" kit (Thermo Scientific, St. Leon-Rot, Germany).

Table 2.3 cDNA synthesis

Sl. No.	Components	Quantity
1.	Total DNase-digested RNA	1.5 µg
2.	Oligo (dT) 18- Primers	1 µl
3.	Random hexamer primers	1 µl
Incubated at 65 ° C for 5 min and cooling over ice.		
1.	dNTPs and reaction buffer	According to the manufactures' protocol
2.	RiboLock RNase inhibitor	1 µl
3.	RevertAid H minus reverse transcriptase	2 µl
The Reaction mixture kept 5 min at 25 °C, 60 min at 45 °C and incubated at 70 ° C for 5 min.		

The cDNA was stored at -20 °C for further experiments.

2.3.1.2 DNA fragment amplification

PCR was used for fragment amplification from cDNA templates. Different polymerases were used as listed. The reactions were performed in a thermocycler, "Thermocycler MWG-Biotech primus 96 plus" (MWG-Biotech, Ebersbach, Germany) or a "peqSTAR Universal 96" (PeqLab, Erlangen, Germany). A list of oligonucleotides was mentioned in Table S1 in supplementary section.

Table 2.4 Polymerases used for PCR reactions

Sl. No.	Polymerases	Company
1.	Go Taq Polymerase (5 U/µl	Promega Corporation, Madison, WI, USA
2.	Advantage 2 Polymerase Mix	Clontech Laboratories Inc. Mountain View, CA, USA
Proof-reading polymerase		
1.	Phusion™ High-Fidelity DNA Polymerase	New England Biolabs Inc.
2.	Q5® High-Fidelity DNA Polymerase	New England Biolabs Inc.

Table 2.5 Reaction temperature for polymerases

Polymerase	Denaturation or Elongation temperature
GoTaq DNA-Polymerase	95 °C/72 °C
Advantage 2 Polymerase Mix	95 °C/68 °C
Phusion High-Fidelity DNA Polymerase	98 °C/72 °C
Q5 High-Fidelity DNA Polymerase	98 °C/72 °C

2.3.1.3 General conditions for PCR reactions

Table 2.6 Reaction mixture in a thermocycler

Components	Vol. to 25 µl	Vol. to 50 µl
DNA template (5-200 ng)	0.5-1 µl	1-2 µl
Polymerase buffer (5-10X)	2.5-5 µl	5-10 µl
DNA polymerase	0.25-0.5 µl	0.5 – 1 µl
Deoxynucleoside triphosphates (dNTPs) (10 mM)	0.25-1 µl	0.5 – 1 µl
Oligonucleotide (primer) forward (10 µM)	0.5-1 µl	1-2.5 µl
Oligonucleotide (primer) reverse (10 µM)	0.5-1 µl	1-2.5 µl
H_2O (Nuclease free)	Up to 25µl	Up to 50 µl

General amplification programme

	95 °C		2 min	Initial denaturation of DNA template
20-35x	95 °C		30 s	Denaturation of DNA template
	50-70 °C		30 s	Primer binding (Annealing)
	68-72 °C		1.5-10 min	Synthesis of complementary DNA strand
	68-72 °C		5 min	Completion of DNA synthesis
	Store the reaction mixture at 4 °C			

2.3.1.4 Agarose gel electrophoresis and isolation of DNA fragments

The agarose plates were made with 1.5% (w/v) agarose (Bio-Rad Laboratories, Hercules, USA) in 0.5x TAE buffer (40 mMTris-HCl, pH 8.0, 20 mM sodium acetate), 1µg / ml ethidium bromide was added. The amplified DNA sample (7-10 µl) was mixed with 3-5 µl of loading buffer (New England Biolabs (Ipswich, MA, USA) for the electrophoresis in a "Mupid^{TM}-One" (Biozym Scientific, HessischOldendorf, Germany) with a constant voltage of 100 Volt for 20-30 mins. The analysis of the DNA fragments was carried out under UV light (Syngene G: Box, Syngene, Cambridge, United Kingdom). 3 µl of a 1 kb or 100 bp DNA markers (Carl Roth GmbH, Karlsruhe, Germany) were used for the confirmation of the size of the DNA band of interest. DNA bands were cut and DNA extraction performed with the Nucleo Spin Extract Kit (Macherey-Nagel, Düren, Germany) according to the manufacturer's instructions.

2.3.1.5 Quantification of DNA and RNA

DNA and RNA were quantified using the 'NanoQuant infinite M200' (Tecan, Männedorf, Switzerland) by using the 'NanoQuantPlate^{TM}' (a quartz sample plate). Absorbance was measured at 260 – 280 nm.

2.3.1.6 Quantitative real-time PCR

CFX96 Real-Time System from Bio-RAD, Munich, Germany was used for transcript quantification.

Table 2.7 qPCR reaction mixture

Sl. no	Reaction components	Vol. for the reaction (µl)
1.	SYBR Green/ROX qPCR Master Mix (Thermo Scientific, Waltham, MA, USA)	10 µl
2.	Forward primer	0.5 µl
3.	Reverse primer	0.5 µl
4.	cDNA template (1:5 diluted)	5 µl
5.	PCR-grade water	to 25 µl

Table 2.8 Real time PCR conditions

	95 °C	10 min	Denaturation of DNA template
30-40x	95 °C	30 s	Denaturation of DNA template
	50-70 °C	30 s	Primer binding (Annealing)
	68-72 °C	40 s	Synthesis of complementary DNA strand (Extension)
After each run, a melting curve analysis from 60 °C to 95 °C was performed.			

For real time PCR, first primer specificity was confirmed by sequence analysis. The amplification plots were analyzed with the "BioRAD CFX manager" (Bio-Rad CFX Manager 3.1 (3.1.1517.0823) with cycle threshold (Ct) values. Actin was employed as a housekeeping gene. The relative quantification of the gene copy number in each cDNA sample was conducted using a standard curve. A standard curve with cDNA containing the respective gene was generated in a dilution series from 3 to 1/27-fold and three technical replicates were performed.

2.3.1.7 Restriction digest of DNA

Restriction digests of DNA fragments or plasmids were done using restriction enzymes from Fermentas (Thermo Fisher Scientific Inc., Waltham, MA, USA) or New England Bio-labs (Ipswich, MA, USA) according to a standard protocol.

Table 2.9 Standard protocol

Sl. No.	Components	Quantity
1.	Restriction enzyme	1-10 U
2.	10x enzyme buffer	10-30 µl
3.	DNA or PCR product	1-10 ng
The restriction mix was incubated in ice – water (3:2) for 18 h.		

2.3.1.8 Cloning Techniques (Ligation mixture)

Table 2.10 Reaction for ligation mixture

Sl. No	Content	Quantity
1.	Vector and DNA fragment	In 1:5 ratio
2.	T4 DNA ligase (Thermo Fisher Scientific Inc., Waltham, MA, USA)	1U
3.	10x ligase buffer	$1/5^{th}$ of total volume
4.	ATP	$1/5^{th}$ of total volume
The pCR4-TOPO sequencing vector (Invitrogen, Carlsbad, CA, USA) was used		

2.3.1.9 DNA Plasmid transformation and purification

Plasmids carrying the desired DNA fragment were transformed into *E. coli.* Positive *E. coli* clones were grown overnight at 37 °C and 220 rpm. The "NucleoSpin Plasmid Kit" (Macherey-Nagel, Düren, Germany) was used for the isolation of plasmid DNA according to the manufacturer's instructions.

2.3.1.10 Analysis of DNA and protein sequences

Plasmids were diluted and submitted according to the specifications of the companies. Sequencing of plasmids was performed by Eurofins MWG Operon (Ebersberg, Germany). DNA and protein sequences were analyzed with the software "BioEdit" (weblink Hall, T.A. 2004; Ibis Biosciences, Carlsbad, CA, USA) and 'Geneious prime' (Biomatters, Inc. Newark, USA). BLAST searches of isolated sequences were performed via the NCBI sequence service (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequence alignments were done with the free online version of "Multalin" (http://multalin.toulouse.inra.fr/multalin/). The protein

sequences identities and similarities were done with the online tools of http://imed.med.ucm.es/Tools/sias.html. For protein homology model Programme 'MOE' and YASARA 3.6.1 (http://www.yasara.org/changelog.htm) was used. For docking, Protein-Ligand ANT System (PLANT) was also used.

2.3.1.11 Overexpression of α-bisabolol synthases in bacteria (*E. coli*) and protein extraction

The expression vector pASK-IBA37+ was chosen, it contained a 6x His-tag and a tet-promotor. The bacteria clones including the insert were selected, cultured in 5 ml of Luria broth (LB) medium with 100 µg/ml ampicillin and grown at 37 °C for 15 h. 4 ml of starter culture were used to inoculate 100 ml of LB medium with ampicillin (100 µg/ml). Bacteria were grown at 37 °C and 220 rpm for 3-4 h until to an optical density (OD) of 0.5-0.8. After that an-hydro-tetracycline (final concentration: 200 µg/L) was added to the culture to continue for 20 h at18 °C, 220 rpm. Cells were harvested by centrifugation for 10 min at 4 °C and 5,000 x g. The pellet was resuspended in 3 ml extraction buffer.

Table 2.11 Composition for extraction and assay buffer

Extraction buffer	Assay Buffer
50 mM Tris-HCl pH 7.5 10 % glycerol 5 mM $MgCl_2$ 5 mM dithiothretol(DTT) 5 mM sodium ascorbate pH 7.0. 0.5 mM Phenyl methyl sulphonyl fluoride (PMSF)	10 mM Tris-HCl pH 7.5 10 % glycerol 1 mM DTT

The cell pellet was sonicated in a "Branson Sonifier 250" (Dietzenbach, Germany) for 3 x 30s (30s interval after each cycle) at 50 % of the maximum power. After centrifugation with 20,800 x g at 4 °C for 20 min the supernatant (crude protein extract) was transferred into an assay buffer by using 10DG columns (Bio-Rad Laboratories, Hercules, CA,USA) according to the instructions of the manufacturer.

2.3.1.12 Purification of α-bisabolol synthase

A large volume (1 Lit) of LB media was set up for the α-bisabolol synthase purification. The Profinity IMAC Ni-charged resin (Bio-Rad Laboratories, Hercules, CA, USA) column was used for elution of the protein solution. Before elution the column was packed with resin. 2 ml resin (column volume) was washed twice with wash buffer (20 ml) while slowly shaken over ice. The washed resin was incubated for 1 h with 6 ml of protein extract (section 2.3) at 4 °C with gentle shaking. The resin with the protein extract was transferred to a Poly-Prep Chromatography column (Bio-Rad Laboratories, Hercules, CA, USA). Initially, two washing steps were done with elution buffer, the next three steps were performed after checking the immediate Bradford reaction. Subsequently, six or more elution steps were performed and every fraction was collected in a separated tube and the concentration was measured with the method of Gill and von Hippel (Gill and Hippel, 1989).

Table 2.12 Buffers used for enzyme purification.

Lysis puffer	Wash buffer	Elution buffer
50 mM Tris-HCl, pH 8.0 300 mM NaCl 10 mM Imidazol, pH 8.0 1 % (v/v) Tween 20 10 % (v/v) Glycerol Final adjusted pH 8.0	50 mM Tris-HCl, pH 8.0 300 mM NaCl 20 mM Imidazol, pH 8.0 10 % (v/v) Glycerol Final adjusted pH 8.0	50 mM Tris-HCl, pH 8.0 300 mM NaCl 500 mM Imidazol, pH 8.0 10 % (v/v) Glycerol Final adjusted pH 8.0

2.4 Microbiological method

2.4.1 Medium

Table 2.13 Nutrient medium for bacterial culture.

LB-Medium	20 g LB-Medium in 1 L H_2O
LB-Agar	25 g LB-Agar in 1 L H_2O
SOC-Medium	2 % (w/v) Trypton (Sigma-Aldrich, St. Louis, USA)

	0.5 % (w/v) Yeast extract (Sigma-Aldrich) 0.5 % (w/v) Glucose 10 mM Sodium chloride 2.55 mM Calcium chloride. 21.6 mM Magnesium chloride 20 mM Magnesium sulphate

2.4.2 Antibiotics

Antibiotics were used to select positive bacterial clones.

Table 2.14 Antibiotics used for selective markers

Antibiotics	Conc.in the stock solution (mg/ml)	Conc. in the medium (µg/ml)
Ampicillin	100	100
Carbenicillin	50	50

2.4.3 Bacteria

Escherichia coli (*E. coli*) strain TOP10 (Invitrogen, Carlsbad, USA) was used. For overexpression experiments *E. coli* BL-21 was used.

2.4.4 Cultivation of Bacteria

E. coli cells were cultivated in LB medium at 37 °C and 220 rpm shaking or on LB agar plates at 37 °C in an incubator. For *E. coli* stock cultures, 500 µl culture were supplied with 150 µl sterile glycerol and stored at -80 °C. Incubators used were CERTOMAT IS Sartorius, Göttingen, Germany) and Heraeus, Function Line, Type T6, Thermo Scientific, Waltham, MA, USA.

2.4.5 Transformation of TOP 10 cells

1-10 µl of the ligation or plasmid mix and 50-80 µl of chemically competent *E. coli* cells were kept on ice for 30 mins. After a heat shock at 42 °C for 45 s cells were immediately cooled for 2 min on ice. The cells were mixed with 200 µl of SOC medium and incubated for 2.5-3 h at 37 °C, 220 rpm. An aliquot of 50-200 µl was plated on LB agar plates including with antibiotics. Positive clones were confirmed by the PCR with gene-specific or vector-specific primers.

2.5 Extraction procedures

SPME (100 µm polydimethylsiloxane; Supelco, Belafonte, PA, USA) and solvent (*n*-hexane extraction) methods were employed for the extraction of plant terpenes.

2.5.1 SPME for plant materials analysis

The finely powdered plant samples (leaves, stems, buds or flowers) of 80-100 mg were poured into GC glass vials. The SPME fibers were inserted through the cap (did not touch with samples) of the vials and kept over a water bath (37 °C) for 45 min hold with a stand. The fibers were inserted into the GC port for chromatographic analysis.

2.5.2 *n*-hexane extraction

For each sample, in liquid nitrogen finely ground powder of 80 to 100 mg material from 3-5 flower heads was used. To each sample 200–400 µl of *n*-hexane was added, followed by overnight incubation with shaking at 25 °C, 220 rpm. A total of 100 µl was transferred to a clean micro-tube fitted for GC vials. Terpene analysis was carried out with a gas chromatograph (GC 2010, Shimadzu, Duisburg, Germany).

2.6 Chromatographic Method

Hexane extraction and Analysis of chemical compounds by gas chromatography-mass spectrometry (GC-MS). The analysis by GC-MS was done with modifications according to (Irmisch et al., 2012; Schnee et al., 2004). Terpenes analysis was carried out with a gas chromatograph equipped with a splitless injector (injector temperature, 220 °C; injection volume, 1 µl) and coupled to a quadrupole mass selective detector (Shimadzu). Hydrogen was used as carrier gas at a flow rate of 1 ml min^{-1}. Samples were analyzed on a Supreme-5 ms column (30 ml ength × 0.25 mm inner diameter × 0.25 µm film thickness, Chromatographie Service GmbH, Germany). The coupled mass spectrometer was operated with a transfer line temperature of 230 °C, a source temperature of 230 °C, a quadrupole temperature of 150 °C, an ionization potential of 70 eV and a scan range of 50–350 amu. A flame ionization detector (FID) was used for the quantification of the compounds (µg/g of fresh weight) and operated at 250 °C. Compounds were identified by comparison of retention times and mass spectra to those of authentic reference substances obtained from SigmaAldrich (Steinheim, Germany), Roth (Karlsruhe, Germany) or by reference spectra in the NIST05, Wiley and Shimadzu libraries (Adams, 2007).

2.6.1 SPME: Enzyme activity reaction

The 70 µl of reaction mix (10 mM Tris-HCl pH 7.5, 10 % glycerol, 1 mM DTT, 60 µM FPP (Echelon Research Laboratories, Salt Lake City, USA), and 10 mM $MgCl_2$). 30 µl crude enzyme extract were used for the terpene synthase activity assay. The reaction volatile product was collected by a SPME kept over water bath at 42 °C for45 min. The fiber was inserted into the injection port of the GC-MS with the following parameters.

Table 2.15 GC-MS parameters

GC-MS parameter	Value or ranges
Transfer line temperature.	230 °C
Source temperature	230 °C
A quadrupole temperature	150 °C
An ionization potential	70 eV
Scan range	50–350 amu.

2.7 Kinetic study of α-bisabolol synthase

2.7.1 Determination the linear phase of the reactions

The enzyme substrate reaction time in the presence of metal ion was evaluated for the determination of kinetic study of the enzyme. The reaction mixtures as follows.

Table 2.16 Reaction mixture for enzyme assay

Sl. No.	Components	Reaction quantity
1.	Crude enzyme	20-30 µl (1800 ng)
2.	$MgCl_2$ (10mM)	10 mM
3.	FPP	3 µM
4.	Assay buffer	157µl

The reaction mixture was kept for 5-120 min. at 37 °C with 200 µl of *n*-hexane (previously added with nonyl acetate as internal standard) overlaid. The reaction was stopped after each time interval (5, 10, 20, 30, 40, 60 and 90 mins.) with vigorous vortexed of the reaction mix. Separated the organic layer and the peak integration evaluated with GC-FID. The product was identified with the retention time and the MS fragmentation spectra of the authentication standard (supplementary Fig. S9 and S10).

2.7.2 K_m values

The reaction mixtures as follows

Table 2.17 Reaction mixture for purified enzyme assay reactions

Sl. No.	Components	Reaction quantity
1.	Purified enzyme	20-30 µl (1800 ng)
2.	$MgCl_2$ (10mM)	10 mM
3.	FPP	0.1 - 30 µM
4.	Assay buffer	157 µl
The reaction mixture kept between 15-17 min (depending upon linear phase of terpene synthase) at 30 °C. All assays were overlaid with 1 ml *n*-hexane (previously added nonyl-acetate as internal std.) The reaction was stopped after two minutes with vigorous shaking of the reaction mixture. Using the micro-pipette organic layer was separated for GC-FID. All assays were performed in triplicate. The K_m values were determined by using the Lineweaver-Burk plot method.		

2.8 Determination of substrate concentration

2.8.1 Solvents used

Table 2.18 Solvents used for extraction of crude substrate of α-bisabolol synthase

Name of the solvent	Composition	pH and ratio
Organic extraction phase	Dichloromethane: ethanol (HPLC grade)	2:1 (v/v)
Polar extraction Buffer	50mM ammonium formate	Adjusted to pH 3 with formic acid (room temp)

2.8.2 Extraction procedures

Finely powdered frozen samples (120-150 mg) were taken in an eppendorf tube and two steel beads (diameter: 3 mm, previously cooled in liquid nitrogen) and 3/4th of silica beads were added into each tube and dropped immediately into liquid nitrogen. In each tube 900 µl of organic extraction buffer followed by 200 µl of polar extraction solvent and internal standard solution (sinigrin (Sigma Aldrich) 2 µg / 100 µl of solvent) were added and homogenized the samples with ball mill (Retsch MM 400, Retsch GmbH, Haan (DE) at 25 units in 3 x 30s (30s interval after each cycle). The polar extraction phase was separated into clean glass vials by centrifugation at 16000 rcf, 4 °C for 4.5 min and immediately frozen in liquid nitrogen. Samples were extracted again with 100 µl of polar extraction buffer and homogenized and centrifuged to separate a clear phase. The glass vials containing the polar buffer phase were lyophilized overnight (Christ Alpha 2-4 LD plus, Martin Christ freeze-drying GmbH, Harz (DE)) and used for LC-MS analysis after dissolving the powder into buffer A.
The substrate (FPP) in the enriched extract was directly analyzed by LC/MS/MS as follows

Table 2.19 LC-MS specifications

LC Pump A and B	Shimadzu LC30AD	Shimadzu Corporation, Kyoto (JPN)
LC Auto sampler	Shimadzu SIL30AC	
LC Column	Nucleoshell RP 18 HPLC column (150 mm × 2 mm × 2.7 µm)	Macherey Nagel, Düren (DE)
QTRAP LCMS	AB Sciex QTRAP 5500	AB Sciex, Foster City (USA)

Table 2.20 Specifications for chromatographic and spectrometric details (LC-MS)

Chromatographic details	
Solvent A	10 mM aqueous tributylamine (pH = 6.2, adjusted by acetic acid)
Solvent B	Acetonitrile
Solvent gradient flow	0 min: 2, 2 min: 2, 18 min: 36, 21 min: 95, 25.5 min: 95, 25.51 min: 2, 31 min: 2

Equilibration time	4 min
Injection volume	25 µL
Flow rate	0.4 mL/min
Column temperature	40 °C
Auto-sampler temperature	4 °C
Needle cleaning solvent	Methanol : water (1:1, v/v)

Mass spectrometric details	
Source temperature	450 °C
GS1/2	60/70 arbitrary units
CUR	40 arbitrary units
IS	-4500 V
Integrated Valco valve position (A to MS and B to waste)	0 min: A, 23 min: B, 27 min: A

2.9 Statistical Analysis

The significance of the stated values was determined using "One Way ANOVA" (Analysis of Variance). Software ‘R’ (Ri386 3.4.2.lnk) was used for the significance test. GraphPad Prism 8 was used for the non-linear kinetic analysis.

3 Results

3.1 Analysis of volatile metabolite diversity in chamomile accessions

3.1.1 Metabolite concentrations during seedling development

Seedlings of seven chamomile accessions (Bodegold, Bona, Germania, Lutea, Manzana, Pnos and Zloty Lan) were grown in a growth chamber between 10 and 40 days. The essential oil composition of these seedlings was analyzed by *n*-hexane extraction to monitor oil production in the young plant. In all evaluated accessions, the developmental stage of the seedlings had a significant effect on terpenes production. Figure 3.1 shows the concentrations of the most abundant terpenes and their distribution during 40 days of development. These compounds were the monoterpenes β-ocimene, citronellene as well as the sesquiterpenes β-caryophyllene, longipinene oxide, germacrene D, β-farnesene. In addition, hexacosane (a C-26, volatile oil component) and spiroether were abundant throughout the chamomile seedlings.

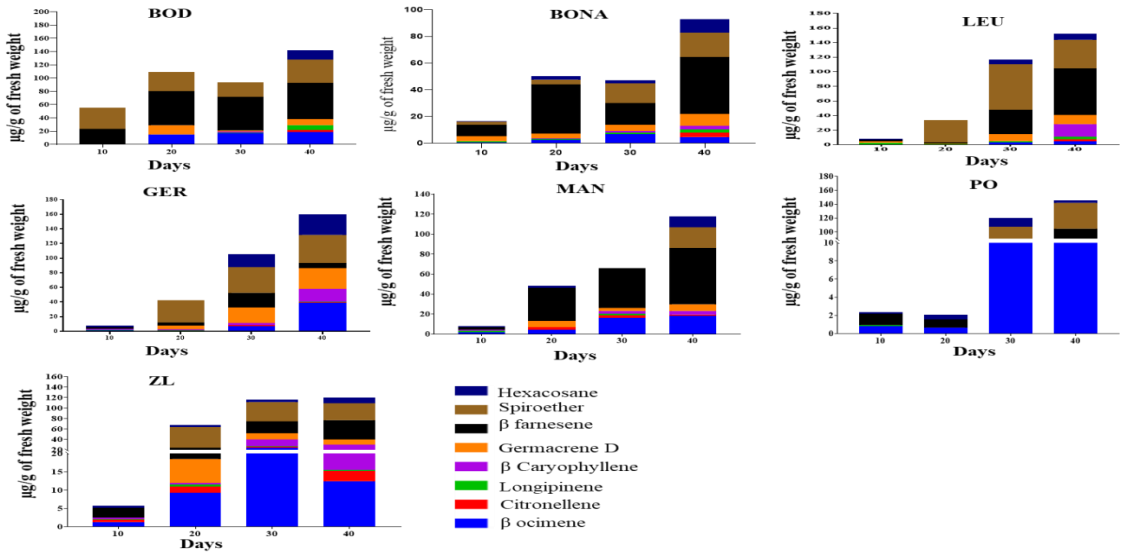


Fig. 3.1 Metabolite concentrations of chamomile accessions during seedling development.
The seedlings were grown in a growth chamber, samples were collected 10, 20, 30 and 40 days interval and analyzed by GC-MS. The accessions are: BOD: Bodegold, BON: Bona, GER: Germania, LEU: Lutea, MAN: Manzana, PO: Pnos and ZL: Zloty Lan. The three biological replicates were chosen for the analysis and average concentrations are shown in Fig. 3.1.

In general, all metabolites were lower in juvenile seedlings and increased with seedling age. The accessions Bodegold, Bona, Lutea and Manzana were dominated by β-farnesene and spiroether while the accessions Pnos and Zloty Lan contained large concentrations of β-ocimene. In Germania, both compounds β-farnesene and β-ocimene were formed in similar amounts at 40 days of age. The sesquiterpene, germacrene D was detected in high concentrations in the accessions Germania and Zloty Lan. Significant concentrations of the compound citronellene were also detected at 20 and 40 days of seedling development in Zloty Lan. However, a high variation of metabolite concentrations was observed among the accessions. β-farnesene and β-caryophyllene were the most highly produced metabolites in all accessions except Pnos. The result indicates that the age of the seedling had a significant effect on the production of volatile oil components in all tested accessions.

3.1.2 Metabolite concentrations in stems and leaves of adult chamomile plants

The metabolite composition of the stems and leaves was identified in approximately eight weeks old chamomile plants of thirteen accessions (Table 3.1). The plant tissues were freshly harvested, extracted with *n*-hexane and analyzed by both GC-MS and GC-FID. The concentration of the eleven most highly produced volatile metabolites are shown in Figure 3.2. Among these compounds are monoterpenes, sesquiterpenes, oxidative derivatives of terpenes, spiroether and En-In-dicycloether. Additional compounds such as flavonoids and di-terpenoids which were previously described in chamomile have not been included in this analysis.

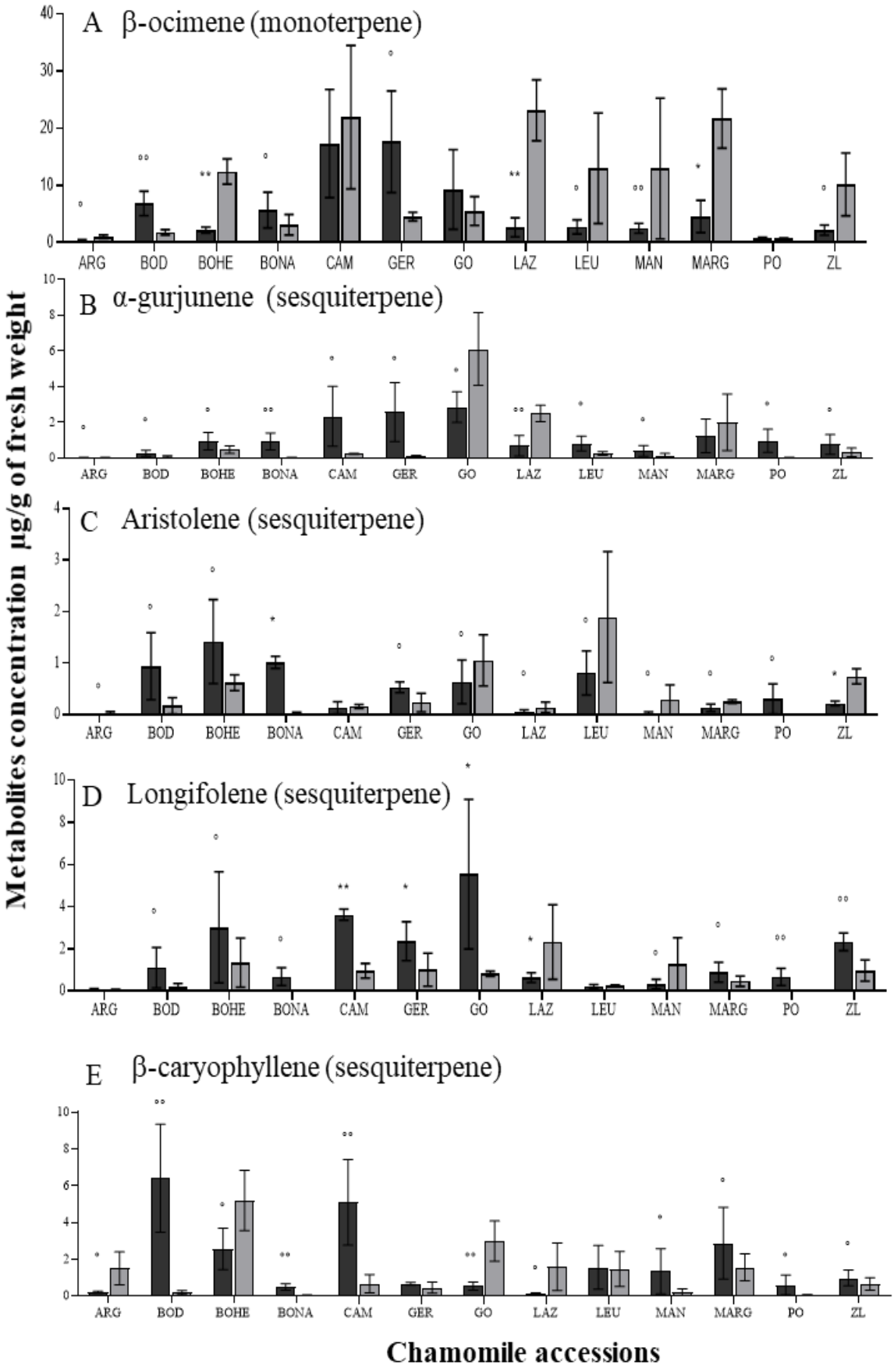
Table 3.1 List of chamomile accessions and their common metabolites in stem and leaf

Accessions	Argenmilla, Bodegold, Bohemia, Bona, Camoflora, Germania, Goral, Lazur, Lutea, Manzana, Margaritar, Pnos, Zloty Lan,
Monoterpenes	β-ocimene
Sesquiterpenes	α-gurjunene, Aristolene, Longifolene, β-caryophyllene,β-farnesene Germacrene D, Bi-cyclogermacrene, α-farnasene.
Others	Spiroether, En-In-dicycloether

The metabolite profile of stems and leaves was dominated by sesquiterpenes in this study. The concentration of the only major monoterpene β-ocimene, varies widely

among the accessions. Its' concentration is often higher in the leaves than in the stems, with the exception of Bodegold and Germania. Each of the eight sesquiterpenes has a unique expression profile and no co-regulation between compounds was observed among the accessions. The regulation of sesquiterpenes appears to be controlled independently in stems and leaves. While α-gurjunene, aristolene, longifolene were often found in higher concentrations in the stem, bi-cyclogermacrene, α and β farnesene, the most prominent sesquiterpenes in this analysis, were often found in higher concentrations in the leaves. Bona seemed to contain higher concentration of sesquiterpenes in the stem. The accessions Goral and Lazur appear to produce more sesquiterpenes in the leaves. Concentration of β-caryophyllene, β-farnesene and germacrene D was higher in stems than leaves within three accessions of Bodegold, Camoflora, and Manzana. There was a high variation of bicyclogermacrene noticed between stem and leaf within accessions Bohemia, Camoflora, Lazur, Margaritar and Zloty Lan. Presence of α-farnesene was higher in leaves compared with stems with exception of Bona, Germania, Lutea and Pnos, while high variation was noticed within stem and leaves in accession Zloty Lan.

Spiroether concentrations were generally high in the leaves except for Bodegold, Bohemia and Goral. In Lazur, Lutea and Margaritar, the spiroether concentrations were significantly lower in stems than leaves. The three compounds α-gurjunene, aristolene and longifolene varied most between stems and leaves in the accessions of Bohemia, Goral, Lazur and Lutea. En-in-dicycloether appears to vary more between the accessions. Leaves contain a higher concentration than stems with the exception of Bodegold, Bohemia, Bona and Zloty Lan. In general, variation of metabolite quantities is very high among the accessions.


A β-ocimene (monoterpene)
B α-gurjunene (sesquiterpene)
C Aristolene (sesquiterpene)
D Longifolene (sesquiterpene)
E β-caryophyllene (sesquiterpene)
Metabolites concentration µg/g of fresh weight
Chamomile accessions
ARG
BOD
BOHE
BONA
CAM
GER
GO
LAZ
LEU
MAN
MARG
PO
ZL

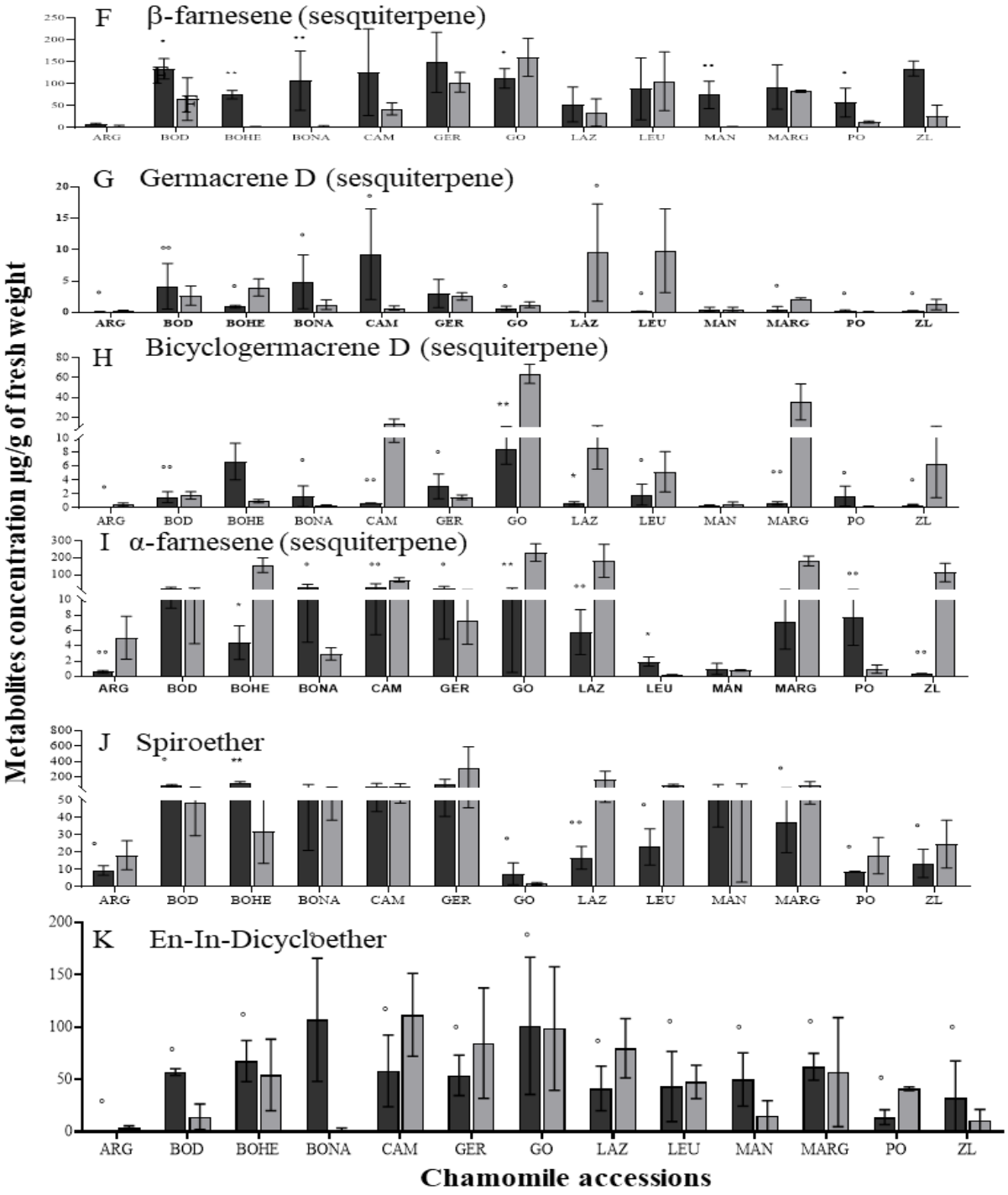


Fig. 3.2 Metabolite concentrations in stems and leaves of adult chamomile plants
In stems (black bar) and leaves (gray bar) of eight weeks old chamomile plants, metabolites were extracted with *n*-hexane, identified by GC-MS and quantified by GC-FID. The compounds were identified as A β-ocimene, B.α-gurjunene, C. aristolene, D. longifolene, E. β-caryophyllene, F. β-farnesene G. germacrene D, H. bicyclogermacrene, I. α-farnesene, J. spiroether, K. en-in-dicycloether. Error bar denoted as (±) SEM, n = 3 and p values were calculated within individual accession between stems and leaves by post ad hoc test where p °≤ 0.5, °° ≤ 0.1, * ≤ 0.05 and ** ≤ 0.01.

3.1.3 Metabolite concentrations in buds and flowers of adult chamomile plants

The chamomile flower is the source of the chamomile essential oil and therefore in the center of most efforts to determine the qualitative and quantitative terpene composition. In this study, the essential oil of buds and flowers was analyzed from fourteen accessions. An additional two accessions (MAR and MO) were included for the analysis from full flowers (no materials for buds were available). The essential oil was extracted with *n*-hexane from respective bud or flower, identified by GC-MS and quantified through FID (Fig.3.3). The three monoterpenes sabinene, β-myrcene andα-pinene were not found in all of the investigated accessions. Irrespective of the accession, sesquiterpenes were most abundant within the fifteen analyzed compounds. β-caryophyllene, α-farnesene, β-farnesene and α-bisabolol as well as its' oxides were most abundantly distributed among all accessions.

Table 3.2 Chamomile accessions and their common metabolites in bud and flower

Accessions	Argenmilla (ARG), Bodegold (BOD), Bohemia (BOHE), Bona (BON), Camoflora (CAM), Germania (GER), Goral (GO), Lazur (LAZ), Lutea (LEU), Manzana (MAN), Margaritar (MARG), Pohorelicky velkokvety (MAT), Pnos (PO), Zloty Lan (ZL),Pop Germany (MAR)*M. discoidea* (MO).
Monoterpenes	Cineole, β-Ocimene, γ-terpinene, sabinene, β-myrcene and α-pinene
Sesquiterpenes	Artemisia ketone, Ipsdienol, β-caryophyllene, β-farnesene, Germacrene D, Bi-cyclogermacrene, Germacrene A, α-farnesene, α-bisabolol and its' oxides.
Others	Spiroether, En-In-dicycloether

The volatile oil components of buds and flowers were dominated by sesquiterpenes in this study. The three monoterpenes cineole, β-ocimene and γ-terpinene were present within all investigated chamomile accessions, while a very low concentration of sabinene, β-myrcene, and α-pinene were found only in few accessions (Argenmilla, Camoflora, Germania, Pnos and Pop Germany (Supplementary Table S5). The concentrations of monoterpenes (especially cineole, β-ocimene and γ-terpinene) varied

within accessions and were often higher in the flowers. Even higher concentrations were found in buds of Bohemia, Bona, Pnos and Zloty Lan. There was a variation between individual accessions.

In six accessions (Bohemia, Camoflora, Germania, Manzana, Pnos and Zloty Lan), big differences of monoterpene contents were found between buds and flowers. In Goral, concentration of cineole and γ-terpinene were similar between buds and flowers and the same was noticed for β-ocimene in Pnos. In Argenmilla, the concentration of all three monoterpenes within buds and flowers was significantly lower than others accessions, while a lesser amount of cineole was found in Margaritar. A low concentration of β-ocimene was found in Lazur.

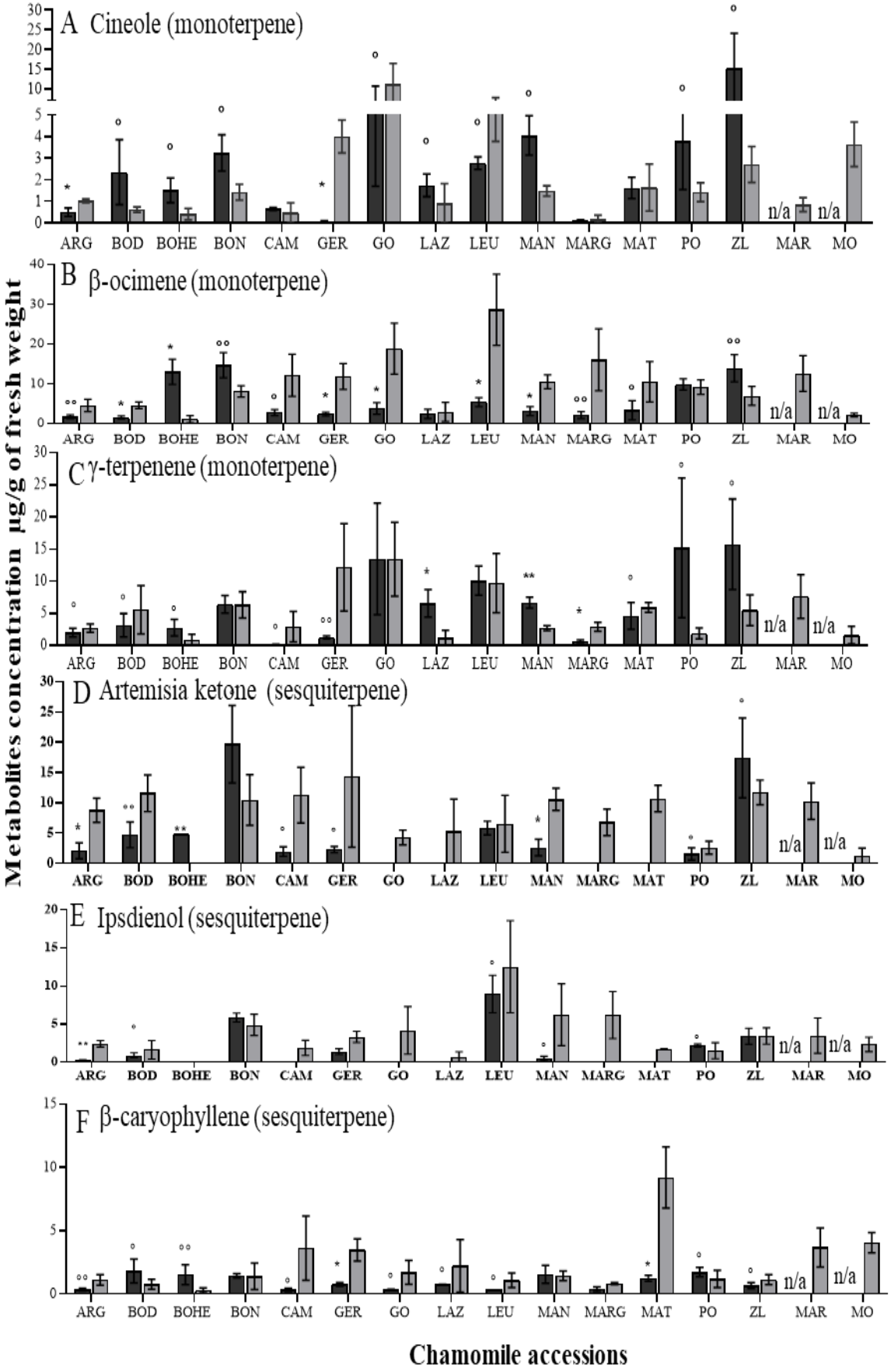
The abundance of eight sesquiterpenes (excluding bisabolol and its' oxides in Table 3.2) was higher in flowers compared to buds, while high variation was noticed within all accessions. The artemisia ketone, a sesquiterpene lactone derivative of chamomile volatile oil, was found in higher quantities in flowers than in buds with the exception of Bona and Zloty Lan.

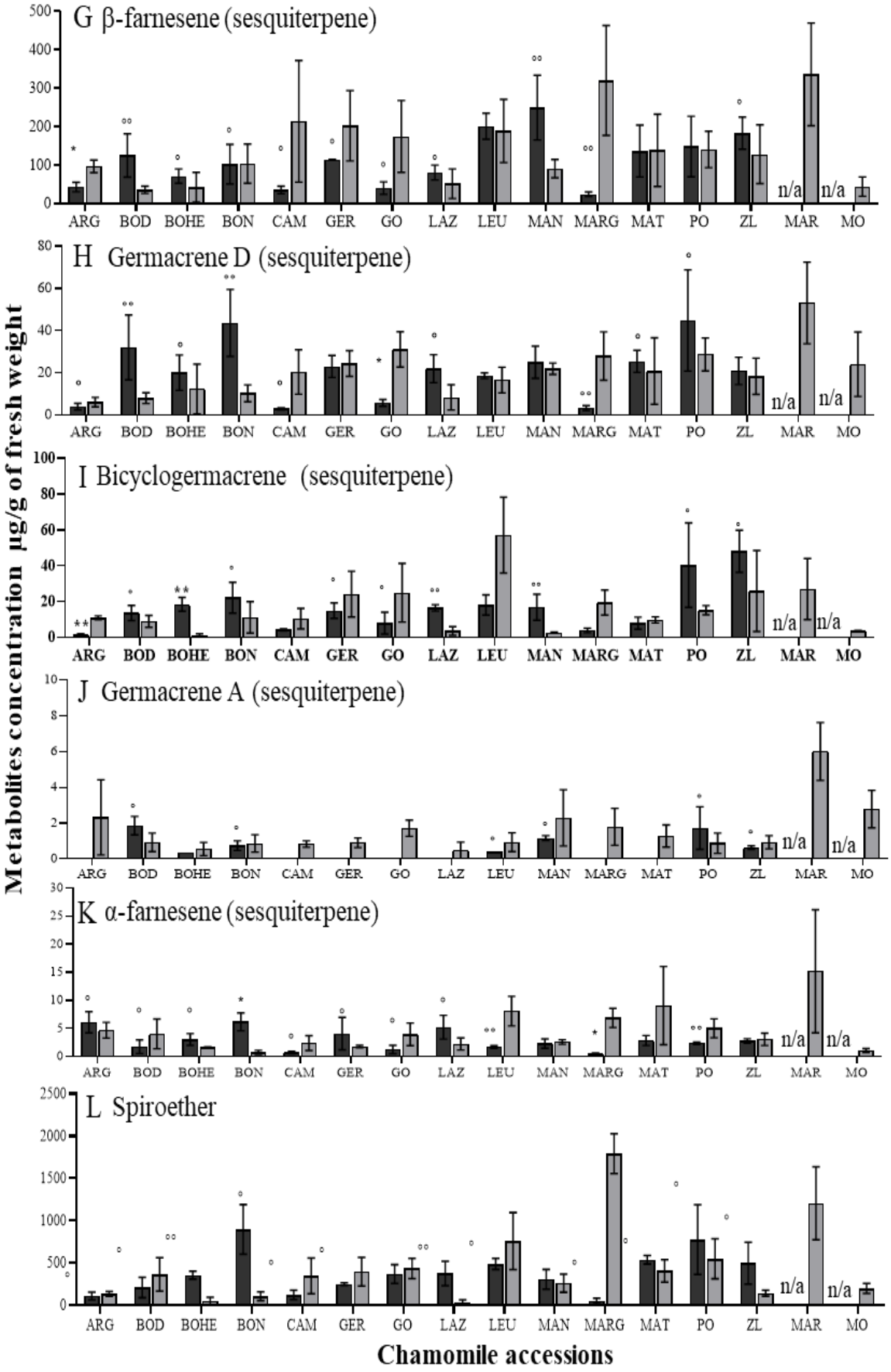
Additionally, the abundances of β-caryophyllene, germacrene D, bi-cyclogermacrene and β-farnesene were higher in buds of the four accessions Bodegold, Bona, Bohemia and Pnos. In Argenmilla and Margaritar the distribution of sesquiterpenes was highly varied between flowers and buds. Concentration of Ipsdienol was very low within all analyzed accessions and no product was found in buds of Bohemia, Camoflora, Goral, Lazur, Margaritar and in Pohorelicky velkokvety. The presence of β-caryophyllene within buds and flowers in the accessions of Bona, Manzana and Pnos was closely matched and the same was found in Lutea, MAT and in Pnos for β-farnesene. Additionally, distribution of germacrene D within buds and flowers was more uniform although variations were found within accessions. The flower contained a higher amount of germacrene A compared to buds with the exception of Bodegold and Pnos. The amount of α-farnesene within flowers and buds varied and its equal distribution was found in Argenmilla, Manzana and in Zloty Lan.

The concentration of the two ethers (spiroether and en-in-dicycloether) was higher in both buds and flowers compared to other monoterpenes and sesquiterpenes. Margaritar was the accession with the highest amount of these two ethers. Irrespective of

accessions, the flower contained higher concentrations of spiroether and en-in-dicycloether than buds.

The variation of bisabolol and its oxides (bisabolol oxide A and bisabolol oxide B) within the analyzed accession were high. The concentration of bisabolol oxide A was mostly higher in flowers than in buds with exception of Bohemia and Bona. We did not find bisabolone oxide A in the six accessions Bona, Camoflora, Germania, Goral, Lazur and Lutea in buds. Small amounts of bisabolol oxide B were found in Argenmilla, Goral and Pohorelicky velkokvety. The concentration of α-bisabolol was higher in flowers than in buds with the exception of Bohemia.


A Cineole (monoterpene)
B β-ocimene (monoterpene)
C γ-terpenene (monoterpene)
D Artemisia ketone (sesquiterpene)
E Ipsdienol (sesquiterpene)
F β-caryophyllene (sesquiterpene)
Metabolites concentration µg/g of fresh weight
ARG BOD BOHE BON CAM GER GO LAZ LEU MAN MARG MAT PO ZL MAR MO
n/a
Chamomile accessions


G β-farnesene (sesquiterpene)
H Germacrene D (sesquiterpene)
I Bicyclogermacrene (sesquiterpene)
J Germacrene A (sesquiterpene)
K α-farnesene (sesquiterpene)
L Spiroether
Metabolites concentration µg/g of fresh weight
ARG BOD BOHE BON CAM GER GO LAZ LEU MAN MARG MAT PO ZL MAR MO
n/a
Chamomile accessions

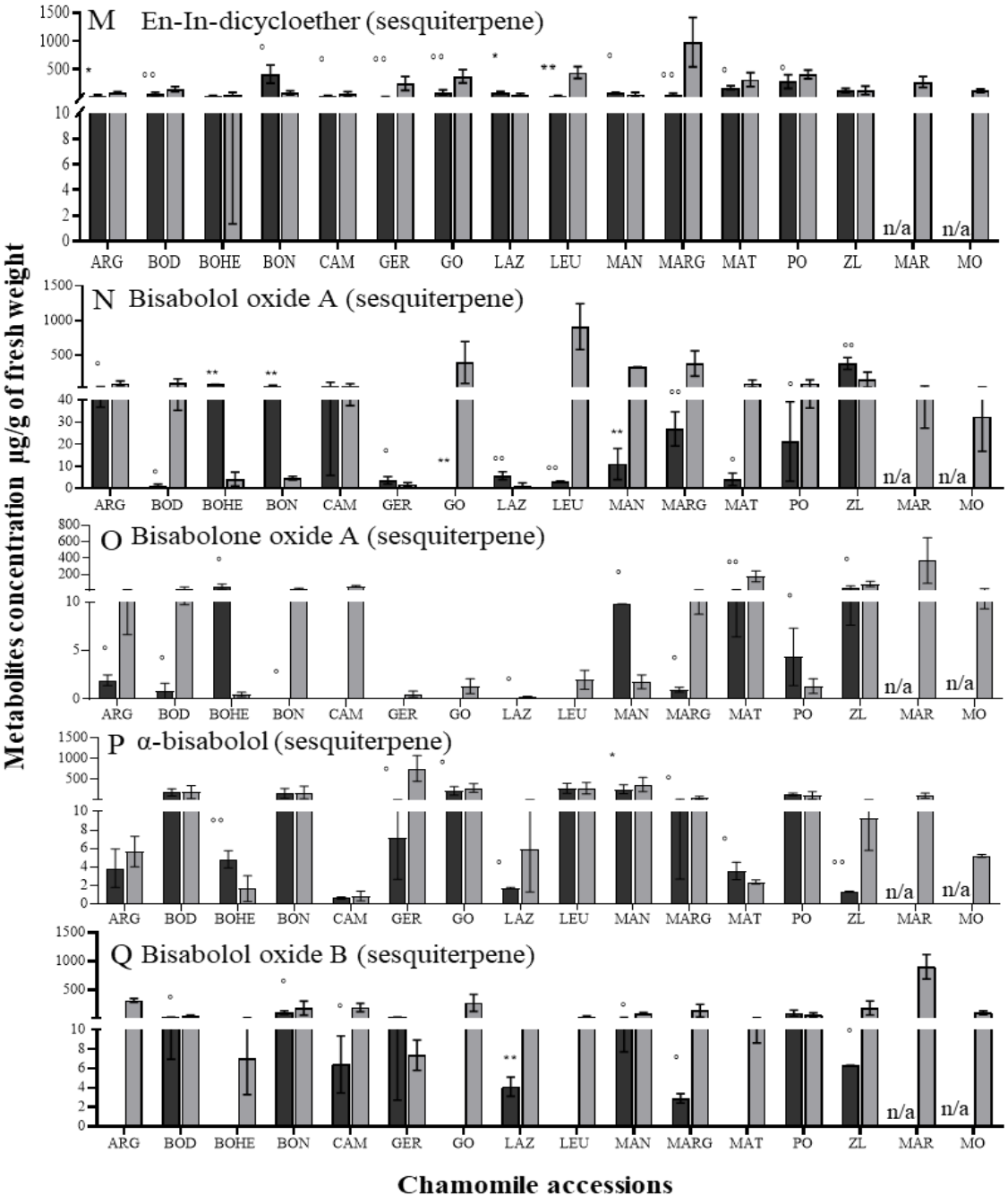


Fig. 3.3 Metabolite concentrations in buds and flowers of adult chamomile plants
In buds (black bar) and flowers (gray bar) of chamomile plants, metabolites were extracted with *n*-hexane, identified by GC-MS and quantified by GC-FID. The compounds were defined as A. cineole, B. β-ocimene, C. γ-terpinene D. artemisia ketone, E. ipsdienol, F. β-caryophyllene, G. β-farnesene, H. germacrene D, I. bi-cyclogermacrene, J. germacrene A, K. α-farnesene, L. spiroether, M. en-In-dicycloether, N. bisabolol oxide A, O. bisabolone oxide A, P. α-bisabolol, Q. bisabolol oxide B.
Error bar denoted as (±) SEM, n = 3, p values were calculated within individual accession between buds and flowers by post ad hoc test where p °≤ 0.5, °° ≤ 0.1, * ≤ 0.05 and ** ≤ 0.01. n/a: not available

3.2 α-bisabolol and its oxides are localized only in buds and flowers and their concentration varied with bisabolone oxide A

Chamomile buds and flowers were enriched with various volatile compounds. The presence of α-bisabolol was specially found in buds and flowers, mostly flowers contained a higher level of this compound than buds analyzed within sixteen chamomile accessions (in case of Pop Germany (MAR) and *M. discoidea* (MO) buds sample were unavailable). Germania containedthe highest concentration of α-bisabolol in flowers and Camoflora contained the least amount of α-bisabolol in buds. A close level of α-bisabolol was found in buds and flowers of five chamomile accessions Bodegold, Bona, LuteaPohorelicky velkokvety and Camoflora. A significant difference of concentration was found in Germania, Goral, Manzana, Margaritar and Zloty Lan within buds and flowers.

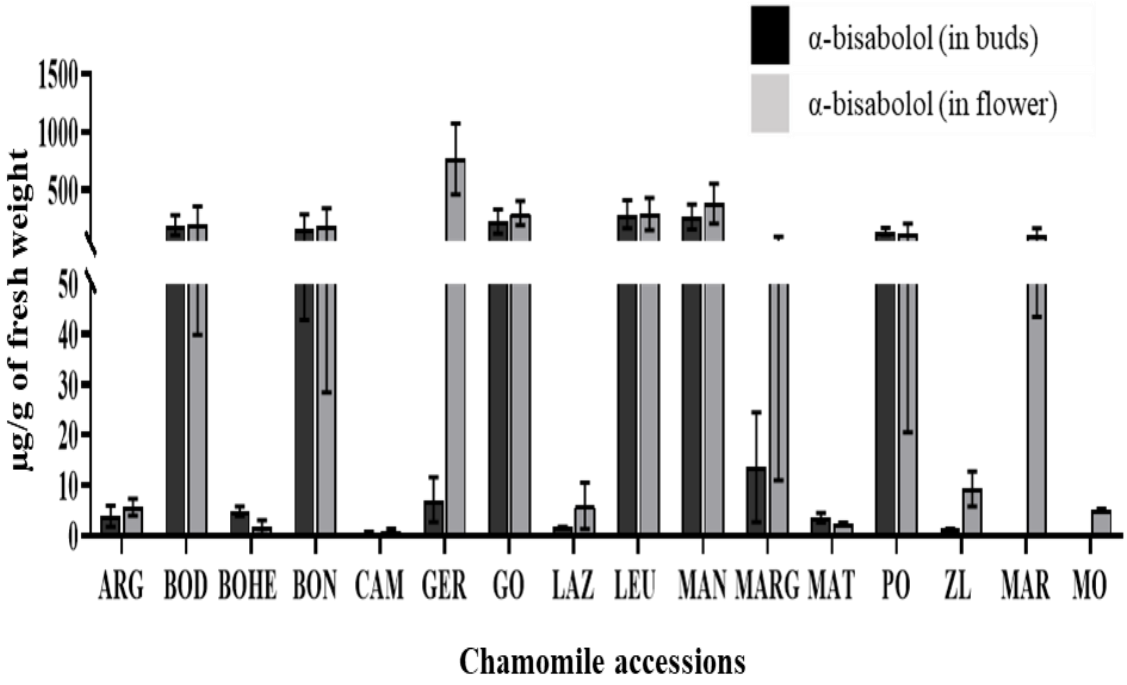


Fig. 3.4 Concentration of α-bisabolol in buds and flowers of chamomile accessions.
The α-bisabolol in chamomile flowers are abundant throughout the all analyzed accessions (fourteen for buds and sixteen for flowers). During the developmental stages, α-bisabolol and its oxides highly varied. Data showed Bodegold, Bona, Germania, Goral Lutea and Manzana contained high levels of bisabolol and α-bisabolol concentration was less with the accessions of Bohemia, Camoflora, MAT and *M. discoidea* in flowers (Fig. 3.3 N, O, P and Q).

Besides α-bisabolol, a significant quantity of bisabolone oxide A was found in flowers of all chamomile accessions (Fig. 3.5). In ten out of sixteen accessions, we found a higher concentration of α-bisabolol than bisabolone oxide A in flowers. There was a close relation between the presence of bisabolone oxide A and α-bisabolol in flowers. The concentration of bisabolone oxide A was relatively low when a significant high quantity of α-bisabolol was present. However, some exception was found with Argenmilla (ARG, Camoflora (CAM), Pohorelicky velkokvety (MAT), Zloty Lan (ZL) and *M. discoidea* (MO), may be due to slow oxidative reactions which can happen towards α-bisabolol formation in flowers compared with others chamomile accessions (Fig. 3.5).

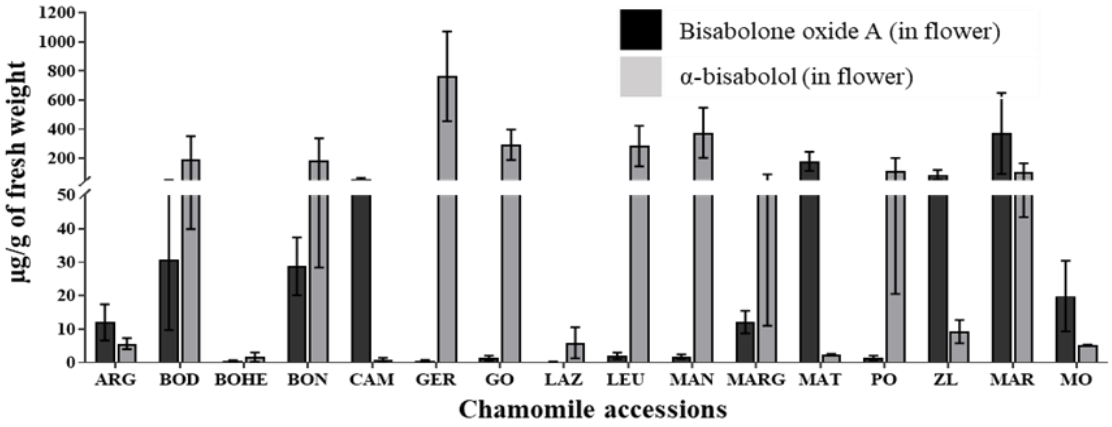


Fig. 3.5 Content of α-bisabolone oxide A and α-bisabolol in chamomile accessions.
Bisabolone oxide A (black bar) and α-bisabolol (gray bar). Metabolites of flowers from individual accession were extracted with *n*-hexane, identified by GC-MS and quantified by GC-FID. Error bars denoted as (±) SEM, n = 3

The GC-MS chromatograms of α-bisabolol were identified with retention time of authentic standard. We found a high concentration of α-bisabolol when bisabolone oxide A was lower within individual accession in flowers.

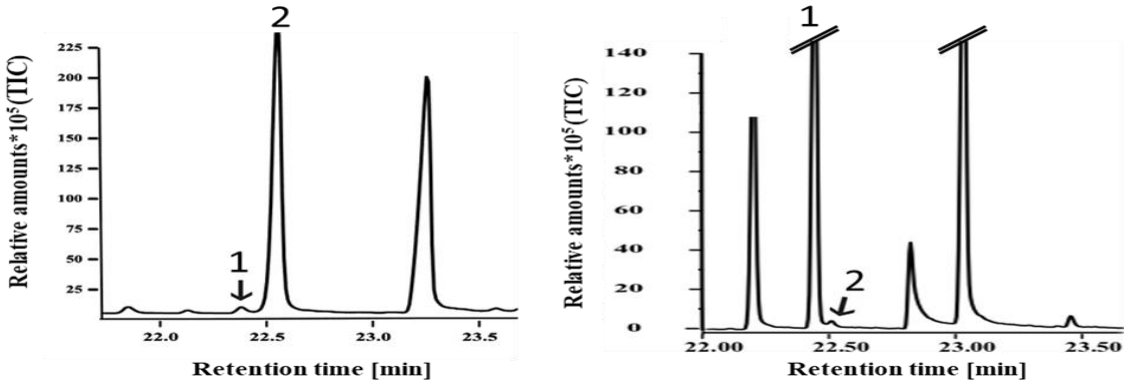


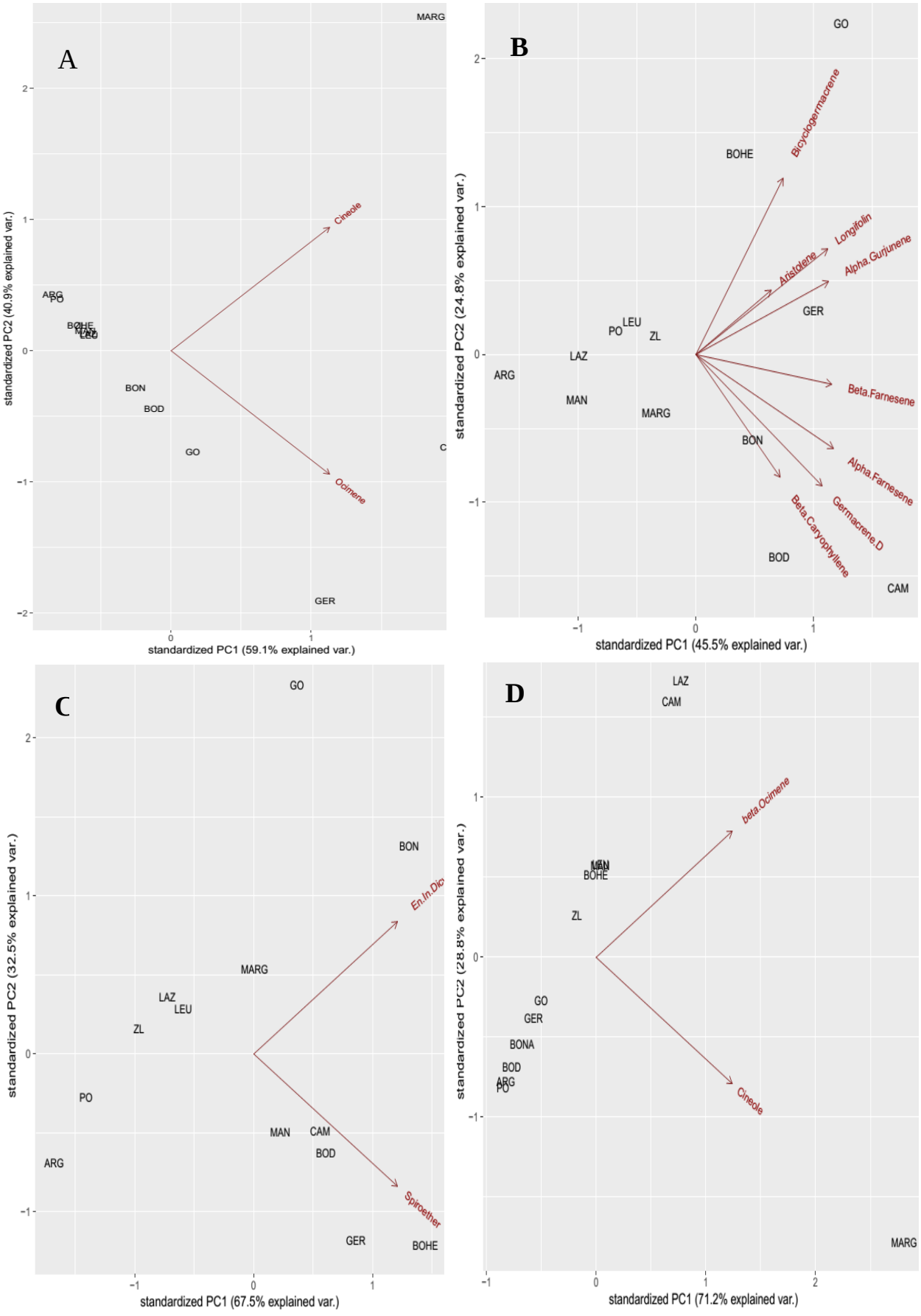
Fig. 3.6 Spectrum of bisabolone oxide A and α-bisabolol
n-hexane extraction was used for the GC-MS analysis and the products were identified as 1. α-bisabolone oxide A and 2. α-bisabolol

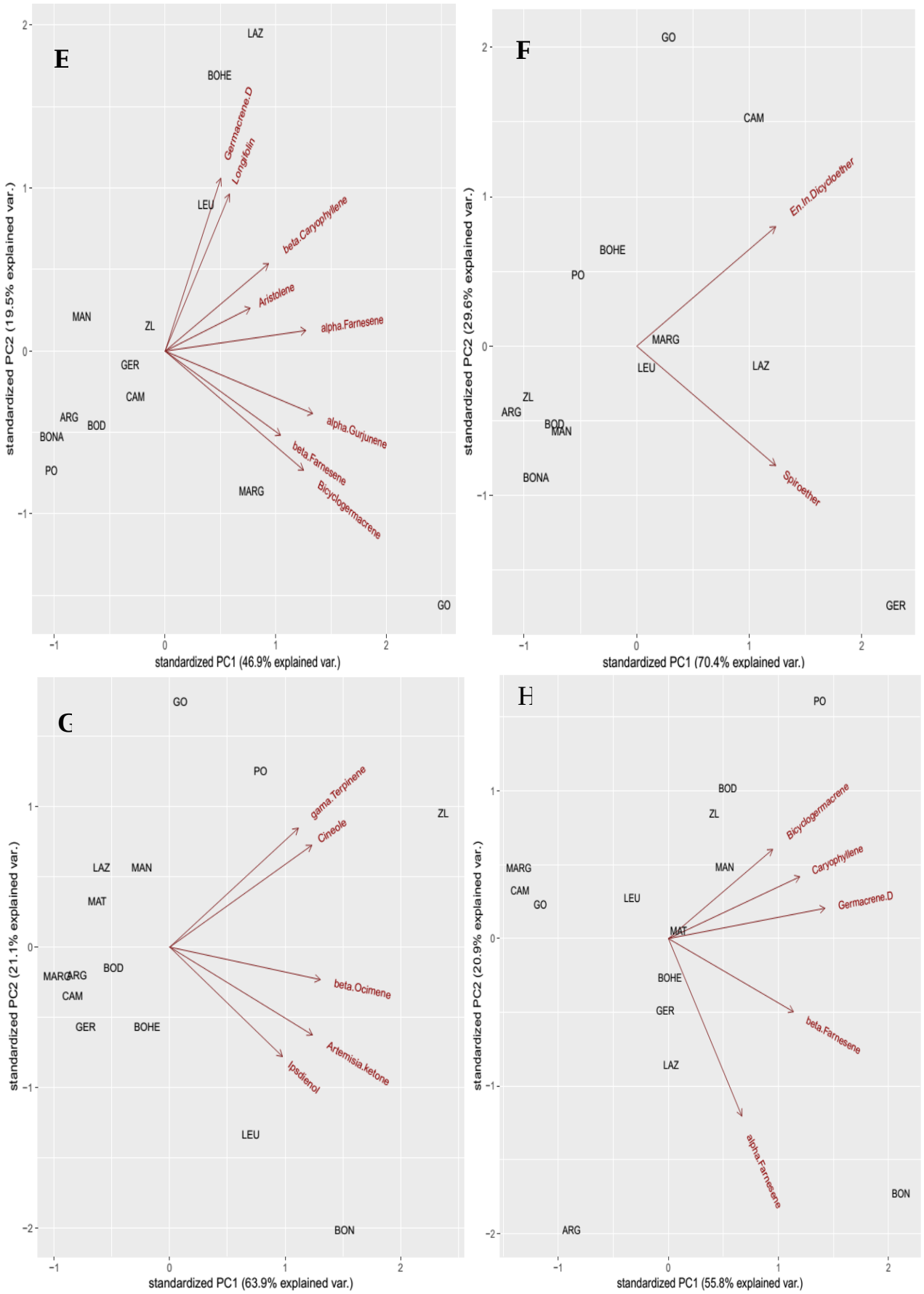
3.3 Principal component analysis (PCA) of terpene metabolites in chamomile accessions

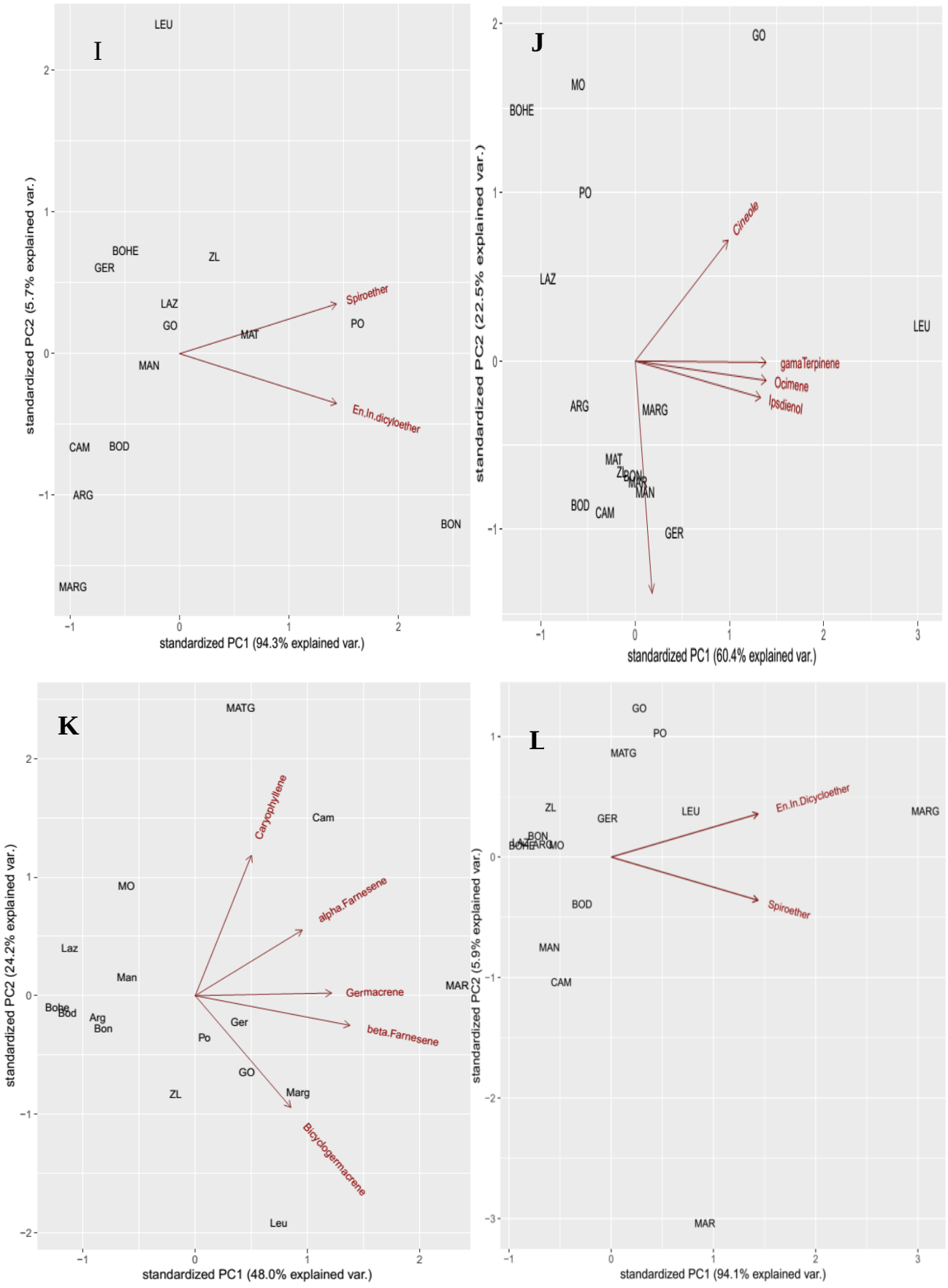
To determine group associations between the metabolite compositions of the essential oils of the chamomile accession, and to weigh relative contributions of compounds to the separation of the group, a principal component analysis (PCA) was conducted. The metabolites content of sixteen accessions was correlated with their distribution in stem, leaf, bud and flower through PCA analysis. A stepwise filtering procedure provided the best separation between the sample groups. The result is illustrated by the PCA score is shown in Fig.3.4 (Fig. 3.4 A, B, C: analysis of monoterpenes, sesquiterpenes and ethers from stem; D, E, F: analysis of monoterpenes, sesquiterpenes and ethers from leaves; G, H, I: buds' monoterpenes, sesquiterpenes and ethers; J, K, L: monoterpenes, sesquiterpenes and ethers from flowers and M, N: bisabolol and its' oxides of buds and flowers of chamomile accessions). In this step, PCA was used as a quality control tool to provide a visual representation of how the data cluster and how the closely related accessions vary with respect to the concentrations of terpenoids. Sample collection and handling techniques can also influence the chemical fingerprints especially for volatile oil constituents hence the chemical compositions of chamomile essential oils may vary dramatically for samples collected from different accessions.

In this analysis two monoterpenes, cineole and β-ocimene for stems and leaves and additional one monoterpene γ-terpinene were included for PCA analysis for buds and flowers. The scores of PCA1 are nearly 60% (stem), leaf (71%) bud (78%) and in full flower (76%) respectively. We found a cluster formation of Argenmilla with Pnos and Bohemia with Manzana in stem and leaves although no significant correlation was found within the others nine accessions. Additionally, PCA analysis of monoterpenes within fourteen accessions of buds and sixteen accessions of flowers did not form any strong cluster within the accessions, only between Germania and Manzana a cluster formation was observed (Fig. 3.4 G).

The abundance of eight sesquiterpenes (excluding bisabolol and its oxides) in all accessions was wildly distributed in analysis of stem to leaves, and bud to flower. The PCA1 scores with stem and leaves were nearly 47% (Fig. 3.4 B and E), while 50% with buds and 38 % with full flowers (Fig. 3.4 H and K). The accessions Pnos, Lutea, and Lazur displayed less difference in stems and leaves regarding the PCA1 scores than all other accessions. A scattered distribution of accessions in case of flower sesquiterpenes was found. However, five accessions Germania, Bodegold, Argenmilla, Manzana and Pnos were closely located near the baseline and strong clusters were found (Fig.3.4 K).


A
MARG
Cineole
Ocimene
BON
BOD
GO
GER
standardized PC2 (40.9% explained var.)
standardized PC1 (59.1% explained var.)
B
GO
BOHE
Bicyclogermacrene
Longifolin
Aristolene
Alpha.Gurjunene
GER
PO
LEU
ZL
LAZ
ARG
MAN
MARG
Beta.Farnesene
BON
Alpha.Farnesene
Germacrene.D
Beta.Caryophyllene
BOD
CAM
standardized PC2 (24.8% explained var.)
standardized PC1 (45.5% explained var.)
C
GO
BON
MARG
LAZ
LEU
ZL
PO
MAN
CAM
BOD
ARG
Spiroether
GER
BOHE
standardized PC2 (32.5% explained var.)
standardized PC1 (67.5% explained var.)
D
LAZ
CAM
beta.Ocimene
BOHE
ZL
GO
GER
BONA
BOD
ARG
PO
Cineole
MARG
standardized PC2 (28.8% explained var.)
standardized PC1 (71.2% explained var.)


E
LAZ
BOHE
Germacrene.D
Longifolin
LEU
beta.Caryophyllene
Aristolene
alpha.Farnesene
MAN
ZL
GER
CAM
ARG
BOD
BONA
alpha.Gurjunene
beta.Farnesene
PO
MARG
Bicyclogermacrene
GO
standardized PC2 (19.5% explained var.)
standardized PC1 (46.9% explained var.)
F
GO
CAM
En.In.Dicycloether
BOHE
PO
MARG
LEU
LAZ
ZL
ARG
BOD
MAN
BONA
Spiroether
GER
standardized PC2 (29.6% explained var.)
standardized PC1 (70.4% explained var.)
G
GO
PO
gama.Terpinene
ZL
Cineole
LAZ
MAN
MAT
MARG
ARG
BOD
CAM
beta.Ocimene
GER
BOHE
Artemisia.ketone
Ipsdienol
LEU
BON
standardized PC2 (21.1% explained var.)
standardized PC1 (63.9% explained var.)
H
PO
BOD
ZL
Bicyclogermacrene
MARG
MAN
Caryophyllene
CAM
GO
LEU
Germacrene.D
MAT
BOHE
GER
beta.Farnesene
LAZ
alpha.Farnesene
BON
ARG
standardized PC2 (20.9% explained var.)
standardized PC1 (55.8% explained var.)


I
LEU
BOHE
ZL
GER
LAZ
GO
MAT
MAN
Spiroether
PO
En.In.dicyloether
CAM
BOD
ARG
BON
MARG
standardized PC2 (5.7% explained var.)
standardized PC1 (94.3% explained var.)
J
GO
MO
BOHE
PO
Cineole
LAZ
LEU
gamaTerpinene
Ocimene
Ipsdienol
ARG
MARG
MAT
ZL
BON
MAN
BOD
CAM
GER
standardized PC2 (22.5% explained var.)
standardized PC1 (60.4% explained var.)
K
MATG
Caryophyllene
Cam
MO
alpha.Farnesene
Laz
Man
Germacrene
MAR
Bohe
Bod
Arg
Bon
Ger
beta.Farnesene
Po
GO
ZL
Marg
Bicyclogermacrene
Leu
standardized PC2 (24.2% explained var.)
standardized PC1 (48.0% explained var.)
L
GO
PO
MATG
ZL
GER
LEU
En.In.Dicycloether
MARG
BON
BOHE
MO
BOD
Spiroether
MAN
CAM
MAR
standardized PC2 (5.9% explained var.)
standardized PC1 (94.1% explained var.)

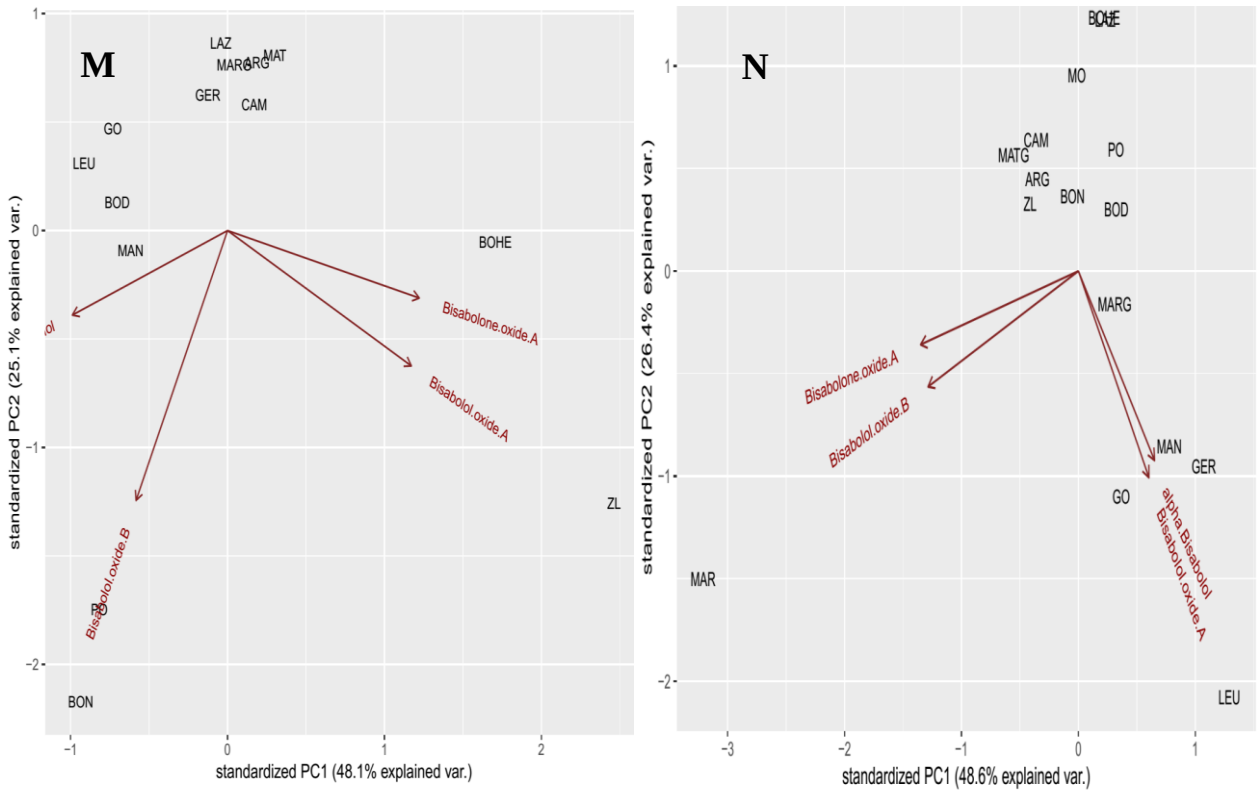


Fig. 3.7 PCA analysis of metabolities from stems, leaves, buds and flowers in sixteen chamomile accessions.

In stems:
A. PCA diagram of monoterpenes B. PCA diagram of sesquiterpenes and C. PCA diagram of ethers.
In leaves:
D. PCA of leaves monoterpenes, E. PCA of leaves sesquiterpenes and F. PCA diagram of leaves ethers. In buds:
G. PCA diagram of monoterpenes H. PCA diagram of sesquiterpenes and I. PCA diagram of ethers.
In flowers:
J. PCA diagram of monoterpenes, K. PCA diagram of sesquiterpenes and L. PCA diagram of ethers.
M. PCA analysis of bisabolol and its' oxides in buds.
N. PCA analysis of bisabolol and its' oxides in flowers.
For accessions abbreviation see Table 3.2.

There was a significant PCA1 score found with two ethers analyzed with chamomile accession in stem, leaves, buds and flowers. The PCA1 level was nearly 70% in stem and 94% when analyzed with buds and flowers. In stem, a cluster of Zloty Lan with Argenmilla was found, while Bodegold and Manzana formed a cluster in leaves and Bona, Bohemia, Lazur and *M. discoidea* formed a cluster in flowers. The concentrations of these two ethers were the highest in Margaritar within flowers and located separately

see Fig. 3.4 L. The PCA1 score with bisabolol and its oxides was nearly 50% corresponding to an overall medium level of correlation (Fig. 3.4 M and N). Additionally, there was a close correlation with Argenmilla, Lazur, Margaritar and MAT in buds (Fig. 3.4 M). A second correlation with Argenmilla, Camoflora, MAT and Zloty Lan in flowers was observed. Two accessions, Bohemia and Lazur formed a cluster when flowers were analyzed (Fig. 3.4 N).

3.4 Identification of putative α-bisabolol synthases (*MrTPS 7*) from chamomile accessions

The previous experiments demonstrated a high variability of metabolite production not only between chamomile accessions but also between developmental stages and plant organs. To understand the origin of this variability, we wanted to study the molecular mechanism underlying the production of these metabolites. We focus on a single sesquiterpene compound, α-bisabolol that is typical for chamomile and is associated with anti-inflammatory activity. A α-bisabolol synthase of chamomile was previously described as one of the key enzymes of α-bisabolol biosynthesis (Son et al., 2014.). We isolated the corresponding alleles of α-bisabolol synthases from cDNA of each of the accessions.

From twelve of the sixteen chamomile accessions, we identified a putative α-bisabolol synthase. While most of these genes displayed a high sequence similarity on the amino acid level, some accessions (BOD, MO, PO, and ZL) encoded alleles with a large degree of sequence variation (Table 3.3). All identified terpene synthases are displayed in the supplementary section. All amino acid sequences of the twelve accessions contained the sequence motives typical of terpene synthases, including the RRxR and the DDxxD motifs that are crucial for the structure-function relationship of an active terpene synthase.

Table 3.3 Similarity (%) of amino acid sequence of MrTPS 7 alleles of chamomile accessions

	ARG	BOHE	BOD	BON	CAM	GER	GO	LAZ	LEU	MAN	MAR	MARG	MAT	MO	PO	ZL
ARG	100															
BOHE	98.61	100														
BOD	98.44	98.05	100													
BON	62.13	62.13	62.18	100												
CAM	98.24	97.85	99.02	62.06	100											
GER	98.44	99.13	98.24	61.94	98.05	100										
GO	99.30	98.61	98.44	62.52	98.24	98.44	100									
LAZ	99.22	98.44	98.44	62.52	98.24	98.64	99.61	100								
LEU	98.95	98.26	98.63	62.33	98.44	98.09	99.30	99.41	100							
MAN	96.36	96.18	95.51	60.00	95.33	96.01	96.70	95.92	96.35	100						
MAR	98.78	99.48	98.24	61.94	98.05	99.3	98.78	98.64	98.43	96.4	100					
MARG	98.44	98.05	99.22	62.33	99.02	98.25	98.44	98.44	98.64	95.5	98.25	100				
MAT	98.78	98.44	98.24	61.55	98.05	98.26	99.13	98.64	98.78	96.5	98.61	98.25	100			
MO	10.22	10.22	10.13	10.29	10.11	10.22	9.87	10.09	9.89	9.70	10.05	10.09	9.87	100		
PO	74.78	74.6	81.28	61.55	80.93	74.78	75.13	81.16	75.3	72.86	74.78	80.97	74.95	10.78	100	
ZL	99.65	98.95	98.83	62.52	98.63	98.77	99.65	99.61	99.82	96.7	99.12	98.83	99.12	10.12	15.39	100

Three of the putative chamomile α-bisabolol synthases from the accessions GER, GO, and MAT were compared to other α-bisabolol synthases which were previously identified. An allele with a very high similarity of 98.95 % (to GER) was previously described as MrTPS1, a *Matricaria chamomilla* var. recutita terpene synthase1(Son et al., 2014b). However, most other α-bisabolol synthases only exhibited a very low level of similarity (Table 3.4). This suggests that α-bisabolol synthases cannot be easily identified by the similarity of their sequences.

Table 3.4 Similarity (%) of amino acid sequence of three selectied alleles of MrTPS7 with α-bisabolol synthases from other species
GER: Germania, GO: Goral, MAT: Pohorelicky velkokvety, MrTPS1 - *Matricaria chamomilla* var. recutita terpene synthase 1, AaBOS - *Artemisia annua* bisabolol synthase , AkBOS- *Artemisia kurramensis* bisabolol synthase, AmBOS - *Artemisia maritima* bisabolol synthase, EeBOS – *Eremanthus erythropappus* bisabolol synthase , PdBOS - *Phyla dulcis* bisabolol synthase, McBOS - *Matricaria recutita* bisabolol synthase, Mrbbs - *Matricaria chamomilla* var. recutita α-bisabolol synthase.

	Gene code	Gene Bank accession no.	In Percentage (%)										
			1	2	3	4	5	6	7	8	9	10	11
1	GER	--	100										
2	GO	--	98.26	100									
3	MAT	--	98.09	99.13	100								
4	MrTPS1	KJ020283.1	98.95	98.95	98.77	100							
5	AaBOS	AFV40969.1	5.53	5.23	5.53	5.53	100						
6	AkBOS	BAW34955.1	3.34	3.34	3.03	3.34	26.46	100					
7	AmBOS	BAW34954.1	4.58	4.58	4.58	4.58	27.07	31.19	100				
8	EeBOS	MH048990.1	3.63	3.63	3.81	3.49	4.61	5.47	6.11	100			
9	PdBOS	AFR23372.1	7.50	7.69	7.5	7.5	4.3	4.25	6.72	6.41	100		
10	McBOS	AIG92846.1	5.71	5.42	5.14	5.71	8.3	12.46	9.17	6.28	8	100	
11	Mrbbs	KM259907.1	4.85	4.85	4.85	4.89	5.84	4.86	4.58	3.96	4.21	5.42	100
			GER	GO	MAT	MrTPS1	AaBOS	AkBOS	AmBOS	EeBOS	PdBOS	McBOS	Mrbbs

3.5 Dendrogram analysis of α-bisabolol synthase

To visualize the relationships of the *MrTPS7* alleles from sixteen chamomile accessions a dendrogram analysis was performed. In addition, ten α-bisabolol synthases from other species were included. The analysis revealed several sub-groups that appear to share closely related alleles. The alignment (Fig.3.5) displayed a high similarity between BOHE, GER and MAR, which are fairly close to *Matricaria chamomilla* var. recutita terpene synthase 1, TPS1 (*Matricaria chamomilla* var. recutita sesquiterpene synthase 1) and McBOS (*Matricaria recutita* bisabolol synthase). The alleles of MAT, PO, CAM, MARG, BOD are somewhat related to the first group. A second group of alleles consists of BON, LAZ, GO, ZL, ARG, LEU, and Matricaria α-bisabolol synthase (*M recutita* α-bisabolol synthase).

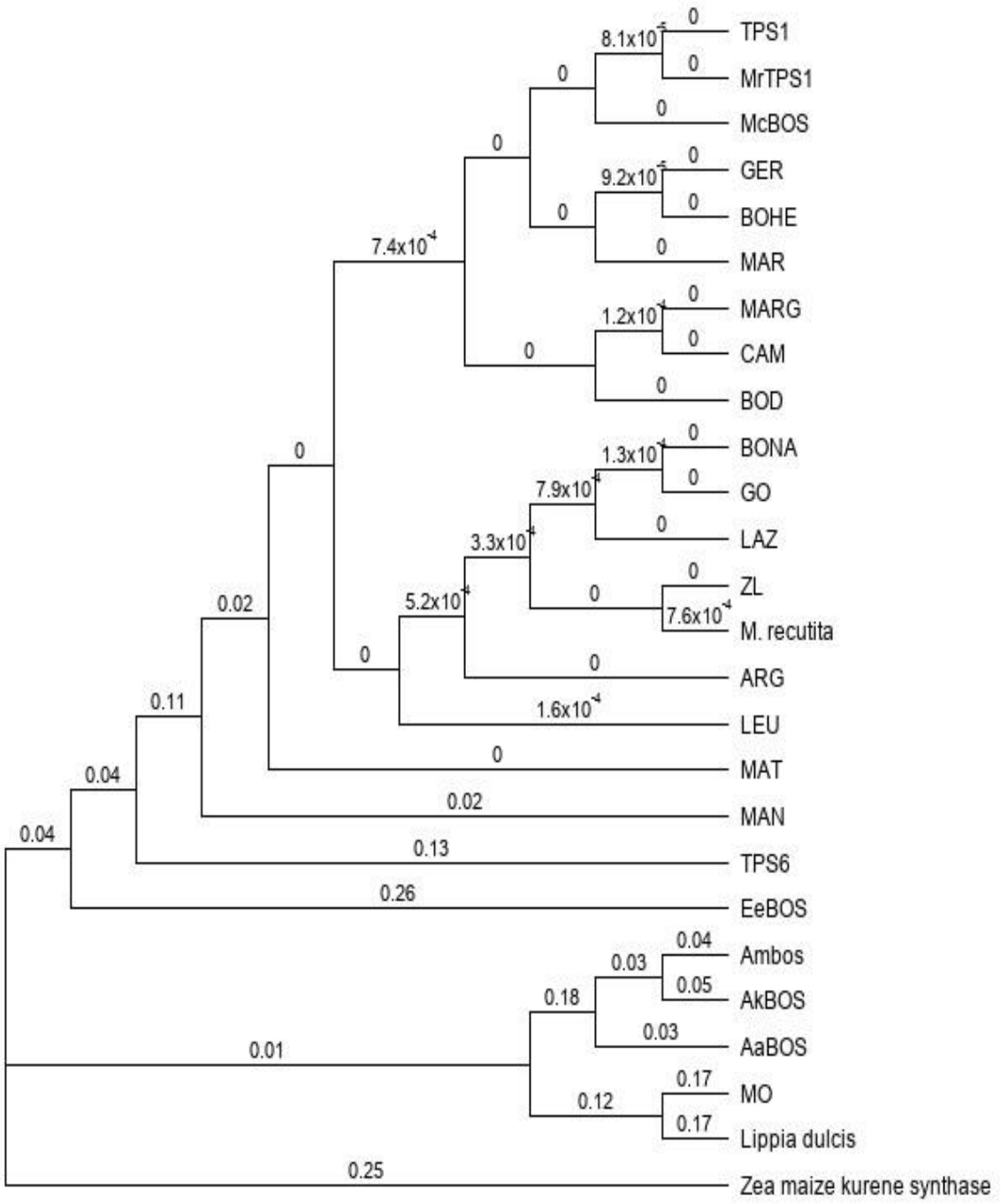


Fig. 3.8 Dendrogram of α-bisabolol synthases from chamomile and related species
The sequences were aligned and analysed by maximum likelihood method. Branch lengths are measured in the number of substitutions per site. Each branch is marked with the abbreviation of the species (in few the main substrate of enzyme). Cost matrix: >65% similarity and tree building method: neighbor joining was used. The tree was rooted on kaurene synthase of *Zea maize* (GenBank ID L37750.1). GenBank IDs in parentheses: *Lippia dulcis-Phyla dulcis* bisabolol synthase (AFR23372.1), AaBOS - *Artemisia annu*a bisabolol synthase (AFV40969.1), AkBOS - *Artemisia kurramensis* bisabolol synthase (BAW34955.1), AmBOS - *Artemisia maritima* bisabolol synthase (BAW34954.1), EeBOS–*Eremanthus erythropappus* bisabolol synthase MH048990.1), McBOS - *Matricaria recutita* bisabolol synthase AIG92846.1), MrTPS1 - *Matricaria chamomilla* var. recutita terpene synthase 1 (JQ255375.1), TPS 1 - *Matricaria chamomilla* var. recutita sesquiterpene synthase 1 (KJ020282.1), Mc recutita α-bisabolol synthase (AIW00681),TPS6 - *Matricaria chamomilla* var. recutita sesquiterpene synthase 6 (KJ020283.1).

Distant alleles are those of MAN and MO. Metabolites from MAN and MO have differences with others accessions. Presence of monoterpenes (cineole, β-ocimene and γ- terpinene) was correlated within MAN and MO although a high variation was noticed with β-ocimene. The presence of two ethers were relatively high within two accessions while eight sesquiterpenes distributed uniformly within these two accessions but a variation was noticed when compared with the other fourteen plant accessions.

The concentration of α-bisabolol and bisabolol oxide B was more or less same while bisabolol oxide A and bisabolone oxide A concentration varied within these two accessions. The α-bisabolol concentration in MO was lower compared with others accessions see Fig. 3.3 (P).The α-bisabolol synthase from MO is similar to several other plants including AaBOS - *Artemisia annua* bisabolol synthase, AkBOS- *Artemisia kurramensis* bisabolol synthase, AmBOS - *Artemisia maritima* bisabolol synthase, EeBOS - *Eremanthus erythropappus s*and displayed less than 20% sequence similarity to those of Matricaria (Table 3.3, Fig 3.8).

3.6 The transcript levels of α-bisabolol synthases of the chamomile accessions do not correlate with α-bisabolol production

Terpene production in plants is often regulated by transcriptional regulation (Crocoll et al., 2010b). To test whether the α-bisabolol content of the chamomile accessions corresponds to the transcript levels of the respective α-bisabolol synthases, we conducted quantitative real-time PCR (qRT-PCR) analysis. We investigated the flowers of sixteen accessions since this organ contains the maximum quantity of α-bisabolol. Transcript levels were determined with primers against the respective alleles of the α-bisabolol synthase. Although each of the alleles are transcribed in their respective accessions, the transcript levels differed significantly. Germania was the accession with the highest transcript abundance of *MrTPS7*alleles, while Lutea (LEU), Margaritar (MARG) and *M. discoidea* (MO) had medium transcription levels and Pop Germany (MAR) had the lowest transcript abundance. Little difference of transcription was noticed within the accessions of Argenmilla (ARG), Bona (BON), Camoflora (CAM),

Goral (GO) and Pnos (PO), while Bodegold (BOD), Bohemia (BOHE) and Zloty Lan (ZL) had very similar transcription levels (Fig. 3.9).

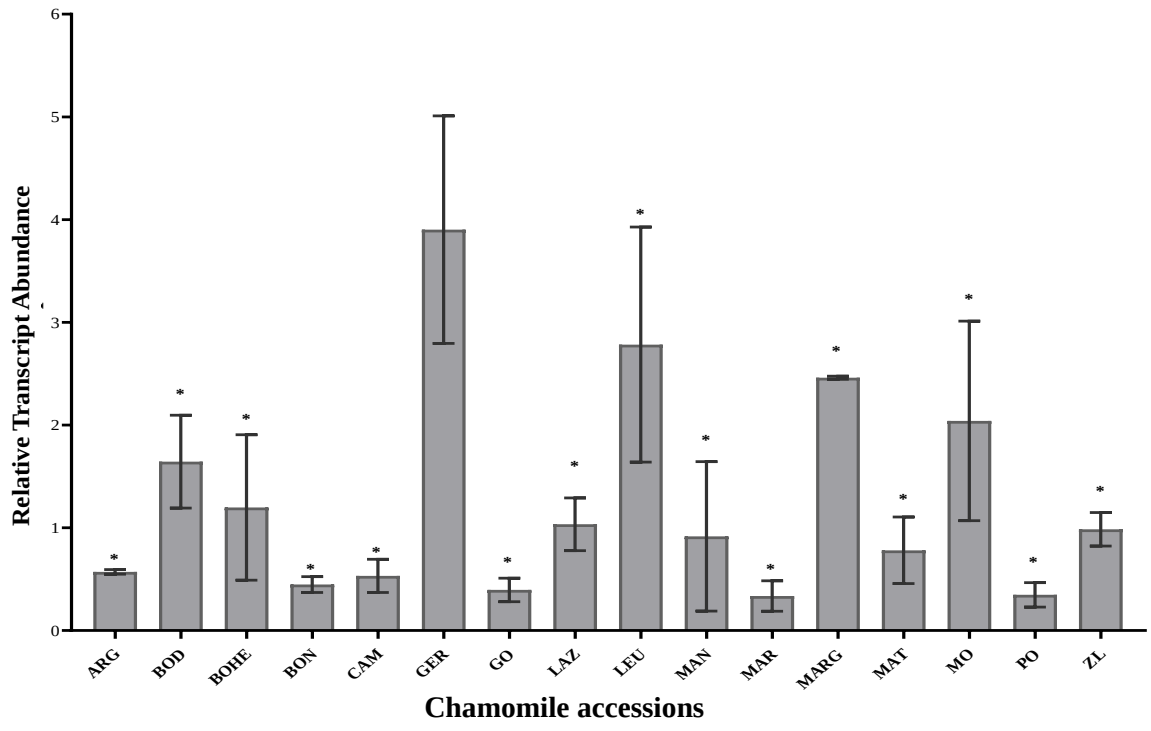


Fig. 3.9 Relative transcription of putative α-bisabolol synthases in flowers of chamomile accessions. The reference gene used was actin each bar represents the average of at least three biological replicates ± SEM (post ad hoc test was performed where * $p < 0.05$ respect with GER)

If α-bisabolol production were regulated on the level of transcription, we would expect a correlation between the transcript levels and the α-bisabolol production in flowers. However, we found no significant correlation between terpene synthase (*MrTPS7* alleles) transcript levels and the respective α-bisabolol amounts in flower. The nonparametric version of Spearman's co-relation coefficient (r_s) test showed only a very low correlation ($r_s = 0.11$) with our dataset (Fig. 3.10). Germania had the highest α-bisabolol concentration and the highest level of gene transcription. Variation of α-bisabolol concentration within few accessions *M. discoidea* (4), Bohemia (6), Lazur (7), Zloty Lan (8), MAT (10), Argenmilla (11) and Camoflora (12) were small. The lowest transcription was noticed with MAR (16). Accession, Camoflora (12) contained the lowest amount of α-bisabolol. However, significant concentration differences were found within all sixteen accessions.

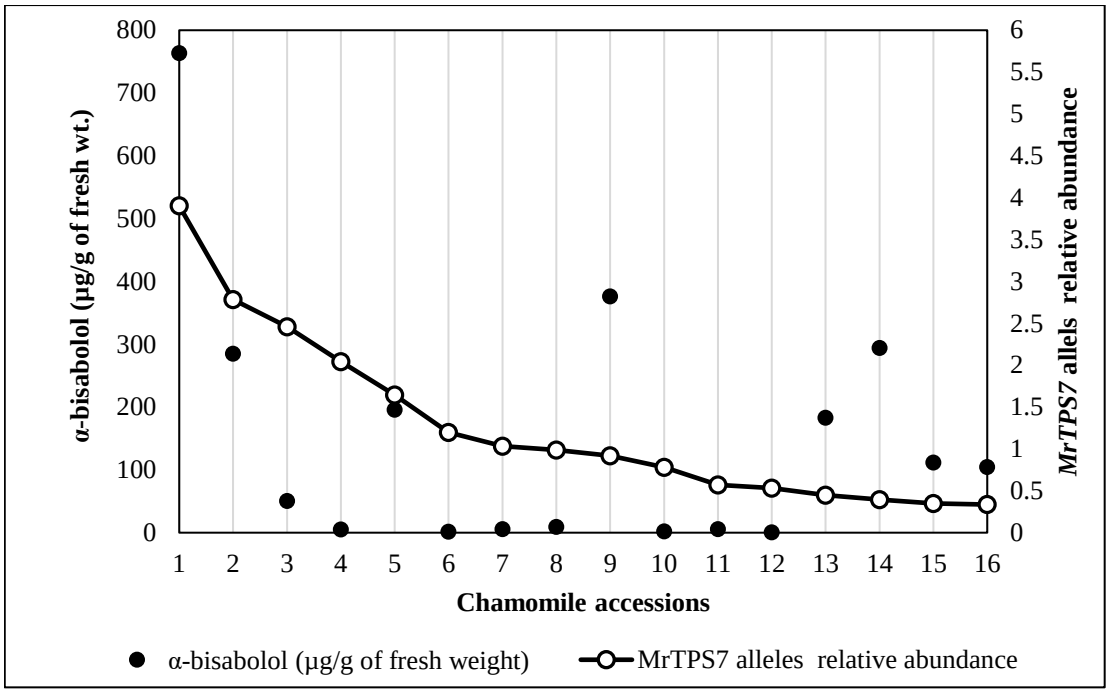


Fig. 3.10 Correlation between *MrTPS7* alleles relative transcript levels and α-bisabolol concentrations in the flower.
The nonparametric version of Spearman's co-relation coefficient (r_s) test was calculated between the relative transcript levels and the content of product. The chamomile accessions are: 1 GER, 2. LEU, 3. MARG, 4. MO, 5. BOD, 6. BOHE, 7. LAZ, 8. ZL, 9. MAN, 10. MAT, 11. ARG, 12. CAM, 13. BON, 14. GO, 15. PO, 16. MAR.
Spearman's correlation (r_s) = 0.11

3.7 Biochemical characterization of putative α-bisabolol synthases from chamomile accessions

For the biochemical characterization of the MrTPS alleles, we expressed twelve putative α-bisabolol synthases in a bacterial expression system and performed terpene synthase assays with the extracted crude enzymes. After incubation with the FPP substrate, the products were analyzed by SPME in GC-MS.

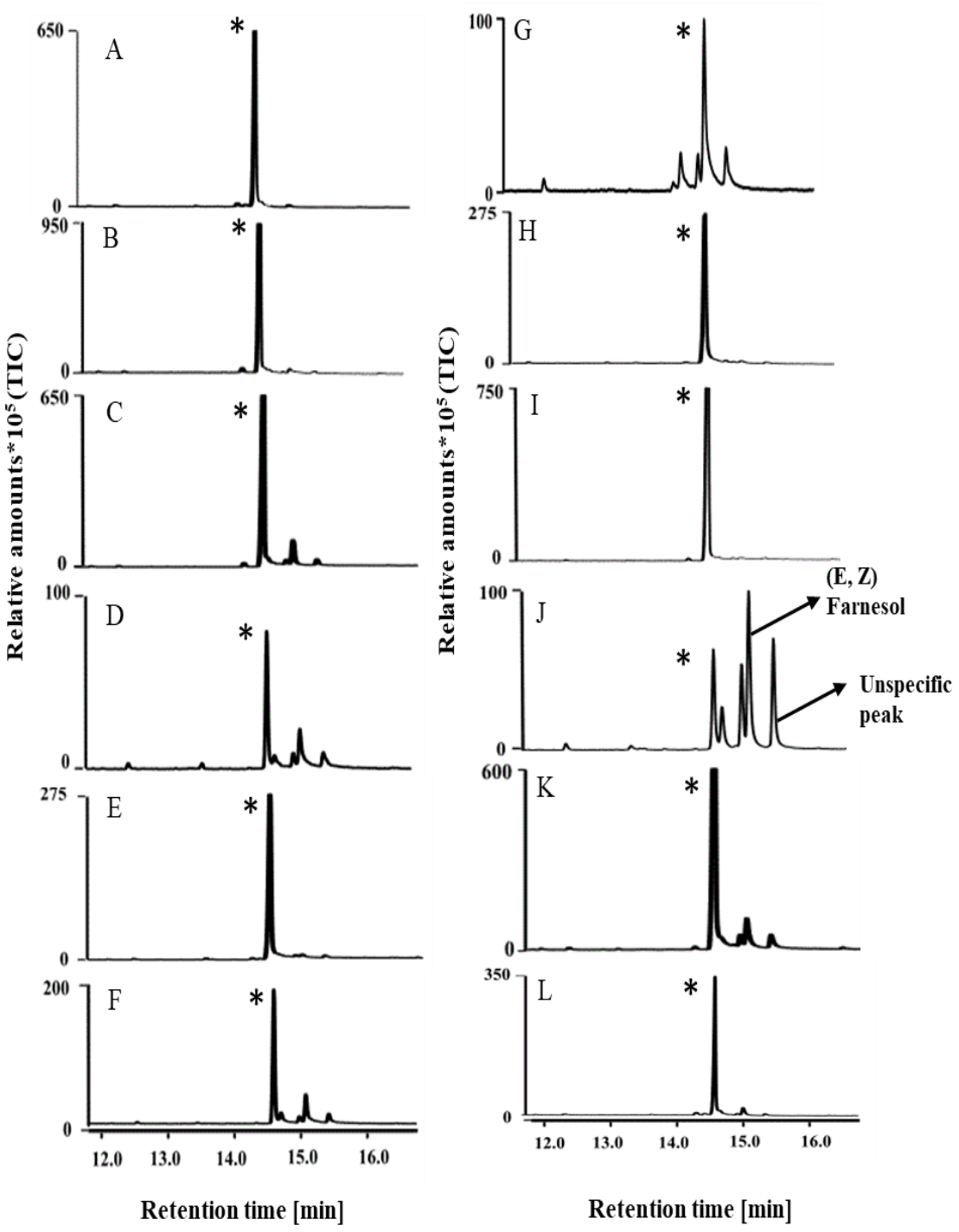


Fig. 3.11 *In vitro* product specificity of MrTPS7 alleles after heterologous expression in a bacterial system.
Product peaks of the enzyme assay (by SPME, GC-MS) after heterologous expression of MrTPS7 alleles in *E. coli*. and incubation with FPP substrate. Peaks were identified by SPME, GC-MS as α-bisabolol as the major peaks where A. ARG, B. BOHE, C. BOD, D.GER, E. GO, F. LEU, G. MAT, H. MAR, I. .MAN, J.MO, K.PO, L.ZL as accessions.

The principal product of all heterologously expressed enzymes from the twelve chamomile accessions was identified as α-bisabolol. In the accession MO, an additional *E, Z* farnesol peak was found. Since all α-bisabolol synthases are capable of producing the product, we hypothesized that accessions with higher α-bisabolol production might contain an MrTPS7 allele with higher activity under physiological circumstances. To test this hypothesis, we purified three enzymes from low, medium and high-expressing

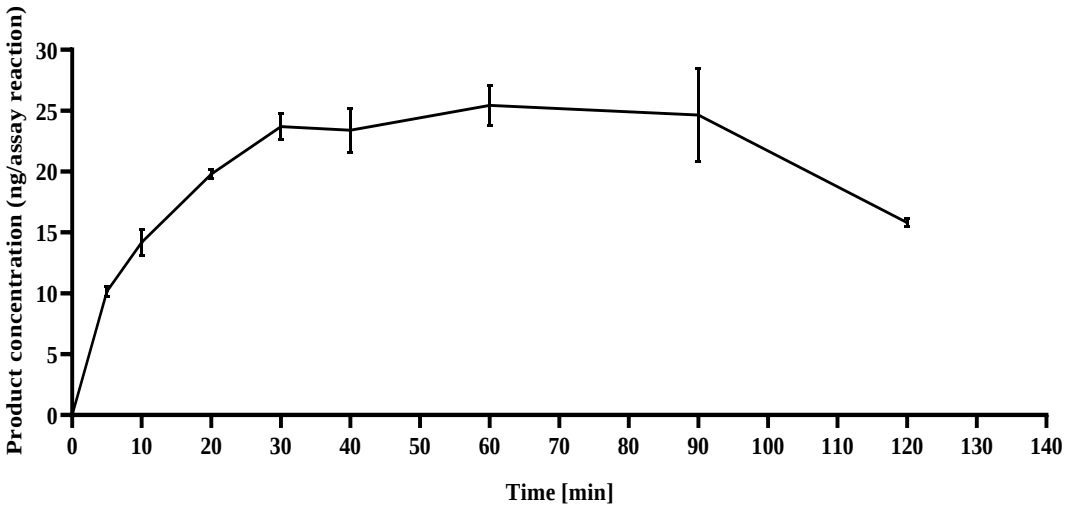


Fig. 3.12 Product formation of MrTPS7 allele of Germania (heterologous expression)
The amount of extracted terpenes from the assay reaction to the organic solvent phase has decreased over time. The master mix assay condition was set up and the reaction was stopped through vortex. Error bar represent (±) SEM (n = 3).

accessions (GER, GO, and MAT) to determine their K_m values. First, linear assay conditions were established for the purified enzymes. In the presence of FPP substrate and divalent metal ions Mg^{2+}, a time interval of 0 to 120 minutes was chosen for the monitoring of product formation (Fig. 3.12). After 15 minutes, the reaction reaches a linear phase which was used for further enzymatic characterization. Three enzymes of MrTPS7 alleles from accessions GER, GO and MAT, respectively, were purified and assayed to determine their K_m (Fig. 3.12). The value for the GER enzyme was 3.59µM, the value for the GO enzyme was 8.30 µM and the value for the MAT enzyme was 12.32µM (see table 3.5).

Table 3.5 Enzymes from three diferent plant accessions, with their K_m value and respective α-bisabolol concentration in flowers.

Accession	K_m value (µM)	α-bisabolol conc. (µg/g of fresh weight)
GER	3.59	763.96
GO	8.30	294.33
MAT	12.32	2.37

These values appears to correlate with the α-bisabolol contents of the respective accessions. This suggests that the biochemical properties of the enzymes have some impact on the regulation of α-bisabolol production. Although, only relatively small differences of K_m were found among the selected accessions.

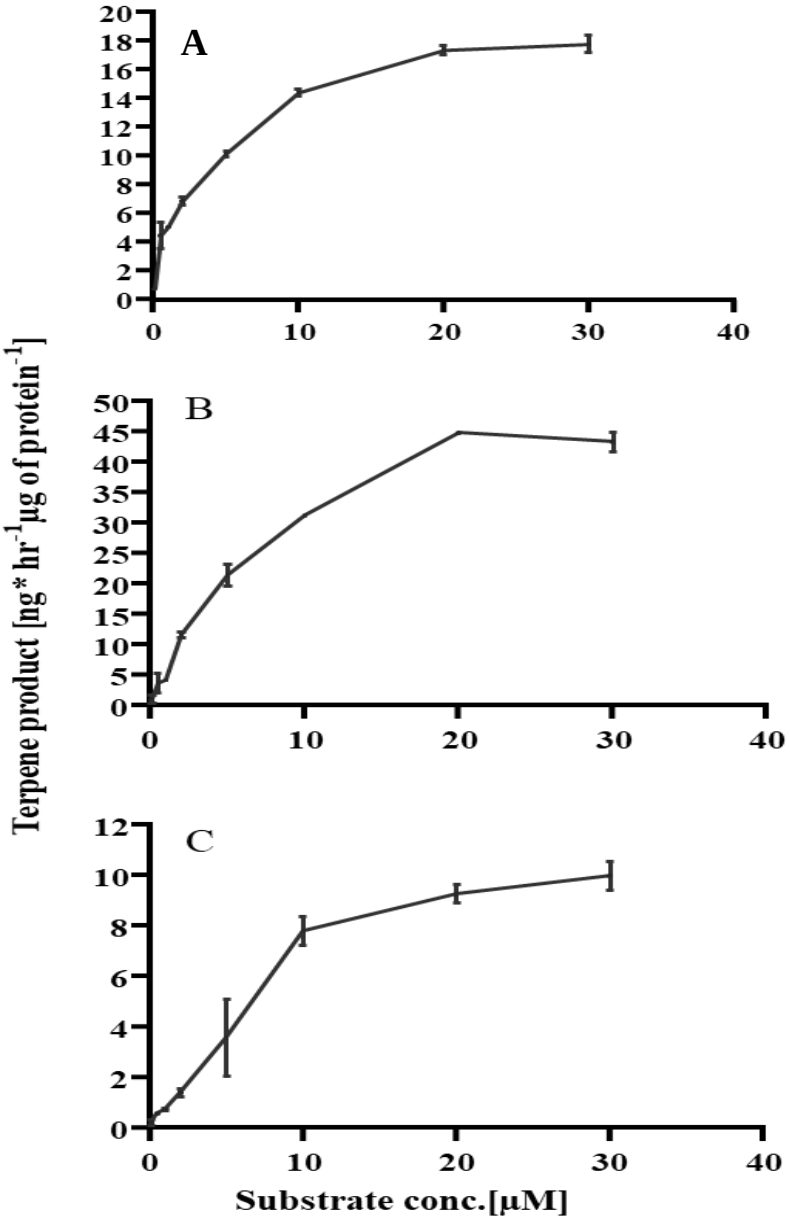


Fig. 3.13 Determination of K_m values for three MrTPS7 alleles
The MrTPS7 enzymes of three accessions (GE, GO and MAT) were heterologously expressed, and purified utilizing a histidine tag. The kinetic properties were examined in an in vitro assay. The K_m of the GER enzyme was 3.59 µM (A) the K_m of the GO enzyme was 8.30 µM (B), and the K_m value of the MAT enzyme was 12.32µM (C).

3.8 Determination of substrate concentration in flowers

The previous experiments indicated that α-bisabolol formation in the chamomile accessions was only to a small extend regulated by the biochemical activities of the MrTPS7 alleles themselves or their transcriptional regulation. This left the hypothesis

that α-bisabolol production is determined by the availability of the FPP substrate for the enzymes of the MrTPS7 alleles. We measured the FPP substrate concentrations of flowers of each accession with LC-MS. For quantification a known quantity of internal standard was added to the samples subsequently lyophilized and dissolved in a mobile phase (specified in section 2.8). The peak of the FPP substrate was identified by retention time and correlated with sinigrin as an internal standard. The peak at RT 7.59 minutes was the internal standard which was previously added in the extraction buffer (2 µg / 100 µl of solvent). Additionally, we observed a significant peak at RT 20.07 minutes which seemed to be FPP substrate peak (Fig. 3.14 C). To confirm this, we use the standard FPP (50µM) and identified the same at RT 19.85 (Fig 3.14 B), simultaneously, the peak at RT 20.07 was from the solvent and marked as extraction blank (Fig. 3.14 C). We identified two peaks with slightly higher retention time than FPP (RT 19.85). We confirmed that these are probably effects from background or instrumental noises (Fig. 3.14 D).

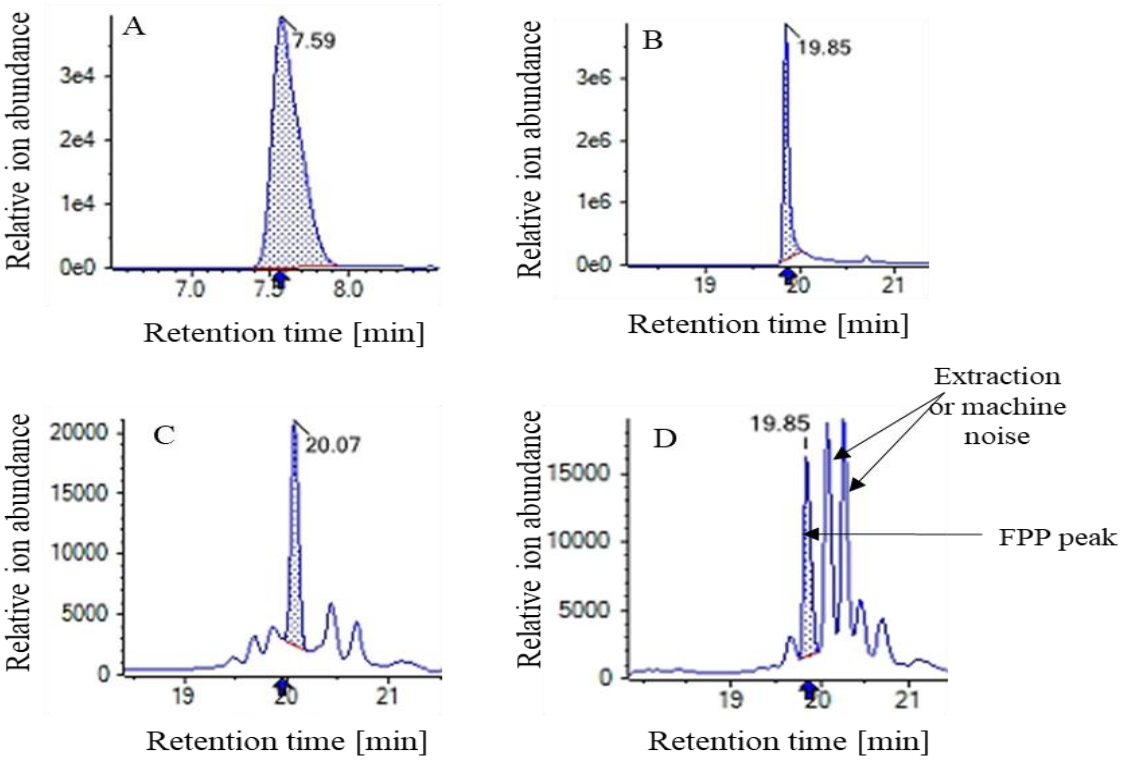


Fig. 3.14 LC-MS peak integration for FPP concentration measurments in chamomile extracts
The peaks are identified as A. Sinigrin as internal standard (RT 7.59) B. Standard. FPP (RT 19.85 min) C. extraction blank and D. FPP peak where septum contamination or instrumentational noises were separated with arrows.

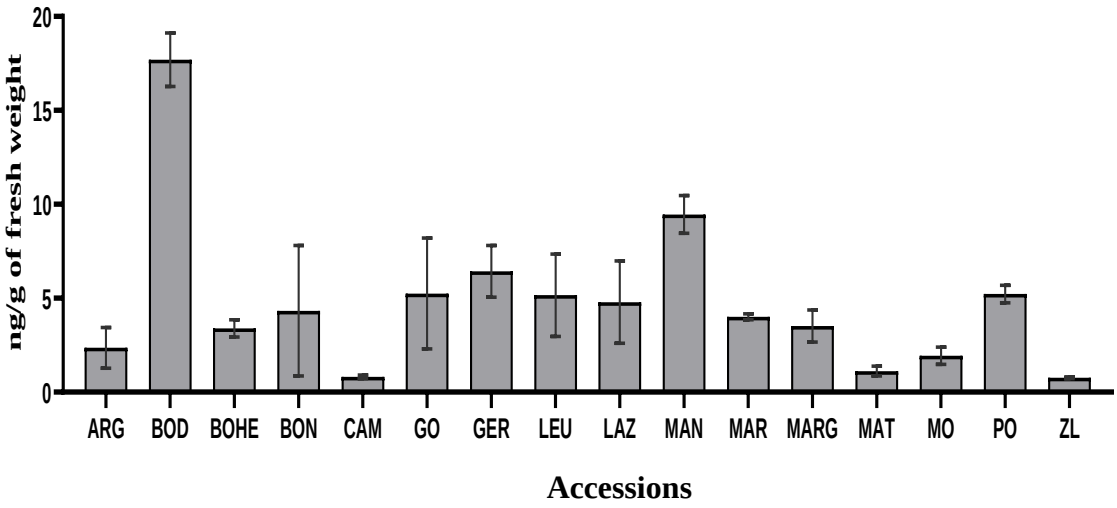


Fig. 3.15 Concentration of the FPP substrate (ng/g of fresh weight) in sixteen accessions of Chamomile.
The FPP concentrations were determined by the correlation with a known standard and with the internal standard sinigrin. Each bar is represented by mean value of the quantified sample and error bar represented the (±) SEM.

The FPP substrate availability within sixteen chamomile accessions varied largely. Bodegold was the accession with the highest FPP substrate concentration. In the accessions, Bohemia, Bona, Goral, Germania, Lutea, Lazur, Manzana, Pop Germany, Margaritar and Pnosa medium level of substrate concentration was found, with some variation within the accessions. Zloty Lan was the accession with the lowest FPP concentration and Argenmilla, Camoflora and MAT had similar low levels of FPP with little differences (Fig. 3.15).

3.8.1 Correlation between α-bisabolol contents and substrate availability in flowers within sixteen plant accessions

To test whether the FPP substrate in flowers might control α-bisabolol production, we correlated the FPP concentration with α-bisabolol production in the chamomile accessions. The test was performed with sixteen different chamomile accessions with flowers. The α-bisabolol concentration in Germania (3) was the highest combined with a medium FPP value. Bodegold (1) contains the highest concentration of FPP substrate while displaying a medium level of α-bisabolol. In three other accessions Bodegold (1),

Manzana (2) and Germania (3) a higher substrate availability was noticed. In Goral and Lutea a higher concentration of α-bisabolol was found (Fig. 3.16).

The Spearman's co-relation coefficient (r_s) was calculated. The resulting r_s= 0.84, denoted that there was a strong correlation between the concentration of α-bisabolol and FPP substrate availability in the flowers.

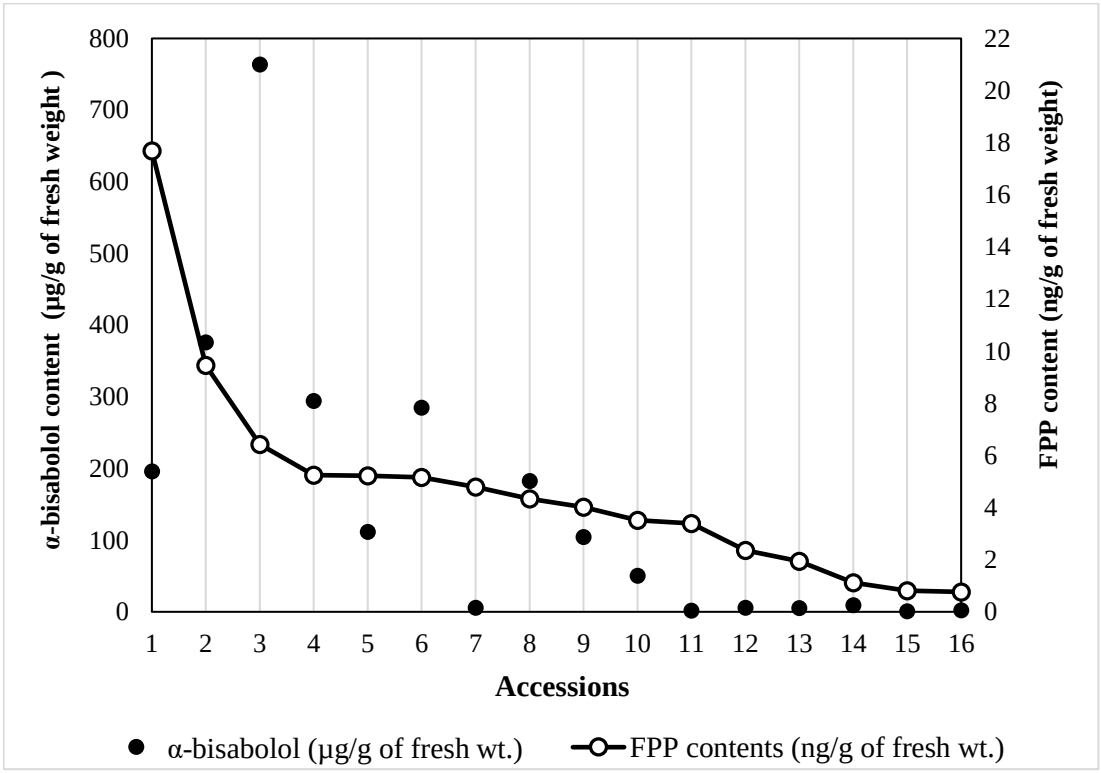


Fig. 3.16 Correlation between substrate concentration and α-bisabolol contents in flower
The accessions were: 1 BOD, 2. MAN, 3. GER 4. GO, 5. PO, 6.LEU, 7. LAZ, 8. BON, 9. MAR, 10. MARG, 11. BOHE, 12. ARG, 13. MO, 14. MAT, 15.CAM, 16. ZL
(Spearman's co-relation coefficient: 0.84)

3.9 Computational modeling of MrTPS7 alleles

The biochemical characterization of three selected MrTPS7 alleles from accessions (GER, GO and MAT) revealed varying levels of activity (in terms of K_m).

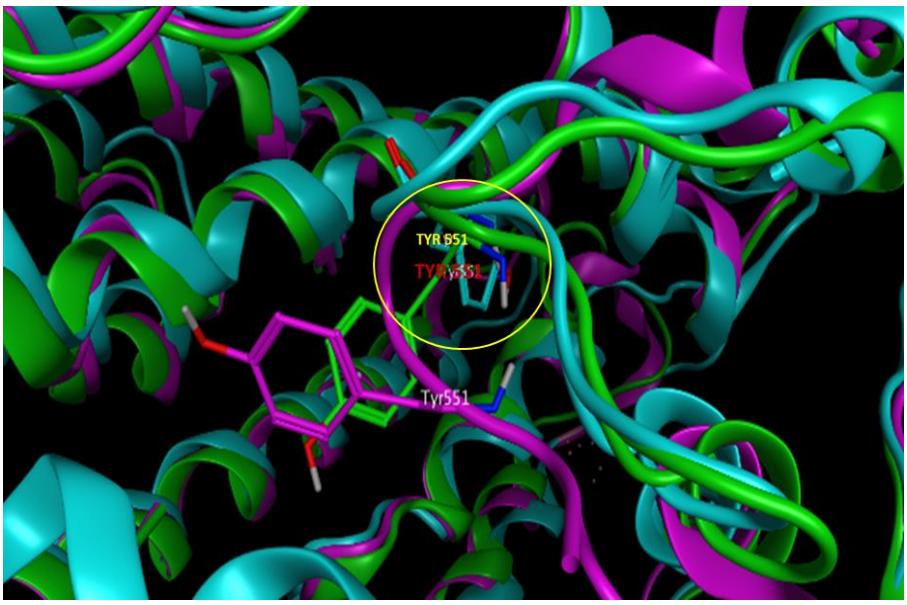


Fig. 3.17 Superimposed *in silico* models of the active site of three MrTPS7 alleles
Color code: green–GER, cyan–GO and magenta – MAT. Active site mentioned with yellow Color circle.

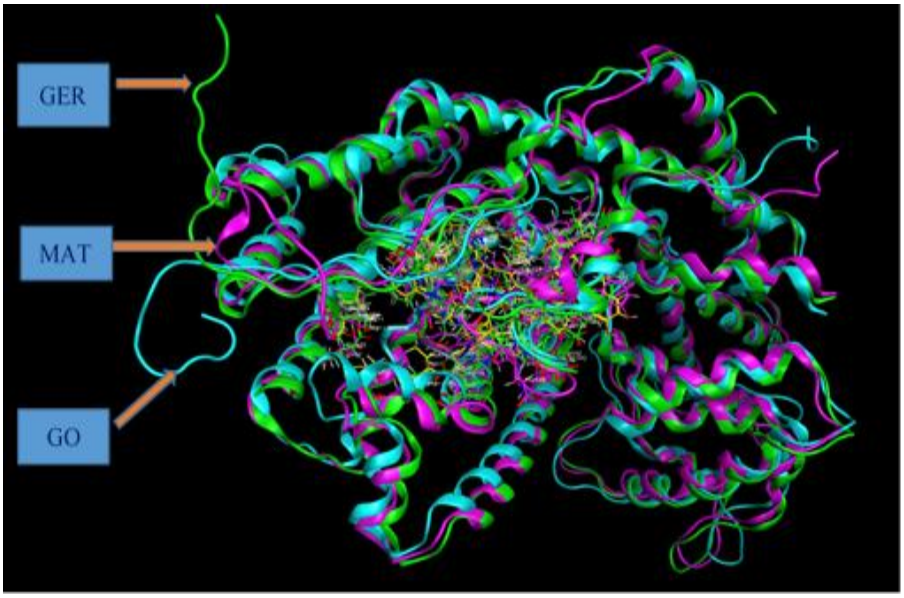


Fig. 3.18 C-terminal active site of superimposed *in silico* model of MrTPS7 alleles
Color code: green–GER, cyan–GO and magenta – MAT.

To understand how the sequence alterations between the alleles result in different biochemical activity, we employed computational modeling. Amino acid alterations in the active site influence most likely the catalytical properties of the enzyme. *In silico* modeling of MrTPS7 alleles of the GER, GO, and MAT accessions displayed significantly different positions of a tyrosine side chain within the three active sites (Fig. 3.17).

This conserved tyrosine residue near the active center of the enzyme is important for the protonation process (Tyr 551). The superimposed models showed that Tyr 551 in GER and GO were situated closer to the active site. While the peptide backbones of GER and GO enzyme models are relatively similar, it is moved to the side in the model of MAT.

The C terminal domain has a hydrophobic pocket at the active site and is closed off towards the outside by two loops which are located on the protein surface. Previous studies demonstrated that the hydrophobic moiety of FPP reaches into the active cavity while the diphosphate function interacts with magnesium ions at the entrance of the active site (Degenhardt et al., 2009). At the N-terminal end, a rigid secondary structure (structural folding) was noticed (showed by arrows in Fig. 3.18) in MAT. This structure prevents the N-terminal end from curving back to the front of the substrate binding site, as it is observed in the GER and GO enzymes (Fig. 3.18). The inability of the MAT allele model to cover the entrance of the active site may contribute to the lower substrate binding activity that was observed in the biochemical characterization (Fig.3.13).

4 Discussion

4.1 High variation of volatile metabolites in chamomile accessions

4.1.1 Seedling age determines the composition and quantity of the volatile oil

Along with the genetic background and environmental conditions, the developmental state of the plants influences its essential oil content and quality (Mohammad et al., 2010). Hence, the developmental state of the plant for the commercial harvest of essential oils is very important. To monitor developmental changes in the essential oil of chamomile, the production of volatile oil components were measured in the seedlings of the seven chamomile accessions Bodegold, Bona, Lutea, Germania, Manzana, Pnos, and Zloty Lan (Figure 3.1).

The sesquiterpenes β-caryophyllene and β-farnesene were generally dominant during the first 30 days of seedling development. The concentration of other terpenes like β-ocimene, citronellene, longipinene oxide and Germacrene D varied highly during seedling growth. We did not find α-bisabolol and its oxides during the seedlings growing period. This finding conflicts with Mohammad et. al., 2010 where localization of bisabolol oxide was detected in 30 day-old seedlings. This could be due to the selection of different chamomile varieties during that study or other environmental conditions, for example long day light exposure which can alter plant morphology. Accumulation of spiroether was consistently high throughout the first 40 days in five accessions of chamomile, but Manzana and Pnos started producing this compound only after 30 days.

Sesquiterpenes have diverse functional roles in plants, some acting as hormones, regulators of wound induced responses and anti-oxidants. We found a significant quantity of sesquiterpenes formed right from the beginning of seedling development that might function to minimize the reactive oxygen in physiological conditions during seedling growth. Also β-caryophyllene and β-farnesene are known as stress modulators

and might protect the plant from herbivory attack and environmental stresses (Köllner et al., 208; Schnee et al., 2006). Generally, juvenile stages express stronger defenses than mature stages (Bryant JP, 1995) and this may allow plants to adapt to different kinds of stress. The young seedlings produced a significantly lower concentration of total metabolites than mature individuals in their stems, leaves, buds or flowers. As the seedling develops, primary metabolism might be more focused on growth and development and the production of secondary metabolites might be compromised at this juvenile stage. However, our study was restricted to terpenes, the C26 compound hexacosane, and spiroether. Our investigation indicated that the synthesis of terpene metabolites in chamomile accessions varies greatly. The variability in the production of metabolites could affect all ecological interactions. The vegetative parts of chamomile produce complex terpene mixtures without α-bisabolol.

Previous studies highlighted the large number of intraspecific variations in the composition of terpenoids in mature individuals of chamomile and other aromatic plants (Das, 2014). The Asteraceae family is an especially diverse group which produces an exceptionally high number of terpenes and their derivatives. In other plants from the Asteraceae family like daisy, sunflower, marigold, chrysanths, the distribution of essential oil was also found to be versatile in developmental organs like stem and leaves (Alok Krishna Gopal R. Mallavarapu, 2004, Ceccarini et al., 2004). Stems and leaves are the primary aerial tissues, and the accumulation of volatile oil is always affected by their variety and growing condition. We analyzed eleven volatile metabolites (mono, sesquiterpenes and ethers) in thirteen different chamomile accessions for their distribution throughout the stem and leaves shown in Table 3.1 and Figure 3.2.

The overall concentrations of sesquiterpenes in leaves and stems is far higher than that of monoterpenes (Table 3.1). The principal sesquiterpenes in leaves and stems were β-caryophyllene, β-farnesene, bicyclogermacrene, and germacrene D, together with minor amounts of α-gurjunene, aristolene and longifolene. In addition, the monoterpenes β-ocimene was found (Fig. 3.2 A). In the thirteen analyzed accessions, β-ocimene and β-farnesene were most commonly produced. We observed that biomass and essential oil yield was rising with the vegetative growth of the plant. Localizations of bisabolol and its' oxides were absent in stems and leaves regardless of the plant accessions. However, there is one report of a detection of α-bisabolol oxide in leaves of in thirty days old

seedlings of chamomile growing in Belgium (Mohammad et al., 2010). The amounts of secondary metabolites in young organs were greater than in mature leaves or stems. Our data matched with the investigation of Sarrouet. al., 2016 in sage plants where secondary metabolites were varied qualitatively and quantitatively in populations of the same species.

4.1.2 Buds and flowers produce a high concentration of essential oils

In chamomile, the flowers contain the highest concentration of essential oils. This blend of monoterpenes, sesquiterpenes and ether type compounds is not only stored in the flower but also released as a volatile scent in low concentrations. The variation of the essential oils of sixteen chamomile accessions contains the major monoterpenes β-ocimene, cineole and γ-terpinene which are present in buds of all accessions (Fig. 3.3 A, B and C). Our analysis showed also that the sesquiterpenes β-caryophyllene, β-farnesene, germacrene D and two ethers, spiroether and en-in-dicycloether, were predominant in most of the accessions (Fig. 3.3 F, G, H, L, M). The β-caryophyllene and β-farnesene which are considered as defensive floral scents were accumulating in the flowers in higher quantity irrespective to accessions. Additionally, bisabolol and its oxides (Fig. 3.3 N, O, P and Q) were distributed within accessions and their concentration varied with organ development (buds to flower).

German chamomile is a rich source of bisabolol and its' oxides (bisabolol oxide A, and bisabolol oxide B) are found in buds and flowers (Fig. 3.3). The concentration of α-bisabolol was higher in flowers than in the buds irrespective of accessions (with the exception of Bohemia and Pohorelicky velkokvety). Nine chamomile accessions (Bodegold, Bona, Germania, Goral, Lutea, Manzana, Margaritar, Pnos and Pop Germany) showed a high content of α-bisabolol (Figure 3.4). Our analysis of Lazur (southeastern Bulgarian origin) differed from the data of Franke and Schilcher (Rolf Franke et al., 2005) as we detected high levels of bisabolol oxide A but less α-bisabolol. Also, the accession Zloty Lan was reported with high levels of bisabolol oxides but we did not found significant levels of bisabolol oxide A and B in this accession (Fig. 3.3 N and Q).

These differences could be explained by polymorphic varieties of this outcrossing species. The resulting diversity is potentially increased by introgression with other chamomile populations during cultivation. This variability should be taken into account

when maintaining a genetic population with high levels of specific compounds serving as a medicinal drug. Our data with the presence of α-bisabolol in Bona, Goral, Lutea and Manzana matched with the findings of Franke and Schilcher (Rolf Franke, 2005). In the accessions of Camoflora, they found high levels of bisabolol oxides which we did not find. Argenmilla, Bohemia and Bodegold were rich in bisabolol oxides which agreed with previous findings. The plant chamomile is a rich source of spiroether and en-in di cycloether in their organs (stems-leaves, buds and flowers). We analyzed essential oil metabolites from the leaves of thirteen different chamomile accessions. A total of sixteen accessions were analyzed for flowers (Fig.3.2 and Fig. 3.3.).

The levels of oxides increased during the flowering phase, maybe due to the enzymatic oxidative reaction that converts the α-bisabolol over time. Ghasemi et. al., 2016 described that contents of α-bisabolol is increasing towards the full flowering stage and the di-cycloether concentration is reduced (Ghasemi et al., 2017; Letchamo, 2015). To test for such a correlation, we have analyzed the buds of thirteen accessions and the flowers of sixteen accessions. In buds, there is a very low correlation (r_s= 0.168, Fig 4.1) between the concentrations of the two compounds found. However, the correlation in the full flowering stage is a little higher (r_s = 0.352, Fig. 4.2 and supplementary Table S7) than in buds. This result is in conflict with the findings of Ghasemi et al., 2017. This may be due to an accelerated oxidative degradation during maturation of the flower. It might also depend on harvesting procedures or changes of the environmental parameters, which could directly influence the oil content.

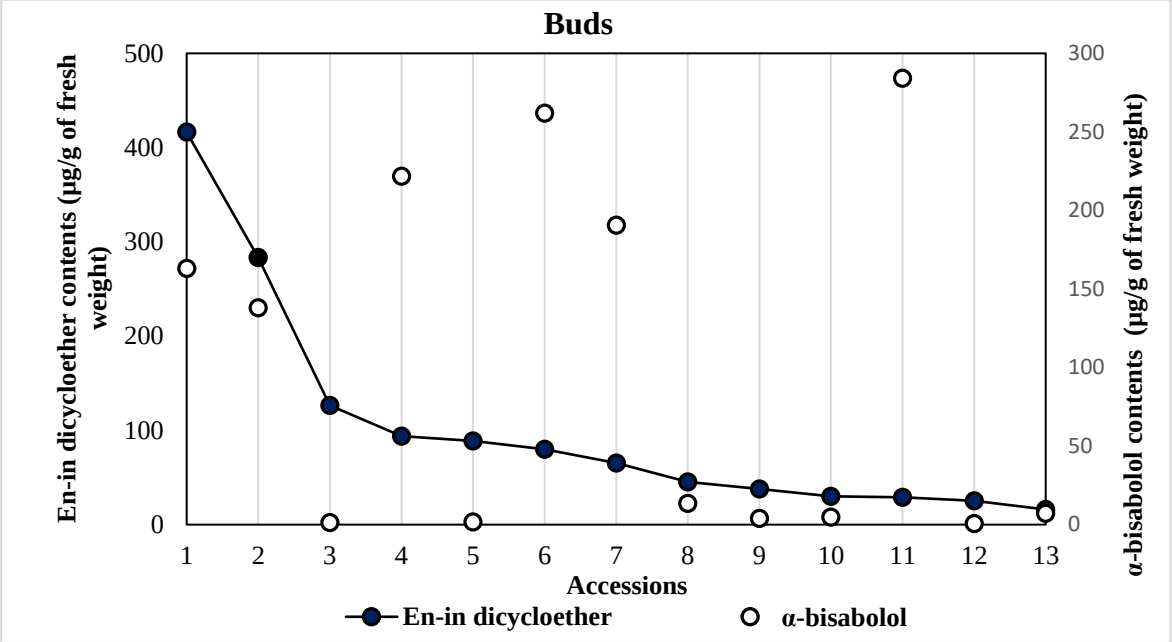


Fig. 4.1 Correlation between content of en-in dicycloether and α-bisabolol in buds of plant accessions
1 BON,2. PO, 3. ZL, 4. GO, 5. LAZ, 6. MAN, 7. BOD, 8. MARG, 9. ARG, 10. BOHE, 11. LEU, 12. CAM, 13. GER
(Spearman's coefficient rank: 0.168)

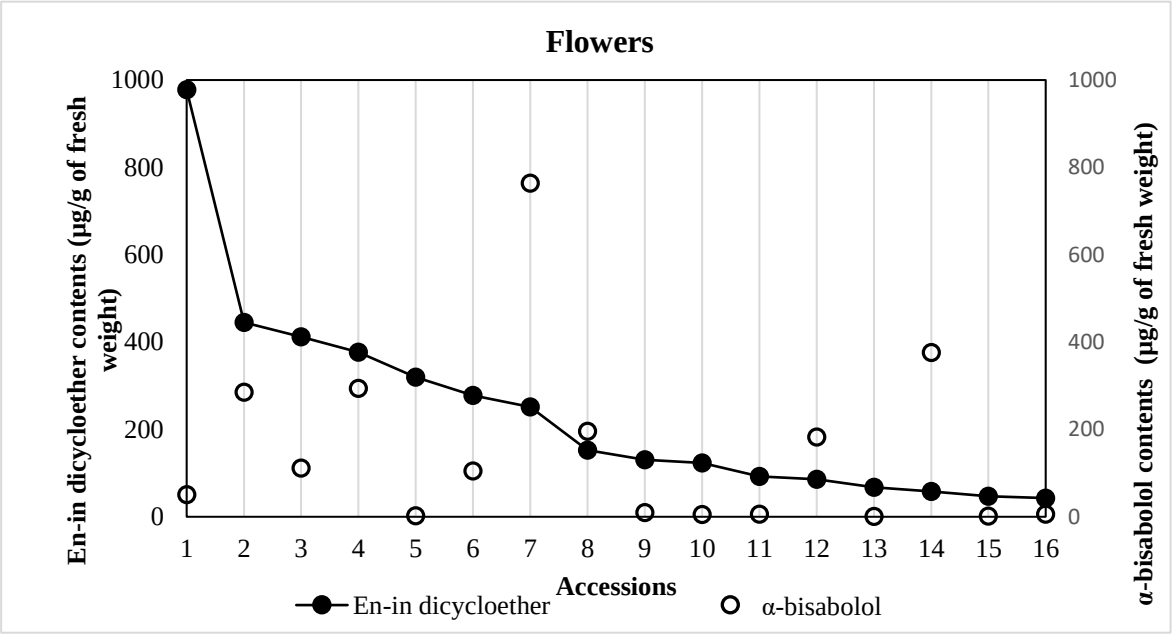


Fig. 4.2 Correlation between content of en-in dicycloether and α-bisabolol in flowers of plant accessions
1 MARG, 2. LEU, 3. PO 4. GO, 5. MAT, 6.MAR, 7. GER, 8. BOD, 9. ZL, 10. MO, 11. ARG, 12. BON, 13. CAM, 14. MAN, 15. BOHE and 16. LAZ
(Spearman's coefficient rank: 0.352)

4.1.3 Essential oil production is affected by the ploidy status of the chamomile accessions

Breeding programs of chamomile cultivars are usually centered on the composition and quantity of the essential oil in the context of regional cultivation techniques and ecological conditions. Hence, the regional chamomile cultivars Degumille in Spain, Bona in Slovakia, Quedlinburger Großblütige Kamille in Germany have been subjected to breeding for better oil production. Heinz Schilcher, Peter Imming, and Susanne Goeters (Rolf Franke and collaborators, 2005), suggested for breeding targets several aspects like resistance against diseases, the ability for quick germination, high flower yield, ripening of all flower heads at the same time, a homologous flower horizon, stability of the flower head and stem strength for the suitability of mechanical harvesting procedures. Unfortunately, not all these criteria were not fulfilled completely in a single cultivar even within the sixteen chamomile accessions in our experiments. Generally, most of the harvesters focused on the accessions with high yield of oil production. As a breeding strategy, tetraploid chamomile varieties were generated several times by polyploidization of diploid plants (Otto et al., 2017).

Among our accessions, nine tetraploid varieties are likely to originate from a polyploidized diploid genetic background, with minor genetic differences arising via selection and introgression during breeding (Table 1.2). As an example, the tetraploid variety ‘Manzana’, was generated from diploid ‘Degumille’ (Franz C, 1986). ‘Bona’ was generated by cross-breeding of wild chamomile from Spain and the variety Bohemia. The basic breeding material for the variety Lutea was Goral. The tetraploid ‘Bodegold’ is genetically distinct from the diploid ‘Camoflora’ but both were similar in “STRUCTURE fingerprints” assuming that the tetraploid originates from the diploid Camoflora genotype (Otto et al., 2017).

Within the sixteen accessions, seven accessions were diploid (Table 1.2). The production of the three monoterpenes cineole, β-ocimene and γ-terpinene was higher in flowers in the tetraploids accessions. However, an exception was found with Germania, a diploid which contains significant level of β-ocimene and γ-terpinene. The concentration of artemisia ketone is high in both the tetraploid Zloty Lan and the diploid

Bona, suggesting that the ploidy status does not affect the production of this compound. Ipsdienol and bicyclogermacrene were much higher in the tetraploid Lutea, and β-caryophyllene in the diploid accession in Pohorelicky velkokvety (MAT). Both β-farnesene and germacrene D concentration was higher in tetraploids compared to diploids accessions. However, Pop Germany (MAR), a diploid cultivar, contains the highest level of germacrene D in chamomile accessions. We could not find any significant correlation in terms of production of monoterpenes and sesquiterpenes except bisabolol and its oxides with diploids and tetraploids accessions. The concentration of these terpenes within diploids and tetraploids individuals are also varied (Fig. 3.2 and 3.3). Therefore, it is difficult to conclude that tetraploids varieties produce more secondary terpenes metabolites than diploids varieties. Sometimes, populations of chamomile exhibit variable ploidy since tetraploids were crossed with diploids (Otto et al., 2015). Variable ploidy can also be caused by contamination with wild growing, diploid chamomile through seeds or (unreduced) pollen (Otto et al., 2015). Tetraploid seeds are slightly bigger than diploid seeds but a separation in soil not possible. We could not exclude that this variation of ploidy status might affect the essential oil yield in laboratory and field experiments to a small extend.

4.2 Mechanisms of regulation of α-bisabolol production in chamomile

4.2.1 Identification of MrTPS7 alleles in chamomile accessions

To find the regulatory mechanism that is responsible for the variation of the α-bisabolol formation among the chamomile accessions, we identified the alleles of the α-bisabolol synthase *MrTPS7* in each accession. We characterized the α-bisabolol synthase *MrTPS7* from a set of twelve chamomile accessions. The similarity of these amino acid sequences from most accessions is above 90%, which is typical for allelic variation among plants (Table 3.3). Although these sequences are likely alleles of *MrTPS7*, we cannot exclude the presence of another closely related α-bisabolol synthase. While the diploid accessions contain two alleles of *MrTPS7,* the tetraploid varieties will contain four alleles. The isolation of genes from cDNA libraries by PCR will favor genes that

are highly transcribed. Therefore, we can assume that the identified genes contribute to much of the α-bisabolol biosynthesis in the respective lines. Two accessions, BON and PO only have a 60 - 70% sequence identity to the other putative α-bisabolol synthases, but they have 61% identity with each other. Here, no assumption about an allelic relationship can be made. One sequence, of the accession MO, displays a sequence identity below 20%. This accession is the only one of *Matricaria discoidea*, the α-bisabolol synthase seems to be a very different one in this species and may not be related to *MrTPS7*. A detailed analysis of the genomic sequences of several accession might be necessary to clearly establish the allelic relationship between all the putative α-bisabolol synthases in this study.

The class of chamomile TPSs including *MrTPS1*, *MrTPS3* and *MrTPS5* which do not produceα-bisabolol showed less than 20% of amino acid similarity with GER (Irmisch et al., 2012). This low identity was mainly found close to the conserved motifs (see supplementary Table S9, Fig.S7). Report from Schnee et. al., 2002 (Schnee et al., 2002) demonstrate that terpene synthases with very low sequence identity might catalyze similar reactions. But genes with high (80%) identity can encode enzymes with very different products specificity. In our study, the product specificity of MrPTS7 seems to be highly conserved, even if the gene sequence is not. In the case of MrTPS7 from accession *M. discoidea* which had an extremely low sequence identity, convergent evolution might have played a role. Convergent evolution causes unrelated terpene synthases to form similar products (Bohlmann et al., 1998; Steele et al., 2002). The putative MrTPS7 alleles encode α-bisabolol synthases in all accessions.

During investigation on α-bisabolol biosynthesis in arabidopsis and sandalwood, α-bisabolol was the minor or co-product of other TPS genes (Jones et al., 2011; Ro et al., 2006). Our heterologous expression and *in vitro* characterization of *MrTPS7* alleles showed mostly a single product with GC-MS analysis (Fig.3.11) which is in agreement with Son et al., 2014). However, we noticed an additional product of sesquiterpene alcohol (*E,E*)-farnesol was formed in a single accession of 'MO' (Fig.3.11 J). We tested whether (*E,E*)-farnesol was a product of the enzyme, or an impurity of the substrate, or a product of unspecific phosphatases in the bacterial extract. To check the latter hypothesis, we have expressed the enzyme and purified it over a pro-infinity IMAC Ni-charged resin column. The sesquiterpene product formed by the purified, His-tagged

protein was only α-bisabolol when examined with both SPME and *n*-hexane solvent extract with GC-MS (Fig. S6 in supplementary). Therefore, (*E*,*E*)-farnesol was not an enzyme product. So, further experiments using phosphatase inhibitors sodium tungstate or sodium fluoride were not needed as described by Schnee et al (Schnee et al., 2002). Hence, MrTPS7 is a monoproduct enzyme and produces exclusively α-bisabolol with the substrate FPP in the chamomile flowers.

4.2.2 Regulation of *MrTPS7* transcript levels does not account for the amount of α-bisabolol product formed

The transcript levels of *MrTPS7* were analyzed in each accession (Fig. 3.9) and correlated with the respective α-bisabolol content. We noticed a wide range of transcript level within the sixteen accessions. Germania (GER) contained the highest levels of transcripts and Pop Germany (MAR) the lowest levels in flowers. When comparing the transcriptional levels of low α-bisabolol-producing accessions (MAR, GO, PO) and those of higher production accessions (GER, LEU, MARG) with their content of α-bisabolol no clear correlation (r_s = 0.11) was observed. This indicates that transcriptional regulation of *MrTPS7* is not sufficient to explain the regulatory mechanisms by itself. Other modes of regulation appear to be more important than transcription. This did not match with the investigations in *Eremanthus erythropappus*. There, the transcript levels of the gene EeBOS encoding an (–)-α-bisabolol synthase correlated well with the α-bisabolol levels in the analyzed tissues (twigs, roots, leaf, wood) (Alves Gomes Albertti et al., 2018). Also, a direct correlation with terpene content and transcript level was described by Bohlmann et al., 2004 when they worked with *(E)*-β-caryophyllene synthases, α-terpineol synthase (Vvter) and α-farnesene synthase (VvCsaFar) in plant *vitalis vinifera*. The transcript levels of the genes (VvCsaFar, Vvter) had a direct correlation with the respective product content (Martin and Bohlmann, 2004). Although, exceptions are not rare as described by VasilikiFalara et.al.2011, in the plant tomato (*Solanum lycopersicum*). There, *TPS8 which* catalyzes the biosynthesis of cineole was expressed in leaves and roots but no cineole was found in that organ (Falara et al., 2011).

Nevertheless, transcripts of *MrTPS7* were found only in the flowers of the chamomile accessions, indicating that transcriptional regulation prevents the production of α-bisabolol in any other part of the plant. Some chamomile accessions possess high transcript levels but a low α-bisabolol content (Margaritor:2.46/50.72, *M. Discoidea*: 2.04/5.20). This might be due to an oxidation of α-bisabolol to some of its oxides as shown in supplementary Table S8. In our study, we assume that the oxidative transformation of α-bisabolol to its oxides might occur during the maturation of flowers (Margaritar (oxide A 382.66, oxide B 133.45 µg/g) and in *M. Discoidea* (oxide A 32.47, oxide B 98.16 µg/g). A similar oxidative transformation was observed for the α-pinene synthase VvPNaPin1 (*Vitis vinifera*) which is expressed in roots. While little α-pinene was detected in this organ, oxidative products of α-pinene like borneol, pinocarveol and myrtenol were observed (Matarese et al., 2014; Prema and Bhattacharyya, 1962).

4.2.2.1 Kinetic properties of putative MrTPS7 alleles do not account for the differences observed in α-bisabolol production of the accession

To determine possible mechanisms regulating α-bisabolol production of MrTPS7, we characterized the some of the enzyme kinetics. We chose three alleles of MrTPS7 from the accessions Germania (GER), Goral (GO) and Pohorelicky velkokvety (MAT) to determine their substrate K_m values. We expected to find the highest K_m value in the low α-bisabolol producer and the lowest value in those producing high concentrations of this compound. Measurement of K_m values were described in (Table 3.5). The availability of the substrate in the three accessions (GER, GO and MAT) differ with each other as 6.4291, 5.2497and 1.121 ng/g of samples respectively (Fig. 3.15 and supplementary Table S6). The spearman's correlation was r =0.84 when considered with substrate concentration and α-bisabolol contents within all accessions (Fig. 3.16). We concluded that the biochemical properties of the α-bisabolol synthases might contribute to the α-bisabolol content in the respective accession, but other regulatory mechanisms appear to have a bigger influence.

4.2.3 FPP Substrate concentration directly correlates with α-bisabolol production in the chamomile accession

We investigated substrate level concentration in chamomile flowers within accessions (Fig. 3.15) to test whether this factor might affect the amount of α-bisabolol formed in chamomile. To confirm the specific substrate contribution to the biochemical reaction with MrTPS7 alleles, we used both GPP and FPP and found no product with GPP (supplementary Fig. S8). Hence, we investigated FPP substrate concentration in the flowers. Fig. 3.15 showed the highly variable FPP concentration within sixteen individual accessions. A correlation of FPP concentrations with α-bisabolol production showed a Spearman's correlation coefficient of 0.84. This correlation value was much higher than the correlations with m-RNA transcript levels (Fig. 3.10). Therefore, the substrate availability and catalytic activity in the biosynthesis pathway of α-bisabolol has the largest impact on product biosynthesis. The high substrate concentration seems to directly influence the production of α-bisabolol by the MrTPS7 alleles. The K_m values of these MrTPS alleles do not seem as important for the regulation of α-bisabolol production.

4.3 Structure-function analysis of MrTPS7 alleles

We purified recombinant MrTPS7 proteins from a bacterial system. The isolated TPS enzyme was found to convert FPP to the sesquiterpene product (α-bisabolol). The ionization of farnesyl diphosphate to farnesyl cation and isomerization of nerolidyl cation is the basic initiation steps for the bisabolol synthesis. The reaction mechanisms was evaluated in earlier studies with multiple-labeled FPP and nerolidyl-di-phosphate substrates (Alchanati et al., 1998; Cane and Tandon, 1995). The arginines of the Rxx8W motif (Fig. 4.4) play a role in diphosphate migration during the formation of carbocation intermediates. The reaction mechanisms of sesquiterpene synthases are initiated by the formation of carbocationic intermediates after de-phosphorylation of the FPP substrate. Cyclizations can involve a carbocationic attack on either of the two remaining double bonds giving a wide range of different carbon skeletons. The nerolidyl cation undergoes a 1, 6 cyclization of to form the bisabolyl cation as an intermediate. Sometimes there are

only small cyclization differences between the minor and in the main product, due to the premature termination of the reaction mechanism of the intermediate product (Iijima, 2004).

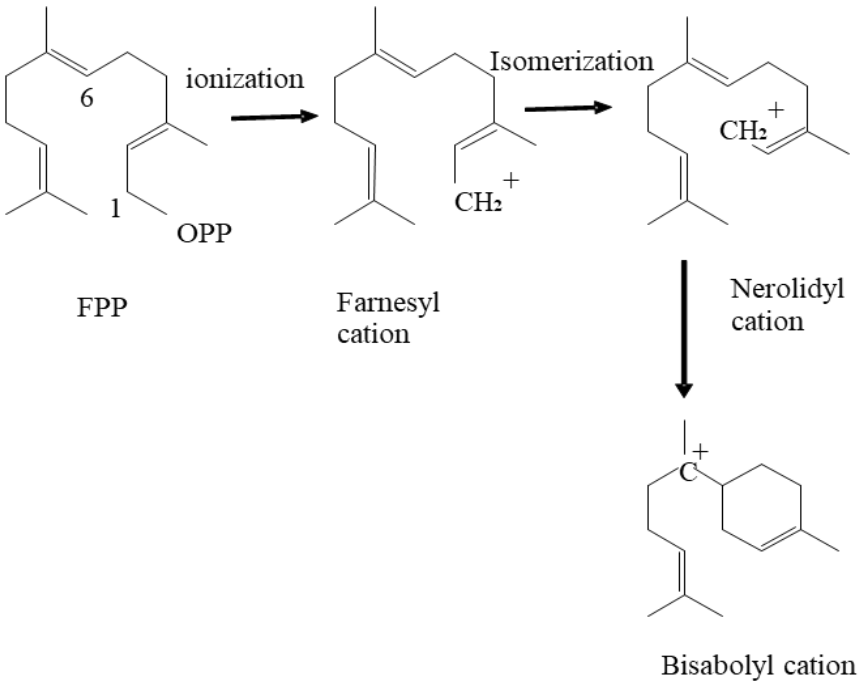


Fig. 4.3 The α-bisabolol synthase catalyzed the generation of α-bisabolol from FPP.
The reaction mechanism stats with the ionization of FPP. The resulting carbocation undergoes isomerization to a nerolidyl intermediate. The cation is quenched by a solvent molecule that is trapped in the active site along with the substrate. (Reaction mechanism adopted from Degenhardt et al., 2009).

The N and C-terminal sites are crucial for the enzymatic activation of the substrate. The conserved aspartate-rich region (DDxxD), (amino acids 324-328 at Fig. 4.4) is the site for substrate binding. The NSE/DTE metal binding motif is also important for the activity. The substitution of metal- ligand serine with non-ligand binding residue glycine in this motif is occasionally observed in plant terpenoid cyclases, so the central serine or threonine residue in the NSE/DTE metal binding motif is occasionally dispensable for Mg^{2+} binding (Christianson, 2017).

4.3.1 Does the N-terminal motif influence MrTPS7 alleles catalytic activity?

To understand the role of N-terminal peptide (positioned 29 bp upstream of Rxx8W) of MrTPS7 allele of GER, we expressed an enzyme with 29 aa N-terminal deletion. After expression and purification of the recombinant protein, we did not find any enzyme

activity (supplementary Fig. S5a and S5b). The result suggested that the predicted N-terminal transit peptide motif directly influences the activity or stability of the MrTPS7 allele. The finding is in agreement with Son et al., 2014 where two N-terminal deletions of MrBBS were tested. After an N-terminal deletion of 23 aa, no recombinant protein was detected after expression in a bacterial system. But an N-terminal deletion of 52 aa let to the very low activity of MrBBs (Son et al., 2014) .

4.3.2 Specific amino acid residues at the active site influences enzyme activity

There are different conserved motifs which are influencing the enzymatic activity and proper folding of the protein. *In silico* model of three MrTPS7 alleles might provide a proper explanation for the substrate binding sites and reaction mechanisms of the enzymes (Fig.3.17).

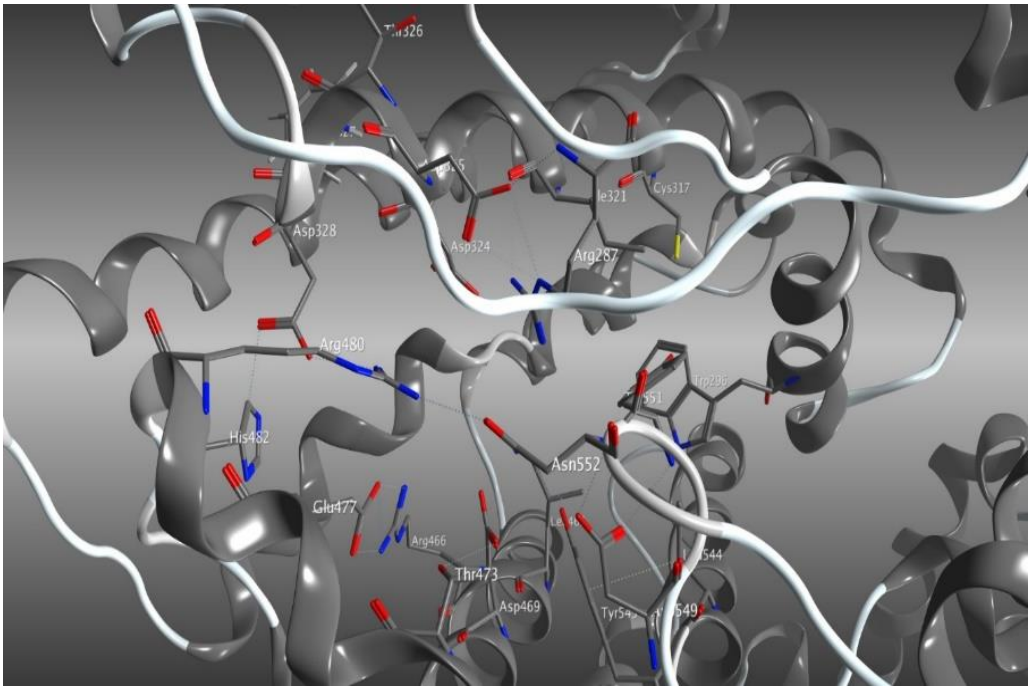


Fig. 4.4 Computational modeling of MrTPS7 allele of accession GER

The C terminal domain has a hydrophobic pocket at the active site and is closed off towards the outside by two loops which are located on the protein surface. The conserved tyrosine residue near to the active center of the enzymes is important for the protonation process (Tyr 551, Fig. 3.17). In the computational model, Tyr 551 in GER

and GO were closer to the active site. This may increase the chances of the formation of the farnesyl cation compared with MAT (in MAT, it is distantly located from the active site Fig. 3.17).

Additionally, a rigid secondary structure is noticed (Fig. 3.18) in front of substrate binding site of MAT. It is completely, differently outstretched than the other two GER and GO. This might create a barrier for the substrate to smooth mobilization at the active site and the enzymatic reaction would not be as fast as in GER and GO. Therefore, we found a higher substrate K_m value for the enzyme of the MAT (K_m 12.32 µM) compared to the other two GER (K_m 3.59 µM) and GO (K_m 8.30 µM). We assume that tyrosine residue (tyr-551) near to the active site might have a significant role for the catalytic activities. The role of 'tyrosine' residue at the active site to form (*S*)-bisabolene intermediate in maize with TPS4 and TPS11 was studied earlier by (Köllner et al., 2008). A site specific mutagenesis investigation was not included here to explain the role of this tyrosine residue. However, a superimposed computational model can already predict the structural differences within the enzymes of the different MrTPS7 alleles especially focusing on the active binding site.

5 Summary

The essential oil of chamomile (*Matricaria recutita* L.) has numerous applications in cosmetics, food, and the pharmaceutical industry. The flower heads of chamomile are especially rich in terpenes and their oxidative derivatives which have anti-inflammatory and antispasmodic properties. The pharmacological effect is mostly attributed to α-bisabolol, the major sesquiterpene of the essential oil. Therefore, α-bisabolol and its oxides are the main interest of this study. To determine the key factors in the regulation of α-bisabolol production, the biosynthesis and regulation of terpenes were analyzed in a group of sixteen chamomile accessions. First, the variation of terpenes and other volatile compounds was studied during the development from seedling to flowering plant. Each accession produced a complex blend of monoterpenes, sesquiterpenes and ethers that changed throughout the development of the plant. This blend of compounds also differed substantially among the accessions. An analysis of flower development showed the highest levels of α-bisabolol and its derivatives is in mature flowers where concentrations between 763.96 µg/g (Germania) and 0.882 µg/g (Camoflora) were found.

To investigate the molecular mechanisms controlling α-bisabolol production in flowers of chamomile, we identified cDNA sequences with high sequence identity to α-bisabolol synthase from the sixteen chamomile accessions. These terpene synthases catalyze the formation of the α-bisabolol basic terpene skeletons from the FPP substrate. While most of the putative a-bisabolol synthases had above 90% amino acid sequence identity and are very likely alleles of the previously identified α-bisabolol synthase, two had a sequence identity around 70% and one sequence from *Matricaria discoidea* only showed less than 15% identity. To test whether the identified genes encode active α-bisabolol synthases, the genes from twelve accession were expressed in a bacterial system. All encoded proteins were able to convert FPP to α-bisabolol and most of these enzymes were highly specific for this product. To assess whether the biochemical properties of these terpene synthases determine the amount of α-bisabolol produced in the respective accession, three putative MrTPS7 alleles were characterized from accessions producing low, medium, and high levels of α-bisabolol. The enzyme activities revealed a significant variation of K_m values for FPP from 12.32 µM to 3.59

µM which did not appear to coincide with levels of *in vivo* α-bisabolol production. The *in silico* protein modeling was also performed from three selective MrTPS7 alleles. The superimposed computational model revealed a structural difference in the active binding site of the substrate, as well as in the peptide backbone of the enzymes.

To test to what extend the transcript levels of the *MrTPS7* alleles determine the amount of α-bisabolol production in the respective accessions, the transcript levels of the accessions were correlated with α-bisabolol concentration. A significant correlation could not be found with α-bisabolol (Spearman's rank correlation coefficient $r_s = 0.11$). Finally, FPP levels were quantified in the flowers of all accessions. These data indicated that the concentration of FPP is between 17.69 ng/g and 0.76 ng/g. Correlation analysis indicates that FPP concentration is one of the key regulatory factors for α-bisabolol production in flowers of *M. recutita* ($r_s= 0.84$).

These results indicate that the early steps of terpene biosynthesis leading to FPP formation are the most important for the amount of α-bisabolol production. Further studies will enable us to pinpoint these early biosynthetic steps. This will facilitate to design trait-specific molecular markers for direct breeding strategies to obtain chamomile with high α-bisabolol content.

6 Zusammenfassung

Das ätherische Öl der Kamille M. recutita hat eine Vielfalt von Anwendungen in der pharmazeutischen, komsmetischen und Nahrungsmittelindustrie. Besonders die Köpfchenblüte der Kamille ist reich an Terpenen und ihren Oxidationsprodukten, mit antiseptischen, antiphlogistischen, spasmolytischen und beruhigenden Eigenschaften. Der pharmakologische Effekt wird vornehmlich α-Bisabolol zugeschrieben, dem Hauptsesquiterpen des ätherischen Öls. Diese Arbeit beschäftigt sich vorrangig mit α-Bisabolol und seinen Oxiden. Um die Schlüsselfaktoren der Regulation der α-Bisabolol-Produktion aufzuklären, wurde die Terpenbiosynthese und ihre Regulation von 16 verschiedenen Kamille Herkünften (Akzessionen) analysiert. Zuerst wurde die Variaton der Terpene und anderer volatiler Inhaltstoffe während der Entwicklung vom Sämling zur blühenden Pflanze untersucht. Jede Herkunftssorte produzierte dabei eine komplexe Mischung aus Monoterpenen, Sesquiterpenen und Ethern, die sich über den Entwicklungsverlauf der Pflanze änderte. Die Zusammensetzung unterschied sich auch deutlich zwischen verschiedenen Herkünften. Die Analyse der Blütenentwicklung zeigte, dass der höchste Gehalt von α-Bisabolol und seinen Derivaten in den reifen Blütenköpfchen zu finden ist, mit Konzentrationen von 0.882 µg/g in Camoflora bis 763,96 µg/g α-Bisabolol in Germania.

Zur Untersuchung der molekularen Mechanismen, die die Produktion von α-Bisabolol in der Kamillenblüte kontrollieren, wurden cDNA Sequenzen mit hoher Sequenzidentität zu α-Bisabolol Synthasen in 16 Kamilleakzessionen isoliert und identifiziert. Diese Terpensynthasen katalysieren die Bildung der Terpengrundstruktur für α-Bisabolol ausgehend vom Substrat FPP. Während die meisten putativen α-Bisabolol Synthasen eine über 90% Aminosäuressequenzidentität aufweisen und mit großer Wahrscheinlichkeit Allele der zuvor isolierten α-Bisabolol Synthase MrTPS7 sind, haben zwei von ihnen Sequenzenidentitäten von 70% und eine Sequenz von Matricaria discoidea nur 15% Identität. Um zu überprüfen, ob die identifizierten Gene aktive α-Bisabolol Synthasen kodieren, wurden die Gene von 12 Akzessionen in einem bakteriellen Expressionssystem getestet. Alle kodierten Proteine waren enzymatisch aktiv und konnten FPP zu α-Bisabolol umwandeln. Die meisten dieser Enzyme waren hochspezifisch für dieses Produkt. Um festzustellen, ob die biochemischen

Eigenschaften dieser Terpensynthasen aus den jeweiligen Akzessionen auch für die jeweils produzierten Mengen von α-Bisabolol in der Blüte verantwortlich sind, wurden drei MrTPS7 Allele von Akzessionen, die niedrige, mittelere oder hohe Werte an α-Bisabolol aufwiesen, untersucht. Die Enzymaktivitäten ergaben eine signifikante Variation der Km - Werte von 12,32 µM bis 3,59 µM für die Bindung von FPP. Die Km -Werte korrelierten allerdings nicht mit der in-vivo Produktion von α-Bisabolol in den Blüten. Eine in-silico Proteinmodellierung dieser drei MrTPS7 Allele zeigte strukturelle Differenzen an der Bindestelle des Substrates am aktiven Zentrum und im Peptidrückgrat der Proteine.

Eine Transkriptionsanalyse sollte zeigen, ob die Transkriptmenge der betreffenden MrTPS7 Allele mit der α-Bisabolol-Produktion in den jeweiligen Akzessionen korreliert. Es konnte keine signifikante Korrelation der Transkription mit der Konzentration von α-Bisabolol oder dessen oxidativen Derivaten (rs = 0.11) festgestellt werden. Zum Schluß wurden die Konzentrationen von FPP in den Blüten aller Akzessionen gemessen. Die FPP-Konzentrationen lagen zwischen 0,76 ng/g und 17,69 ng/g. Eine Korrelationsanalyse ergab, dass die FPP-Konzentration eine Korrelation von rs= 0.84 mit der α-Bisabolol-Konzentration in den Blüten hat. Die Konzentration des FPP-Substrates in der Blüte ist somit der ausschlagebende Faktor für die Menge der α-Bisabolol Produktion.

Insgesamt zeigen diese Ergebnisse, dass die frühen Schritte in der Terpenbiosynthese, die zur Bildung von FPP führen, entscheidend für die Produktion von α-Bisabolol in der Kamillenblüte sind. Weitere Studien an den frühen Schritten des Biosyntheseweges können Aufschluss darüber bringen, welcher Schritt die α-Bisabolol-Produktion kontrolliert. Diese Erkenntnisse dienen zur Entwicklung spezifischer molekularer Marker für gerichtete Züchtungsstrategien, um Kamillesorten mit einem hohen α-Bisabolol zu erhalten.

7 Bibliography

Adams, R.P., 2007. Identification of Essential Oil Components By Gas Chromatography / Mass Spectrometry e, 4th ed. Allured Pub Corp, USA.

Andreucci, A.C., Ciccarelli, D., Desideri, I., Pagni, A.M., 2008. Glandular hairs and secretory ducts of Matricaria chamomilla (Asteraceae): morphology and histochemistry 3847, 11–18.

Aharoni, A., 2004. Gain and Loss of Fruit Flavor Compounds Produced by Wild and Cultivated Strawberry Species. Plant Cell Online 16, 3110–3131. doi:10.1105/tpc.104.023895

Alchanati, I., Patel, J.A.A., Liu, J., Benedict, C.R., Stipanovic, R.D., Bell, A.A., Cui, Y., Magill, C.W., 1998. The enzymatic cyclization of nerolidyl diphosphate by δ-cadinene synthase from cotton stele tissue infected with Verticillium dahliae. Phytochemistry 47, 961–967. doi:10.1016/S0031-9422(97)00644-4

Alok Krishna Gopal R. Mallavarapu, S.R., 2004. Composition of the Essential Oils of the Leaves and Flowers of Tagetes erecta L. J. Essent. Oil Res. 16. doi:https://doi.org/10.1080/10412905.2004.9698786

Alonso, W.R., Rajaonarivony, J.I.M., Gershenzon, J., Croteau, R., 1992. Characterization and Mechanism of (4S)-Limonene Synthase, A Monoterpene Cyclase from the Glandular Trichomes of peppermint. J. Biol. Chem. 267, 7582–7587.

Alves Gomes Albertti, L., Delatte, T.L., Souza de Farias, K., Galdi Boaretto, A., Verstappen, F., van Houwelingen, A., Cankar, K., Carollo, C.A., Bouwmeester, H.J., Beekwilder, J., 2018. Identification of the Bisabolol Synthase in the Endangered Candeia Tree (Eremanthus erythropappus (DC) McLeisch). Front. Plant Sci. 9. doi:10.3389/fpls.2018.01340

an Southwell, R.L., 1999. Tea Tree: The Genus Melaleucatle. CRC press.

Andreas Plescher, 2004. Abiotic and biotic stress affecting the common Chamomile (Matricaria recutia L.) and the Roman chamomile (Chamaemelum nobile L. syn. Anthemis bobilis L.), in: Rolf Franke and Heinze Schilcher (Ed.), Chamomile Industrial Profile. NW, pp. 169–172.

Andrzejewska, J., Woropaj-Janczak, M., 2014. German chamomile performance after stubble catch crops and response to nitrogen fertilization. Ind. Crops Prod. 62, 350–358. doi:10.1016/j.indcrop.2014.09.004

Applequist, W.L., 2002. A reassessment of the nomenclature of Matricaria L. and Tripleurospermum Sch. Bip. (Asteraceae). Taxon 51, 757–761. doi:10.2307/1555032

Ara Der Marderosian, L.E.L., 1989. Natural Product Medicine. A Scientific Guide to Foods, Drugs, Cosmetics. G.F. Stickley; n edition (1988). doi:https://doi.org/10.1086/416216

Atta-ur-Rahman, Nakatsu [xc*], T., Lupo, A.T., Chinn, J.W., Kang, R.K.L., 2000. Biological activity of essential oils and their constituents. Stud. Nat. Prod. Chem. 21, 571–631. doi:10.1016/S1572-5995(00)80014-9

Avallone, R., Zanoli, P., Puia, G., Kleinschnitz, M., Schreier, P., Baraldi, M., 2000. Pharmacological profile of apigenin, a flavonoid isolated from Matricaria chamomilla. Biochem. Pharmacol. 59, 1387–1394. doi:10.1016/S0006-2952(00)00264-1

Azizi, M., Bos, R., Woerdenbag, H.J., Kayser, O., 2007. A comparative study of four chamomile cultivars cultivated in Iran. Acta Hortic. 749, 93–96. doi:10.17660/ActaHortic.2007.749.7

Babenko, N.A., Shakhova, E.G., 2006. Effects of Chamomilla recutita flavonoids on age-related liver sphingolipid turnover in rats. Exp. Gerontol. 41, 32–39. doi:10.1016/j.exger.2005.08.008

Baczek-Kwinta, R., Adamska, A., Seidler-Łozykowska, K., Tokarz, K., 2010. Does the rate of German chamomile growth and development influence the response of plants to soil drought? Biologia (Bratisl). 65, 837–842. doi:10.2478/s11756-010-0084-9

Baghalian, K., Haghiry, A., Naghavi, M.R., Mohammadi, A., 2008. Effect of saline irrigation water on agronomical and phytochemical characters of chamomile (Matricaria recutita L.). Sci. Hortic. (Amsterdam). 116, 437–441. doi:10.1016/j.scienta.2008.02.014

Banchio, E., Zygadlo, J., Valladares, G.R., 2005. Quantitative variations in the essential oil of Minthostachys mollis (kunth.) griseb. in response to insects with different

feeding habits. J. Agric. Food Chem. 53, 6903–6906. doi:10.1021/jf051157j

Biochemistry, A.O.F., 1987. metabolism of monterpenes: specificity of the Dehydrogenases responsible for the biosynthesis of camphor, 3 thujone and 3 isothujone 258, 287–291.

Bohlmann, J., Meyer-Gauen, G., Croteau , R., 1998. Plant terpenoid synthases: Molecular biology and phylogenetic analysis. Proc. Natl. Acad. Sci. 95, 4126–4133. doi:10.1073/pnas.95.8.4126

Bohlmann, J., Steele, C., 1997. Monoterpene synthases from grand fir (Abies grandis). J. Biol. Chem. 272, 21784–21792. doi:10.1074/jbc.272.35.21784

Brabandt, H., Ehlert, D., 2011. Chamomile harvesters: A review. Ind. Crops Prod. 34, 818–824. doi:10.1016/j.indcrop.2011.03.017

Bryant JP, J.-T.R., 1995. Ontogenic development of chemical defense by seedling resin birch: Energy cost of defense production. J Chem Ecol. 21, 883–896. doi:10.1007/BF02033796

Buchanan, B., Gruissem, W., Jones, R., 2000. Biochemistry and Molecular Biology of Plants 52–108.

Cane, D.E., Tandon, M., 1995. Epicubenol Synthase and the Stereochemistry of the Enzymic Cyclization of Farnesyl and Nerolidyl Diphosphate. J. Am. Chem. Soc. 117, 5602–5603. doi:10.1021/ja00125a029

Cane, D.E., Xue, Q., Van Epp, J.E., Tsantrizos, Y.S., 1996. Enzymatic formation of isochamigrene, a novel sesquiterpene, by alteration of the aspartate-rich region of trichodiene synthase. J. Am. Chem. Soc. 118, 8499–8500. doi:10.1021/ja961897w

Ceccarini, L., Macchia, M., Flamini, G., Cioni, P.L., Caponi, C., Morelli, I., 2004. Essential oil composition of Helianthus annuus L. leaves and heads of two cultivated hybrids “Carlos” and “Florom 350.” Ind. Crops Prod. 19, 13–17. doi:10.1016/S0926-6690(03)00076-1

Christianson, D.W., 2017. Structural and Chemical Biology of Terpenoid Cyclases. Chem. Rev. 117, 11570–11648. doi:10.1021/acs.chemrev.7b00287

Colby, S.M., Alonso, W.R., Katahira, E.J., McGarvey, D.J., Croteau, R., 1993. 4S-limonene synthase from the oil glands of spearmint (Mentha spicata). cDNA isolation, characterization, and bacterial expression of the catalytically active

monoterpene cyclase. J. Biol. Chem. 268, 23016–23024.

Crocoll, C., Asbach, J., Novak, J., Gershenzon, J., Degenhardt, J., 2010. Terpene synthases of oregano (Origanum vulgare L.) and their roles in the pathway and regulation of terpene biosynthesis. Plant Mol. Biol. 73, 587–603. doi:10.1007/s11103-010-9636-1

Croteau RB, Wheeler CJ, Cane DE, Ebert R, H.H., 1987. Isotopically sensitive branching in the formation of cyclic monoterpenes: proof that (-)-alpha-pinene and (-)-beta-pinene are synthesized by the same monoterpene cyclase via deprotonation of a common intermediate. Biochemistry 26, 5383–5389.

Das, M., 2014. Genetics and Breeding of Chamomile, in: Chamomile Medicinal, Biochemical and Agricultural Aspects. Taylor & Francis, New York, pp. 217–219.

David C. Williams, Douglas J. McGarvey, Eva J. Katahira, and R.C., 1998. Truncation of limonene synthase preprotein provides a fully active "pseudomature" form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochemistry 37, 12213–12220. doi:10.1021/bi980854k

David E. Cane, Qun Xue, and B.C.F., 1996. Trichodiene Synthase. Probing the Role of the Highly Conserved Aspartate-Rich Region by Site-Directed Mutagenesis. Biochemistry 35, 12369–12376. doi:10.1021/bi961344y

De Pascual Teresa, J., Ovejero, J., Anaya, J., Caballero, E., Hernández, J., Caballero, M., 2007. Chemical Composition of the Spanish Spike Oil. Planta Med. 55, 398–398. doi:10.1055/s-2006-962044

de Souza, A.T., Benazzi, T.L., Grings, M.B., Cabral, V., Antônio da Silva, E., Cardozo-Filho, L., Ceva Antunes, O.A., 2008. Supercritical extraction process and phase equilibrium of Candeia (Eremanthus erythropappus) oil using supercritical carbon dioxide. J. Supercrit. Fluids 47, 182–187. doi:10.1016/j.supflu.2008.08.001

Degenhardt, J., Köllner, T.G., Gershenzon, J., 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70, 1621–1637. doi:10.1016/j.phytochem.2009.07.030

Douglas, D.J.J., John, J.J.J., Newton, A.P.N.A.P., 2018. 1991 , English , Book , Illustrated edition : Eucalyptus leaf oils : use , chemistry , distillation and marketing / D . J . Boland , co-ordinating editor ; J . J . Brophy , chemical editor ; A . P . N . House , technical editor . 4–6.

Dudareva, N., 2003. (E)-beta-Ocimene and Myrcene Synthase Genes of Floral Scent Biosynthesis in Snapdragon: Function and Expression of Three Terpene Synthase Genes of a New Terpene Synthase Subfamily. Plant Cell Online 15, 1227–1241. doi:10.1105/tpc.011015

Dudareva, N., 1996. Evolution of Floral Scent in Clarkia: Novel Patterns of S-Linalool Synthase Gene Expression in the C. breweri Flower. Plant Cell Online 8, 1137–1148. doi:10.1105/tpc.8.7.1137

Evans, W.C., 2009. Pharmacognosy 16th edition, SIXTEENTH. ed. Elsevier, New York, USA.

Evans, W.C., 2009. Pharmacognosy 16th edition, SIXTEENTH. ed. Elsevier, New York, USA.

Falara, V., Akhtar, T.A., Nguyen, T.T.H., Spyropoulou, E.A., Bleeker, P.M., Schauvinhold, I., Matsuba, Y., Bonini, M.E., Schilmiller, A.L., Last, R.L., Schuurink, R.C., Pichersky, E., 2011. The Tomato Terpene Synthase Gene Family. Plant Physiol. 157, 770–789. doi:10.1104/pp.111.179648

Fidler, P., Loprinzi, C.L., O'Fallon, J.R., Leitch, J.M., Lee, J.K., Hayes, D.L., Novotny, P., Clemens-Schutjer, D., Bartel, J., Michalak, J.C., 1996. Prospective evaluation of a chamomile mouthwash for prevention of 5-FU-induced oral mucositis. Cancer 77, 522–525. doi:10.1002/(SICI)1097-0142(19960201)77:3<522::AID-CNCR14>3.0.CO;2-6

Franz C, I.O., 1986. Tetraploid, bisabolol rich chamomile. p. 64.

Frost, R.G., West, C.A., 1977. Properties of Kaurene Synthetase from Marah macrocarpus. Plant Physiol. 59, 22–9.

Garms, S., Köllner, T.G., Boland, W., 2010. A multiproduct terpene synthase from medicago truncatula generates cadalane sesquiterpenes via two different mechanisms. J. Org. Chem. 75, 5590–5600. doi:10.1021/jo100917c

Ghasemi, M., Jelodar, N., Modarresi, M., Bagheri, N., Jamali, A., 2017. Increase of Chamazulene and α-Bisabolol Contents of the Essential Oil of German Chamomile (Matricaria chamomilla L.) Using Salicylic Acid Treatments under Normal and Heat Stress Conditions. Foods 6, 11. doi:10.3390/foods6020011

Gill, C., Hippel, P.H. Von, 1989. Calculation of Protein Extinction Coefficients from Amino Acid Sequence Data 326, 319–326.

Grayer, R.E.J., Kite, G.C., Goldstone, F.J., Bryan, S.E., Patont, A., Putievsky, E.L.I., Gardens, R.B., Tw, S., 1996. Infraspecific taxonomy and essential oil chemotypes in sweet basil, Ocimum basilicum. Phytochemistry 43, 1033–9. doi:10.1016/S0031-9422(96)00429-3

Guitton, Y., Nicolè, F., Moja, S., Valot, N., Legrand, S., Jullien, F., Legendre, L., 2010. Differential accumulation of volatile terpene and terpene synthase mRNAs during lavender (Lavandula angustifolia and L. x intermedia) inflorescence development. Physiol. Plant. 138, 150–163. doi:10.1111/j.1399-3054.2009.01315.x

H. Vuorela, Y. Holm, R. Hiltunen, T. Harvala, A.L., 1990. Extraction of the Volatile Oil in Chamomile Flowerheads Using Supercritical Carbon Dioxide. Flavour Fragr. J. 5, 81–84. doi:10.1002/3527606726

Hallahan, T.W., Croteau, R., 1988. Monoterpene biosynthesis: Demonstration of a geranyl pyrophosphate: Sabinene hydrate cyclase in soluble enzyme preparations from sweet marjoram (Majorana hortensis). Arch. Biochem. Biophys. 264, 618–631. doi:10.1016/0003-9861(88)90328-1

Hänsel, R., Sticher, O., Steinegger, E., 1999. Pharmakognosie - Phytopharmazie Inhaltsverzeichnis.

Han, G.H., Kim, S.K., Yoon, P.K.-S., Kang, Y., Kim, B.S., Fu, Y., Sung, B.H., Jung, H.C., Lee, D.-H., Kim, S.-W., Lee, S.-G., 2016. Fermentative production and direct extraction of (−)-α-bisabolol in metabolically engineered Escherichia coli. Microb. Cell Fact. 15, 185. doi:10.1186/s12934-016-0588-2

Hanlidou, E., Kokkini, S., Bosabalidis, A.M., Bessière, J.M., 1991. Glandular trichomes and essential oil constituents of Calamintha menthifolia (Lamiaceae). Plant Syst. Evol. 177, 17–26. doi:10.1007/BF00937823

Hardman, R., 1998. Caraway. The Genus Carum 200.

Huchelmann, A., Boutry, M., Hachez, C., 2017. Plant Glandular Trichomes: Natural Cell Factories of High Biotechnological Interest. Plant Physiol. 175, 6–22. doi:10.1104/pp.17.00727

Hüsnü, K., Başer, C., Demirci, F., 2007. Chemistry of essential oils, in: Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability. Springer, Berlin Heidelberg New York, pp. 43–86. doi:DOI 10.1007/b136889

Iijima, Y., 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, 3724–3736. doi:10.1104/pp.104.051318

Irmisch, S., Krause, S.T., Kunert, G., Gershenzon, J., Degenhardt, J., Köllner, T.G., 2012. The organ-specific expression of terpene synthase genes contributes to the terpene hydrocarbon composition of chamomile essential oils. BMC Plant Biol. 12. doi:10.1186/1471-2229-12-84

J. D. Connolly and R. A. Hill, 1991. Dictionary of terpenoids. Chapman & Hall, London. doi:https://doi.org/10.1002/ffj.2730070418

Jones, C.G., Ghisalberti, E.L., Plummer, J.A., Barbour, E.L., 2006. Quantitative co-occurrence of sesquiterpenes; a tool for elucidating their biosynthesis in Indian sandalwood, Santalum album. Phytochemistry 67, 2463–2468. doi:10.1016/j.phytochem.2006.09.013

Jones, C.G., Moniodis, J., Zulak, K.G., Scaffidi, A., Plummer, J.A., Ghisalberti, E.L., Barbour, E.L., Bohlmann, J., 2011. Sandalwood fragrance biosynthesis involves sesquiterpene synthases of both the terpene synthase (TPS)-a and TPS-b subfamilies, including santalene synthases. J. Biol. Chem. 286, 17445–17454. doi:10.1074/jbc.M111.231787

Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, W., Ninga, E., Katerelos, N.A., Anssour, S., Dunwell, J.M., Degenhardt, J., Makris, A.M., Goodenough, P.W., Johnson, C.B., 2007. Rational Conversion of Substrate and Product Specificity in a *Salvia* Monoterpene Synthase: Structural Insights into the Evolution of Terpene Synthase Function. Plant Cell 19, 1994–2005. doi:10.1105/tpc.106.047779

Kenmogne, M., Prost, E., Harakat, D., Jacquier, M.J., Frédérich, M., Sondengam, L.B., Zèches, M., Waffo-Téguo, P., 2005. Five labdane diterpenoids from the seeds of Aframomum zambesiacum. Phytochemistry 67, 433–438. doi:10.1016/j.phytochem.2005.10.015

Kessler, A., Baldwin, I.T., 2002. PLANT RESPONSES TO INSECT HERBIVORY: The EmergingMolecular Analysis. Annu. Rev. Plant Biol. 53, 299–328. doi:10.1146/annurev.arplant.53.100301.135207

Kintzios, S.E., 2002. Oregano: The genera Origanum and Lippia, First. ed. CRC Press.

Köllner, T.G., Held, M., Lenk, C., Hiltpold, I., Turlings, T.C.J., Gershenzon, J., Degenhardt, J., 2008. A Maize (E)- -Caryophyllene Synthase Implicated in Indirect Defense Responses against Herbivores Is Not Expressed in Most American Maize

Varieties. Plant Cell Online 20, 482–494. doi:10.1105/tpc.107.051672

Köllner Schnee, C., Gershenzon, J., Degenhardt, J., T.G., 2004. The Variability of Sesquiterpenes Emitted from Two Zea mays\nCultivars Is Controlled by Allelic Variation of Two Terpene\nSynthase Genes Encoding Stereoselective Multiple\nProduct Enzymes. Plant Cell 16, 1115–1131. doi:10.1105/tpc.019877.

Köllner, T.G., Schnee, C., Li, S., Svatoš, A., Schneider, B., Gershenzon, J., Degenhardt, J., 2008. Protonation of a neutral (S)-β-bisabolene intermediate is involved in (S)-β-macrocarpene formation by the maize sesquiterpene synthases TPS6 and TPS11. J. Biol. Chem. 283, 20779–20788. doi:10.1074/jbc.M802682200

Kováčik, J., Bačkor, M., 2008. Oxidative status of Matricaria chamomilla plants related to cadmium and copper uptake. Ecotoxicology 17, 471–479. doi:10.1007/s10646-008-0200-6

Kováčik, J., Tomko, J., Bačkor, M., Repčák, M., 2006. Matricaria chamomilla is not a hyperaccumulator, but tolerant to cadmium stress. Plant Growth Regul. 50, 239–247. doi:10.1007/s10725-006-9141-3

Krause, S., 2016. Biosynthesis of oxygenated monoterpenes in Thyme , Melaleuca , and Eucalyptus species.

Kusano, M., Fukushima, A., Redestig, H., Saito, K., 2011. Metabolomic approaches toward understanding nitrogen metabolism in plants. J. Exp. Bot. 62, 1439–1453. doi:10.1093/jxb/erq417

Lanka, S., Lanka, S.R.I., Heller, K., Abbildung, L.W., Adulti, D., Tiere, D., Weibchen, D., Ovipositor, D., 2014. Zur Genetik der Bisaboloide bei der Kamille 312, 3–10.

Lanznaster, N., Croteau, R., 1991. Dye-ligand and immobilized metal ion interaction chromatography for the purification of enzymes of prenyl pyrophosphate metabolism. Protein Expr. Purif. 2, 69–74. doi:10.1016/1046-5928(91)90013-9

Lawrence B.M., 1987. Progress in essential oils. Perfum. Flavorist 12, 35–52.

Lesburg, C.A., Zhai, G., Cane, D.E., Christianson, D.W., 2007. Crystal Structure of Pentalenene Synthase : Mechanistic Insights on Terpenoid Cyclization Reactions in Biology Crystal Structure of Pentalenene Synthase : Mechanistic Insights on Terpenoid Cyclization Reactions in Biology 1820, 1820–1825. doi:10.1126/science.277.5333.1820

Letchamo, W., 2015. A comparative Study of Camomile Yield, Essential Oil and

Flavonoids Content Under Two Sowing Seasons and Nitrogen Levels. Acta Hortic. doi:10.17660/actahortic.1992.306.51

Lewinsohn, E., Gijzen, M., Croteau, R., 1992. Wound-inducible pinene cyclase from grand fir: Purification, characterization, and renaturation after SDS-PAGE. Arch. Biochem. Biophys. 293, 167–173. doi:10.1016/0003-9861(92)90380-F

Liao, P., Hemmerlin, A., Bach, T.J., Chye, M.L., 2016. The potential of the mevalonate pathway for enhanced isoprenoid production. Biotechnol. Adv. 34, 697–713. doi:10.1016/j.biotechadv.2016.03.005

Little, D.B., Croteau, R.B., 2002. Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases d-selinene synthese and g-humulene synthase. Arch. Biochem. Biophys. 402, 120–135.

Lucknow, C.I. of M. and A.P., 1995. CIAMP Newsletter. CIAMP Newsl. 22.

Mahmoud, A.A., Al-Shihry, S.S., Son, B.W., 2005. Diterpenoid quinones from Rosemary (Rosmarinus officinalis L.). Phytochemistry 66, 1685–1690. doi:10.1016/j.phytochem.2005.04.041

Maksymiec, W., 2007. Signaling responses in plants to heavy metal stress. Acta Physiol. Plant. 29, 177–187. doi:10.1007/s11738-007-0036-3

Mann, C., Staba, E.J., 1986. The chemistry, pharmacology and commercial formulations of chamomile. Herbs, Spices Med. Plants Recent Adv. Bot. Hortic. Pharmacol. Vol. 1 235–280.

Matarese, F., Cuzzola, A., Scalabrelli, G., D'Onofrio, C., 2014. Expression of terpene synthase genes associated with the formation of volatiles in different organs of Vitis vinifera. Phytochemistry 105, 12–24. doi:10.1016/j.phytochem.2014.06.007

Martin, D.M., Bohlmann, J., 2004. Identification of Vitis vinifera (-)-α-terpineol synthase by in silico screening of full-length cDNA ESTs and functional characterization of recombinant terpene synthase. Phytochemistry 65, 1223–1229. doi:10.1016/j.phytochem.2004.03.018

Matt, P., Krapp, A., Haake, V., Mock, H.P., Stitt, M., 2002. Decreased Rubisco activity leads to dramatic changes of nitrate metabolism, amino acid metabolism and the levels of phenylpropanoids and nicotine in tobacco antisense RBCS transformants.

Plant J. 30, 663–677. doi:10.1046/j.1365-313X.2002.01323.x

McCaskill, D., Croteau, R., 1997. Prospects for the bioengineering of isoprenoid biosynthesis. Adv. Biochem. Eng. Biotechnol. 55, 107–46. doi:10.1007/BFb0102060

Misawa, N., 2010. Comprehensive Natural Products II. Compr. Nat. Prod. II 733–753. doi:10.1016/B978-008045382-8.00009-5

Mohammad, R., Hamid, S., An, A., Norbert, D.K., Patrick, V.D., 2010. Effects of planting date and seedling age on agro-morphological characteristics, essential oil content and composition of German chamomile (Matricaria chamomilla L.) grown in Belgium. Ind. Crops Prod. 31, 145–152. doi:10.1016/j.indcrop.2009.09.019

Morehouse, B.R., Kumar, R.P., Matos, J.O., Olsen, S.N., Entova, S., Oprian, D.D., 2017. Functional and Structural Characterization of a (+)-Limonene Synthase from Citrus sinensis. Biochemistry 56, 1706–1715. doi:10.1021/acs.biochem.7b00143

Mucositis, C., 2018. Wild Chamomile Mouth Rinse and Oral Glutamine Suspension for 1–9.

Mukherjee, P.K., Kumar, V., Mal, M., Houghton, P.J., 2007. Acorus calamus: Scientific validation of ayurvedic tradition from natural resources. Pharm. Biol. 45, 651–666. doi:10.1080/13880200701538724

Nagegowda, D.A., 2010. Plant volatile terpenoid metabolism: Biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett. 584, 2965–2973. doi:10.1016/j.febslet.2010.05.045

Naghdi Badi, H., Yazdani, D., Ali, S.M., Nazari, F., 2004. Effects of spacing and harvesting time on herbage yield and quality/quantity of oil in thyme, Thymus vulgaris L. Ind. Crops Prod. 19, 231–236. doi:10.1016/j.indcrop.2003.10.005

Nieuwenhuizen, N.J., Wang, M.Y., Matich, A.J., Green, S.A., Chen, X., Yauk, Y.K., Beuning, L.L., Nagegowda, D.A., Dudareva, N., Atkinson, R.G., 2009. Two terpene synthases are responsible for the major sesquiterpenes emitted from the flowers of kiwifruit (Actinidia deliciosa). J. Exp. Bot. 60, 3203–3219. doi:10.1093/jxb/erp162

Oka, N., Ikegami, A., Ohki, M., Sakata, K., Yagi, A., Watanabe, N., 1998. Citronellyl disaccharide glycoside as an aroma precursor from rose flowers. Phytochemistry 47, 1527–1529. doi:10.1016/S0031-9422(97)00526-8

Orav, A., Kailas, T., Ivask, K., 2001. Volatile Constituents Of Matricaria recutita L . From Estonia. Proc. Est. Acad. Sci. Chem. 50, 39–45.

Otto, L.G., Junghanns, W.R., Plescher, A., Sonnenschein, M., Sharbel, T.F., 2015. Towards breeding of triploid chamomile (Matricaria recutita L.) - Ploidy variation within German chamomile of various origins. Plant Breed. 134, 485–493. doi:10.1111/pbr.12285

Otto, L.G., Mondal, P., Brassac, J., Preiss, S., Degenhardt, J., He, S., Reif, J.C., 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. doi:10.1186/s12864-017-3991-0

Padovan, A., Keszei, A., Hassan, Y., Krause, S.T., Köllner, T.G., Degenhardt, J., Gershenzon, J., Külheim, C., Foley, W.J., 2017. Four terpene synthases contribute to the generation of chemotypes in tea tree (Melaleuca alternifolia). BMC Plant Biol. 17. doi:10.1186/s12870-017-1107-2

Pichersky, E., Noel, J., Dudareva, N., 2006. Biosynthesis of plant volatiles: Nature diversity and ingenuity. Science (80-.). 311, 808–811.

Pintore, G., Usai, M., Bradesi, P., Juliano, C., Boatto, G., Tomi, F., Chessa, M., Cerri, R., Casanova, J., 2002. Volatile constituents of different plant parts of Chamomilla recutita L. Rausch grown in the Indo-Gangetic plains. Flavour Fragr. J. 17, 9–12. doi:10.1002/ffj.1035

Prema, B.R., Bhattacharyya, P.K., 1962. Microbiological Transformation of Terpenes: II. Transformation of alpha-Pinene. Appl. Microbiol. 10, 524–8.

Prokop'ev, I.A., Filippova, G. V., Shein, A.A., Gabyshev, D. V., 2014. Impact of urban anthropogenic pollution on seed production, morphological and biochemical characteristics of chamomile, Matricaria chamomila L. Russ. J. Ecol. 45, 18–23. doi:10.1134/s106741361306009x

Raal, A., Orav, A., Arak , E., 2007. Composition of the essential oil of Salvia officinalis L. from various European countries. Nat. Prod. Res. 21, 406–411. doi:10.1080/14786410500528478

Raal, A., Orav, A., Püssa, T., Valner, C., Malmiste, B., Arak, E., 2012. Content of essential oil, terpenoids and polyphenols in commercial chamomile (Chamomilla

recutita L. Rauschert) teas from different countries. Food Chem. 131, 632–638. doi:10.1016/j.foodchem.2011.09.042

Rafieiolhossaini, M., Adams, A., Sodaeizadeh, H., Van Damme, P., De Kimpe, N., 2019. Fast Quality Assessment of German Chamomile (Matricaria chamomilla L.) by Headspace Solid-Phase Microextraction: Influence of Flower Development Stage . Nat. Prod. Commun. 7, 1934578X1200700. doi:10.1177/1934578x1200700133

Rajaonarivony, J.I.M., Gershenzon, J., Miyazaki, J., Croteau, R., 1992. Evidence for an essential histidine residue in 4S-limonene synthase and other terpene cyclases. Arch. Biochem. Biophys. 299, 77–82. doi:10.1016/0003-9861(92)90246-S

Repčák, M., Imrich, J., Franeková, M., 2001. Umbelliferone, a stress metabolite of Chamomilla recutita (L.) Rauschert. J. Plant Physiol. 158, 1085–1087. doi:10.1078/0176-1617-00284

Richter, A., Schaff, C., Zhang, Z., Lipka, A.E., Tian, F., Köllner, T.G., Schnee, C., Preiß, S., Irmisch, S., Jander, G., Boland, W., Gershenzon, J., Buckler, E.S., Degenhardt, J., 2016. Characterization of Biosynthetic Pathways for the Production of the Volatile Homoterpenes DMNT and TMTT in *Zea mays*. Plant Cell 28, 2651–2665. doi:10.1105/tpc.15.00919

Ro, D., Ehlting, J., Keeling, C.I., Lin, R., Mattheus, N., Bohlmann, J., 2006. Microarray expression pro W ling and functional characterization of AtTPS genes : Duplicated Arabidopsis thaliana sesquiterpene synthase genes At4g13280 and At4g13300 encode root-speci W c and wound-inducible (Z) - -bisabolene synthases 448, 104–116. doi:10.1016/j.abb.2005.09.019

Rolf Franke, 2005. Cultivation, in: Rolf franke, H.S. (Ed.), Chamomile Industrial Profile. Taylor & Francis, pp. 125–126.

Rolf Franke and collaborators, 2005. Cultivarion, in: Rolf Franke amd heinz Schilcher (Ed.), Chamomile Industrial Profile. Taylor and francis, p. 104.

Rosenthal, G.A., Berenbaum, M.R., 1992. Herbivores: their interactions with secondary plant metabolites. Physiol. Entomol. 17, 308. doi:10.1016/B978-0-12-597183-6.50015-2

Russell, Kathryn; Jacob, S.E., 2010. Bisabolol. Dermatitis 21, 57–58. doi:10.2310/6620.2010.09072

Rutering, J., Ilmer, M., Recio, A., Coleman, M., Vykoukal, J., Alt, E., Orleans, N., 2016. Functional Characterization and Mechanism of (4S) Limonene 5, 1–8. doi:10.4172/2157-7633.1000305.Improved

Rynkiewicz, M.J., Cane, D.E., Christianson, D.W., 2002. X-ray crystal structures of D100E trichodiene synthase and its pyrophosphate complex reveal the basis for terpene product diversity. Biochemistry 41, 1732–1741. doi:10.1021/bi011960

Sarrou, E., Martens, S., Chatzopoulou, P., 2016. Metabolite profiling and antioxidative activity of Sage (Salvia fruticosa Mill.) under the influence of genotype and harvesting period. Ind. Crops Prod. 94, 240–250. doi:10.1016/j.indcrop.2016.08.022

Sangwan, N.S., Farooqi, A.H.A., Shabih, F., Sangwan, R.S., 2001. Regulation of essential oil production in plants. Plant Growth Regul. 34, 3–21. doi:10.1023/A:1013386921596

Savage, T.J., Hatch, M.W., Croteaufl, R., 1994. Monoterpene Synthases of Pinus contortu and Related Conifers. J. Biol. Chem. 269, 4012–4020.

Schnee, C., Köllner, T.G., Gershenzon, J., 2002. The Maize Gene terpene synthase 1 Encodes a Sesquiterpene Synthase Catalyzing the Formation of (E)-β-Farnesene, (E)-Nerolidol, and (E,E)-Farnesol after Herbivore Damage. Society 130, 2049–2060. doi:10.1104/pp.008326.the

Schnee, C., Köllner, T.G., Held, M., Turlings, T.C.J., Gershenzon, J., Degenhardt, J., 2006. The products of a single maize sesquiterpene synthase form a volatile defense signal that attracts natural enemies of maize herbivores. Proc. Natl. Acad. Sci. U. S. A. 103, 1129–1134. doi:10.1073/pnas.0508027103

Scilcher, H., Novotny, L., Ubik, K., Motl, O., Herout, V., 1976. No Title. Arch. Pharm 189, 309.

Sharon-Asa, L., Shalit, M., Frydman, A., Bar, E., Holland, D., Or, E., Lavi, U., Lewinsohn, E., Eyal, Y., 2003. Citrus fruit flavor and aroma biosynthesis: Isolation, functional characterization, and developmental regulation of Cstps1, a key gene in the production of the sesquiterpene aroma compound valencene. Plant J. 36, 664–674. doi:10.1046/j.1365-313X.2003.01910.x

Shimada, T., Endo, T., Fujii, H., Hara, M., Ueda, T., Kita, M., Omura, M., 2004. Molecular cloning and functional characterization of four monoterpene synthase

genes from Citrus unshiu Marc. Plant Sci. 166, 49–58. doi:10.1016/j.plantsci.2003.07.006

Slovak, T., 2017. The Slovak gene pool of German chamomile (Matricaria recutita L.) and comparison in its parameters 1–4.

Son, Y.-J., Kwon, M., Ro, D.-K., Kim, S.-U., 2014. Enantioselective microbial synthesis of the indigenous natural product (−)-α-bisabolol by a sesquiterpene synthase from chamomile (*Matricaria recutita*). Biochem. J. 463, 239–248. doi:10.1042/BJ20140306

Srivastava, J.K., Shankar, E., Gupta, S., 2010. Chamomile: A herbal medicine of the past with a bright future (review). Mol. Med. Rep. 3, 895–901. doi:10.3892/mmr.2010.377

Steele, C.L., Crock, J., Bohlmann, J., Croteau, R., 2002. Sesquiterpene Synthases from Grand Fir (Abies grandis) . J. Biol. Chem. 273, 2078–2089. doi:10.1074/jbc.273.4.2078

Starks, C.M., Back, K., Chappell, J., Noel, J.P., Guedes, M.E.M., Kuc, J., Hammerschmidt, R., Bostock, R., Morse, M.A., Toburen, A.L., Cane, D.E., Croteau, R., Chappell, J., Back, K., Yin, S., Chappell, J., Back, K., Chappell, J., Back, K., Chappell, J., Navaza, J., Lamzin, V.X., Wilson, K.S., Aleshin, A.E., Hoffman, C., Firsov, L.M., Honzatko, R.B., Juy, M., Tarshis, L.C., Yan, M., Poulter, C.D., Sacchettini, J.C., Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C., Poulter, C.D., Cane, D.E., Xue, Q., Fitzsimons, B.C., Dougherty, D.A., Wildung, M.R., Croteau, R., Lin, X., Hezari, M., Koepp, A.E., Floss, H.G., Croteau, R., Carson, M., Bugg, C.E., Nicholls, A., Sharp, K.A., Honig, B., Jones, T.A., Zou, J.Y., Cowan, S.W., Kjeldgaard, M., 1997. Structural basis for cyclic terpene biosynthesis by tobacco 5-epi-aristolochene synthase. Science 277, 1815–1820. doi:10.1126/science.277.5333.1815

Sueramaniant, G., 1988. Volatile aroma constituents of Orange 27, 2185–2188.

Švehlíková, V., Bennett, R.N., Mellon, F.A., Needs, P.W., Piacente, S., Kroon, P.A., Bao, Y., 2004. Isolation, identification and stability of acylated derivatives of apigenin 7-O-glucoside from chamomile (Chamomilla recutita [L.] Rauschert). Phytochemistry 65, 2323–2332. doi:10.1016/j.phytochem.2004.07.011

T. C. J. Turlings, J. H. Tumlinson, W.J.L., 1990. Exploitation of hervibore-plant odors by host-seeking parasitic wasps. Science (80-.). 250, 1251–1253.

Tarshis, L.C., Yan, M., Poulter, C.D., Sacchettini, J.C., 1994. Crystal Structure of Recombinant Farnesyl Diphosphate Synthase at 2.6-.ANG. Resolution. Biochemistry 33, 10871–10877. doi:10.1021/bi00202a004

Taura, F., Morimoto, S., Shoyama, Y., 1996. Purification and characterization of cannabidiolic-acid synthase from Cannabis sativa L. Biochemical analysis of a novel enzyme that catalyzes the oxidocyclization of cannabigerolic acid to cannabidiolic acid. J. Biol. Chem. 271, 17411–17416. doi:10.1074/jbc.271.29.17411

Tetali, S.D., 2019. Terpenes and isoprenoids: a wealth of compounds for global use. Planta 249. doi:10.1007/s00425-018-3056-x

Tiwari, P., 2016. Recent advances and challenges in trichome research and essential oil biosynthesis in Mentha arvensis L. Ind. Crops Prod. 82, 141–148. doi:10.1016/j.indcrop.2015.11.069

Uno, M., Kokuryo, T., Yokoyama, Y., Senga, T., Nagino, M., 2016. α-Bisabolol Inhibits Invasiveness and Motility in Pancreatic Cancer Through KISS1R Activation. Anticancer Res. 36, 583–9.

Vernet, P., Gouyon, R.H., Valdeyron, G., 1986. Genetic control of the oil content in Thymus vulgaris L: a case of polymorphism in a biosynthetic chain. Genetica 69, 227–231. doi:10.1007/BF00133526

Vial, M. V., Rojas, C., Portilla, G., Chayet, L., Pérez, L.M., Cori, O., Bunton, C.A., 1981. Enhancement of the hydrolysis of geranyl pyrophosphate by bivalent metal ions. A model for enzymic biosynthesis of cyclic monoterpenes. Tetrahedron 37, 2351–2357. doi:10.1016/S0040-4020(01)88888-8

Wagner, C., Friedt, W., Marquard, R.A., Ordon, F., 2005. Molecular analyses on the genetic diversity and inheritance of (-)-α-bisabolol and chamazulene content in tetraploid chamomile (Chamomilla recutita (L.) Rausch.). Plant Sci. 169, 917–927. doi:10.1016/j.plantsci.2005.06.010

Wagschal, K., Savage, T.J., Croteau, R., 1991. Isotopically sensitive branching as a tool for evaluating multiple product formation by monoterpene cyclases. Tetrahedron 47, 5933–5944. doi:10.1016/S0040-4020(01)86486-3

Wang, G., 2014. Recent progress in secondary metabolism of plant glandular trichomes. Plant Biotechnol. 31, 353–361. doi:10.5511/plantbiotechnology.14.0701a

Wang, G., Tian, L., Aziz, N., Broun, P., Dai, X., He, J., King, A., Zhao, P.X., Dixon, R.A., 2008. Terpene Biosynthesis in Glandular Trichomes of Hop. Plant Physiol. 148, 1254–1266. doi:10.1104/pp.108.125187

Weber, B., Herrmann, M., Hartmann, B., Joppe, H., Schmidt, C.O., Bertram, H.J., 2008. HPLC/MS and HPLC/NMR as hyphenated techniques for accelerated characterization of the main constituents in Chamomile (Chamomilla recutita [L.] Rauschert). Eur. Food Res. Technol. 226, 755–760. doi:10.1007/s00217-007-0587-y

Weiss RF Fintelmann V, 2000. Herbal medicine. Thieme.

Wendt, K.U., Schulz, G.E., Corey, E.J., Liu, D.R., 2000. Enzyme Mechanisms for Polycyclic Triterpene Formation. Angew. Chem. Int. Ed. 39, 2812–2833. doi:10.1002/1521-3773(20000818)39:16<2812::AID-ANIE2812>3.0.CO;2

Whittington, D.A., Wise, M.L., Urbansky, M., Coates, R.M., Croteau, R.B., Christianson, D.W., 2002. Nonlinear partial differential equations and applications: Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. Proc. Natl. Acad. Sci. 99, 15375–15380. doi:10.1073/pnas.232591099

Wildung, M.R., Croteau, R., 1996. A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. J. Biol. Chem. 271, 9201–9204. doi:10.1074/jbc.271.16.9201

William Charles Evans, 2009. Trease and Evans Pharmacognosy, 16th ed. Elsevier, New York, USA.

Yu, F., Harada, H., Yamasaki, K., Okamoto, S., Hirase, S., Tanaka, Y., Misawa, N., Utsumi, R., 2008. Isolation and functional characterization of a β-eudesmol synthase, a new sesquiterpene synthase from Zingiber zerumbet Smith. FEBS Lett. 582, 565–572. doi:10.1016/j.febslet.2008.01.020

Yuba, A., Yazaki, K., Tabata, M., Honda, G., Croteau, R.B., 1996. cDNA Cloning , Characterization , and Functional Expressionof 4S-(-)-limonene synthase from Perilla frutescens. Arch. Biochem. Biophys. 332, 280–287.

7 Supplementary Materials

Table S1 Oligonucleotides used for isolation, cloning and qRT-PCR analysis of MrTPS 7 genes

Name	Sequence (5'-3')	usages
MrBBS		isolation
MrBBS1fwd	ATGTCAACTTTATCAGTTTC	isolation
MrBBS1rev	CTAGACAATCATAGGGTGAAC	isolation
MrBBS2fwd	ATGTCAACTTTATCAGTTTCTACTCCTTCC	isolation
MrBBS2rev	CTAGACAATCATAGGGTGAACC	isolation
MrBBS3fwd	TTATGTTGCCCGGGATGAC	isolation
MrBBS3 rev	TTGTTCCTCCTTGTGGGTGG	isolation
MrBBS4fwd	TCGCAACAGTGTCATCTTCC	QPCR
MrBBS4 rev	GCCAAGGCGTTCAACTACAT	QPCR
MrBBS5fwd	TGGTTTCGACTGCTACGAC	QPCR
MrBBS5 rev	TGCCAAGGTGTAGTTTGGTG	QPCR
MrBBS6fwd	CCCACAAGGAGGAACAAG	QPCR
MrBBS6rev	GTATCACTCACACGCGCAAG	QPCR
MrBBS7fwd	TGTGGTTTCGACTGCTACG	QPCR
MrBBS7rev	TGCCAAGGTGTAGTTTGGTG	QPCR
MrBBS8fwd	ATCCTCAAGGGATGCTTGC	QPCR
MrBBS8rev	CGCCTCTAACCTTGGCAAC	QPCR
MrBBS9fwd	CCACGATAGTATATGGGGGG	QPCR
MrBBS9rev	GCAAGGATTCCTCGATCTCC	QPCR
MrBBS10fwd	CAGCAGAGGAACATATGGC	QPCR
MrBBS10rev	CTGGTGGAATACTTTATGGAC	QPCR
MrBBS11fwd	GTGATGCAATTACAGACGAG	QPCR
MrBBS11rev	GATCACCCGCATGATAACAGG	QPCR
MrBBS12fwd	CTTAGACAAAGTGGTTACCG	QPCR
MrBBS12rev	GCCTCAACATCCTACGATAC	QPCR

MrBBS13fwd	CACTTCCGCTTATAATGTCATTTC	QPCR
MrBBS13rev	CTTGGCTTTAGACATCCTTCG	QPCR
Hk-actfwd	GCTAACAGGGAAAAGATGACTC	HK-QPCR
Hk-actrev	ACTGGCATAAAGAGAAAGCACG	HK-QPCR
HkUbiq-fwd	ATGATAACTCGACGGATCGC	HK-QPCR
HkUbiq-rev	CTTGGATGTGGTAGCCGTTT	HK-QPCR
IBA37-fwd	ATGGTAGGTCTCAGCGCATGTCAACTTTATCAGTTTCTACTCC	IBA primer
IBA37-rev	ATGGTAGGTCTCATATCAGACAATCATAGGGTGAACGAAGAG	IBA primer
MrBBS16fwd	TGGTTTCGACTGCTACGACA	QPCR
MrBBS16rev	CACCAAACTACACCTTGGCA	QPCR
MrBBS17fwd	CCCACAAGGAGGAACAAGAA	QPCR
MrBBS17rev	CTTGCGCGTGTGAGTGATAC	QPCR
MrBBS18fwd	TGTATTTTGAGCCAAAATATTCCCGAGCTAGGATC	Isolation
MrBBS18rev	GAATGGTTGAATGTTATTTTTGGGCACTAGGTGTG	Isolation
MrBBS19fwd	AAATAA CATTCAA CCATT CTATCT CTTGC ATACGG	Isolation
MrBBS19rev	GGAATATTTTGGCTCAAAATACACACCTAGTGCCC	Isolation
MrBBS20fwd	TCGCAACAGTGTCATCTTCC	IBA 37+
MrBBS20rev	ACCTTCCAACCATCTTCGAC	IBA 37+
MrBBS21fwd	CAGCTAATCGAGGAGCTCAAA	IBA 37+
MrBBS21rev	TATTACTCACACGCGCAAGG	IBA 37+
MrBBS22fwd	ATGTGGCTCGCAACAGTGTA	IBA 37+
MrBBS22rev	CCTTCCAACCATCTTCGACT	IBA 37+
MrBBS23fwd	TGGCCTAGCCTATCATTTCG	IBA 37+
MrBBS23rev	TTCGCCCTTCTAGACAATCA	IBA 37+
MrBBS24fwd	ATGGTAGGTCTCAGCGCCAATTTCTTGAATATAAGGAGAAATCC	IBA 37+
MrBBS24rev	ATGGTAGGTCTCATATCACTTCCAACCATCTTCGACTTGTTT	IBA 37+

MrBBS25fwd	ATGGTAGGTCTCAGCGCGAATTTCTTGAATATAAGGAGAAATCC	IBA 37+
MrBBS25rev	ATGGTAGGTCTCATATCACTTCCAACCATCTTCGACTTGCT	IBA 37+
MrBBS26fwd	ATGGTAGGTCTCAGCGCCAATTTCTTGAATATAAGGAGAAATCC	IBA 37+
MrBBS26rev	ATGGTAGGTCTCATATCAATCCATGCATGCTTCCTCCTCTG	IBA 37+
MrBBS27fwd	ATGGTAGAAGACAAGCGCCAATTTCTTGAATATAAGGAGAAATCC	IBA 37+
MrBBS27rev	ATGGTAGAAGACAATATCAACCAGTTTCCTTTCGGTAGCATTC	IBA 37+
MrBBS28fwd	ATGGTAGGTCTCAGCGCATGTCAACTTTATCAGTTTCTACTCC	IBA 37+
MrBBS28rev	ATGGTAGGTCTCATATCAAACGAAGAGCGATTTGATGTAACC	IBA 37+

Degenerate primers used for isolation of TPS fragments; IBA37+, primers used for cloning of the ORF into the vector of pASK-IBA37+; QPCR, primers used for QRT-PCR analysis, HK denoted for housekeeping.

cDNA sequences of the isolated α-bisabolol synthases (*MrTPS7* alleles)

>MrTPS7 allel (ARG)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTCCATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAGGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACTAGTTACTTGTGGA

TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTTCAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCACGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

> *MrTPS7* allel (BOD)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGGAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTCTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGACGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGCGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGGGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAA

> *MrTPS7* allel (BOHE)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAGGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCCATA
AATGGACCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCGCTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAGTGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGGGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG

CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCAAGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCGCGTGTGAGTGATACCTTAT
ATAAAGACAATGACGGCTACAATCATGCAGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

> *MrTPS7* allel (BON)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATGAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCGAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGGCAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACAATTCTC
GTACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGT
ACTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCC
AGAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGG
AGGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAG
GCCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACCAGTTACTTGTGG
ATATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAAT
GGGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACC
CACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATC
AGAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGT

>*MrTPS7* allel (CAM)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGGAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCTT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAATAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGCGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCCGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGGGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGT

>*MrTPS7* allel (GER)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCCATA
AATGGACCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGGGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCGCTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACACGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAGTATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCAAGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAGCCTTGCGCGTGTGAGTAATACCTTAT
ATAAAGACAATGACGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

> *MrTPS7* allel (GO)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGCTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATGAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAGG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCACGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

>*MrTPS7* allel (LAZ)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTACATGGGGGGATGAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA

ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAGCTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAGG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTTTTAAAGCAAGTCGAAGATGGTTGGAAGGTA

>MrTPS7 allele(LEU)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGCTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCACGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

>MrTPS7 allele(MAN)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGCTAATAAAAATAGCAC
GAAGCAACATGTGGCTCGCAACAGTGTAATATTCCACGATAGTATATGGGGGGATCAGTTTCTTGAATATA
AGGAGAAAGCCAATGTAGCAACTGAGAAACAGCTAATTGAGGAGCTCAAAGAAGAGGTGAGAAAGGAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGAGATATAAAGCTTATACAACTCATTGATGTAGTTCAACGCCT

TGGCCTAGCCTATCATTTCGAAAAAGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGTTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGAATGCTTGTTTTATACGAAGCAGTATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TTGAGTTCACCAAAATACACCTTGGCATCATATCCAAAGATTCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAGAACAAGCGCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCTGTGCGCTACATACCAAT
CTACCAACAAAAAGCTTCCCACAATGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCT
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAGCATTCTCG
TACAAGAATGTTCGTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTCACTCAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTACCA
AAGTACATGAAACTAATATATCAGGAACAGTTCCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTCTTATTAGAAG
CCAAATGGTTAAAAGAGGGGTATATGCCAACATTAGACGAGTACATGTCTAATTCACTGGTTACTTGTGGC
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGAGATGACGGTATAGTCACCGAGGATACCTTTAAATG
GGTGGCCACGCATCCTCCTATCGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACTC
ACAAGGAGGAACAAGAACGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCAACA
GAGGAGGAGGCATGCGAGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCACGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCCGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

>MrTPS7 allele(MAR)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGGGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCCATA
AATGGACCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCGCTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCAAGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATCCCTGCAATCAACCTTGCGCGTGTGAGTGATACCTTAT
ATAAAGACAATGACGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

>MrTPS7 allele(MARG)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA

AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGGAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTAGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGCGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAGGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCCCCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATAGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGGGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTAA

>MrTPS7 allele(MAT)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCCTCTGCTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGGTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACACTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAATATGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGAGATGACGGCATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCGCGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCTCTCTTCGTTCAC
CCTATGATTGTCTAGAAGGGCGAATTC

>MrTPS7 allele(MO)

TTATGTTGCCCGGGATGACGCTTGGCCTCCATGGTCTTGGAGGATAACTCATCGGTATCCACTCCGCCCAG
AGTCACCTCGGCGGTACGATAACCCTCGGTGCCGCCGGGTTGGATCTGCCAGCAGTGGAAATAACTGTCCA
GCGCAGCTATATCAGCCTTGGACAGTTGGTTAAGGTTCTTGCTGGGTAATTCACCCAGCTCCAGCAGTCTC
TCAACCAAGCGCTTAGGTAACAACCGAGATAATCCATTCTTAAGCGACAATTTTGGTGATTCAGCAGTAAT
TGTTTTGCACCAATCGACAAAATCCAGGTCCGGTAGTAAATCAAAAAGCACACTCTGGCCTGGCTGCCAGT
AGGAGGAGATCTGCAGCACAGCCGGCCCGGACAAACCTCTGTGGGTGAATAACATGGCTTCTCTAAACTGC

CTGCCGCACTCGGCTTTGGCGATAACCGGCAAGCTGATACCGGACAGGCCTTCAAACCTGGCCTTGTCTTC
AGGTTGCAAGGTAAAGGGAACCAGCCCGGCAGTGGTCGGCAGCACCTTGATGGCAAACTGCTCGGCGATGC
GGAATCCAAACGGCGTAGCCCCGAGCTTGGGCATGGACAAGCCACCAGTGGCTATCACCAGTGATTCACAA
CGGTATTCCCCTTCACTGGTTTGTAGGCTGTAACCGTCATCTGTGGCTGCCACAGACAGAATTTCGGTACG
CAGCTGAATTTTGACTCCGGCCCAATCGCACTCTGTGGTCAGCATATCAACAATGTCTTTAGCCGAGTCAT
CACAAAACAGCTGTCCGAGAGTCTTCTCGTGATAATCGATGCCGTGCCGCTCGACCATGGCGATAAAGTCC
CACTGGCTGAAACGAGCCAGCGCCGATTTGACGAAGTGTGGGTTACTGCAGAGATAGGCGCCCGGTTCAAT
GTTCTGATTGGTGAAGTTGCAGCGACCGCCGCCACTGATCAGAATTTTGCGCCCCAGTTGTTTGGCGTTAT
CCAGCAAGAGGACATCCCGGCCCCGATATCCGGCAGTGAGGGCACACATTAACCCTGCGGCGCCAGCGCCG
ATAATGATCACCTGATGCTGTTTAACTGACATAACGCTCACTCTCGGGCCCAACAAGGGGCAATAAAAAAG
GACGCTATTCTAGCATCCTTTTCCGGGATTGGAGATGGCTCAGTCTGCCGAGGTTTCCGACTTCTTCTCTC
TGGCCAGCAGTTCTTTGGCCGCATCAATAGGCGATTTGCCCTGATACAACACCTGATAGATCTGCTCGGTG
ATCGGCATCTCTACCCCCAAGCGCTTGGCCAGGGTGTAGACCTCTTTGGTGTTGCGGTAACCTTCAACCAC
CTGACCGATTTCAGCCTGGGCAGTATCGACATCACAACCCTTACCCAATGCCAAACCAAAGCGACGGTTGC
GCGATTGGTTATCGGTACAGGTCAGCACCAGGTCCCCAAGACCCGCCATTCCCATAAAGGTAGAAGCATGG
GCGCCCAAAGCTTCCCCCAGACGGCTCAACTCCACCAGACCTCGGGTGATCAGGGCTGTACGGGCATTGGC
ACCAAAGCCTATACCGTCGGACATGCCGGCACCTATGGCGATGACATTCTTGACCGCGCCACCGAGTTGCA
GGCCGATAAAATCGTCATTGGCATAGACTCTCAGACGTTTGGGGCTGTGCAACAGCTCCACCAGATCCTGG
GTAAACTCGGCATCGGTACCAGCCACTGAGATGGCGGTTGGTAAACCTGCCGCCAGCTCTTTGGCAAAAGT
CGGTCCGGACAATACCGCCAGCGGATACTTCTCCCCCAGCACATCCCGGGCAACATAA

>MrTPS7 allele(PO)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGCTAATAAGGATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACGAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGaAACAAGTTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAACCATGCATGTGGTTAATTGTTTTAGATGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATGCCAACATTAGACGAGTACCTGTCTAATTCACTAGTTACTTGTGGATATG
CATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATGGGTG
GCCACACATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCCACAA
GGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCAGAGG
AGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCTCATG
CCTACAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCGCGTGTGAGTGATACCTTATATAA
AGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCACCCTA
TGATTGTCTAGAAGGGCGAATTC

>*MrTPS7* allele(ZL)

ATGTCAACTTTATCAGTTTCTACTCCTTCCTTTTCTTCATCTCCATTGTCTTCTGTTAATAAGAATAGCAC
GAAGCAACATGTTACTCGCAACAGTGTCATCTTCCACGATAGTATATGGGGGGATCAATTTCTTGAATATA
AGGAGAAATCCAATGTAGCTACTGAGAAACAGCTAATCGAGGAGCTCAAAGAAGAAGTGAGAAACAAACTA
ATGATAAGAGCTTGTAATGAAGCAAGCCGATATATAAAGCTTATACAACTCATTGATGTAGTTGAACGCCT
TGGCCTAGCCTATCATTTTGAAAAGGAGATCGAGGAATCCTTGCAACATATCTATGTTACATATGGCGATA
AATGGATCAACTATAACAACATTGAAAGCCTTTCGCTGTGGTTTCGACTGCTACGACAAAATGGCTTCAAC
GTATCATCTGATATATTCGAGAACCATATAGATGAGAAGGGAAACTTTCAGGAATCTTTATGTAATGATCC
TCAAGGGATGCTTGCTTTATACGAAGCAGCATATATGAGGGTGGAAGGAGAAATAATACTAGATAAGGCAC
TCGAGTTCACCAAACTACACCTTGGCATCATATCCAATGATCCTTCTTGTGACTCTTCTCTAAGAACAGAA
ATAAAACAAGCTCTAAAGCAGCCACTTCGTAGAAGGTTGCCAAGGTTAGAGGCGGTGCGCTACATAGCAAT
CTACCAACAAAAAGCTTCTCACAGTGAGGTCTTGTTAAAGCTTGCAAAGTTAGACTTCAACGTGCTTCAAG
AAATGCACAAAGACGAGCTTAGCCAAATCTGCAAATGGTGGAAAGATTTGGACATTCGAAACAAGCTACCA
TATGTTCGAGACAGATTGATTGAAGGCTACTTTTGGATATTGGGAATCTATTTCGAGCCTCAACATTCTCG
TACAAGAATGTTCTTAATGAAAACATGCATGTGGTTAATTGTTTTAGACGATACATTTGATAATTATGGTA
CTTATGAGGAACTCGAGATATTTACACAAGCTGTCGAAAGATGGTCCATAACCTGCTTGGATGAGCTGCCA
GAGTACATGAAACTAATATATCATGAACAGTTTCGTGTTCACCAAGAAATGGAGGAATCACTTGAGAAGGA
GGGAAAAGCATATCAAATCCATTATATTAAGGAGATGGCGAAAGAGGGCACACGCAGCCTTTTATTAGAAG
CCAAATGGTTGAAAGAGGGATACATGCCAACATTAGACGAGTACCTGTCTAACTCACTAGTTACTTGTGGA
TATGCATTGATGACAGCAAGATCTTATGTTGCCCGGGATGACGGTATAGTCACCGAGGATGCCTTTAAATG
GGTGGCCACGCATCCTCCTATTGTGAAAGCTGCATGTAAAATTTTAAGACTTATGGATGATATTGCCACCC
ACAAGGAGGAACAAGAAAGAGGCCATATTGCTTCAAGCATTGAATGCTACCGAAAGGAAACTGGTGCATCA
GAGGAGGAAGCATGCATGGATTTCTTAAAACAAGTCGAAGATGGTTGGAAGGTTATAAATCAGGAGTCGCT
CATGCCTTCAGATGTACCATTTCCTCTCCTTATTCCTGCAATCAACCTTGCACGTGTGAGTGATACCTTAT
ATAAAGACAATGATGGCTACAATCATGCTGATAAAGAAGTCATTGGTTACATCAAATCGCTCTTCGTTCAC
CCTATGATTGTTTAG

Table S2 Metabolites concentration in stems of thirteen chamomile accessions

Chamomile accessions	Compounds (µg g^{-1} of fresh weight)							
	Cineole ± SEM	β-ocimene ± SEM	α-gurjunene ± SEM	Aristolene ± SEM	Longifolene ± SEM	β-caryophyllene ± SEM	β-farnesene ± SEM	Germacrene D ±SEM
ARG	0.000 ± 0.000	0.4208 ±0.0952	0.0270 ± 0.0184	0.0138 ± .0004	0.0737 ± 0.0451	0.2085 ± 0.0550	8.3029± 2.1489	0.1246± 0.0350
BOD	0.000 ±0.000	6.8419 ± 2.1442	0.2758 ± 0.1692	0.9365 ± 0.6514	1.1065 ± 0.9517	6.4086± 2.9484	134.4877± 23.2604	4.1729± 3.6421
BOHE	0.0000 ±0.000	2.1774 ± 0.5115	0.9529 ± 0.4800	1.4175 ±0.8160	3.0148± 2.6357	2.5594± 1.1275	75.3932± 9.6289	0.8870± 0.2090
BON	0.0000 ±0.000	5.6631 ± 3.1062	0.9321 ± 0.4811	1.0157 ± 0.1188	0.6784± 0.4230	0.4872± 0.1702	107.5927± 67.5902	4.8763± 4.2774
CAM	0.2479 ±0.000	17.2975 ± .4702	2.3466 ± 0.6771	0.1325 ± 0.1184	3.6157± 0.2621	5.0992± 2.3306	126.6098± 98.9074	9.2919± 7.2491
GER	0.0000 ±0.000	17.6411 ± .8772	2.5826 ± 0.6479	0.5295 ± 0.1033	2.3554± 0.9237	0.6568± 0.0761	149.2649± 68.7640	3.0043± 2.2306
GO	0.0000 ± 0.000	9.2754 ± 6.9682	2.8579 ± 0.8575	0.6360 ± 0.4247	5.5424± 3.5549	0.5422± 0.2163	112.5213± 22.1177	0.6386± 0.3535
LAZ	0.0000 ± 0.000	2.6622 ± 1.6546	0.6928 ± 0.5679	0.0533 ± 0.0414	0.6299± 0.2328	0.1210± 0.0482	53.0456± 39.6056	0.0624± 0.0491
LEU	0.0000 ± 0.000	2.7247 ± 1.2343	0.8256 ± 0.4145	0.8104 ±0.4286	0.2153± 0.0907	1.5578± 1.1900	88.8940± 70.7665	0.1719± 0.0557
MAN	0.0000 ± 0.000	2.4874 ± 0.8488	0.4123 ± 0.3011	0.0294 ± 0.0202	0.3286± 0.2224	1.3424± 1.2309	74.8396± 31.3374	0.4658± 0.3528
MARG	0.5923 ± 0.000	4.5726 ± 2.8094	1.2522 ± 0.9391	0.1297 ± 0.0726	0.8848± 0.4746	2.8813± 1.9510	92.8978± 50.5420	0.5062± 0.4132
PO	0.0000 ± 0.000	0.6920 ± 0.2085	0.9720 ± 0.6500	0.3036 ±0.2915	0.6651± 0.4065	0.5983± 0.5480	57.4693± 32.8922	0.2526± 0.1205
ZL	0.0000 ± 0.000	2.1647 ± 0.8888	0.7723 ± 0.5526	0.2130 ±0.0458	2.3261± 0.4140	0.9801± 0.4295	134.9906± 17.0457	0.2528± 0.0436

Chamomile accessions	Compounds (µg g^{-1} of fresh weight)				
	Bi-cyclogermacrene ± SEM	α-farnesene ±SEM	α- bisabolol ± SEM	Spiroether ±SEM	En-in dicycloether ± SEM
ARG	0.0274 ± 0.0028	0.5790±0.1890	0.1167 ±0.0000	9.4305±2.7130	0.3555 ± 0.1092
BOD	1.5409 ± 0.7906	17.3200 ± 8.4179	0.0000 ±0.0000	83.6123±18.1746	57.0017 ±3.2298
BOHE	6.6570 ± 2.6320	4.4003 ± 2.2108	0.0000 ±0.0000	124.8848±12.4646	67.4532±19.6180
BON	1.6691 ± 1.4875	23.1979 ± 18.7356	0.0000 ±0.0000	62.1535±41.1411	106.8842±58.8101
CAM	0.6743 ± 0.0346	26.2022 ± 20.7451	0.0880 ±0.0000	78.9493±35.5309	58.1434±34.3183
GER	3.0756 ± 1.7711	17.6389±12.7597	0.0000 ±0.0000	104.6409 ±63.9185	53.8134±19.2695
GO	8.4440 ± 2.1700	10.4642±9.9518	0.2861±0.0000	7.5580±6.2571	101.1091±65.5535
LAZ	0.6021 ± 0.2832	5.7831±2.9323	0.0000±0.0000	16.7851±6.5501	41.3177±21.3595
LEU	1.8948 ± 1.5202	1.9247±0.6098	0.0000±0.0000	23.0200±10.3745	43.1736±33.5090
MAN	0.3053 ± 0.1354	0.9625±0.7363	0.0000±0.0000	67.8524±33.3611	49.9391±25.4411
MARG	0.6194 ± 0.2611	7.2061±3.6575	0.0000±0.0000	37.1327±17.4302	62.0718±12.8665
PO	1.7033 ± 1.4327	7.6661±3.6018	0.0000±0.0000	8.7287±0.3471	13.8901±7.1297
ZL	0.4198 ± 0.0643	0.3288±0.0456	0.0000±0.0000	13.5997±8.1371	32.0253±35.5333

Table S3 Metabolites concentration in leaves of thirteen chamomile accessions.

Chamomile accessions	Compounds (μg g^{-1} of fresh weight)							
	Cineole ± SEM	β-ocimene ±SEM	α-gurjunene ±SEM	Aristolene ±SEM	Longifolene ±SEM	β-caryophyllene ±SEM	β-farnesene ±SEM	Germacrene D ±SEM
ARG	0.000 ± 0.000	1.0396±0.2986	0.0424±0.0046	0.0354±0.0226	0.0547±0.0241	1.5027±0.8985	2.8296±2.8122	0.2632±0.0685
BOD	0.000 ± 0.000	1.7998±0.4661	0.0875±0.0453	0.1708±0.1591	0.2011±0.1451	0.1764±0.1098	65.3819±48.8015	2.6647±1.5554
BOHE	0.000 ± 0.000	12.3966±2.1958	0.4813±0.2129	0.6171±0.1552	1.3496±1.1641	5.2041±1.6342	2.7972±0.3869	3.9555±1.3850
BONA	0.000 ± 0.000	3.0974±1.7884	0.0380±0.0156	0.0238±0.0211	0.0265±0.0143	0.0381±0.0128	3.0115±1.6344	1.2143±0.7727
CAM	0.000 ± 0.000	21.9235±12.5435	0.2428±0.0345	0.1544±0.0387	0.9573±0.3494	0.6693±0.4953	43.0343±13.6075	0.6941±0.3354
GER	0.000 ± 0.000	4.5370±0.7482	0.1133±0.0486	0.2333±0.1821	1.0116±0.7850	0.4514±0.3049	103.7370±22.5274	2.5496±0.5900
GO	0.000 ± 0.000	5.5076±2.5147	6.1117±2.0344	1.0503±0.4972	0.8123±0.1242	2.9873±1.0942	160.7517±43.2283	1.2075±0.4659
LAZ	0.000 ± 0.000	23.1016±5.3278	2.5034±0.4592	0.1389±0.1003	2.3250±1.7704	1.5913±1.2960	34.8573±31.1224	9.5297±7.7730
LEU	0.000 ± 0.000	12.9920±9.6607	0.2598±0.1083	1.8914±1.2682	0.2465±0.0358	1.4637±0.9507	105.9588±66.9404	9.8388±6.6707
MAN	0.000 ± 0.000	12.9363±12.2997	0.1569±0.1123	0.2910±0.2846	1.2642±1.2562	0.2028±0.1918	2.2091±1.0902	0.4583±0.3725
MARG	1.9864 ±1.1136	21.6934±5.1871	2.0114±1.5796	0.2521±0.0337	0.4593±0.2422	1.5615±0.7287	83.0640±2.8484	2.1505±0.1696
PO	0.000 ± 0.000	0.6895±0.1275	0.0468±0.0038	0.0073±0.0013	0.0185±0.0159	0.0507±0.0297	13.3202±2.3796	0.1051±0.0368
ZL	0.000 ± 0.000	10.1575±5.4870	0.3324±0.2392	0.7413±0.1468	0.9633±0.5012	0.6618±0.3389	26.6977±24.7743	1.2467±0.8505

Chamomile accessions	Compounds (µg g^{-1} of fresh weight)			
	Bicyclogermacrene ± SEM	α-farnesene ±SEM	Spiroether ±SEM	En-in dicycloether ±SEM
ARG	0.4421±0.2472	5.0530±2.8131	18.2424±8.4386	3.8809±1.7038
BOD	1.7928±0.5281	12.7802±8.4976	48.7783±19.2245	14.2685±12.2282
BOHE	0.9395±0.2472	158.0500±44.2283	31.8682±18.3980	54.2089±34.1718
BONA	0.3239±0.1014	2.9325±0.8112	53.6492±15.1271	1.8462±1.6056
CAM	13.9124±4.4913	69.2402±12.8180	79.9586±31.6804	111.6042±39.6215
GER	1.5061±0.3412	7.3194±3.1085	318.9909±273.4562	84.5721±52.7600
GO	63.6770±9.5442	233.1825±51.8353	1.8971±0.7516	98.6692±59.0261
LAZ	8.5785±3.0477	182.6115±99.0311	160.5837±111.8818	79.7552±28.4030
LEU	5.1719±2.9047	0.2382±0.0215	89.4493±16.6499	47.5073±15.9387
MAN	0.4565±0.3735	0.7974±0.0850	55.4474±52.7811	15.2476±14.5183
MARG	35.4263±18.0219	182.6405±29.2603	92.7047±44.9998	56.8170±52.1374
PO	0.1542±0.0874	0.9496±0.5515	18.0342±10.4528	41.0119±1.7311
ZL	6.2292±4.7760	113.0908±55.9743	24.6833±13.7596	10.9657±10.2813

Table S4 Metabolites concentration in buds of fourteen chamomile accessions

Chamomile accessions	Compounds (μg g^{-1} of fresh weight)							
	Sabinene ±SEM	β-myrcene ± SEM	α- pinene ±SEM	Cineole ±SEM	β- ocimene ±SEM	γ-terpinene ±SEM	Artemisia ketone± SEM	Ipsdienol ±SEM
ARG	0.0000 ± 0.0000	0.0000±0.0000	0.3464±0.0000	0.4920±0.1953	1.6321±0.4682	2.0244±0.6692	2.1077±1.3143	0.2886±0.0544
BOD	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	2.3447±1.5044	1.4256±0.3602	3.1813±1.8169	4.7259±2.1177	0.9244±0.3423
BOHE	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	1.4997±0.5712	12.9074±3.1669	2.7842±1.2744	4.8452±0.0000	0.0000±0.0000
BON	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	3.2411±0.8381	14.5661±3.1605	6.4265±1.3721	19.6989±6.3813	5.8795±0.5885
CAM	6.2735±4.3904	0.0000±0.0000	0.4050±0.0614	0.6474±0.0655	2.6975±0.7561	0.1495±0.0419	1.9552±0.7674	0.0000±0.0000
GER	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	0.0801±0.0174	2.3132±0.4481	1.2825±0.2353	2.3087±0.5072	1.4087±0.3811
GO	0.0000±0.0000	0.0000±0.0000	1.4448±0.9210	6.1821±4.4892	3.7297±1.4278	13.4838±8.6765	0.0000±0.0000	0.0000±0.0000
LAZ	0.0000±0.0000	0.0000±0.0000	1.0783±0.3613	1.7351±0.5281	2.3165±1.1509	6.5842±2.1137	0.0000±0.0000	0.0000±0.0000
LEU	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	2.7621±0.2900	5.3082±1.1048	10.1305±2.2702	5.8599±1.1223	8.9535±2.4444
MAN	43.1396±22.8685	3.7630±0.0000	2.0800±0.3067	4.0430±0.9137	3.0320±1.0976	6.6921±0.8659	2.6298±1.3520	0.5031±0.2888
MARG	0.0000±0.0000	0.0000±0.0000	0.2410±0.1082	0.1165±0.0264	2.0419±0.8620	0.6659±0.2176	0.0000±0.0000	0.0000±0.0000
MAT	0.0000±0.0000	0.0000±0.0000	0.9844±0.1260	1.6147±0.4913	3.2735±2.4292	4.6177±2.0997	0.0000±0.0000	0.0000±0.0000
PO	1.0645±0.0201	2.5223±0.0000	1.1020±0.0508	3.7845±2.2410	9.7259±1.4173	15.2107±10.8701	1.5801±1.0071	2.2065±0.1957
ZL	0.0000±0.0000	0.0000±0.0000	3.1293±1.3557	15.1560±8.9649	13.8417±3.3761	15.7666±7.0417	17.4306±6.5781	3.4064±1.0532

Chamomile accessions	Compounds (µg g^{-1} of fresh weight)							
	Caryophyllene ±SEM	β- farnesene ±SEM	Germacrene D ±SEM	Bicyclogermacrene ±SEM	α-farnesene ±SEM	Germacrene A ± SEM	Bisabolol oxide A ± SEM	Bisabolone oxide A ± SEM
ARG	0.3581±0.1009	43.0631±12.2456	3.9689±1.4974	1.4547±0.6174	6.1117±1.8940	0.0000±0.000	43.8961±7.1502	1.9545±0.5511
BOD	1.8079±0.9318	125.1050±56.6642	31.9812±15.3922	13.5957±4.2151	1.7586±1.2245	1.8665±0.5151	1.2040±0.6858	0.8578±0.7970
BOHE	1.5344±0.7763	71.2323±18.5792	20.0736±8.3530	18.4675±3.8578	3.0416±1.0388	0.3646± NA	80.7137±0.3888	52.5995±31.5106
BON	1.4383±0.1622	102.6608±51.3304	43.5957±15.8366	22.1845±8.5802	6.1979±1.5983	0.7456±0.2680	56.3978±14.1228	0.0000±0.0000
CAM	0.3320±0.1280	35.2197±9.8692	2.9714±0.4620	4.1934±0.4924	0.7168±0.2029	0.0000±0.0000	59.6449±53.7601	0.0000±0.0000
GER	0.7577±0.1455	114.5375±0.3053	22.9688±5.1977	14.9103±4.3011	4.0827±2.8969	0.0000±0.0000	3.6751±1.6753	0.0000±0.0000
GO	0.3275±0.0404	40.0371±16.3139	5.7162±1.5974	7.9299±6.1597	1.2518±0.7393	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000
LAZ	0.7666±0.0103	80.6285±19.1697	21.9473±6.6307	16.5091±1.7882	5.2145±2.1153	0.0000±0.0000	5.7536±1.7759	0.0000±0.0000
LEU	0.3285±0.0081	201.3413±33.8087	18.6046±1.3377	18.0665±5.6067	1.6937±0.2237	0.3991±0.0017	3.0964±0.3031	0.0000±0.0000
MAN	1.5501±0.7110	249.0887±84.2514	24.9500±7.6179	16.8084±7.2802	2.2975±0.8301	1.1653±0.1352	11.0474±7.0699	9.7989±0.0412
MARG	0.3810±0.1742	24.2671±6.1343	3.2704±1.1749	3.9340±1.2944	0.4506±0.1748	0.0000±0.0000	26.9495±7.7097	0.9796±0.2594
MAT	1.2452±0.2414	136.3502±67.2521	25.4227±5.2113	7.9581±3.3814	2.8400±0.8893	0.0000±0.0000	4.1588±2.7522	12.2788±5.8537
PO	1.7382±0.3600	148.3164±78.7752	44.8094±23.9477	40.3855±23.6154	2.3663±0.2366	1.7314±1.1928	21.2781±17.9725	4.3679±2.9529
ZL	0.6642±0.2380	183.1947±42.0273	20.8582±6.4580	48.0702±11.7254	2.7706±0.4023	0.6380±0.0998	382.5773±84.2190	35.5105±27.8637

Chamomile accessions	Compounds (µg g^{-1} of fresh weight)			
	α-bisabolol ± SEM	Bisabolol oxide B ±SEM	Spiroether ±SEM	En-in dicycloether ±SEM
ARG	3.8725±2.0738	NA	109.8182±45.2971	37.9734±15.4192
BOD	190.6083±86.2276	15.9294±8.9849	208.8654±121.3273	65.4422±29.8963
BOHE	4.8422±0.9360	NA	355.3610±47.8565	30.0624±6.8551
BON	163.1191±120.3141	101.5816±32.3100	896.7694±293.0537	416.3527±161.0606
CAM	0.6656±0.1335	6.4184±2.9499	121.3530±54.4204	25.1688±10.8453
GER	7.1424±4.4576	13.5316±10.8155	249.1776±16.2028	16.0660±1.1667
GO	221.8086±104.0792	NA	368.3005±111.2797	93.7581±43.4998
LAZ	1.7220±0.0816	4.1114±0.9981	375.7868±144.1632	88.8473±19.4508
LEU	284.0807±121.4874	NA	488.0930±65.4067	29.1591±8.7355
MAN	261.9542±106.3021	11.6455±3.9324	306.3737±116.0766	79.9448 ± 15.2367
MARG	13.5574±10.8709	2.8965±0.4937	49.8131±31.0815	45.4351±32.6300
MAT	3.5697±0.9386	NA	537.9418±50.1099	176.1956±35.4520
PO	138.1076± 30.3300	90.5409±52.9302	776.2459±412.1004	283.2761±120.1606
ZL	1.3697±0.0121	6.3004±0.0221	495.3949±247.1165	126.4669±36.0090

NA: Not avaiable

Table S5 Metabolites concentration in flowers of sixteen chamomile accessions

Chamomile accessions	Compounds (μg g^{-1} of fresh weight)							
	Sabinene ±SEM	β-myrecene ±SEM	α-pinene ± SEM	Cineole ±SEM	β-ocimene ±SEM	γ-terpinene ±SEM	Artemisia ketone ±SEM	Ipsdienol ±SEM
ARG	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	1.0187±0.0873	4.4693±1.5328	2.7250±0.6503	8.7831±1.9954	2.4406±0.4255
BOD	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	0.6107±0.1230	4.4819±0.8265	5.5884±3.7744	11.5979±3.0228	1.6431±1.2211
BOHE	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	0.3954±0.2617	0.9942±0.9158	0.9360±0.8266	0.0247±0.0247	0.0105±0.0105
BON	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	1.4182±0.3697	7.9955±1.4588	6.3395±2.0493	10.4761±4.1844	4.9216±1.3886
CAM	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	0.4633±0.4633	12.0749±5.2804	2.9238±2.3645	11.2882±4.5897	1.8941±0.9882
GER	62.5628±50.5231	0.0000±0.0000	0.0000±0.0000	3.9966±0.7636	11.7519±3.2274	12.1789±6.8177	14.3697±11.6928	3.3406±0.7319
GO	0.0000±0.0000	1.7317±0.8685	0.0000±0.0000	11.0297±5.4342	18.7488±6.4514	13.4619±5.7453	4.2988±1.2151	4.1941±3.1116
LAZ	0.0000±0.0000	0.0000±0.0000	0.4549±0.2399	0.9072±0.9072	2.8435±2.4235	1.1787±1.1787	5.3127±5.3127	0.6864±0.6864
LEU	0.0000±0.0000	0.0000±0.0000	0.0000±.0000	5.7424±1.9718	28.5680±8.9589	9.7431±4.6160	6.5449±4.6959	12.5386±6.0438
MAN	15.0695±1.0170	0.0000±0.0000	0.0000±0.0000	1.4731±0.2387	10.4297±1.7361	2.7124±0.3856	10.5927±1.8318	6.2467±4.0611
MAR	24.3040±2.2405	0.0000±0.0000	0.0000±0.0000	0.8406±0.3225	12.5024±4.4924	7.6486±3.4210	10.2751±3.0094	3.5012±2.3241
MARG	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	0.1741±0.1741	15.9441±7.7907	2.9159±0.6879	6.8001±2.1935	6.2040±3.0767
MAT	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	1.6326±1.0871	10.4398±5.0687	5.9365±0.7722	10.7225±2.1655	1.7333±0.0964
MO	0.0000±0.0000	0.2913±0.1263	0.0000±0.0000	3.6353±1.0282	2.0406±0.4856	1.6187±1.3662	1.2689±1.2631	2.3499±0.9456
PO	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	1.4185±0.4346	9.0927±1.8135	1.9025±0.8347	2.6169±1.0533	1.5140±1.0583
ZL	0.0000±0.0000	0.0000±0.0000	0.0000±0.0000	2.7026±0.8358	6.8754±2.3581	5.5164±2.3829	11.7294±2.0269	3.4511±1.0803

	Compounds (μg g^{-1} of fresh weight)						
Chamomile accessions	β-caryophyllene ±SEM	β- Farnesene ±SEM	Germacrene D ±SEM	Bi-cyclogermacrene ±SEM	α- farnesene ±SEM	Germacrene A ±SEM	Bisabolol oxide A ±SEM
ARG	1.1172±0.4188	96.4823±16.4692	6.0758±2.2648	10.9394±1.0473	4.6895±1.4060	2.3279±2.1042	97.1353±33.8706
BOD	0.7692±0.3986	35.5563±9.1915	7.9277±2.5465	8.9395±3.2925	4.0463±2.6479	0.9247±0.5143	98.1356±62.8357
BOHE	0.2790±0.1929	42.3216±38.5321	12.2927±11.7375	1.1061±0.9276	1.6647±0.0758	0.5639±0.3692	4.2236±3.1078
BON	1.3886±1.0382	103.6588±51.0888	10.2796±4.0134	11.1499±8.7702	0.7352±0.3731	0.8689±0.4894	4.6251±0.8266
CAM	3.6142±2.5282	213.8049±158.3467	20.3317±10.5427	10.5288±5.7412	2.3915±1.3071	0.8356±0.1792	64.5707±27.1245
GER	3.4687±0.8892	202.3083±91.5179	24.4283±6.0774	24.1625±12.7697	1.7970±0.2110	0.9053±0.2678	1.8775±0.8056
GO	1.7163±0.9348	174.3290±93.3395	31.0755±8.3491	24.9542±16.3446	3.9572±2.0283	1.7202±0.4510	397.2462±302.3700
LAZ	2.2237±2.0784	51.4401±38.4432	8.3512±5.9603	3.5667±2.4567	2.2416±1.0690	0.4728±0.4728	1.3405±1.1514
LEU	1.0759±0.5738	188.6562±82.1562	16.5555±6.1259	57.1927±21.1561	8.0952±2.6264	0.9362±0.5229	913.4216±331.7324
MAN	1.4306±0.3825	90.6977±23.7343	21.9240±2.6859	2.7754±0.2275	2.5931±0.3969	2.3001±1.5701	332.4824±0.6956
MAR	3.6721±1.5441	335.9560±133.5866	53.0661±19.3304	27.0802±17.0484	15.1993±10.9638	6.0112±1.6170	43.0250±15.7461
MARG	0.8297±0.0826	320.1348±142.8727	27.9002±11.4509	19.4519±7.0217	6.8721±1.7037	1.7972±1.0304	382.6642±181.6296
MAT	9.2070±2.4287	138.6080±94.1785	20.7853±15.7293	9.7897±1.7054	9.0845±6.9780	1.2862±0.6197	93.9241±51.6906
MO	4.0351±0.8019	43.9480±25.2078	24.0247±15.2675	3.5414±0.2423	1.0805±0.3435	2.7867±1.0429	32.4776±15.6882
PO	1.1840±0.6775	140.6748±47.4065	28.7054±7.7699	15.2096±2.5141	5.0275±1.6850	0.8809±0.5587	91.7231±55.3854
ZL	1.1463±0.3888	128.3805±76.4540	18.3170±8.6067	25.8814±22.6072	3.0741±1.1115	0.9327±0.3743	150.4195±108.7124

	Compounds (µg g^{-1} of fresh weight)				
Chamomile accessions	Bisabolone oxide A ±SEM	α- bisabolol ±SEM	Bisabolol oxide B ±SEM	Spiroether ±SEM	En-in dicycloether ±SEM
ARG	12.0252±5.3705	5.6641±1.6441	312.0502±33.8647	135.0752±27.1391	92.2242±10.5292
BOD	30.7261±20.9964	196.1256±156.2723	37.7896±18.9594	364.1136±197.9664	152.6322±45.0167
BOHE	0.4822±0.2300	1.6640±1.3898	7.0307±3.7380	49.7515±44.6169	47.1884±45.8545
BON	28.7610±8.6243	182.8915±154.4870	180.0485±122.8590	106.1205±52.7238	85.5050±37.1398
CAM	58.7685±7.7695	0.8818±0.5314	189.3048±73.6613	345.0390±210.8139	67.7396±38.1523
GER	0.5329±0.3043	763.9611±307.9159	7.3678±1.5741	396.1155±168.7021	251.5653±122.8132
GO	1.3471±0.7568	294.3329±104.8370	270.2811±149.0190	433.8008±120.0873	377.1178±116.6647
LAZ	0.2141±0.1035	5.9239±4.6135	11.4952±0.2095	35.3117±26.7626	42.8043±31.7404
LEU	2.0027±0.9878	285.2185±139.0368	30.5377±8.8748	757.4494±338.1471	444.8035±103.4118
MAN	1.7944±0.7070	376.0756±171.7781	82.6261±20.8799	260.2942±106.9424	57.7478±35.9998
MAR	371.9219±277.8452	104.6335±61.1774	903.0291±214.1232	1204.4127±431.8789	278.1000±95.9077
MARG	12.1116±3.3579	50.7276±39.7110	133.4597±109.3198	1792.3288±236.1889	977.6586±436.3878
MAT	180.0725±65.9776	2.3782±0.2216	10.6038±1.9736	406.0486±133.2952	319.7417±120.9621
MO	19.8792±10.5439	5.2079±0.1600	98.1608±32.8756	197.2517±59.0860	123.3649±30.8118
PO	1.3499±0.7534	111.7692±91.2527	60.2056±35.7607	548.4000±236.1889	412.0802±74.7834
ZL	87.6199±33.8620	9.2480±3.4637	182.0162±123.1576	141.1149±34.6807	130.8929±74.7874

Table S6 Data for transcript abundance of MrTPS7 alleles and substrate concentration in flowers of chamomile accessions

Chamomile accessions	Relative gene expression	(±) SEM	FPP (mean value) ng g^{-1} of fresh weight.
ARG	0.571373	0.022777	2.358
BOD	1.643876	0.452058767	17.694
BOHE	1.198283	0.708359833	3.392
BON	0.447782	0.077470841	4.332
CAM	0.532115	0.1617115	0.811
GER	3.902561	0.990391049	6.429
GO	0.395272	0.114385106	5.250
LAZ	1.034238	0.256579443	4.792
LEU	2.783294	1.144251438	5.160
MAN	0.916751	0.727628214	9.457
MAR	0.335811	0.148120656	4.012
MARG	2.460811	0.0149575	3.518
MAT	0.780904	0.324352111	1.121
MO	2.0410244	0.971690263	1.939
PO	0.347582	0.11927779	5.226
ZL	0.985703	0.200081205	0.765

Table S7 Data for α-bisabolol and en-in dicycloether concentration in flowers of chamomile accessions

Chamomile accessions	α-bisabolol in flower heads (µg g^{-1} of fresh weight)	(±) SEM	En-in dicycloether in flower heads (µg g^{-1} of fresh weight)	(±) SEM
ARG	5.664	1.644	92.2242	10.5292
BOD	196.126	156.272	152.6322	45.0167
BOHE	1.664	1.390	47.1884	45.8545
BON	182.891	154.487	85.5050	37.1398
CAM	0.882	0.531	67.7396	38.1523
GER	763.961	307.916	251.5653	122.8132
GO	294.333	104.837	377.1178	116.6647
LAZ	5.924	4.613	42.8043	31.7404
LEU	285.218	139.037	444.8035	103.4118
MAN	376.076	171.778	57.7478	35.9998
MAR	104.633	61.177	278.1000	95.9077
MARG	50.728	39.711	977.6586	436.3878
MAT	2.378	0.222	319.7417	120.9621
MO	5.208	0.160	123.3649	30.8118
PO	111.769	91.253	412.0802	74.7834
ZL	9.248	3.464	130.8929	74.7874

Table S8 Sixteen chamomile accessions with their *MrTPS7* alleles trascription and concentration of α-bisabolol and its oxides

Chamomile accessions	*MrTPS7* alleles transcription level	α-bisabolol in flower (µg g^{-1} of fresh weight)	Bisabolol oxide A (µg g^{-1} of fresh weight)	Bisabolol oxide B (µg g^{-1} of fresh weight)
ARG	0.571373	5.664	97.135	312.050
BOD	1.643876	196.126	98.135	37.789
BOHE	1.198283	1.664	4.223	7.030
BON	0.447782	182.891	4.625	180.048
CAM	0.532115	0.882	64.570	189.304
GER	3.902561	763.961	1.877	7.367
GO	0.395272	294.333	397.246	270.281
LAZ	1.034238	5.924	1.340	11.495
LEU	2.783294	285.218	913.421	30.537
MAN	0.916751	376.076	332.482	82.626
MAR	0.335811	104.633	43.025	903.029
MARG	2.460811	50.728	382.664	133.459
MAT	0.780904	2.378	93.924	10.603
MO	2.0410244	5.208	32.477	98.160
PO	0.347582	111.769	91.723	60.205
ZL	0.985703	9.248	150.419	182.016

Table S9 Amino acid similarity (%) of GER with others *MrTPS* in chamomile
GER: Germania, The encoded main products are *MrTPS1*: (−)-(*E*)-β-caryophyllene, *MrTPS3*: (+)- germacrene A and *MrTPS5*: (−)-germacrene D.

Sl. No	Code	Gene bank accession no.	GER	MrTPS1	MrTPS3	MrTPS5
1.	GER	N/A	100			
2.	*MrTPS1*	JQ255375	12.22	100		
3.	*MrTPS3*	JQ255377	12.76	15.14	100	
4.	*MrTPS5*	JQ837261	15.63	57.84	15.81	100

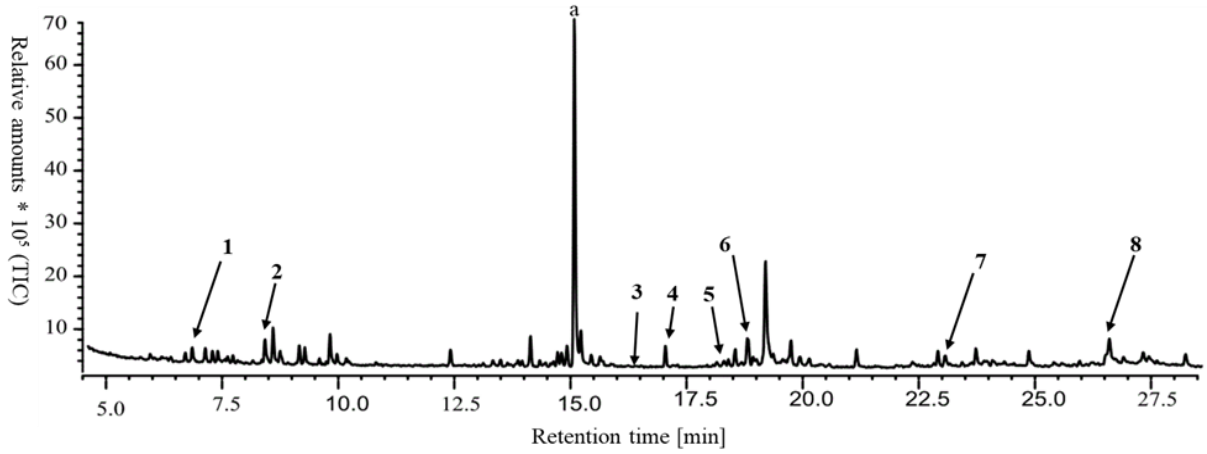


Fig. S1 A typical spectrum of metabolities in *n*-hexane extraction and analyzed by GC-MS of chamomile seedlings

The samples extracted with n-hexane as a solvent. The peak identifed as 1 Citronellene, 2. β-Ocimene, 3. Longipinene, 4. β- Caryophyllene, 5. Germacrene D, 6. β- Farnesene, 7. Hexacosane, 8. Spiroether and a. nonayl acetate (as internal std.).

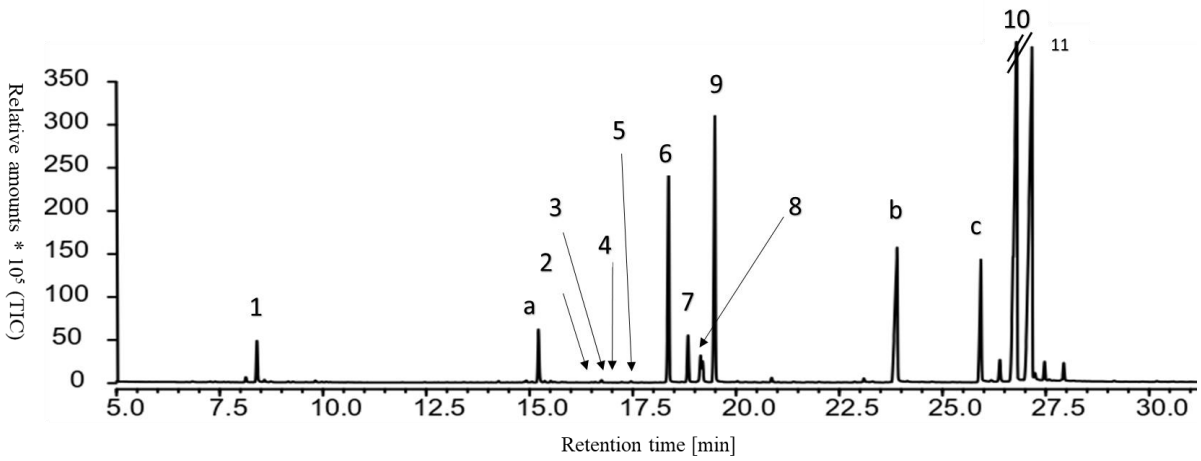


Fig. S2 A typical spectrum of metabolities analysis of stems and leaves in *n*-hexane and analyzed by GC-MS

The samples extracted with n-hexane as a solvent. The peak identifed as 1 β- Ocimene, 2 α-guanojene, 3 Aristolene 4 Longifolin, 5 β- Caryophyllene, 6 β- Farnesene, 7 Germacrene D, 8 Bicyclogermacrene, 9 α-farnesene, 10 Spiroether, 11 En-in-dicycloether, a. nonayl acetate (as internal std.), b and c unspecific peaks.

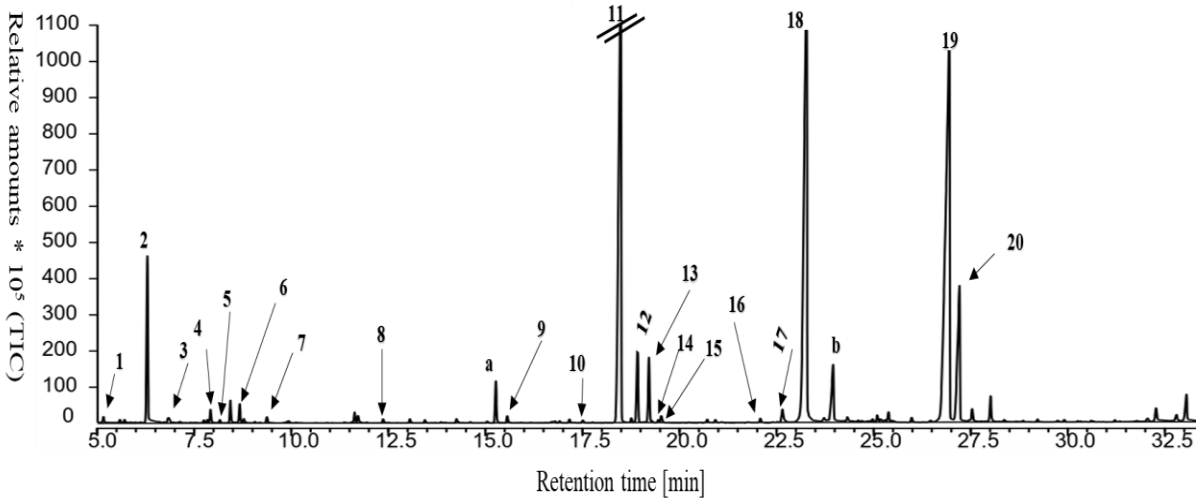


Fig. S3 A typical spectrum of terpene analysis of buds and flowers were extracted with *n*-hexane and analyzed by GC-MS

The peaks are identified as 1 α-pinene, 2 sabinene, 3 β Myrcene, 4 1,8 cineole, 5 β-ocimene, 6 γ-terpinene, 7 Ipsadinol, 8 α-´terpenol, 9. Bicycloelemene, 10 β- Caryophyllene, 11 β-farnesene, 12 Germacrene D, 13 Bicyclogermacrene, 14 Germacrene A 15 α-farnesene, 16 Bisabolol oxide A, 17 Bisabolone oxide A, 18 α-bisabolol, 19 Spiroether, 20 En-in-dicycloether, a Nonayl acetate (as internal std.) b unspecific peak.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
               1         10         20         30 RX(X8)W 40         50
ARG             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
BOHE            MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
GER             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
GO              MSTLSVSTPS FSSSPLSSAN KNSTKQHVTRNSVIFHDSIW GDEFLEYKEK
LEU             MSTLSVSTPS FSSSPLSSAN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
MAT             MSTLSVSTPS FSSSPLSSAN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
PO              MSTLSVSTPS FSSSPLSSAN KDSTKQHVTRNSVIFHDSIW GDQFLEYKEK
MAR             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
BOD             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
BON             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDEFLEYKEK
CAM             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
LAZ             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSTW GDEFLEYKEK
MAN             MSTLSVSTPS FSSSPLSSAN KNSTKQHVARNSVIFHDSIW GDQFLEYKEK
MAR             MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
MARG            MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
ZL              MSTLSVSTPS FSSSPLSSVN KNSTKQHVTRNSVIFHDSIW GDQFLEYKEK
MO              LCCPG.RLAS MVLEDNSSVS TPPRVTSAVR.PSVPPGSIW QQWK.LSSAA
```

Fig. S4 RX(X8)W conserved motifs of MrTPS7 alleles

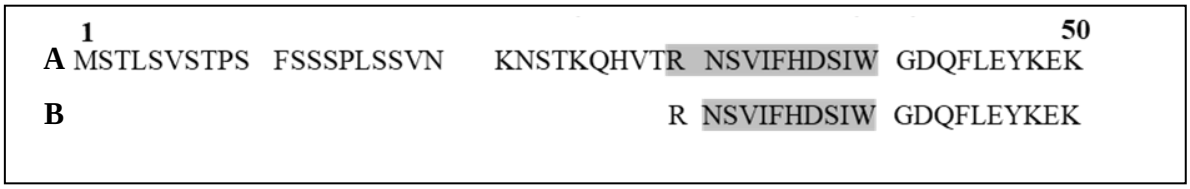


Fig. S5a Amino acids sequence of MrTPS7 allel (GER) with and without of signal peptide motif
A. MrTPS7 allele of GER (with signal peptide). B. s-MrTPS7 allele of GER without 29 aa (N-terminal signal peptide).

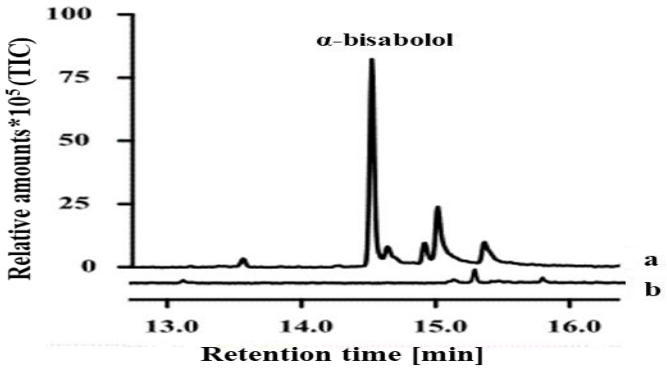


Fig. S5b Product peak of of MrTPS7 enzymes (GER) with and without of signal peptide motif a: products peak of A and b. products peak of B (no product).

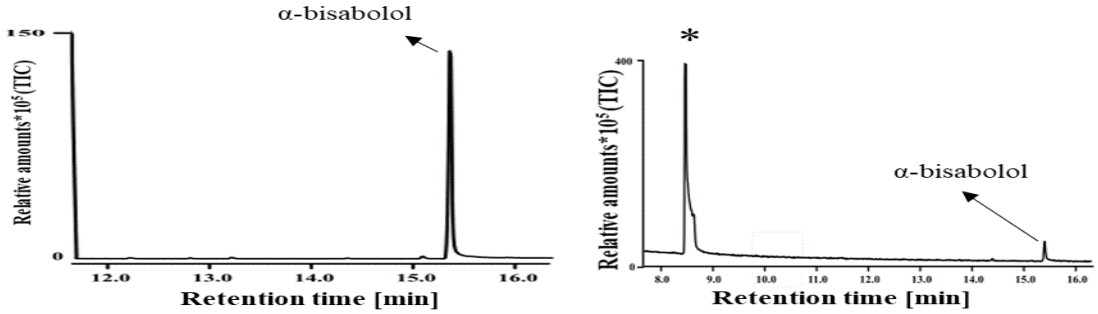
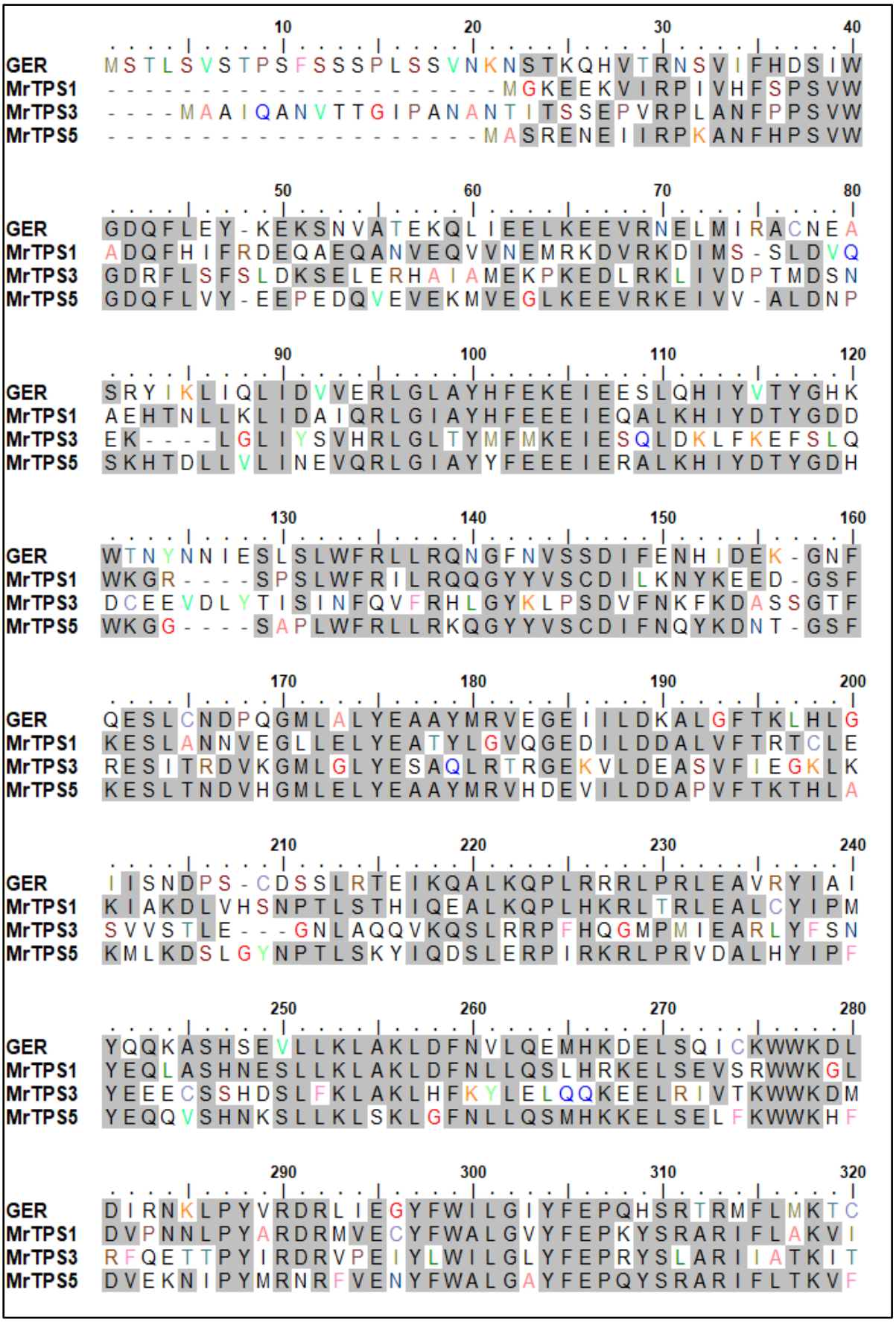


Fig. S6 The product peak of MrTPS7 allels of MO after purification over a pro-infinity IMAC Ni-charged resin column

On the left side is the single product of α-bisabolol tested with SPME and on right side is the *n*-hexane purified extract in GC-MS. * Unspecific peak from solvent (*n*-hexane).


10 20 30 40
GER MSTLSVSTPSFSSSPLSSVNKNSTKQHVTRNSVIFHDSIW
MrTPS1 ----------MGKEEKVIRPIVHFSPSVW
MrTPS3 ----MAAIQANVTTGIPANANTITSSEPVRPLANFPPSVW
MrTPS5 ----------MASRENEIIRPKANFHPSVW
50 60 70 80
GER GDQFLEY-KEKSNVATEKQLIEELKEEVRNELMIRACNEA
MrTPS1 ADQFHIFRDEQAEQANVEQVVNEMRKDVRKDIMS-SLDVQ
MrTPS3 GDRFLSFSLDKSELERHAIAMEKPKEDLRKLIVDPTMDSN
MrTPS5 GDQFLVY-EEPEDQVEVEKMVEGLKEEVRKEIVV-ALDNP
90 100 110 120
GER SRYIKLIQLIDVVERLGLAYHFEKEIEESLQHIYVTYGHK
MrTPS1 AEHTNLLKLIDAIQRLGIAYHFEEEIEQALKHIYDTYGDD
MrTPS3 EK----LGLIYSVHRLGLTYMFMKEIESQLDKLFKEFSLQ
MrTPS5 SKHTDLLVLINEVQRLGIAYYFEEEIERALKHIYDTYGDH
130 140 150 160
GER WTNYNNIESLSLWFRLLRQNGFNVSSDIFENHIDEK-GNF
MrTPS1 WKGR----SPSLWFRILRQQGYYVSCDILKNYKEED-GSF
MrTPS3 DCEEVDLYTISINFQVFRHLGYKLPSDVFNKFKDASSGTF
MrTPS5 WKGG----SAPLWFRLLRKQGYYVSCDIFNQYKDNT-GSF
170 180 190 200
GER QESLCNDPQGMLALYEAAYMRVEGEIILDKALGFTKLHLG
MrTPS1 KESLANNVEGLLELYEATYLGVQGEDILDDALVFTRTCLE
MrTPS3 RESITRDVKGMLGLYESAQLRTRGEKVLDEASVFIEGKLK
MrTPS5 KESLTNDVHGMLELYEAAYMRVHDEVILDDAPVFTKTHLA
210 220 230 240
GER IISNDPS-CDSSLRTEIKQALKQPLRRRLPRLEAVRYIAI
MrTPS1 KIAKDLVHSNPTLSTHIQEALKQPLHKRLTRLEALCYIPM
MrTPS3 SVVSTLE---GNLAQQVKQSLRRPFHQGMPMIEARLYFSN
MrTPS5 KMLKDSLGYNPTLSKYIQDSLERPIRKRLPRVDALHYIPF
250 260 270 280
GER YQQKASHSEVLLKLAKLDFNVLQEMHKDELSQICKWWKDL
MrTPS1 YEQLASHNESLLKLAKLDFNLLQSLHRKELSEVSRWWKGL
MrTPS3 YEEECSSHDSLFKLAKLHFKYLELQQKEELRIVTKWWKDM
MrTPS5 YEQQVSHNKSLLKLSKLGFNLLQSMHKKELSELFKWWKHF
290 300 310 320
GER DIRNKLPYVRDRLIEGYFWILGIYFEPQHSRTRMFLMKTC
MrTPS1 DVPNNLPYARDRMVECYFWALGVYFEPKYSRARIFLAKVI
MrTPS3 RFQETTPYIRDRVPEIYLWILGLYFEPRYSLARIIATKIT
MrTPS5 DVEKNIPYMRNRFVENYFWALGAYFEPQYSRARIFLTKVF

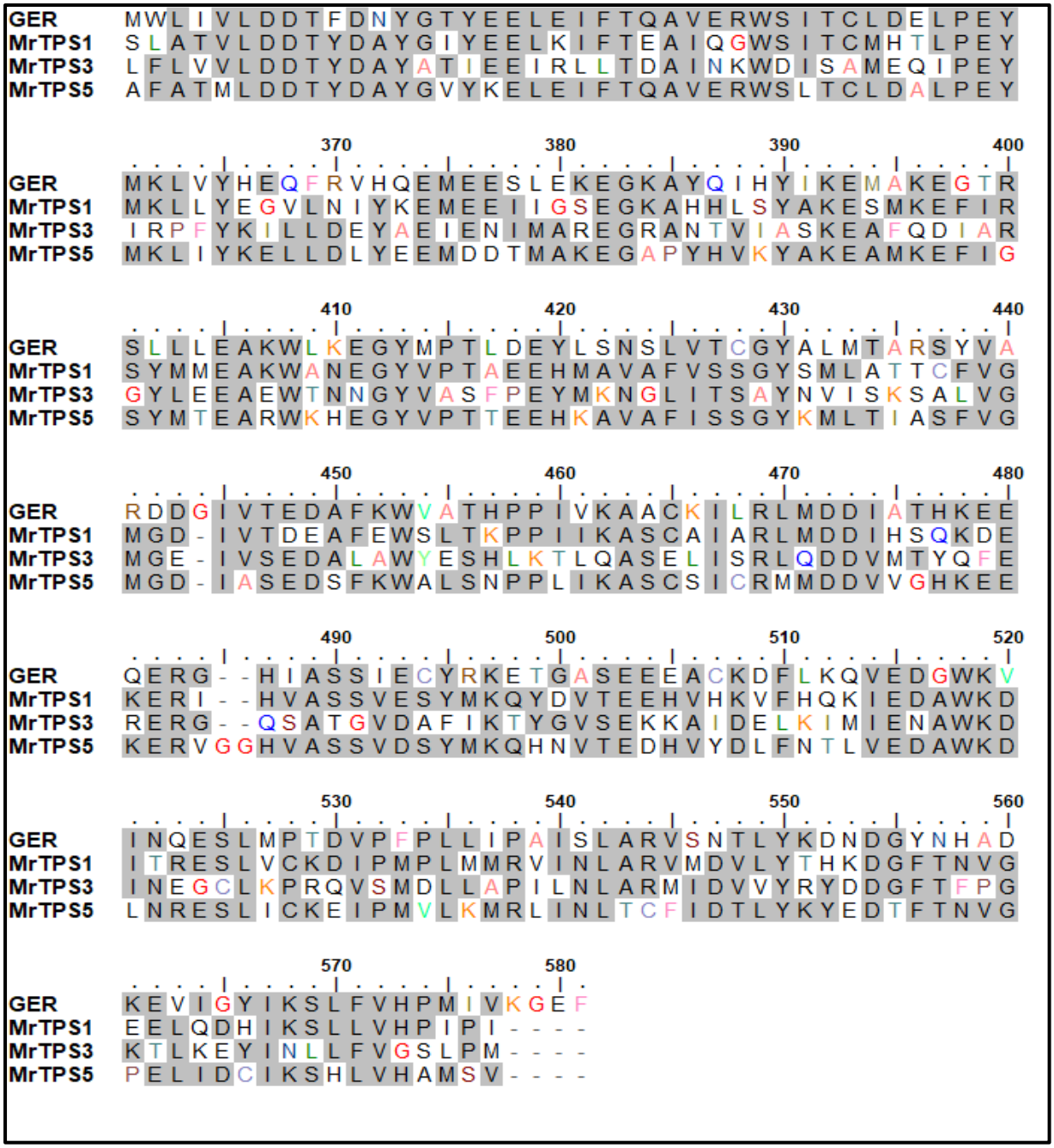


Fig. S7 Amino acid similarity within GER, MrTPS1, MrTPS3 and MrTPS5
Shadded are the identical amino acids. MrTPS1 encoded product: (−)-(*E*)-β-caryophyllene, MrTPS3 encoded product: (+)- germacrene A and MrTPS5 encoded product: (−)-germacrene D.

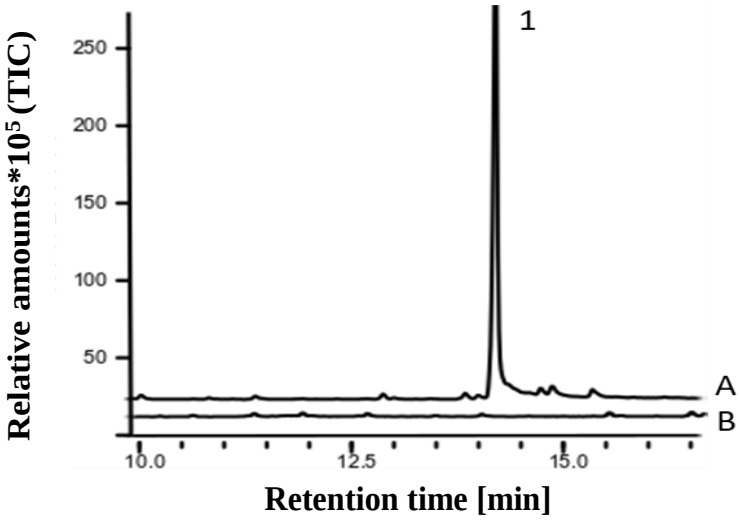


Fig. S8 *In vitro* product formation after heterologous expression enzymes of MrTPS7 in presence of FPP and GPP

A. Spectrum of product formation using FPP as a substrate, B. no product with substrate GPP, Product was identified as (1.) α-bisabolol by SPME in GC-MS.

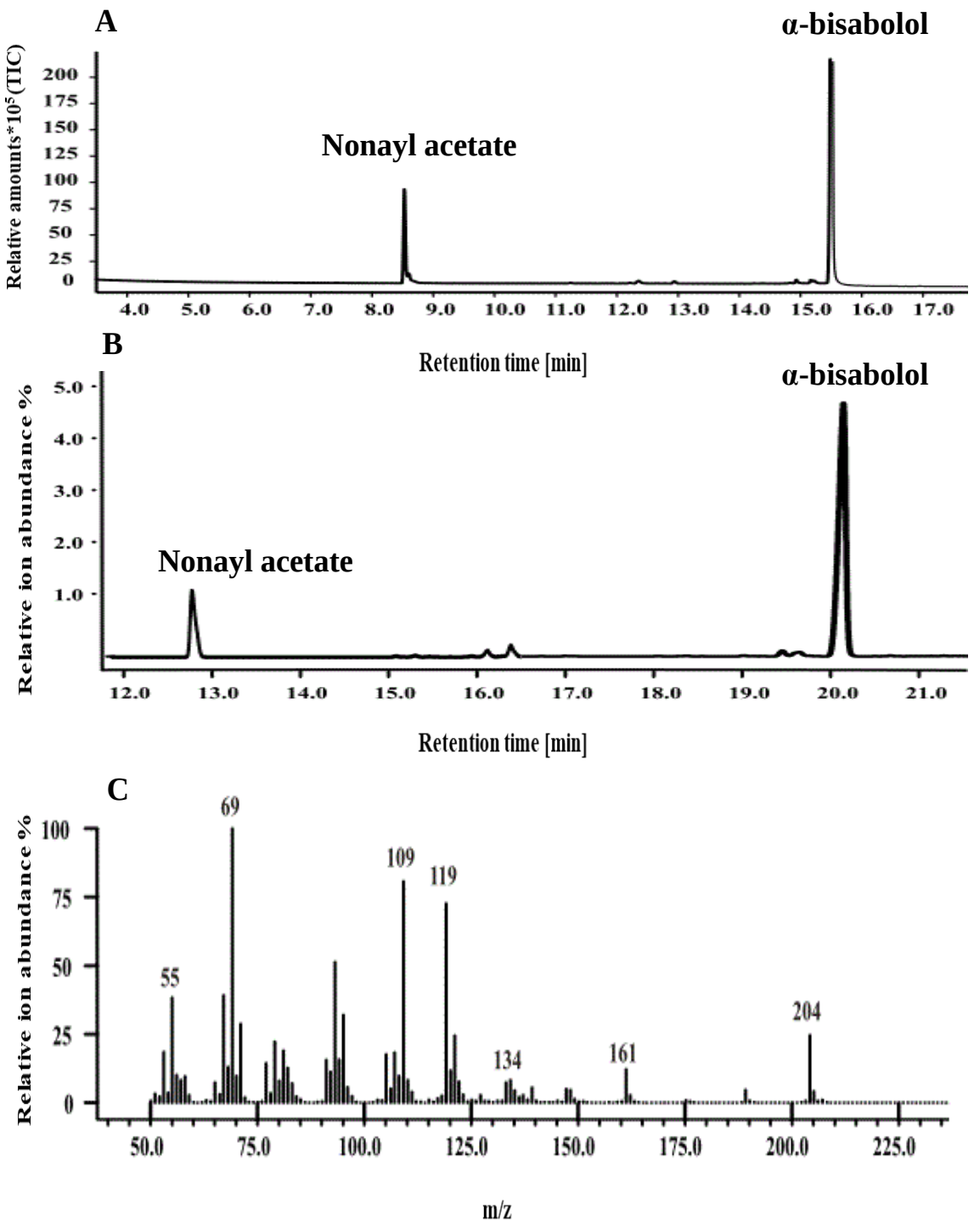


Fig. S9 MS spectrum and corresponding FID spectrum of α-bisabolol of MrTPS7 using FPP as a substrate and Mg^{2+} as cofactor

The assay reaction mixture including nonyl acetate as an internal standard was incubated and the overlaid product was collected and injected in GC-MS. The product peak identified with previously determined retention time and fragmentation of a standard and with post run analysis of GC-MS library software(Wiley and Shimadzu libraries(Adams, 2007). A. MS spectrum, B. FID spectrum and C. MS fragmentation spectrum of α-bisabolol in GC-MS.

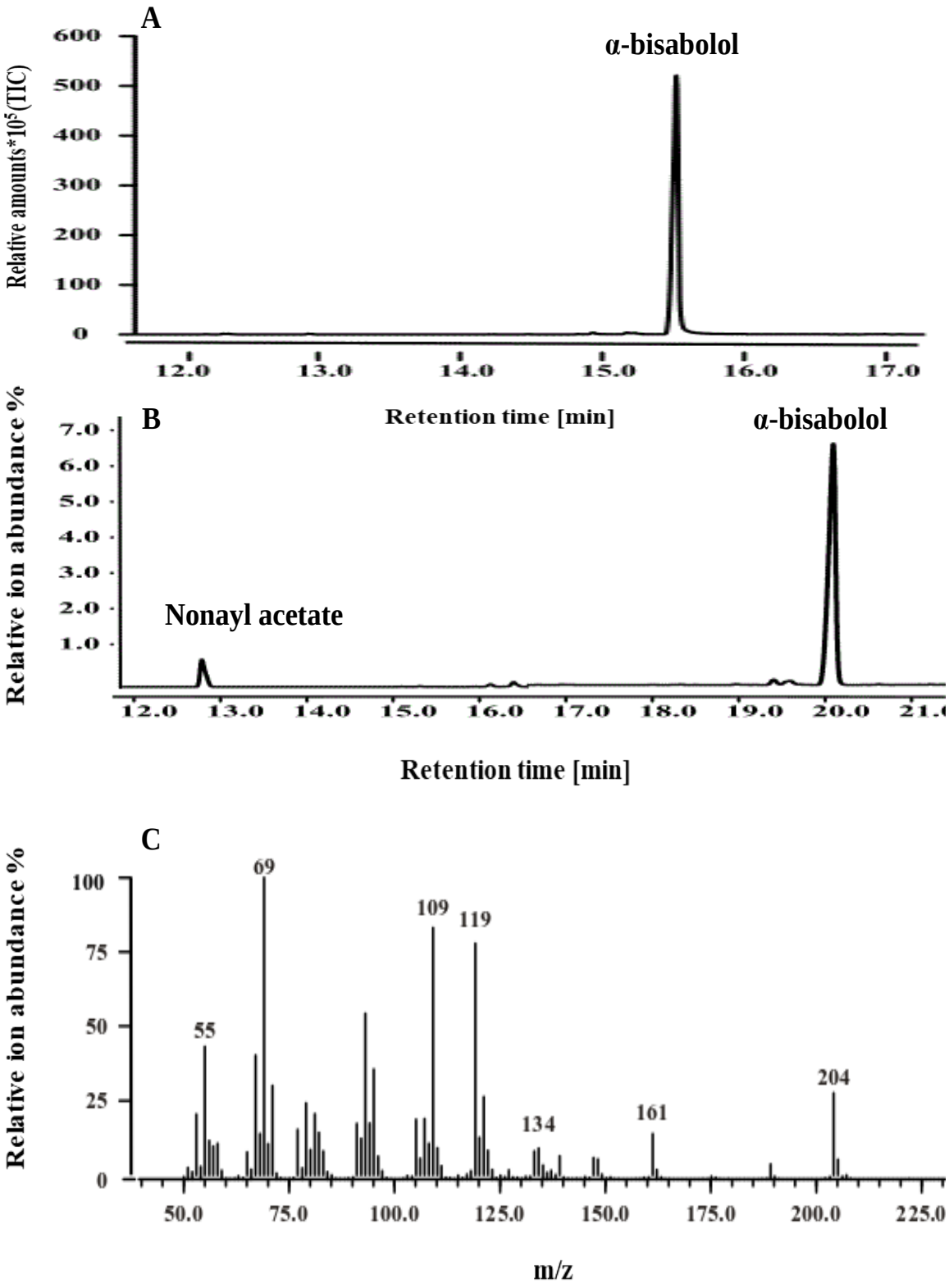


Fig. S10 MS spectrum and corresponding FID spectrum of standard α-bisabolol in GC-MS
A. MS spectrum of standard α-bisabolol, B. Corresponding FID spectrum of A and C. MS fragmentation spectrum of standard α-bisabolol. The fragmentation pattern shows the base peak at 109 m/z for α-bisabolol, 69, 109 and 119 m/z are the unique identified fragments of α-bisabolol.

Acknowledgement

I would like to express my deep gratitude to my "Doktorvater" Prof. Dr. Jörg Degenhardt thanks to him for offering me a chance to join his laboratory and keeping the door open always for ideas and supervision during my PhD work.

I would like to acknowledge the constant support by Dr. Susanne Preiss. I want to thank always Dr. Sandra Krause for her contineous technical support and constructive ideas during the work. A special thanks go to Dr. Claudia Schaff for her expertise in molecular biology. A big thank to all my doctoral colleagues, especially Ms. Jennifer Szczesny for her expertise and support with the *in silico* modeling. A heartful thanks to Sebastian Bley for his help and great companionship throughout my lab tenure. A special thanks to Mrs. Franziska Leidecker for her scientific discussions.

I do thank, Dr. Peter Lindemann especially for technical supports with the GC-MS laboratory and very inspiring scientific discussion. I wish to thanks Mrs. Natalie Wiese for being a good lab colleague. Big thanks to Mrs. Kathrin Thomasch and Mrs. Anja Wodak for their support. My special thanks go to Mrs. Kathrin Kunth for her help with administrative work always. I would also like to acknowledge Dr. Sabine Willscher and Dr. Bettina Rahfeld for scientific talks. A big gratitude to Dr. Heiko Noack and Somnath Koley for technical and instrumentational support. I do thanks Dr. Manish Raorane for his scientific suggestions during this dissertation.

Thanks to Dr. Lars-Gernot Otto for providing the plant and seed materials. I acknowledge Mrs. Marlis Sonnenschein (Pharmaplant GmbH) for providing the seed materials for the experiments.

A big thank goes to Erasmus Mundus BRAVE program of the European Union for the first three years funding of my PhD work. I would like to acknowledge the BRAVE coordinator Dr. Andreas Voloudakis (Agricultural University of Athens) and Prof. Dr. Sven Erik Behrens (Martin Luther University, Germany) for their constant support during the funding period. A very special gratitude goes to Prof. Dr. William Foley from the Australian National University and Prof. Dr. Jonathan Gershenzon (Max Plank Institute, Jena) for a special supporting grant to continue my PhD work. At last, I again want to say thanks to Prof. Dr. Degenhardt for offering the funding to finish my doctoral dissertation.

Acknowledgement

I appreciate the efforts of BRAVEians and some nice friends in Halle. Thanks to Somnath Koley for being a great colleague and friend. Also thanks to Mr. Shiv Meena and Mrs. Priyanka Mishra, Dr. Somur Selvaraj Tamilarasan and his family for being great fellows. I am deeply indebted to Ms. Ranju Chutia and Ms. Manaswita Baruah for being nice companions during the stay in Halle. I wish to thank Mrs. Pratibha Kumari and Ms. Sudeshna Chaki for their consultant in different matters. I wish to mention special thanks to Dr. Mayank Sharma and Dr. Arunava Datta for their very positive inspiring attitude always and at last, I wish to special gratitude to Mr. Sebastian Bley and Ms. Sara Werner for their friendly companions always. I owe a deep sense of gratitude to my parents for their love and patience during my PhD work. Thanks to my elder sisters for their encouragement and support.

Curriculum Vitae

Personal

Name	Prodyut Mondal
Date of Birth	21.11.1983
Place of Birth	Iswarigacha, India
Address	Iswarigacha, Ashokenagar
	North 24 Pgs, West Bengal 743222. India
Citizenship	Indian

Institutional affiliation

Dept. of Pharmaceutical Biology

Institute of Pharmacy

Martin Luther University, Halle-Wittenberg

Hoher Weg 8,

06120 Halle

Germany

Email: mondalprodyut0@gmail.com

Education

2011-2013	Master of Pharmacy (Pharmacognosy) Dibrugarh University, Dibrugarh-Assam, India
2008-2011	Bachelor of Pharmacy Dr. B C Roy College of Pharmacy and Allied Health Sciences. Durgapur, India
2001-2003	Diploma in Pharmacy Institute of Pharmacy, Jalpaiguri West Bengal India
2001	Secondary school (West Bengal Board of secondary Education)

Academic positions

Octboer, 2014 -	Ph.D Fellow (Prof. Dr. Jörg Degenhardt Lab) Dept. of Pharmaceutical Biology, Institute of Pharmacy Martin Luther University Halle-Wittenberg
2013-2014	Junior Research Fellow West Bengal University of Technology, India.

Work Experiences

2013-2014	Quality Control Chemist Greenco Biologicals Private Limited, Kolkata, India
2005-2008	Clinical Pharmacist North City Hospital and Neuro Institute., Kolkata, India

Research grants and academic achievements

2014 Erasmus Mundus 'BRAVE' Ph.D. fellowship by European Union

2013 Medel (Second) of National elocution competition-2013 organized by - State Pharmacy Council, Assam, India.

2011 Qualified GATE*

2011 20th positioned in National Pharmacy Talent search Completion - 2011, India

*Graduate Aptitude Test for Engineers conducted by All India Institute of Technical education (AICTE).

Selected Publications

2019 Kainer, D., Padovan, A., Degenhardt, J., Krause, S., **Mondal, P**., Foley, W.J., Külheim, C., 2019. High marker density GWAS provides novel insights into the genomic architecture of terpene oil yield in Eucalyptus. New Phytol. 223, 1489–1504. doi:10.1111/nph.15887

2018 Junejo, J.A., Gogoi, G., Islam, J., Rudrapal, M., **Mondal, P**., Hazarika, H., Zaman, K., 2018. Exploration of antioxidant, antidiabetic and hepatoprotective activity of Diplaziumesculentum - A wild edible plant from North Eastern India. Futur. J. Pharm. Sci. 4, 93–101. doi:10.1016/j.fjps.2017.10.005

2017 Otto, L.G., **Mondal, P**., Brassac, J., Preiss, S., Degenhardt, J., He, S., Reif, J.C., 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. doi:10.1186/s12864-017-3991-0

2016 **Mondal, P**., Das, S., Mahato, K., Borah, S., Junejo, J.A., Zaman, K., 2016. Evaluation of Anti-Arthritic Potential of the Hydro-Alcoholic Extract of the Stem Bark of Plumeria Rubra in Freund'S Complete Adjuvant-Induced Arthritis in Rats. Int. J. Pharm. Sci. Res. 7, 3675–3688. doi:10.13040/IJPSR.0975-8232.7(9).3675-88

2015 Junejo, J.A., Ghoshal, A., **Mondal, P**., Nainwal, L., Zaman, K., Singh, K.D., Chakraborty, T., 2015. In-vivo Toxicity Evaluation and Phytochemical, Physicochemical Analysis of Diplazium esculentum (Retz.) Sw. leaves a Traditionally used North-Eastern Indian Vegetable. Adv. Biores. 6, 175–181. doi:10.15515/abr.0976-4585.6.5.175181

2014 Junejo, J.A., **Mondal, P**., Verma, V.K., Rudrapal, M., Zaman, M.K., 2014. Anti-diabetic assessment of the hydro-alcoholic leaf extracts of the plant tetrastigma angustifolia (Roxb.) A traditionally used North-Eastern Indian vegetable. Biomed. Pharmacol. J. 7, 635–644. doi:10.13005/bpj/535

2013 **Mondal, P**., Bhuyan, N., Das, S., Kumar, M., Borah, S., Mahato, K., 2013. Herbal medicines useful for the treatment of diabetes in North-East India: A review. Int. J. Pharm. Biol. Sci. 3, 575–589.

Kumar, M., **Mondal, P**., Borah, S., Mahato, K., 2013. Physico-chemical evaluation, preliminary phytochemical investigation, fluorescence and TLC analysis of leaves of the plant Lasia spinosa (Lour) Thwaites. Int. J. Pharm. Pharm. Sci. 5, 306–310.

Eidesstattliche Erklärung

Hiermit versichere ich, dass diese Arbeit in gleicher oder ähnlicher Form bisher keiner anderen Prüfbehörde zur Erlangung des Doktorgrades vorgelegt wurde. Ferner erkläre ich, gemäß der Promotionsordnung der Naturwissenschaftlichen Fakultät I – Biowissenschaften der Martin-Luther-Universität, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen sind, habe ich als solche kenntlich gemacht.

Halle an der Saale, Datum _____ _____

Unterschrift

(PRODYUT MONDAL)