

Cloning of *Lavandula* Essential Oil Biosynthetic Genes

by

Md. Lukman Syed Sarker

B.Sc., The University of Dhaka, 2008

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES
(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA
(Okanagan)

February 2013

© Md. Lukman Syed Sarker, 2013

Abstract

Several varieties of *Lavandula x intermedia* (lavandins) are cultivated for their essential oils (EO) which are extensively used in a wide range of hygiene and personal care products. These EOs are mainly dominated by monoterpenes, including linalool, linalool acetate, borneol, 1,8-cineole, and camphor. Among these, camphor is of particular significance as it adds a sharp overtone to the EO fragrance, reducing its olfactory appeal compared to finer lavender EOs in which linalool and linalool acetate impart a pleasant scent. We have recently constructed a cDNA library from the secretory cells of floral glandular trichomes of *L. x intermedia* plants. In this thesis we describe the cloning of a borneol dehydrogenase (LiBDH), a putative linalool acetyltransferase (LiLAT), and a caryophyllene synthase (LiCPS) cDNA from this library. The 780 bp open reading frame (ORF) of the LiBDH cDNA encoded a 259 amino acid short chain alcohol dehydrogenase enzyme with a predicted molecular mass of ca. 27.5 kDa. The recombinant LiBDH was expressed in *E. coli*, purified by Ni-NTA agarose affinity chromatography, and functionally characterized. The bacterially produced enzyme specifically converted borneol to camphor as the only product with K_m and k_{cat} values of 53 μM and $4.0 \times 10^{-4} \text{ s}^{-1}$, respectively. The *LiBDH* transcripts were detected both in leaf and flower tissues. However, they were concentrated in floral glandular trichomes of mature *L. x intermedia* flowers indicating that like other *Lavandula* monoterpene synthases the expression of this gene is regulated in a tissue-specific manner.

Using the same procedures described above a putative sesquiterpene synthase (LiCPS) was also cloned and functionally characterized. LiCPS produced caryophyllene from farnesyl diphosphate (FPP), and α -terpieol, 1,8-cineole, and a few other monoterpenes from geranyl diphosphate (GPP) and neryl diphosphate (NPP). Further, two additional lavender TPS's, LiLAT and *L. angustifolia* terpene synthase like protien-I (LaTPS-I), were expressed in bacteria and assayed. Purified recombinant LiLAT and LaTPS-I, however, did not produce detectable amounts of any product *in vitro*.

We believe that the cloning of lavender genes including *LiBDH* and LiCPS has far reaching implications for improving the quality of *Lavandula* EOs through metabolic engineering.

Preface

Chapter 2, Section 2.4 “Glandular trichome extraction” has been published in: Alexander Lane, Astrid Boecklemann, Grant N. Woronuk, **Lukman S. Sarker**, Soheil S. Mahmoud (2010). A genomics resource for investigating regulation of essential oil production in *Lavandula angustifolia*. *Planta*. 231:835–845. I collected the tissue and optimized the method for glandular trichome extraction from various lavender tissues.

Chapter 2, Section 2.5.1 “cDNA library construction” and section 2.5.2 “Microarray analysis” has been published in: Zerihun A. Demissie, Monica A. Cella, **Lukman S. Sarker**, Travis J. Thompson, Mark R. Rheault, Soheil S. Mahmoud (2011). Cloning, functional characterization and genomic organization of 1,8-cineole synthases from *Lavandula*. *Plant Mol Biol*. 79:393–411. The cDNA library construction was carried out at Plant Biotechnology Institute (PBI), and microarray experiment was done at University of Toronto. My contribution was to isolate glandular trichomes, extract RNA from glandular trichomes for cDNA library and microarray data experimentation, match the microarray probe ID with cDNA library annotation, and complete the microarray data analysis. In this publication, I also cloned cineol synthase genomic DNA from *L. angustifolia*, *L. x intermedia* and *L. latifolia* flower tissues.

The work presented in Chapter 3, Section 3.2 “Short chain alcohol dehydrogenase” has been published in: **Lukman S. Sarker**, Mariana Galata, Zerihun A. Demissie, Soheil S. Mahmoud (2012). Molecular cloning and functional characterization of borneol dehydrogenase from the glandular trichomes of *Lavandula x intermedia*. *Arch. of Biochem. and Biophysics*. 528: 163–

170. I conducted this work under the supervision of Dr. Soheil Mahmoud. Mariana Galata cloned one medium chain alcohol dehydrogenase and Zerihun Demissie helped me conduct the qRT-PCR experiment.

Chapter 3, Section 3.5 “LaTPS-I” has been published in: Zerihun A. Demissie, **Lukman S. Sarker**, Soheil S. Mahmoud (2011). Cloning and functional characterization of beta-phellandrene synthase from *Lavandula angustifolia*. *Planta*. 233: 685-696.

Table of Contents

Abstract.....	ii
Preface.....	iv
Table of Contents	vi
List of Tables	ix
Table of Figures.....	x
List of Symbols	xii
Acknowledgements	xv
Dedication	xvi
1 Chapter: Introduction	1
1.1 Composition and significance of lavender EO	1
1.1.1 Medicinal and commercial importance	2
1.2 Terpene Biosynthesis.....	4
1.2.1 Terpenes	4
1.2.2 Stages of terpene biosynthesis.....	5
1.2.3 Isoprene biosynthesis	5
1.2.4 Mono- and sesquiterpene synthesis	6
1.2.4.1 Mono- and sesquiterpene synthases	7
1.2.4.2 Short chain alcohol dehydrogenase/reductase (SDR)	9
1.2.4.3 Acetyltransferase.....	10
1.3 Storage and secretion of volatile terpenes	11
1.4 Regulation of Terpene synthesis.....	12
1.5 Research purpose	13
2 Chapter: Materials and Methods	27
2.1 Chemicals and reagents	27
2.2 Bacteria and plasmids	27
2.3 Plant material	27
2.4 Glandular trichome extraction	28

2.5	RNA extraction and reverse transcription	28
2.5.1	cDNA Library construction	28
2.5.2	Microarray data generation	30
2.6	Cloning and functional characterization	30
2.6.1	Primer design & PCR to amplify the gene of interest	30
2.6.2	Cloning of full length GOI and expression into <i>E.coli</i>	31
2.6.3	Recombinant protein purification and SDS-PAGE	31
2.6.4	Enzymatic Assay reaction	32
2.6.4.1	Alcohol dehydrogenase	32
2.6.4.2	Sesquiterpene synthase & Putative LaTPS-I	33
2.6.4.3	Acetyltransferase	33
2.7	Enzyme kinetics study for LiBDH	33
2.7.1	Gas Chromatography / Mass Spectrometry	34
2.8	Transcript analysis	34
2.8.1	Standard PCR reaction	34
2.8.2	Real time quantitative PCR (qPCR)	35
2.9	Multiple sequence alignment and phylogenetic tree construction	35
3	Chapter: Results.....	42
3.1	cDNA library	42
3.2	Short chain alcohol dehydrogenase	43
3.2.1	Candidate selection	43
3.2.2	Functional analysis of recombinant LiBDH	43
3.2.3	Kinetics study	45
3.2.4	Tissue specific regulation of <i>LiBDH</i>	45
3.2.5	Sequence comparison and phylogenetic tree analysis	46
3.3	Multiproduct sesquiterpene synthase	46
3.3.1	Candidate selection	46
3.3.2	Functional analysis of recombinant LiCPS	47
3.3.3	Tissue specific regulation of <i>LiCPS</i>	47
3.3.4	Multiple sequence alignment and phylogenetic tree analysis	47
3.4	Alcohol acetyltransferase	48
3.4.1	Candidate selection	48
3.4.2	Cloning and Functional analysis of recombinant LiLAT	49
3.4.3	Sequence comparison and phylogenetic tree analysis	50

3.5	LaTPS-I	50
3.5.1	Candidate selection	50
3.5.2	Tissue specific regulation of <i>LaTPS-I</i>	51
3.5.3	Sequence comparison and phylogenetic tree analysis.....	51
4	Chapter: Discussion	77
4.1	Cloning and functional characterization of LiBDH.....	77
4.2	Cloning and functional characterization of LiCPS	80
4.3	Cloning and functional characterization of LiLAT	82
4.4	Cloning and functional characterization of LaTPS-I.....	83
4.5	Conclusion	85
	References	87

List of Tables

Table 1.1 Major mono- and sesquiterpenes in lavender species.....	16
Table 2.1 Supplier and Genotype of bacterial strains used in this project.....	36
Table 2.2 <i>L. x intermedia</i> tissues used in this project.	39
Table 2.3 Lists of primers used in this project.....	40
Table 2.4 Microarray experiment design.	41
Table 3.1 GO annotation of <i>L. x intermedia</i> gland cDNA library.	52
Table 3.2 Microarray analysis.....	53

Table of Figures

Figure 1.1 Biosynthesis of IPP and DMAPP.....	17
Figure 1.2 Overview of terpene biosynthesis..	19
Figure 1.3 Schematic presentation of major monoterpene synthesis of lavender.....	20
Figure 1.4 Schematic presentation of major sesquiterpene synthesis of lavender.....	22
Figure 1.5 Proposed pathway for camphor biosynthesis.	24
Figure 1.6 Proposed pathway for linalool acetate biosynthesis.....	25
Figure 1.7 Schematic diagram of a glandular trichome.....	26
Figure 2.1 Map of pET41b(+) vector.....	37
Figure 2.2 Map of pGEX4T1 vector.....	38
Figure 3.1 Extracted glandular trichomes from 30% flowering stage of <i>L. x intermedia</i>	54
Figure 3.2 Protein purification and GC-MS analysis of LiBDH.	55
Figure 3.3 Kinetic assay of LiBDH with borneol as a substrate.....	57
Figure 3.4 Transcriptional activity of <i>LiBDH</i> in different tissues of lavender species.....	58
Figure 3.5 Multiple alignment of the deduced amino acid sequences of LiBDH.....	60
Figure 3.6 Phylogenetic tree analysis of the “Classical Group” of plant SDRs.	61
Figure 3.7 Enzymatic assays of LiCPS.....	62
Figure 3.8 Amplification of <i>LiCPS</i> transcripts by PCR.....	64
Figure 3.9 Multiple alignment of the deduced amino acid sequences of LiCPS.....	65
Figure 3.10 Phylogenetic tree analysis of LiCPS	66
Figure 3.11 Amino Acid sequence comparisons of three LiLAT candidates.....	67
Figure 3.12 Amino acid sequence comparisons of LiLAT-1 and LiLAT-4.	68

Figure 3.13 Restriction enzyme digestion of LiLAT clones after cloning into pET41b(+) vector.....	69
Figure 3.14 Restriction enzyme digestion of LiLAT clones in pGEX4T1 vector.....	70
Figure 3.15 SDS-PAGE analysis of protein samples extracted from bacterial cells expressing LiLAT candidates	71
Figure 3.16 Phylogenetic tree analysis of LiLAT candidates	72
Figure 3.17 SDS-PAGE analysis of LaTPS-I.	73
Figure 3.18 <i>LaTPS-I</i> transcript analysis.....	74
Figure 3.19 Multiple alignment of the deduced amino acid sequence of LaTPS-I.	75
Figure 3.20 Phylogenetic tree analyses of LaTPS-I	76

List of Symbols

aa	Amino acid
BDH	Bornyl dehydrogenase
bp	Base pair
BSA	Bovine serum albumin
CoA	Coenzyme A
Contig	A set of overlapping DNA segments that together represent a consensus region of DNA
cDNA	Complementary DNA
cv.	Cultivar
DMAPP	Dimethyl allyl diphosphate
dNTPs	Deoxynucleotide tri phosphate mix
DOXP	1-deoxy-D-xylulose 5-phosphate
DTT	Dithiothreitol
DXP	Deoxy-xylulose-phosphate
DXS	Deoxy-xylulose-P synthase
EDTA	Ethylenediaminetetraacetic acid
EO	Essential oil
EST	Expressed sequence tag
FPP	Farnesyl diphosphate
GC/MS	Tandem gas chromatography mass spectrometry
GGPP	Geranyl geranyl diphosphate

GO	Gene Ontology
GOI	Gene of interest
GPP	Geranyl diphosphate
GST	Glutathione S-transferase
HMBPP	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate
HMG-CoA	3-hydroxy-3-methylglutaryl CoA
IDT	Integrated DNA technologies
IPP	Isopentyl diphosphate
IPTG	Isopropyl β -D-1-thiogalactopyranoside
ISO	International organization for standardization
LaLINS	<i>L. angustifolia</i> linalool synthase
LaTPS-I	<i>L. angustifolia</i> terpene synthase like-I
LiBDH	<i>L. x intermedia</i> borneol dehydrogenase
LaBERS	<i>L. angustifolia</i> trans α -bergamotene synthase
LiCPS	<i>L. x intermedia</i> caryophyllene synthase
LiLAT	<i>L. x intermedia</i> linalool acetyltransferase
LiSDR	<i>L. x intermedia</i> SDR
MDR	Medium chain alcohol dehydrogenase
MECP	2-C-methylerythritol-2,4-cyclodiphosphate
MEP	2-C-methyl-D-erythritol 4-phosphate
mg/gfWT	mg per gram fresh weight
MOPSO	3-morpholino-2-hydroxypropanesulfonic acid
mRNA	Messenger RNA

MVA	Mevalonic acid
NAD ⁺	Nicotinamide adenine dinucleotide
Ni-NTA agarose	Nickel-charged nitrilotriacetic acid affinity resin column
NPP	Neryl diphosphate
ORF	Open reading frame
PCR	Polymerase chain reaction
PO ₄ buffer	Phosphate buffer
PTV	Programmable temperature vaporizing
PVP-40	Polyvinylpyrrolidone, absorbs upto 40% of its weight in atmospheric water
qRT-PCR	Quantitative reverse transcriptase PCR
RE	Restriction enzyme
SDR	Short chain alcohol dehydrogenase
SDS-PAGE	Sodium dodecyl sulfate- polyacrilamide gel
T _m	Annealing temperature
TPS	Terpenoid synthase
U	Unit- amount of the enzyme that catalyzes the conversion of 1 micro mole of substrate per minute
UV	Ultraviolet
XADS bead	Amberlite XAD polymeric adsorbents

Acknowledgements

All praises to the almighty Allah.

First, I would like to thank all UBC's Okanagan campus faculty and staff who have helped me with my studies. In particular, I am grateful to my supervisor Dr. Soheil Mahmoud for providing me with the opportunity of conducting my research in his laboratory. His guidance and support helped me exceed all my personal expectations, and raised my confidence level to new heights. I would like to express my heartfelt gratitude to my committee members, Dr. Kirsten Wolthers and Dr. David Theilmann. I feel very fortunate to have had them as mentors. My sincere gratitude to Dr. Kirsten Wolthers, Dr. Mark Rheault and their staff (Brienne J. Matier, Carla Meints and Caitlyn Makins) for the use of their facilities and assistance with my project. I would like to thank my coworkers in Dr. Mahmoud's laboratory, in particular my good friend and colleague Zerihun A. Demissie. Zerihun was always there to help me with my questions and experiments. Finally, I would like to thank UBC Okanagan, Natural Sciences and Engineering Research Council of Canada, Investment Agriculture Foundation of British Columbia, NRC Plant Biotechnology Institute, and Genome British Columbia, for supporting my research through grants and/or in-kind contributions to Dr. Soheil S. Mahmoud.

Heartiest appreciation goes out to my parents and brothers for being the source of my inspirations. Special thanks to my beautiful wife for providing me with motivation and tolerating me being an unhelpful person at home ☺.

To my family

1 Chapter: Introduction

Plants produce thousands of primary and secondary metabolites, including volatile compounds, at various developmental stages throughout their life cycle, during flowering, ripening, and maturation. Individual species in *Planta* synthesize a unique blend of volatile compounds which generates the “flavor fingerprint” (Goff and Klee 2006). Lavenders, known for their volatile compounds, are small aromatic shrubs cultivated worldwide for essential oil (EO) (Lane et al. 2010). The world market for lavenders was estimated to be CAD \$200 million in 2005 and has grown substantially since then (www.ienica.net). Bulgaria, England, the United States, and France are the leading countries in the world that produce the most amount of lavender EOs. Considerable amounts are also produced in Canada, with more than 18 farms in British Columbia alone.

1.1 Composition and significance of lavender EO

Like other *Lamiacea* plants, such as common sage (*Salvia sp.*), mint (*Mentha sp.*) and thyme (*Thymus sp.*), the genus *Lavandula* is a member of the family *Lameacea*, which is composed of over 32 morphologically distinct species including *L. angustifolia* (English lavender), *L. latifolia* (Spike lavender), and *L. x intermedia* (Lavandin) (Upson et al. 2004). It is important to note that Lavandin is derived from a natural cross of Spike and English lavender. Lavender EOs- a blend of mono and sesquiterpenoid alcohols, esters, oxides, and ketones- are extensively used in cosmetics, hygiene products, and alternative medicines. Around 50-60 monoterpenes have been identified in different lavender varieties, although only a few components determine the characteristic EO of a given species (Upson et al. 2004). The most abundant monoterpenes found in lavenders include linalool, linalool acetate, borneol, camphor, and 1,8-cineole. Among these, camphor, linalool, and linalool acetate are key

determinants of the lavender EO quality (Lis-Balchin 2002, Upson et al. 2004). EOs with a high linalool and linalool acetate to camphor ratio are considered to be of “high quality”, and thus are used in cosmetic products and aromatherapy (Cavanagh and Wilkinson 2002, Lane et al. 2010). EOs added to alternative medicines are typically rich in camphor and 1,8-cineole. Lavender EO also contains sesquiterpenes such as caryophyllene, bergamotene and nerolidol with trace amounts of other terpenoid compounds such as perillyl alcohol.

Lavender EO composition is greatly influenced by environmental factors, and the species it is collected from (Cavanagh and Wilkinson 2002). Oil composition for the most common lavender species: English lavender, Spike lavender, and Lavandin are listed in Table 1.1. Though it is smaller in size and the oil yield is relatively low, *L. angustifolia* has a better linalool and linalool acetate to camphor ratio compared to *L. x intermedia* and *L. latifolia*. *L. latifolia* contains large quantities of camphor while producing a small amount of linalool and linalool acetate, making it more useful for the alternative medicine industry. On the other hand, Lavandin produces an EO with a less favorable linalool and linalool acetate to camphor ratio; the overall oil yield is much higher, and the plant has better adaptability in cold weather (Interactive European Network for Industrial Crops and their Applications (IENICA) September 27, 2002). The choice of lavender variety is therefore dependent on required oil yield, and quality (higher quality oils for pure EOs, fragrance, and medical application; lower quality oils for soaps and detergents), and the growth environment (Boeckelmann 2008).

1.1.1 Medicinal and commercial importance

Lavender was named the “Herb of the year” in 1999 by the Herb Growing and Marketing Network in the United States because of its use in the therapeutic and cosmetic industry. This use can be traced back to the ancient Roman and Greek era (Cavanagh and Wilkinson 2002).

The lavender flower was used in the mummification process during the ancient Egyptian times. More recently the EO of this plant has been prescribed for treating infection, anxiety, infertility, and fever. It has also been used as an anti-depressant, anti-spasmodic, anti-flatulent agent, anti-emetic remedy and a diuretic (Chu and Kemper 2001). Lavender EO is very popular in aromatherapy and has gained a good reputation for relieving stress, depression and insomnia (Wolfe and Herzberg 1996). In animal trials, lavender EO was found to have positive effects on pulmonary diseases. It also had sedative/hypnotic, anxiolytic, anticonvulsant properties, and helped cognitive function, though there is no conclusive evidence using human trials (Chu and Kemper 2001). Studies suggest that Lavender aroma prevents deterioration of work performance during recesses (Sakamoto et al. 2005), and might improve memory and cognition in Alzheimer's patients (Adersen et al. 2006). Lavender essential oils have numerous other medicinal applications. For example, linalool, linalool acetate, and cineole have antibacterial, antifungal, and insecticide properties (Cosentino et al. 1999, Pattnaik et al. 1997). Lavender oils containing high camphor content are used in inhalants to relieve coughs and colds (Theis and Koren 1995), and as active ingredients in liniments and balms used as topical analgesics (Xu et al. 2005). Camphor has also been considered as a potential radio sensitizing agent, and has been used in oxygenating tumors prior to radiotherapy (Goel and Roa 1988, Guilandcumming and Smith 1979). Perillyl alcohol, a monoterpene found in trace amounts in *L. angustifolia* (Perrucci et al. 1994), has drawn recent interest due to its chemopreventative and chemotherapeutic properties (Hohl 1996, Peffley and Gayen 2003, Schulz et al. 1994). Caryophyllene oxide, found in *L. angustifolia* and *L. latifolia*, has strong anti-inflammatory effects.

Lavender EOs are also extensively used in the cosmetics industry. In Victorian, Medieval and Renaissance periods lavender EO was used for the storage of linens and to disguise objectionable odours (Chu and Kemper 2001). Now a day, we commonly find Lavender EO in a wide variety of perfumes and soaps, and the lavender-based perfumery/cosmetic industry is growing worldwide.

1.2 Terpene Biosynthesis

1.2.1 Terpenes

Lavender EO is a complex mixture of many different aromatic compounds; however, terpenes are its primary constituents. Terpenes are naturally occurring organic hydrocarbons that play a tremendous physiological and developmental role in plants. Terpenes are produced through the polymerization of a five-carbon unit called ‘isoprene’, and are classified based on the number of isoprene units they contain. The smallest terpenes contain only a single five-carbon unit, and are called hemiterpenes. The best known hemiterpene is isoprene itself, which is released from photosynthetically active plant tissues. Monoterpenes are composed of two five-carbon units and predominate in the volatile essences of flowers and the EO of spices and herbs. Sesquiterpenes contain three five-carbon units and, like monoterpenes, they are volatile components in essential oils. In addition, sesquiterpenes act as phytoalexins, antibiotic compounds, and anti-feedants. Di-terpenes contain four five-carbon units and include phytols, gibberellin hormones, and phytoalexins. Some di-terpenes, like taxol and forskolin, are pharmacologically important in the treatment of cancer and glaucoma, respectively. Triterpenes contain six five-carbon units and include phytosterol membrane components, certain phytoalexins, various toxins, and feeding deterrents.

Tetraterpenes that contain eight five-carbon units are the carotenoids accessory pigments, and are essential to photosynthesis (Cavanagh and Wilkinson 2006, Croteau et al. 2000).

1.2.2 Stages of terpene biosynthesis

Terpene biosynthesis can be conveniently divided into four stages: Stage 1 involves the production of isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). In stage 2, IPP and DMAPP are condensed in a “head-to-tail fashion” to form the higher order isoprenoid building blocks: geranyl pyrophosphate (GPP; C₁₀) or neryl diphosphate (NPP; C₁₀), farnesyl pyrophosphate (FPP; C₁₅), and geranylgeranyl pyrophosphate (GGPP; C₂₀). The third stage of terpene synthesis involves the conversion of GPP, FPP and GGPP to respective terpene groups. The final stage of terpene synthesis occurs mostly in the cytosol and involves redox transformations, cyclization, and/or carboxylation, to produce derivatives of the basic terpene groups. For example, the monoterpene linalool is carboxylated to produce linalool/linalyl acetate (Croteau et al. 1978, Lane et al. 2010, Liu et al. 2005, Mahmoud and Croteau 2002, Piel et al. 1998).

1.2.3 Isoprene biosynthesis

In plants, two independent but interactive pathways called Mevalonate (MVA) or cytosolic, and 2-C-methyl-D-erythritol 4-phosphate (MEP) or plastidial pathway, produce the general terpene precursors IPP and DMAPP (Arigoni et al. 1997; Bick and Lange 2003; Gershenzon et al. 2000; Laule et al. 2003). The MVA pathway is the only pathway found in animals and fungi as well as in the cytoplasm of phototropic organisms. Precursors produced through this pathway are mainly converted to FPP to synthesize sesquiterpenes, and triterpenes among others (Chappell 1995a, Chappell 1995b, McGarvey and Croteau 1995). The MEP pathway, present in most bacteria and in plant chloroplasts, provides precursors for the biosynthesis of

GPP and GGPP that are ultimately used to produce monoterpenes and diterpenes, respectively (Mahmoud and Croteau 2001, Mahmoud and Croteau 2003, Mahmoud et al. 2004).

The MVA pathway initiated by the three molecules of acetyl-coenzyme (Co)A yields 3-hydroxy-3-methylglutaryl CoA (HMG-CoA). The enzyme HMG-CoA reductase reduces HMG-CoA to mevalonic acid (MVA), which is then converted to mevalonate 5-diphosphate by mevalonate kinase and mevalonate 5-phosphate kinase. Mevalonate 5-phosphate is subsequently decarboxylated to yield IPP (Figure 1.1) (Liu et al. 2005). The MEP pathway or DXP pathway, inaugurated by the condensation of pyruvate and glyceraldehyde-3-phosphate to 1-deoxy-D-xylulose 5-phosphate (DOXP), which is catalyzed by DOXP synthase (DXPS). DOXP is then reduced to 2-C-methyl-D-erythritol 4-phosphate (MEP) by DX reductoisomerase (DXR). The cytidine 5-phosphate derivative is synthesized from MEP, which then undergoes phosphorylation and cyclization to produce 2-C-methylerythritol-2,4-cyclodiphosphate (MECP). 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate (HMBPP) synthase converts MECP to HMBPP, and is then transformed to IPP and DMAPP (Figure 1.1) (Liu et al. 2005).

1.2.4 Mono- and sesquiterpene synthesis

Mono- and sesquiterpenes are derived from the precursors GPP and FPP, respectively, by the activity of various terpene synthases (sometimes called cyclases) (Figure 1.2). Some monoterpenes, such as camphor and linalool acetate, are further modified through acetylation, oxidation or, reduction reactions. Camphor is produced from borneol by the action of a short chain alcohol dehydrogenase and linalool acetate is produced from linalool by the linalool acetyltransferase enzyme. Monoterpene synthases initiate the reaction by

forming cationic intermediates such as geranyl cation, linalyl diphosphate (transoid and cisoid), linalyl cation, and finally the α -terpinyl cation. These intermediate products individually experience a number of cyclizations, hydride shifts, or other rearrangements until they form a stable component. For example, α -terpinyl cation, a critical branchpoint intermediate, forms all cyclic monoterpenes such as limonene, α -terpineol, pinene, phellandrene, sabinene, terpinene, borneol, camphor, cineol, etc. Geraniol, linalool, myrcene, and β -ocimene are derived from geranyl and linalyl cation (Figure 1.3). Two intermediate cationic forms, such as farnesyl cation and its isomer nerolidyl cation, are produced by sesquiterpene synthases before any rearrangement occurs for stable compounds (Degenhardt et al. 2009) (Figure 1.4).

1.2.4.1 Mono- and sesquiterpene synthases

To date, hundreds of different mono and sesquiterpene synthases have been isolated and cloned from a number of plants including mint, lemon, snapdragon, sage, and Arabidopsis and even gymnosperms like grand fir (for details please follow (Degenhardt et al. 2009)). Interestingly, most terpene synthases possess similar properties, such as a native 50-100 kDa molecular mass range (either monomeric or dimeric), a requirement for a divalent metal ion as cofactor for catalysis (usually Mg^{2+} or Mn^{2+} for angiosperms, K^+ , Mn^{2+} , Fe^{2+} for gymnosperms), a pI near 5.0 and a pH optimum within a unit of neutrality (Bohlmann et al. 1998). In general, plant monoterpene synthases (600-650 amino acids) are larger than sesquiterpene synthases (550-580 amino acids) due to the N-terminal signal peptide sequences which target the protein towards the plastids. The N-terminal signal peptides contain a high frequency of serine and threonine residues with low amounts of acidic amino acids; however, they do not share any common sequence similarities (Bohlmann et al. 1997).

Sequence analysis of terpene synthases from different plant species revealed four conserved motifs, i.e. the RR(X₈)WD motif, LQLYEASFLL motif, DDXXD and (N,D)D(L,I,V)X(S,T)XXXE motifs (Bohlmann et al. 1998). The arginine rich N-terminal RR(X₈)W motif is essential for cyclization of GPP and the enzymatic activity of many monoterpene synthases (Williams et al. 1998b), while the LQLYEASFLL motif is thought to be part of the active site (McGeady and Croteau 1995, Wise et al. 1998). The aspartate rich regions: DDXXD and (N,D)D(L,I,V)X(S,T)XXXE motifs, are responsible for the enzymatic activity and coordination of divalent cations and are thus responsible for substrate binding and ionization, respectively (Christianson 2006, Whittington et al. 2002). The DDXXD motif is highly conserved compared to the less conserved (N,D)D(L,I,V)X(S,T)XXXE motif in almost all plant terpene synthases, and both of these motifs bind a trinuclear magnesium cluster involved in coordination of the pyrophosphate group of the substrate (Zhou and Peters 2009). Site-directed mutagenesis studies revealed that this region is very important for terpene catalysis, as mutations in this region frequently lower or completely abolish the catalytic activity, while other alteration to this region lead to abnormal products (Cane et al. 1996a, Cane et al. 1996b, Degenhardt et al. 2009, Seemann et al. 2002b). However, a NDXXD motif which is a natural variant of DDXXD motif of (+) germacrene synthase from goldenrod has no impact on catalytic activity. This shows that the highly conserved DDXXD motif is not as necessary for catalytic activity in farnesyl diphosphate cyclization (Prosser et al. 2004). Moreover, some sesquiterpene synthases share a repeated DDXXD motif instead of (N,D)D(L,I,V)X(S,T)XXXE motif which is also involved in catalysis (Little and Croteau 2002, Steele et al. 1998).

Phylogenetic analysis has discerned that terpenoid synthases (TPS) are categorized into six gene subfamilies (designated TPSa-TPSf). TPSa is constituted by sesquiterpene and diterpene synthases from angiosperms. TPSd is comprised of 11 gymnosperm mono-, sesqui- and di-terpene synthases. Many monoterpene synthases, including identified monoterpene synthases from *Lamiaceae*, belong to the TPSb family (Bohlmann et al. 1998, Trapp and Croteau 2001). TPSc, TPSe and TPSf are represented by single angiosperm terpene synthases, i.e. the diterpene synthases such as copalyl diphosphate synthase, kaurene synthase, and the angiosperm linalool synthase, respectively.

1.2.4.2 Short chain alcohol dehydrogenase/reductase (SDR)

The SDR family includes various oxido-reductase, isomerase, and lyase enzymes which lead to the production of some new monoterpenes from the regular monoterpenes; this family also exhibits a variety of substrate specificities for steroids, retinoids, prostaglandins, sugars, alcohols and other small molecules (Figure 1.5) (Jornvall et al. 1995, Kallberg et al. 2002, Kallberg et al. 2010, Okamoto et al. 2011). SDRs are usually 750-800 nucleotides, 250-275 amino acid, long with a molecular mass of ca. 28 – 30 kDa (Jornvall et al. 1995). Two conserved motifs have been found in amino acid sequence comparisons between the SDR enzymes, even though pairwise identities are quite low (10%-30%) (Jornvall et al. 1995, Kallberg et al. 2002, Kallberg et al. 2010, Kavanagh et al. 2008). These two motifs are the coenzyme binding motif GXXXGXXG, and an active site pattern of YXXXXK. The first SDR crystal structure revealed that the coenzyme and substrate binding sites fall into a single domain which is clearly distinct from the structures of the medium chain dehydrogenase/reductase (MDR) enzymes that are composed of two separate domains (Tanaka et al. 2001). The active site motif, YXXXXK, positioned at 155-159 in ZSD1 from *Zingiber zerumbet*, is

one of the most commonly conserved motifs in SDR. A Ser142 residue (13 residues upstream from the Tyr) is conserved in most SDR enzymes. The ‘Ser-Tyr-Lys triad’ is responsible for the catalysis of SDR. Again, one Asp residue at the N-terminal end (Asp39 for ZSD1) plays a critical role for determining the coenzyme specificity for NAD(H) over NADP(H) in the SDR enzyme (Okamoto et al. 2011).

1.2.4.3 Acetyltransferase

Acylation is a common and biochemically significant modification of plant secondary metabolites. Plant BAHD acyltransferases comprise a large family of acyl CoA utilizing enzymes which produce volatile esters, as well as constitutive defense compounds and phytoalexins (Figure 1.6) (St-Pierre and De Luca 2000). The BAHD acetyltransferase family was named according to the first letter of each of the first four biochemically characterized enzymes of the family (BEAT-*benzylalcohol O-acetyltransferase*, AHCT-*anthocyanin O-hydroxycinnamoyltransferase*, HCBT- *anthranilate N-hydroxycinnamoyl/benzoyltransferase*, and DAT- *deacetylindoline 4-O-acetyltransferase*) (St-Pierre and De Luca 2000). The BAHD members are cytosolic monomers with an average of 445 amino acids, and molecular masses ranging from 48 to 55 kDa (D'Auria 2006). Analysis of sequence similarities revealed some conserved domains among the BAHD family members. The first one is the HXXXDG domain located near the center portion of each enzyme, which has been shown to be important for a general base catalysis (D'Auria 2006, Shaw 1992, St-Pierre and De Luca 2000). The oxygen or nitrogen atom of the corresponding substrate is deprotonated by the histidine residue of the HXXXDG motif allowing a nucleophilic attack on the carbonyl carbon of the coenzyme A thioester. It forms a tetrahedral intermediate between the coenzyme A thioester and the acceptor substrate. The intermediate tetrahedral is reprotonated

to produce the free CoA and the acylated ester or amide (D'Auria 2006). The second highly conserved region is the DFGWG motif, located in the carboxyl end, which is believed to have a structural role in enzymatic function, but is not strictly required for membership within the BAHD family of enzymes (D'Auria 2006). Nearly all functionally characterized BAHD enzymes contain both of these motifs. Site directed mutagenesis experiments showed that deletion or modification of one or both of these motifs result in highly reduced enzyme activity (D'Auria 2006).

1.3 Storage and secretion of volatile terpenes

There are numerous functions of the volatile compounds produced by plants. For example, they help to attract pollinators during pollination, protect the plant from herbivorous attacks, or act as a pathogen deterrent. In some plants, e.g., lavenders and mints, these volatile compounds are produced and accumulated in a specialized secretion structures called glandular trichomes (Fahn 1988, Lis-Balchin 2002). Glandular trichomes are modified epidermal cells that cover leaves, stems and parts of the flower. There are two forms of glandular trichomes available in mint including capitate and peltate glandular trichomes. The capitate glandular trichomes are smaller in size and simple in form having only a basal cell, a short stalk and a one to two cell head (Fahn 1988). On the other hand, peltate glandular trichomes are complex in structure and consist of secretory cells (usually eight disc cells), a stalk cell and a basal cell anchoring the trichome in the epidermis (Figure 1.7) (Fahn 1988). Essential oils are stored in the subcuticular space between the cuticle and the apical walls of the secretory cells. The exact secretory mechanism is not known yet; however, it is believed that volatile compounds are secreted through a diffusion system through the cuticle (Fahn 1988).

Volatile compound production is related to the size and age of the glandular trichomes as well as to the number of glands per area of tissue. In recent studies, it was shown that monoterpene synthesis and accumulation are directly controlled by the development of the oil glands during the growth season. For example, linalool content in lavender is proportional to the flower developmental stages (Boeckelmann 2008). The gland development process is rapid and their number increases simultaneously while the tissue matures, especially during the vegetative growth.

1.4 Regulation of Terpene synthesis

The biosynthesis of terpenes is mediated by a group of enzymes called terpene synthases, and is regulated by environmental and developmental cues, and through regulation of gene expression. In peppermint, a genus of *Lamiaceae*, plant growth and oil yield was regulated by the amount of daylight. Shorter days resulted in decumbent plants with small leaves, while longer days with high photon flux density and high night temperatures favoured erect plants, large leaves and flowers, and highest EO yields (Clark and Menary 1980). The terpene profile of a plant also changes during different developmental stages of tissue maturation. β -ocimene and myrcene from snapdragon flowers were undetectable in unopened and 1 day old flowers, but were strongly detected in the anthesis or later stages (Dudareva et al. 2003). In peppermint, monoterpene content reached a peak between twelve to twenty days after leaf emergence and then rapidly declined at full leaf expansion stage (Turner et al. 2000). In lavender, linalool production was developmentally regulated. More specifically, linalool content gradually increased during flower developmental stages starting from bud to full bloom (Boeckelmann 2008).

Although environment and developmental regulation can take place, terpene biosynthesis is mainly regulated at the level of gene expression. Numerous studies have shown that the abundance of terpene synthase was highly correlated with the emission level of their respective terpenes (e.g., β -ocimene synthase from snapdragon, S-linalool synthase from *Clarika breweri*, 1,8-cineole synthase from *L. x intermedia*, and menthofuran synthase from peppermint (Boeckelmann 2008, Bohlmann et al. 1997, Demissie et al. 2012, Dudareva et al. 1996, Dudareva et al. 2003, Mahmoud and Croteau 2003). Mahmoud and Croteau (2003) demonstrated that overexpression and co-suppression of menthofuran synthase resulted in the respective increase or decrease in the production of menthofuran in peppermint. In some cases, manipulation of terpene synthase expression did not show any impact on the EO profile for that plant. For example, overexpression of limonene synthase in peppermint leaves failed to increase the limonene synthase transcript in oil glands (Mahmoud et al. 2004). Further, some terpenes such as farnesol and cadinol were present in very minute amounts in lavender EO, although the transcripts for farnesyl synthase and cadinene synthase were highly abundant in glandular trichome (Lane et al. 2010) suggesting that the precursor, FPP, might be the limiting factor in the production of farnesol and cadinol by farnesyl synthase and cadinene synthase, respectively in lavender glandular trichome (Lane et al. 2010).

1.5 Research purpose

There is increasing demand for high quality lavender EOs. As previously discussed, English Lavenders (*L. angustifolia*) produce the finest quality EO, which contains trace amounts of camphor and high concentrations of linalool and linalool acetate; however, the EO yield in these plants is very low. On the other hand, Lavandins (*L. x intermedia*) have a much better

yield of low quality EOs that contain high amounts of camphor. In principle, it should be possible to enhance EO quality in Lavandins by reducing camphor production in these plants through biotechnology once the gene responsible for the production of this compound is cloned. The main objective of the work presented in this thesis was to identify and clone the gene responsible for the production of camphor. Based on previous studies reported in the literature I hypothesized that Lavandins express an alcohol dehydrogenase (borneol dehydrogenase; BDH) enzyme that oxidizes borneol to camphor, and that camphor production is regulated at the level of BDH transcription. The experiments reported in this thesis examined the following predictions of this hypothesis:

- i) Lavandin cDNA library would contain sequences homologous to known alcohol dehydrogenases.
- ii) Lavandin alcohol dehydrogenase homologues expressed in bacteria would convert borneol to camphor
- iii) The steady state levels of mRNA of the Lavandin alcohol dehydrogenase that converts borneol to camphor would correlate with tissues that produce high levels of camphor.

Three alcohol dehydrogenase candidates from a *L. x intermedia* floral glandular trichome cDNA library were identified and cloned into a bacterial expression vector. Recombinant proteins were produced in *E. coli*, purified, and assayed for camphor production using borneol as a substrate. The transcription activity of the clones encoding BDH activity was analyzed in *L. x intermedia* plants by standard and real time PCR.

In a search of EO related genes, I also identified one contig from the cDNA library of glandular trichomes from *L. x intermedia* containing 53 EST members. This was a

sesquiterpene synthase-like sequence, though it was missing the N-terminal signal peptide. I hypothesized that the highly abundant terpene synthase like sequence was a sesquiterpene synthase. To test this hypothesis, I cloned the gene using a bacterial expression system and analyzed its enzymatic function *in vitro*.

Linalool acetate is absent in the EO of *L. latifolia* whereas it is present in *L. angustifolia* and *L. x intermedia*. Based on our cDNA library and microarray data analysis, I hypothesized that a linalool acetyltransferase is involved in the conversion of linalool to linalool acetate. To test this hypothesis, I selected four candidates and cloned them using the bacterial expression vectors (pET41b(+)) and pGEX4T1 vector).

The previously reported cDNA library of leaf and flower tissues from *L. angustifolia* has revealed a highly abundant (higher numbers of ESTs) terpene synthase like sequence (LaTPS-I) which was missing a stretch of 73 amino acid sequence in the middle of the sequence as compared to the other terpene synthases. This missing region contains the most important DDXXD motif which is involved in divalent metal ion binding as well as interacting with the diphosphate moiety of the substrate. I hypothesized that LaTPS-I did not utilize terpene precursors (GPP, FPP, or NPP) to produce any secondary metabolites. To test this hypothesis, I cloned LaTPS-I into a bacterial expression system and used purified recombinant protein for the enzymatic reaction.

Ultimately, this study will lay the groundwork for future research to improve the quality of lavender essential oil which is necessary to meet current demand from the aromatherapy and alternative medicine industries.

Table 1.1 Major mono- and sesquiterpenes in lavender species.

(English lavender, Lavandin and Spike lavender) (Lis-Balchin 2002)

Content (%) of major terpenes in lavender oil			
	English lavender (<i>L. angustifolia</i>)	Lavandin (<i>L. x intermedia</i>)	Spike lavender (<i>L. latifolia</i>)
Linalool	10-50	20-23	26-44
Linalool acetate	12-54	19-26	0-1.5
<i>cis</i> and / or <i>trans</i> -Ocimen	1.0-17	1.0-3.0	0-0.3
Lavandulol and acetate	0.1-14	0.5-0.8	0.2-1.5
Camphor	0-0.2	12	5.3-14.3
1,8-Cineiol	2.1-3.0	10	25-36
α -and β -Pinene	0.02-0.3	0.6-0.9	1.6-3.6
Borneol	1.0-4.0	2.9-3.7	0.8-4.9
Caryophyllene	3.0-8.0	2.7-6.0	0.1-0.3
Myrcene	0.4-1.3	1.2-1.5	0.2-0.4
Farnesene	Trace	1.1	0.2-0.3
Germacrene D	0.2-0.9	1.0-1.2	Trace
Camphene	0.1-0.2	0.3-0.6	0.2-1.8
Limonene	0.2-0.4	0.9-1.5	1.0-2.2

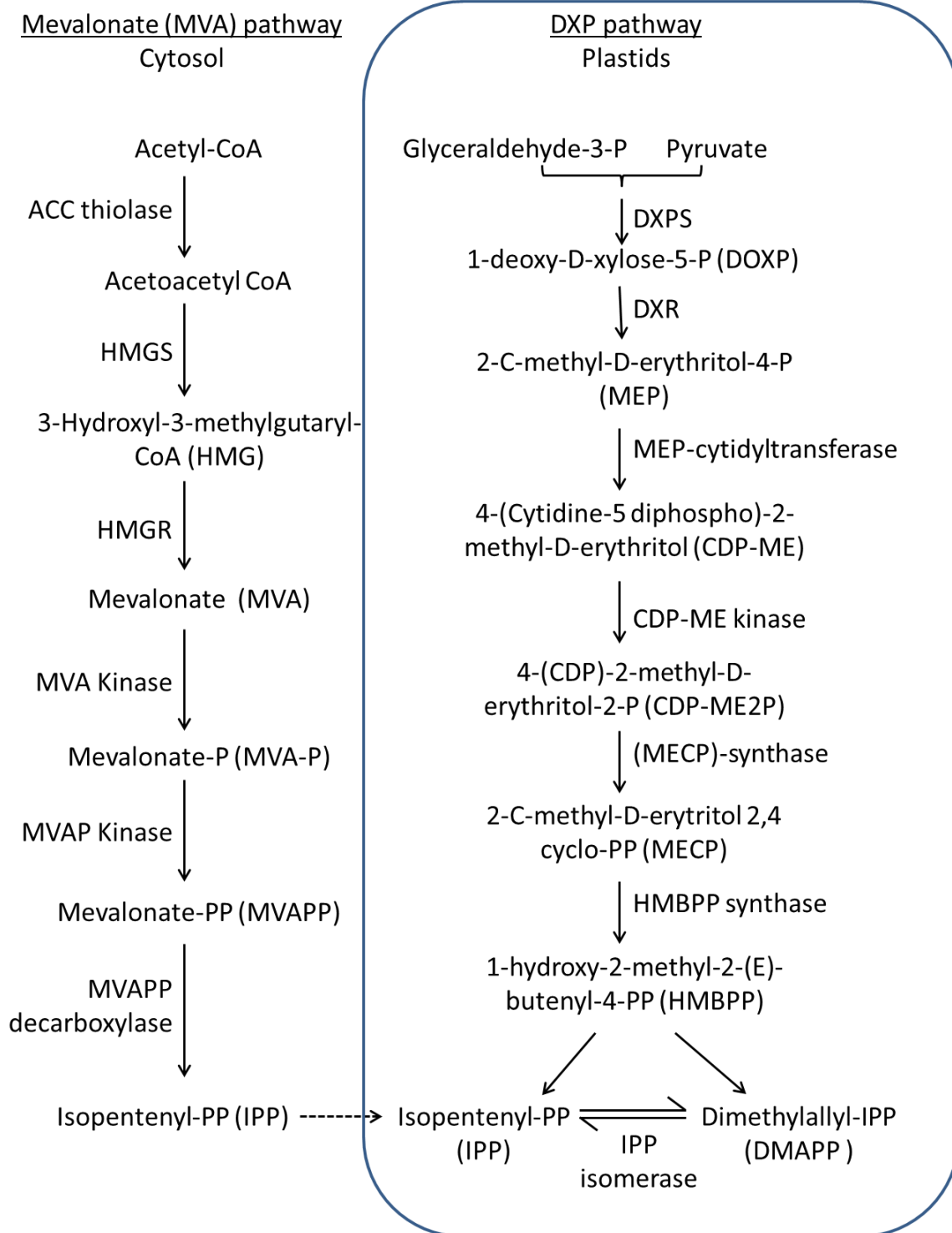


Figure 1.1 Biosynthesis of IPP and DMAPP.

Figure 1.1 Biosynthesis of IPP and DMAPP via the mevalonate pathway (left) and the mevalonate-independent (DXP) pathway (right). The indicated enzymes are: AACT, acetoacetyl-coenzyme A thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MVA kinase, mevalonate kinase; MVAP kinase, mevalonate 5-phosphate kinase; MVAPP decarboxylase, mevalonate-5-diphosphate decarboxylase; DXPS, 1-deoxyxylulose-5-phosphate synthase; DXR, 1-deoxyxylulose-5-phosphate reductoisomerase; MEP cytidyl transferase, 2-C-methylerythritol-4-phosphate cytidyltransferase; CDP-ME kinase, 4-(cytidine-5'-diphospho)-2-C-methylerythritol kinase; MECP synthase, 2-C-methylerythritol-2,4-cyclodiphosphate synthase; HMPPP synthase, 1-hydroxy-2-methyl-E-butenyl-4-diphosphate synthase; HMBPP reductase, 1-hydroxy-2-methyl-E-butenyl-4-diphosphate reductase and IPP isomerase (IPPI). The pathway may give rise to IPP and DMAPP independently of the interconversion catalyzed by IPPI. A transfer of IPP/DMAPP between cytosol and plastid is possible but, as of yet, unproven. Inspired from Mahmoud and Croteau 2002 with permission from Trends in Plant Science.

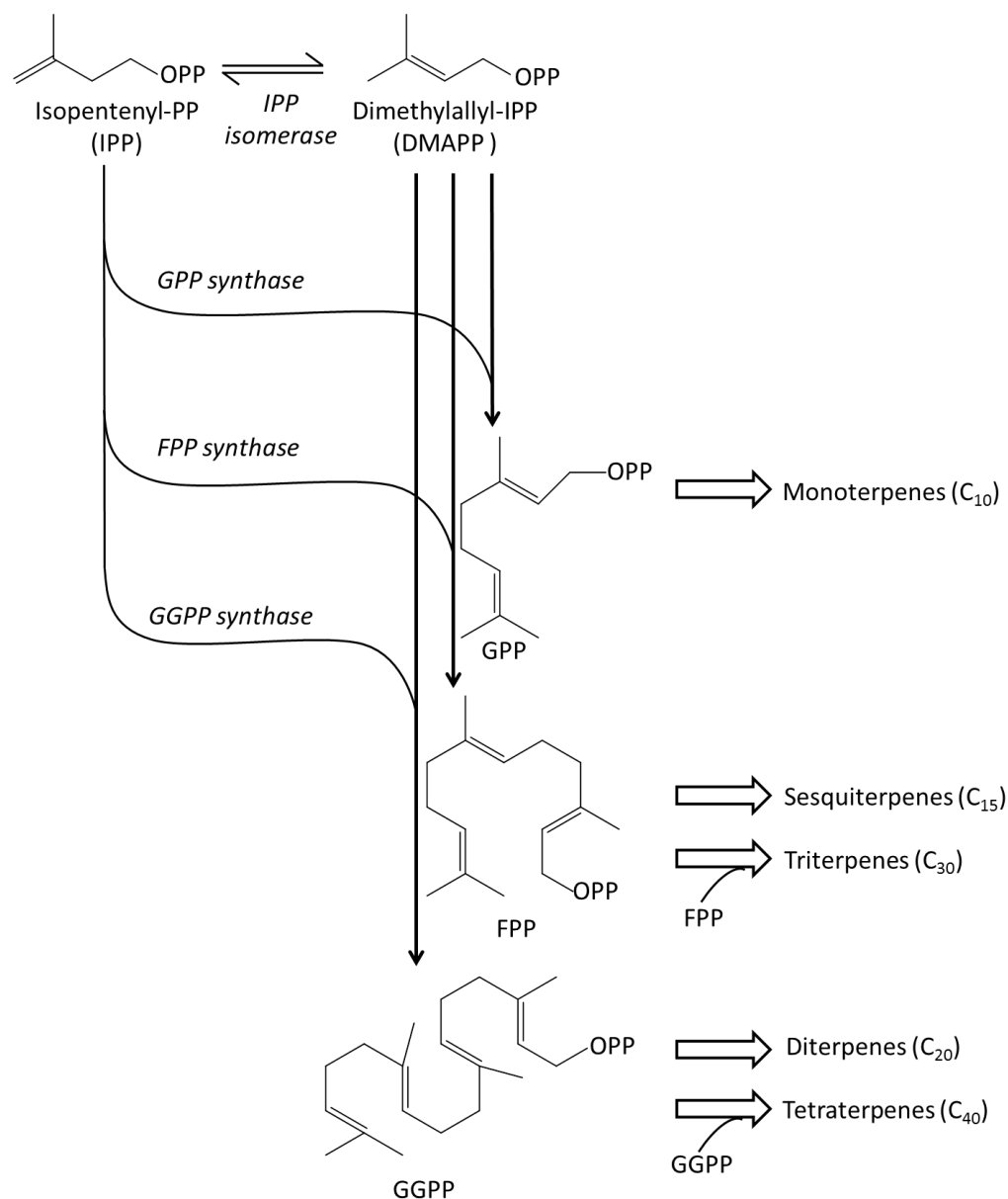


Figure 1.2 Overview of terpene biosynthesis. GPP is synthesized by the condensation of one molecule of IPP and one molecule of DMAPP catalyzed by GPP synthase. FPP is the condensation product of GPP and one molecule of IPP, while GGPP is produced through the condensation of one molecule of GPP and two molecules of IPP. Monoterpenes are a result of the derivatization and rearrangement of GPP, while FPP and GGPP are the precursors to sesqui- and triterpenes, and di- and tetraterpenes, respectively. Inspired from Mahmoud and Croteau 2002 with permission from Trends in Plant Science.

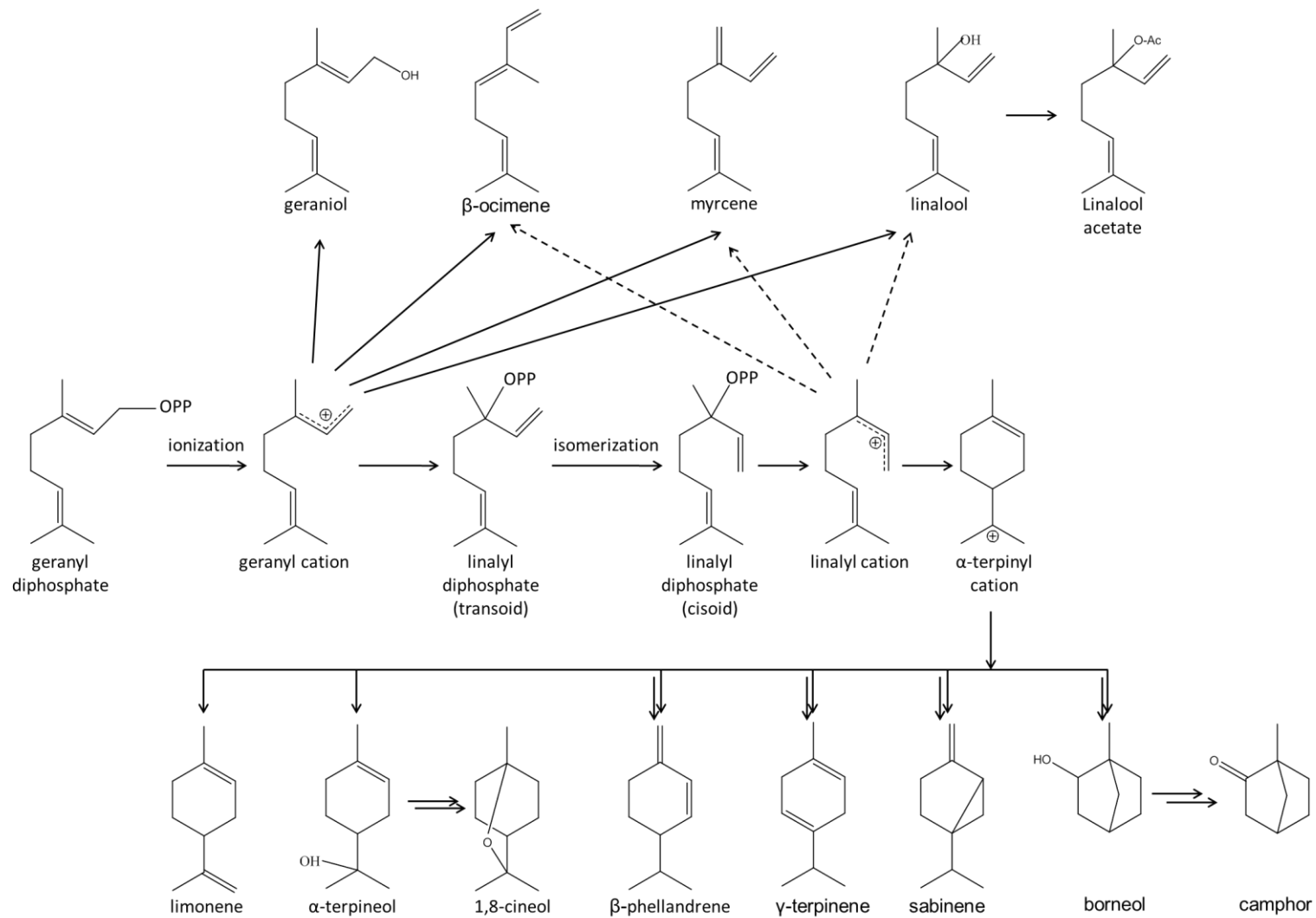


Figure 1.3 Schematic presentation of major monoterpene synthesis of lavender.

Figure 1.3 Schematic presentation of major monoterpene synthesis of lavender. The reaction mechanism starts with the ionization of the geranyl diphosphate substrate. The resulting carbocation can undergo a range of cyclizations, hydrogen shifts and rearrangements before the reaction is terminated by deprotonation or water capture. Cyclic monoterpenes are synthesized from α -terpinyl cation. Acyclic monoterpenes are produced from either geranyl cation or linalyl cation. Reproduced from Degenhardt et al. 2009 with permission from Phytochemistry.

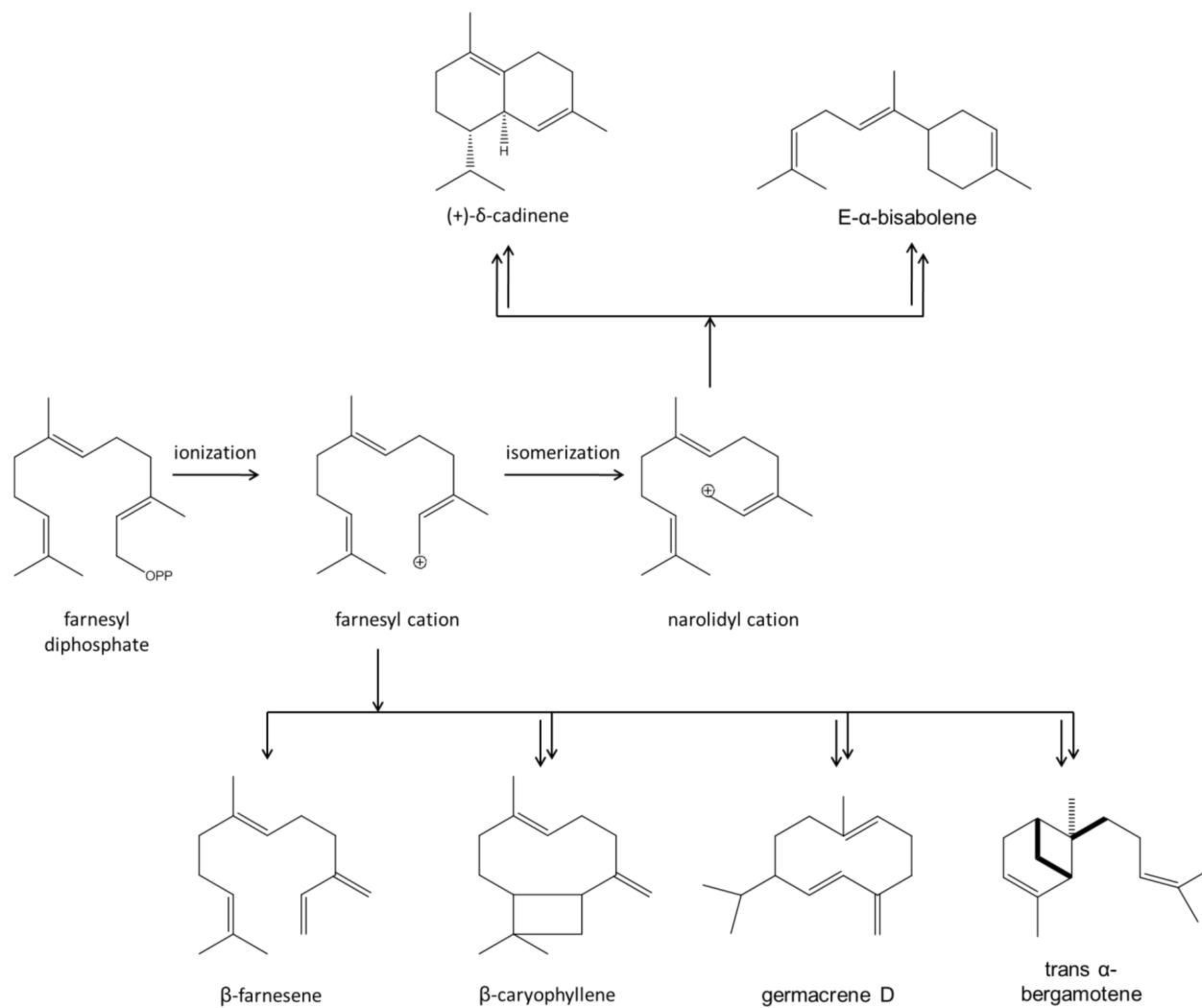


Figure 1.4 Schematic presentation of major sesquiterpene synthesis of lavender.

Figure 1.4 Schematic presentation of major sesquiterpene synthesis of lavender. The reaction mechanism for sesquiterpene synthases in figure 1.4 start with the ionization of the FPP. The resulting carbocation can undergo a range of cyclizations, hydrogen shifts and rearrangements before any stable compound is produced. Reproduced from Degenhardt et al. 2009 with permission from Phytochemistry.

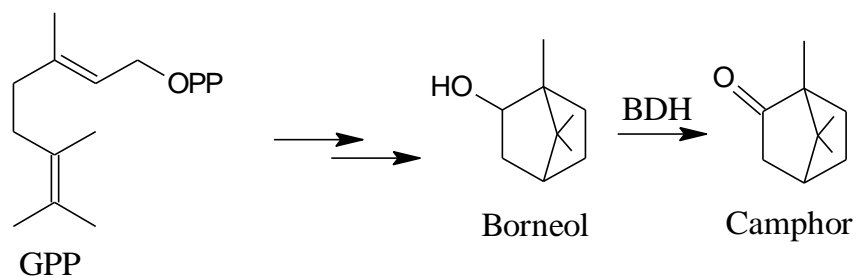


Figure 1.5 Proposed pathway for camphor biosynthesis. Double arrows indicate involvement of multiple enzymes. GPP, geranyl diphosphate; OPP, diphosphate moiety; BDH, borneol dehydrogenase. Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

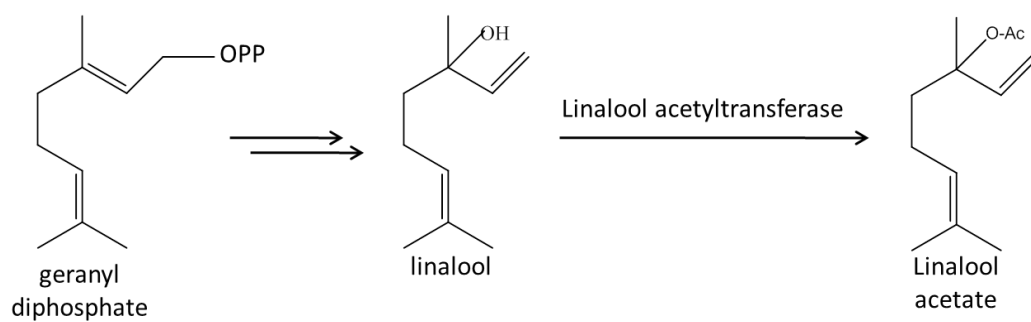


Figure 1.6 Proposed pathway for linalool acetate biosynthesis. Double arrows indicate involvement of multiple enzymes. GPP, geranyl diphosphate; OPP, diphosphate moiety.

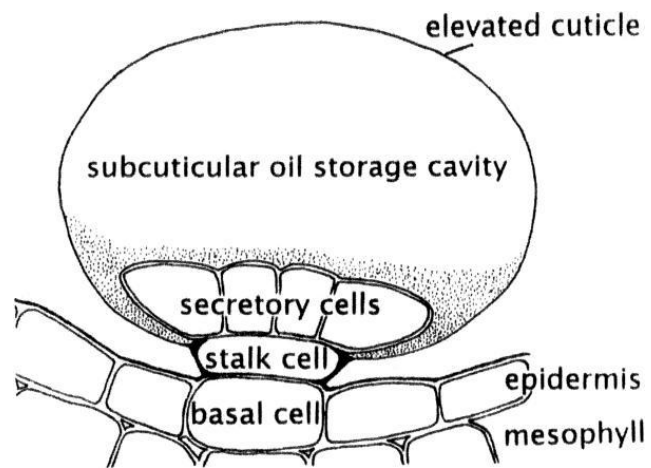


Figure 1.7 Schematic diagram of a glandular trichome. Copied from Turner et al. 2000 with permission from Plant Physiology.

2 Chapter: Materials and Methods

2.1 Chemicals and reagents

Analytical grade chemicals and reagents were used in this study. Unless otherwise specified, biochemical reagents and organic solvents were bought from either Sigma-Aldrich (Canada) or Fisher Scientific (Canada). Restriction enzymes, T4-DNA ligase, Taq-DNA polymerase, Vent-DNA polymerase, M-MuLV reverse transcriptase, and corresponding buffer solutions were purchased from New England Biolab (NEB, Canada). Deoxynucleotide-tri-phosphate mix (dNTPs) was purchased from Invitrogen (Canada), or Promega (USA). The iProof (Bio-Rad, USA), and Kappa hotstar (Kappa biosystem, USA) Hi-fidelity taq polymerase was used to amplify the gene of interest during cloning into bacterial plasmids. Real time PCR, using SsoFast™ Eva- Green® Supermix (Bio-Rad, USA), was used to analyze the transcripts of different gene(s) of lavenders.

2.2 Bacteria and plasmids

The suppliers and genotypes of bacterial strains used in this study are listed in Table 2.1.

Plasmids used were pBluescript-SK for cDNA library construction. pET41b(+) (Figure 2.1) and pGEX4T1 (Figure 2.2) vectors were used for heterologous protein expression.

2.3 Plant material

L. x intermedia cv. Grosso and *L. angustifolia* cv. Lady were grown under natural conditions at a field site at the University of British Columbia, Okanagan campus (Kelowna, BC, Canada) (Boeckelmann, 2008). *L. latifolia* flower and leaf tissue was provided by Dr. Tim Upson from Cambridge University (UK).

2.4 Glandular trichome extraction

Glandular trichome secretory cells were isolated by a modified glass bead abrasion method previously reported (Gershenzon et al. 1992). Briefly, *L. x intermedia* flowers were collected and soaked for 1 hour in ice-cold extraction buffer (200 mM sorbitol, 10 mM sucrose, 25 mM MOPSO, 0.5 mM PO₄ buffer, 10 mM sodium bisulfate, 10 mM ascorbic acid, 1 mM EDTA, 1 % PVP-40, and 0.6 % methylcellulose) containing 2 mM aurinticarboxylic acid, 5 mM thiourea, and 2 mM DTT (for RNA preservation) at pH 6.6. Cells were then transferred to the bead beater chamber including 500 µm glass beads and XADS beads, and were beat for 30 sec four times (4X) on ice. The slurry was passed through a 70 µm mesh membrane and cleaned with wash buffer (10% glycerol, 25 mM phosphate buffer, 1 mM EDTA and 1 mM DTT containing 1 mM aurinticarboxylic acid, 5 mM thiouria and 2 mM DTT). The flow through – containing the glandular cells - was then passed through a 20 µm mesh membrane to collect glandular trichomes. The isolated trichome cells were diluted by adding Rnase free water and used immediately, or were flash frozen in liquid N₂ and stored in a -80 °C freezer until used.

2.5 RNA extraction and reverse transcription

Qiagen RNA extraction kit was used to isolate total RNA from different tissues of lavenders using the manufacturer protocol. Complementary DNA (cDNA) was prepared by using Oligo dT and Random hexamer primers (IDT, USA), and M-MuLV reverse transcriptase enzyme according to the manufacturer protocol.

2.5.1 cDNA Library construction

Total RNA was extracted from the secretory cells by the Qiagen RNAeasy mini Kit in Dr. Mahmoud's Lab, and was sent to Plant Biotechnology Institute's Natural Products Genomic

Laboratory (PBI-NAPGEN) for library construction. The quality of all RNA samples was confirmed using an Agilent 2100 Bioanalyzer. Messenger RNA was purified from total RNA using a Dynabeads® mRNA purification kit (Invitrogen). cDNA synthesis, insert sequencing, bioinformatics, and clone archiving were performed by staff at the PBI as a paid service. The cDNA library was constructed with a ZAP Express® cDNA Synthesis Kit (Agilent Technologies, Palo Alto, CA, USA) as prescribed by the manufacturer. The respective libraries were then cloned into the pBlueScript SK± vector (Stratagene), and transformed into ElectroMAX DH10B T1 Phage Resistant Cells (Invitrogen). Aliquots of each library were plated on LB growth medium containing 100 g/ml of carbenicillin. A total of 10,000 ESTs were isolated and partially sequenced from the 5' end. These colonies were then cultured and archived at -80°C in 96 well micro-titer plates as 10% glycerol stocks.

Raw sequence files were produced using the 'phred' software package (Ewing and Green 1998; Ewing et al. 1998). These sequences were then vector-trimmed and low-quality trimmed using 'crossmatch' (Chou and Holmes 2001; Phrap 2008) and 'lucy' (Ewing and Green 1998; Ewing et al. 1998) software, respectively. Poly-AT regions were then removed using custom scripts and low-complexity regions were masked with 'mdust' (The Computational Biology and Functional Genomic Laboratory 2008). All good quality sequences were then BLASTX searched against the NCBI non-redundant (nr) database. Sequence clustering was performed using 'tgiel' with parameters set as follows: tgiel sequences -p 94 -v 1000 -O '-p 98 -o 49 -t 10000' (The Computational Biology and Functional Genomic Laboratory 2008). Gene ontology (GO) terms were assigned to the ESTs by transferring the GO terms from the TAIR database (TAIR 2008) assigned to their top BLAST hit.

2.5.2 Microarray data generation

To understand the expression of genes represented in our EST database, we conducted a transcript profiling experiment in various tissues of *L. angustifolia*, *L. x intermedia* and *L. latifolia* using microarrays. Tissues used in this study included flowers, leaves, and glandular trichomes from 30% flowering stage of *L. angustifolia* and *L. x intermedia* (See Table 2.4). *L. latifolia* gland tissue was not available for this experiment. Probe generation, array construction, RNA labeling, array hybridization, washing, scanning, signal quantification and data analysis were performed by staff at the University Health Network Microarray Centre (Toronto, Canada) as a paid service. The microarray experiment was designed to allow the comparison of different mRNA transcript variations among the tissues from different lavender species (Table 2.4). For each array, a total of 54,346 probes were used. The probes were filtered to remove those with little or no expression across all samples. Probes that were in the lower 20th percentile across all arrays were also removed.

2.6 Cloning and functional characterization

2.6.1 Primer design & PCR to amplify the gene of interest

Selected candidates were fully sequenced to obtain the sequence for the complete open reading frame (ORF). The ORFs corresponding to respective genes of interest (GOI) were cloned into NdeI (5' end) and EcoRI/XhoI (3' end) restriction sites of the pET41b(+) vector (Fig. 3.2). Cloning primers were designed manually by adding restriction enzyme (RE) sites and a few extra nucleotide sequences were added at the 5' upstream for the RE recognition. The primer sequences were analyzed by using Integrated DNA Technologies [IDT (www.idtdna.com)] PrimerQuest© Software (2002), which incorporates the Primer3© Software (Rozen and Skaletsky 2000). The predicted ORFs of the candidate genes were

amplified by PCR using iProof high fidelity DNA polymerase (Bio-Rad, USA) and specific primer sets containing appropriate restriction enzyme sites (Table 2.3) using the following program: initial denaturation at 95 °C for 5 min, followed by 95 °C for 1 min, $T_m \pm 5$ °C for 30 sec and 72 °C for 1 min for 35 cycles with a final elongation at 72 °C for 5 min PCR program. PCR products were cleaned up by the PCR purification kit from Omega (Omega Bio-Tek, USA).

2.6.2 Cloning of full length GOI and expression into *E.coli*

The Amplicons were digested with Nde I and EcoRI/XhoI restriction enzymes, and ligated into pET41b(+) bacterial expression vector, where it was fused to sequences encoding eight C-terminus histidine residues at the carboxylic terminus end of the protein to facilitate purification by Ni-NTA agarose affinity chromatography (EMD Chemicals, Darmstadt, Germany). For cloning into the pGEX4T1 expression vector, amplicons were digested with EcoRI and XhoI restriction enzymes and the GST fusion tag was used to purify the recombinant protein. To produce the recombinant protein, different *E. coli* cells were transformed with individual constructs, grown to log phase at 20 °C in Luria-Bertani (LB) media supplemented with respective antibiotics, (either ampicillin at 50 mg/L, kanamycin at 30 mg/L, and chloramphenicol at 34mg/L or, combination of two of them) and induced with isopropyl- β -D-thiogalactopyranoside (IPTG) at a final concentration of 0.5 mM. The induced cells were chilled on ice for 15-20 min, collected by centrifugation at 3,220 g and 4 °C for 20 min, and stored at -80 °C overnight.

2.6.3 Recombinant protein purification and SDS-PAGE

The stored cells were resuspended either in Novagen bind buffer (0.3 M NaCl, 50 mM Na₂HPO₄, 10 mM imidazole, pH 8.0; EMD Chemicals, Germany), or in GST purification

buffer (1x PBS) containing 1 mM protease inhibitor phenylmethanesulfonyl fluoride (PMSF), kept on ice for 20 min and sonicated on ice using a Sonic Dismembrator Model 100 (Fisher Scientific, Ottawa, ON, Canada) to complete bacterial membrane disruption. The cell debris were removed by centrifugation at 15,000 g and 4 °C for 15 min (Sorvall, USA), and the recombinant proteins harvested from the soluble fraction by Ni-NTA agarose affinity chromatography (EMD Chemicals, Germany) following the manufacturer's procedure. Protein samples were resolved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by staining with Coomassie Brilliant Blue. Protein concentration was determined by Bradford Assay (Bio-Rad).

2.6.4 Enzymatic Assay reaction

2.6.4.1 Alcohol dehydrogenase

Initially, enzyme assays were performed in 5 ml reaction volume containing 100 mM sodium phosphate buffer (pH 8.0), 40 µg of the enzyme, 1 mM NAD⁺, and 0.5 mM substrate (borneol). After overnight incubation at 30 °C with 150 rpm shaking, assay products were extracted into 1 ml pentane and concentrated ~ 50 times before analysis by GC-MS (see below). For linear kinetics study, assays were performed in 2 ml reaction volume (keeping reagent concentrations as before) at five different time points: 1 hr, 2 hr, 4 hr, 8 hr, and 16 hr. The optimum temperature was determined from a set of reactions performed at 27, 30, 32, 35 and 37°C. The optimum pH was determined by performing assays at pH 7.0, 7.5, 8.0, 8.5, 9.0 and 10.0 using MOPS for pH 7.0 and pH 7.5, sodium phosphate for pH 8.0, TAPS for pH 8.5 and CAPSO for pH 9.0 and 10.0 as a buffer. All assays were performed in duplicate or triplicate.

2.6.4.2 Sesquiterpene synthase & Putative LaTPS-I

In vitro enzyme activity assays were performed in 500 µl reactions volume containing the assay buffer (50 mM Tris/HCl, 5% (v/v) glycerol, 1 mM MnCl₂, 1 mM MgCl₂, 1 mg/ml Bovine Serum Albumin [BSA], pH 7.0), 1 mM DTT, 25 µM substrate (GPP or FPP; Echelon, Salt Lake City, UT, USA), and 10 µg of the purified recombinant protein. The mixture was overlaid by 500 µl of pentane and incubated at 30 °C for 60 min. Purified protein extracts from *E. coli* Rosetta™ (DE3) pLysS cells transformed with expression vector without insert were assayed under the same conditions as controls. Reactions were stopped by vigorous vortexing followed by flash freezing in liquid nitrogen, and stored in a -80 °C freezer until analyzed. To concentrate the assay products, ca. 90% of the pentane was evaporated using a gentle stream of highly purified helium gas.

2.6.4.3 Acetyltransferase

Initially, enzyme assays were performed in 5 ml of 100 mM sodium phosphate buffer (pH 8.0), containing 40 µg of the enzyme, 1 mM acetyl CoA, and 0.5 mM substrate (linalool, sigma aldrich). After overnight incubation at 30 °C with 150 r.p.m shaking, assay products were extracted into 1 ml pentane and concentrated ~ 50 times before analysis by GC-MS.

2.7 Enzyme kinetics study for LiBDH

To construct the Michaelis-Menten saturation curve, alcohol dehydrogenase enzyme assays were performed at optimum temperature (32 °C) and pH (8.0) for 30 minutes (n=5) in 1 ml reaction volume containing 100 mM sodium phosphate buffer, 1 µM enzyme, 1 mM NAD⁺, and substrate concentration of 5 µM, 25 µM, 50 µM, 100 µM, 200 µM, 400 µM and 1 mM. Assay progress was monitored by measuring the conversion of NAD⁺ to NADH at 340 nm

using a Lambda 25 UV–visible spectrometer (Perkin-Elmer). Absorbance results were used with NADH extinction coefficient ($6.22 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$) to calculate the enzyme activity. The kinetic parameters of the enzyme were determined from a Michaelis–Menten saturation curve constructed using SigmaPlot software version v.10.00 (Systat Software, Germany).

2.7.1 Gas Chromatography / Mass Spectrometry

All GC–MS analyses were performed on the Varian GC 3800 Gas Chromatographer coupled with a Saturn 2200 Ion Trap mass detector (Agilent technologies, USA). The instrument was equipped with a 30 m X 0.25 mm capillary column coated with a 0.25 μm film of acid-modified polyethylene glycol (ECTM 1000, Altech, Deerfield, IL, USA) and a CO_2 cooled 1079 Programmable Temperature Vaporizing (PTV) injector (Agilent technologies, USA). The cold on-column mode, where the injector was set to a temperature of 40 °C for injections, was used. The oven temperature was initially maintained at 40 °C for 3 min, followed by a two-step temperature increase, first to 130 °C (at a rate of 10 °C per minute) and then to 230 °C (at a rate of 50 °C per minute), and held at 230 °C for 8 min. The carrier gas (helium) flow rate was set to 1 ml per minute. The identity of product was confirmed by comparing its retention time and mass spectrum to those for an authentic sample.

2.8 Transcript analysis

2.8.1 Standard PCR reaction

Total RNA was extracted from different lavender tissues by using a plant RNA extraction kit and treated with DNase I enzyme to remove genomic DNA using the Omega Bio-Tek kit. Treated total RNA was reverse transcribed with Oligo (dT) (80 μM) and random hexamers (40 μM) (Custom oligos, IDT Canada) by using M-MuLV Reverse Transcriptase enzyme (New England Biolabs, MA, USA) following the manufacturer's protocol. The

transcriptional activity was measured in 30% flowering stage and in young leaf tissues from *L. angustifolia*, *L. x intermedia*, and *L. latifolia* by standard PCR based on the intensity of the bands. The PCR reaction was repeated from Bud-I, Anthesis, 30% flowering stage, and glandular trichome from the 30% flowering stages of *L. x intermedia*.

2.8.2 Real time quantitative PCR (qPCR)

The relative abundance of LiBDH, LiCPS, LiLAT and LaTPS-I were analyzed from tissues described in 2.8.1 by using CFX96TM Real Time detection system (Bio-Rad, USA).

Complementary DNA (cDNA) for relative transcript analysis was synthesized using iScript cDNA synthesis kit from Bio-Rad according to manufacturer's instructions. SsoFastTM EvaGreen® Supermix (Bio-Rad, USA) along with approximately 150 ng of cDNA as a template and 500 nM of each of the primers in 20 µl reaction volume. Gene specific primers (Table 2.3) used in quantitative real-time PCR experiments were designed using the IDT primer quest software (<http://www.idtdna.com/Scitools/Applications/Primerquest/>) targeting 120–200 base-pairs (bp) fragment size. The following program was used for real time PCR: 95 °C for 30 sec followed by 40 cycles of 5 sec at 95 °C and 30 sec at 58 °C. Normalized expression values ($\Delta\Delta CT$) of *LiBDH* and *LaLINS* were calculated by CFX96TM data manager (Bio-Rad, USA) using β -actin as a reference gene.

2.9 Multiple sequence alignment and phylogenetic tree construction

The phylogenetic tree was constructed using the default parameters of PhyML software available at <http://www.phylogeny.fr> (Dereeper et al. 2008). PhyML employs MUSCLE software to generate multiple alignments and the maximum likelihood computational method to construct the phylogenetic tree. Classical SDRs from different plants were employed in the phylogenetic tree construction.

Table 2.1 Supplier and genotype of bacterial strains used in this project.

<i>Escherichia coli</i>	Supplier	Genotype
DH10B	Invitrogen, Canada	F ⁻ <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\Phi80lacZ\Delta M15$ $\Delta lacX74$ <i>recA1</i> <i>endA1</i> <i>araD139</i> $\Delta(ara\ leu)$ 7697 <i>galU</i> <i>galK</i> <i>rpsL</i> <i>nupG</i> λ -
BL21	Novagen, Canada	F ⁻ <i>ompT</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i>
BL21(DE3)	Novagen, Canada	F ⁻ <i>ompT</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i> (DE3)
BL21(DE3)pLysS	Novagen, Canada	F ⁻ <i>ompT</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i> (DE3) pLysS (Cam ^R)
Rosetta(DE3)pLysS	Novagen, Canada	F ⁻ <i>ompT</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻) <i>gal dcm</i> (DE3) pLysSRARE2 (Cam ^R)
C43(DE3)pLysS	Lucigen, USA	F ⁻ <i>ompT</i> <i>gal dcm</i> <i>hsdS_B</i> (r _B ⁻ m _B ⁻)(DE3)pLysS (Cam ^R)

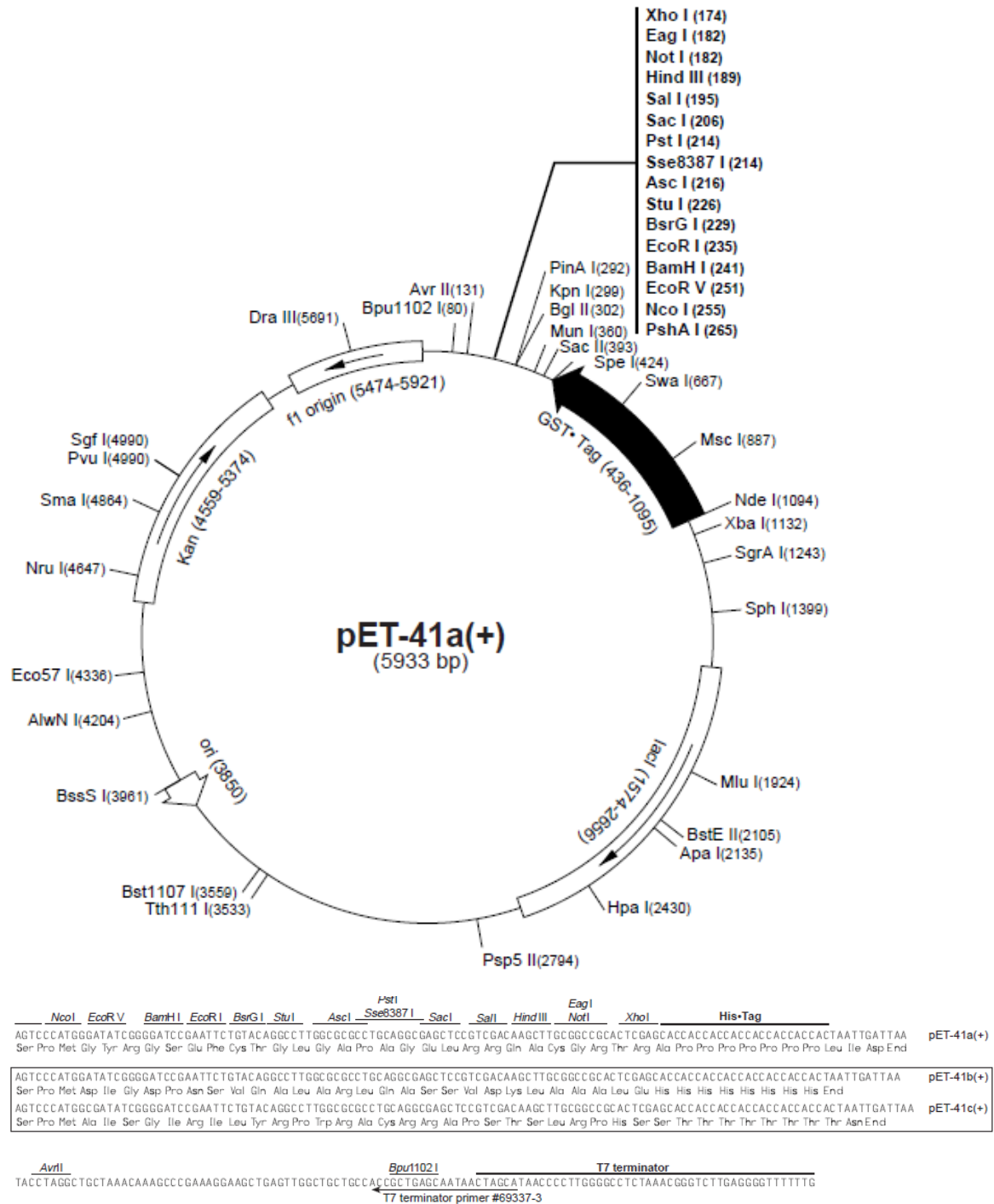


Figure 2.1 Map of pET41b(+) vector. pET41a-c(+) vectors have the same map, although they have a different purification tag at C-terminal end (Novagen, EMD chemicals, Germany).

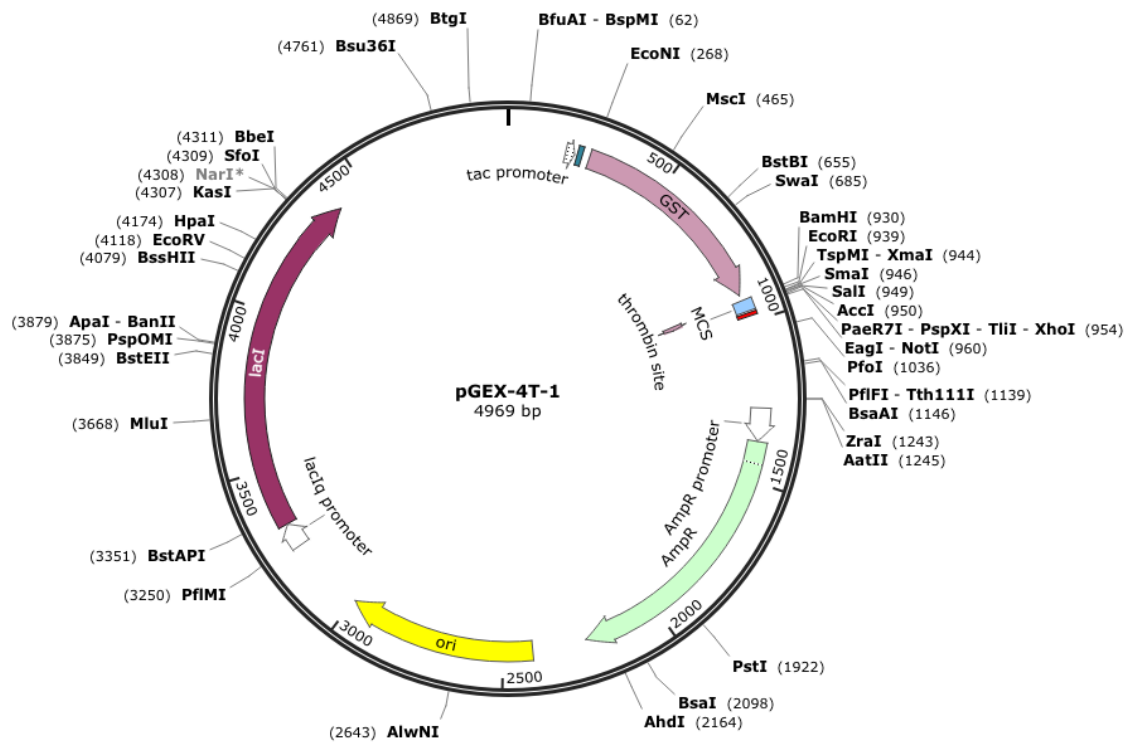


Figure 2.2 Map of pGEX4T1 vector. (GE healthcare, Canada)

Table 2.2 *L. x intermedia* tissues used in this project.

(Flower developmental stages and leaf)





Flower			Leaf
Bud	Anthesis	30% bloomed flower	
1-2 cm, green	5-8 cm, violet, unopened	30% flowers on spike opened	2 cm, young
			

Table 2.3 Lists of primers used in this project.

	Target gene (Vector)	Primers
Full length	<i>LiBDH/LiSDR2</i> (pET41b(+))	F-5'-CCCTCATATGGCTTCAACTGTTTTGAGA-3' R -5'-AGTCTCGAGCGAATCCATCAAATCAAAC-3'
	<i>LiSDR1</i> (pET41b(+))	F-5'-CCGCGCATATGGAGAAAATAGGGAAAAG-3' R-5'-TTTGAATTCATGAGTCTGGAAGGCATTCC-3'
	<i>LiLAT-3</i> (pET41b(+))	F-5'-ATCTCATATGGCATCTACCAAAACCCTAA-3' R -5'-CCCGAATTCAACAATGCTGAAAGATTAGA-3'
	<i>LiLAT-3</i> (pGEX4T1)	F-5'-AATTGAATTCATGGCATCCACCAAAACC-3' R-5'-AAAGCTCGAGCAATGCTGAAAGATTGAAAG-3'
	<i>LiLAT-4</i> (pET41b(+))	F-5'-AGGGCATATGGCGATGATTATTACAAAAC-3' R-5'-GGGGAATTCCCAGTATCCAATTTATTGTAA-3'
	<i>LiLAT-4</i> (pGEX4T1)	F-5'-CCCCGAATTCATGGCGATGATTATTACA-3' F-5'-CCCCCTCGAGAGTATCCAATTTATTGTA-3'
	<i>LiCPS</i> (pET41b(+))	F-5'-GATAACATATGAGGAGGTCAGCGAATTATAG-3' R -5'-AATATGAATTCTTAATATGGAAGGGGTGAAGG-3'
	<i>LaTPS-I</i> (pET41b(+))	F-5'-GGCCATATGGATAGCAAGTCAGTCAATG-3' R -5'-ATACTCGAGCTCTGCAAGATTCGCTGC-3'
	<i>LiBDH</i>	F-5'-AATCGGAGCGGCAGCATAATCT-3' R-5'-TAATACGGCGAGACGCAGTTCA-3'
	<i>LiLINS</i>	F-5'-ACACGCACGACAATTTGCCA-3' R-5'-AGCCCTCCAATGAAGTGGGAT-3'
qPCR	<i>β-actin</i>	F-5'-TGTGGATTGCCAAGGCAGAGT-3' F-5' AATGAGCAGGCAGCAACAGCA 3'
	<i>LaTPS-I</i>	F-5'-AATGAGCAGGCAGCAACAGCA-3' R-5'-AATCTGGAAGCTCGCATTTGGCG-3'

Table 2.4 Microarray experiment design.

Microarray experiment design	
A) <i>L. x intermedia</i> Grosso-Bud (Gland)	B) <i>L. x intermedia</i> Grosso -Anthesis (Gland)
B) <i>L. x intermedia</i> Grosso -Anthesis (Gland)	C) <i>L. x intermedia</i> Grosso -30% flower (Gland)
C) <i>L. x intermedia</i> Grosso -30% flower (Gland)	D) <i>L. anguatifolia</i> -30% flower (Gland)
E) <i>L. x intermedia</i> -PP1-30% flower (Gland)	F) <i>L. x intermedia</i> -PP2-30% flower (Gland)
G) <i>L. x intermedia</i> Grosso-total flower (30% stage)	H) <i>L. anguatifolia</i> -total flower (30% stage)
I) <i>L. x intermedia</i> Grosso-total flower (30% stage)	J) <i>L. latifolia</i> -Total flower
K) <i>L. anguatifolia</i> -total flower (30% stage)	L) <i>L. latifolia</i> -Total flower
M) <i>L. anguatifolia</i> - total flower (30% stage)	N) <i>L. anguatifolia</i> - total leaf (young)

Vs.

PP1=*L.x intermedia* cv. Provance containing peak 1 (linalool)

PP2=*L.x intermedia* cv. Provance containing peak 2 (linalool, linalool synthase)

3 Chapter: Results

3.1 cDNA library

A cDNA library was constructed from the secretory cells of *L. x intermedia* floral oil glands (Figure 3.1) to identify the genes involved in the biosynthesis and storage of lavender EO. A total of 10,000 ESTs were isolated and partially sequenced at 5' end with an average sequence read of 637 bp. The experiment yielded 8,205 high quality reads, which were clustered into 4,116 unigenes. The unigene library contained 3,075 genes that were represented by a single EST (singletons) and 1,041 contigs corresponding to 5,130 ESTs, representing over 62.53 % of the reads. GO annotation showed that 4753 (57.93%) and 3068 (37.39%) sequences were involved in biological processes and cellular components, respectively (Table 3.1). From the biological processes, 809 (9.86%) sequences were predicted to be involved in the isoprenoid biosynthesis process. Most of the terpene synthesis genes contained numerous EST members suggesting their strong transcriptional regulation in secretory cells. For example, the contig corresponding to linalool synthase contained over 68 members indicating that, as expected, linalool synthase is strongly expressed in lavender oil glands (Lane et al. 2010). Given that oil gland cells are specialized to produce large quantities of the EO, the above observations could be well justified. Based on homology to proteins in the Plant Genomic Database (PlantGDB), putative functions could be assigned to 3,903 (approximately 94.83 %) of the unigenes.

3.2 Short chain alcohol dehydrogenase

3.2.1 Candidate selection

A homology-based analysis of our sequences against those in TAIR and UniProt protein databases identified a total of ten ESTs as putative SDRs. Among these, two ESTs corresponded to singletons, while the remaining eight formed three contigs. Two contigs (Contig 1 and 2) included two EST members each, and one contig (Contig 3) contained four members. Only Contigs 1 and 3 produced ESTs that encoded full length SDRs, and thus were selected for further analysis. The full length EST corresponding to Contig 1 (designated LiSDR-1) was 1020 base pairs, with an ORF of 759 nucleotides that encoded a protein of 253 amino acids with a predicted molecular weight of 27 kDa. The other full length EST, Contig 3, (designated LiSDR-2) was 841 nucleotides long and had an ORF of 780 nucleotides that encoded a 259 amino acid protein with a predicted molecular weight of 27.5 kDa. Both proteins bore predicted mitochondrial targeting sequences, as identified by the IPSORT (<http://ipsort.hgc.jp/>), TargetP (<http://www.cbs.dtu.dk/services/TargetP/>) and PREDOTAR (<http://urgi.versailles.inra.fr/predotar/predotar.html>) online protein analysis tools. It is worth noting that all three protein analysis tools identified a mitochondrial targeting sequence for LiSDR-2. However, a mitochondrial targeting sequence for LiSDR-1 was only identified by IPSORT.

3.2.2 Functional analysis of recombinant LiBDH

The LiSDR-1 and LiSDR-2 proteins were expressed in *E. coli* (Rosetta (DE3)pLysS) cells, and purified by Ni-NTA affinity column chromatography. Following purification, the recombinant enzymes were assayed for dehydrogenase activity with borneol as a substrate, and either NAD⁺ or NADP⁺ as a cofactor. Analysis of the assay products by GC-MS revealed

that LiSDR-1 did not produce a detectable product, while LiSDR-2 (subsequently renamed LiBDH) produced camphor from borneol with NAD^+ (but not NADP^+) as a cofactor (Figure 3.2). The negative control assays, which contained all the reagents including recombinant protein extracts obtained from bacterial cells harboring the empty vector, did not produce detectable products. Also, the reverse reduction assay in which camphor was used as a substrate and NADH as a cofactor, did not produce detectable amounts of borneol or other products (data not shown). Furthermore, a recombinant *L. angustifolia* SDR and a medium chain alcohol dehydrogenase (MDR) from *L. x intermedia* (expressed and purified using the same procedures) were not able to produce any detectable amount of camphor from borneol when assayed under identical conditions. The SDR cDNA cloned from *L. angustifolia* 30% flower was obtained from our previously reported *L. angustifolia* floral cDNA library (Lane et al. 2010).

To examine the substrate specificity of the enzyme, LiBDH was also assayed with eight other monoterpenes (α -terpineol, 1,8-cineole, citronellol, linalool, lavandulol, nerol, geraniol, and perillyl alcohol), and one sesquiterpene (farnesol) found in lavenders, and with menthol which is a common monoterpene in other *Lamiaceae* plants. A significant amount of products were not found after 12 hours of incubation in any of the assays. Those assays containing α -terpineol as a substrate produced trace amounts of camphor and isoborneol. Comparable quantities of both camphor and isoborneol were also found in negative control assays, indicating that these monoterpenes were produced either non-specifically, or through the action of a contaminated bacterial protein.

3.2.3 Kinetics study

LiBDH showed linear catalytic activity starting from 30 min up to several hours (Figure 3.3a), while its optimum pH and temperature were determined to be 8.0 (Figure 3.3b) and 32 °C (Figure 3.3c), respectively. Kinetic parameters for the enzyme were obtained from a plot of velocity versus substrate concentration fit to the Michaelis-Menten equation using the hyperbolic enzyme kinetics analysis module of the SigmaPlot software v.10.00 (Systat Software, Erkrath, Germany) (Figure 3.3d). The K_m of LiBDH was found to be 53.6 ± 14.9 μM , while its V_{max} , k_{cat} and k_{cat} / K_m were calculated to be 3.97×10^{-1} pmol sec^{-1} , 4.0×10^{-4} sec^{-1} and 7.5×10^{-6} $\mu\text{M}^{-1} \text{sec}^{-1}$, respectively.

3.2.4 Tissue specific regulation of *LiBDH*

Initially a standard PCR strategy was used to study the expression pattern of *LiBDH* transcript in various *L. x intermedia* tissues, including leaf, bud, anthesis, and mature (30% in bloom) flowers. The results indicated that the transcripts for this gene were highly abundant in floral glandular trichomes (Figure 3.4a). Next, we employed a quantitative PCR (qPCR) approach to quantify the expression of *LiBDH* mRNA in various tissues of *L. x intermedia*, including leaf, anthesis, floral tissues collected at 30% flowering stage and secretory cells isolated from floral tissues at 30% flowering stages. The results of this experiment confirmed that the *LiBDH* transcripts were concentrated (ca. 200 fold higher) in the secretory cells of glandular trichomes (Figure 3.4b). Finally, the transcriptional activity of *LiBDH* in young leaves and floral tissues (30 % flowering) of *L. angustifolia*, *L. x intermedia* and *L. latifolia* plants was determined by qPCR. In this experiment, the abundances of the *L. x intermedia* linalool synthase (*LiLINS*) transcripts were also measured as a control (Lane et al. 2010). As expected, *LiLINS* mRNA was more strongly expressed in flowers compared to leaves (Figure

3.4c). The LiBDH transcripts were detected in both *L. angustifolia* and *L. x intermedia* flowers; however, they were much less abundant than those of *LiLINS*. Further, *LiBDH* mRNA was much less abundant in *L. latifolia* flowers compared to those of *L. angustifolia* and *L. x intermedia* plants (Figure 3.4c). The relatively low expression of *LiBDH* mRNA paralleled the concentrations of borneol and camphor (also relatively low compared to linalool), which amounted to 0.6 mg to 2.0 mg per gram of fresh tissue for both monoterpenes in these tissues (Figure 3.4d).

3.2.5 Sequence comparison and phylogenetic tree analysis

LiBDH exhibited a significant similarity to SDRs from *Camellia sinensis* (61% identity), *Phaseolus lunatus* (61% identity), *Lactuca sativa* (57% identity), *Artemesia annua* (56% identity), and *Zingiber zerumbet* (52% identity) in multiple sequence alignment (Figure 3.5), and was closely rooted with the above SDRs in the phylogenetic tree (Figure 3.6).

3.3 Multiproduct sesquiterpene synthase (LiCPS)

3.3.1 Candidate selection

In search of sesquiterpene synthases expressed in lavenders, the *L. angustifolia* trans α -bergamotene synthase (LaBERS) cDNA sequence was blasted against the *L. x intermedia* EST database. This search identified the LiCPS contig as the best candidate (77% identity to LaBERS) for a sesquiterpene synthase. The LiCPS contig contained 53 ESTs, indicating that the corresponding gene was strongly expressed in glandular trichomes. Microarray data suggested that *LiCPS* is developmentally regulated during flower developmental stages of *L. x intermedia*, and is up-regulated in flower and glandular trichomes of *L. x intermedia* compared to *L. angustifolia* while it is down regulated in total flower compared to *L. latifolia* (Table 3.2).

3.3.2 Functional analysis of recombinant LiCPS

It has previously been demonstrated that exclusion of the transit peptide enhances solubility of recombinant proteins in bacterial hosts (Williams et al. 1998a). In search of the transit peptide, LiCPS was analyzed in Signal IP, Signal 3L, and Target IP. The search did not find a transit peptide. This is consistent with the fact that sesquiterpene synthases do not contain a signal peptide and are predominant in the cytosol. Recombinant LiCPS was expressed, purified and was used for enzymatic assay reaction. The purified recombinant LiCPS produced only one essential oil constituent, caryophyllene, when assayed with FPP (Figure 3.7a). LiCPS produced α -terpeniol, cineole, limonene and alpha pinene from GPP and NPP as substrate (Figure 3.7b, 3.7c). When assayed with GPP as a substrate, LiCPS produced 1,8-cineole as a major product. LiCPS produced alpha terpineol as a major product when assayed with NPP.

3.3.3 Tissue specific regulation of *LiCPS*

LiCPS was found to be developmentally regulated in flowers of *L. x intermedia*. Using standard PCR, we observed a strong band of *LiCPS* in glandular trichome compare to bud, anthesis and flower from 30% flowering stage. We could not detect LiCPS transcripts in young leaves of *L. x intermedia* plants (Figure 3.8).

3.3.4 Multiple sequence alignment and phylogenetic tree analysis

A BLASTP search showed that LiCPS was most closely related to trans-alpha-bergamotene synthase from *L. angustifolia* with 77% amino acid identity level. LiCPS also exhibited sequence identity with α -zingiberene synthase of *Ocimum basilicum* (63%), α -terpineol synthase of *Vitis vinifera* (52%), β -ocimene/myrcene synthase of *Vitis vinifera* (52%), and β -

ocimene/myrcene synthase of *Vitis vinifera* (50%). From the multiple sequence alignment it was found that LiCPS contains all the conserve motifs such as the initial substrate isomerization RR(X₈)W motif, DDXXD and (N,D)D(L,I,V)X(S,T)XXXE motifs which are responsible for the enzymatic activity and coordination of divalent cations and thus responsible for substrate binding and ionization, respectively (Figure 3.9). The phylogenetic analysis revealed that LiCPS is closes related to LaBERS and other terpene synthases from lavenders and to α -zingiberene synthase from *Ocimum basilicum*. Surprisingly, the gene was distantly related to the caryophyllene synthases from *Zea mays* (Figure 3.10).

3.4 Alcohol acetyltransferase

3.4.1 Candidate selection

Around 117 ESTs corresponding to acetyltransferase were found in our cNDA library. Among these, 15 ESTs were singleton while the remaining 102 ESTs formed contigs. GO annotation revealed only four contigs (98 ESTs) and two singletons that were potential alcohol acetyltransferases, while others corresponded to histone acetyltransferase, amino acid acetyltransferase and, dihydrolipoamide acetyltransferase. The two singletons were very short sequences covering only the C-terminal end of acetyltransferase proteins. The four selected contigs LiLAT-1, LiLAT-2, LiLAT-3, and LiLAT-4 candidates have six, thirty five, twenty, and thirty five ESTs, respectively (Table 3.2). According to the EO profile of lavenders, LiLAT should be up regulated in *L. x intermedia* and *L. angustifolia* tissues compared to *L. latifolia*. This was confirmed by the microarray transcript analysis as LiLAT-2 and LiLAT-3 transcripts were up-regulated accordingly (Table 3.2). Initially, LiLAT-1 was selected since it had a higher sequence homology with LiLAT-2 and LiLAT-3 sharing 90% and 84% sequence similarities, respectively. LiLAT-2 and LiLAT-3 share 82% sequence

similarities between themselves though, LiLAT-2 lacks 59 amino acids at the C-terminal end including the highly conserve DFGWG motif (Figure 3.11). LiLAT-1 also lacks a 22 amino acids stretch at the N-terminal end. LiLAT-4 candidate falls into a different clade of alcohol acetyltransferase classification including unaltered conserved motifs; as well, the mRNA transcript of this candidate is down regulated in *L. angustifolia* and *L. x intermedia* compared to *L. latifolia* flower and gland tissues (Figure 3.12).

3.4.2 Cloning and Functional analysis of recombinant LiLAT

LiLAT-3 and LiLAT-4 were cloned into pET41b(+) vector using NdeI and EcoRI, or, EcoRI and XhoI restriction enzymes sites to facilitate the His tag or, GST tag purification, respectively. The initial cloning strategy was confirmed by restriction enzyme digestion (Figure 3.13) and followed by sequencing the clones to address cloning confirmation. The final clones were transformed into *E. coli* Rosetta(DE3)pLysS, Origami(DE3)pLysS, and C43(DE3)pLysS cells for recombinant protein production. After overnight incubation of crude LiLAT recombinant proteins with substrate linalool and coenzyme A did not produce any detectable products.

LiLAT-3 and LiLAT-4 were cloned into the pGEX4T1 vector using EcoRI and XhoI restriction enzyme sites and transformed into BL21 and BL21(DE3)pLysS cells for recombinant protein production (Figure 3.14, Figure 3.15). Total protein samples extracted from the induced bacterial cells for respective clones were analyzed by SDS-PAGE. Both cell types strongly expressed the recombinant LiLAT-GST fusion protein (Figure 3.15). However, protein purification using His tag affinity chromatography was unsuccessful.

3.4.3 Sequence comparison and phylogenetic tree analysis

All four candidates were found to contain all the conserved motifs for alcohol acetyltransferase with few changes (Figure 3.16). The HXXXDG domain located near the center portion of each enzyme has been shown to be important for a general base catalysis mechanism. The second highly conserved region is the DFGWG motif located in the carboxyl end, and is believed to have a structural role in enzymatic functions which is not strictly required for membership within the BAHD family of enzymes (D'Auria 2006). Phylogenetic analysis showed that, LiLAT-1, LiLAT-2 and LiLAT-3 fell into close proximity with alcohol acetyltransferases while LiLAT-4 shared a close relationship with vinorine synthases, which is also an acetyltransferase.

3.5 LaTPS-I

3.5.1 Candidate selection

Two cDNA libraries, including 14,000 ESTs, from *L. angustifolia* leaf and flower were recently reported (Lane et al. 2010). A cDNA sequence (LaTPS-I) deduced from a contig formed by ESTs present in both libraries exhibited strong homology to the *L. angustifolia* limonene synthase. This sequence, however, did not contain a stretch of 73 amino acids, including the DDXXD terpene synthase signature motif, present in other terpene synthases. Given that limonene is produced mainly in lavender leaves, transcriptional activity for LaTPS-I was expected in leaves, but not in flowers. Due to this unique structure (i.e., the lack of an important motif) and unexpected expression pattern LaTPS-I was selected for further analysis to understand its contribution to EO synthesis in lavenders.

Functional analysis of recombinant LaTPS-I

The ORF for LaTPS-I contained 1,482 nucleotides, and encoded a 494 amino acid protein with a predicted molecular weight of 54.3 kDa. LaTPS-I was cloned into the pET41b(+) vector using Nde I and Xho I restriction sites, and the recombinant protein was produced in *E. coli* cells (Figure 3.17). The Ni-NTA affinity purified recombinant LaTPS-I was incubated with GPP, NPP or FPP as substrates for 2 h to overnight, but did not produce any products. Small quantities of linalool and geraniol were found in GC-MS analysis, which were also formed in control assays devoid of LaTPS-I, indicating that they are most likely arose from the hydrolysis of the substrate after long incubation (data not shown).

3.5.2 Tissue specific regulation of *LaTPS-I*

Transcripts of *LaTPS-I* were compared to those of *LaβPHLS* and *LaLINS* by standard PCR. While *LaTPS-I* mRNA was found to be present in both flower and leaf tissue of *L. angustifolia*, those of *LaβPHLS* and *LaLINS* were present only in leaf or flower tissues, respectively (Figure 3.18a). In real-time PCR, *LaTPS-I* was seen to be ten times more abundant in glandular trichomes compared to flower tissue (Figure 3.18b).

3.5.3 Sequence comparison and phylogenetic tree analysis

A BlastP search revealed that LaTPS-I could be a monoterpene synthase, however, it has a 73 amino acid stretch in the middle of the sequence which contains the most important divalent cation binding motif DDxxD (Figure 3.19). Except for that motif it contains all other conserve motifs for terpene synthases including the 58 amino acid signal peptide. From the phylogenetic tree analysis, LATPS-I was found to be closely related to monoterpene synthases from lavenders and fell into the terpene synthase clade TPSb which generally contains TPSs from angiosperms including *Lamiaceae* (Figure 3.20).

Table 3.1 GO annotation of *L. x intermedia* gland cDNA library.

Biological process		Cellular component	
GO annotation	Sequences	GO annotation	Sequences
Cellular process	2812 (34.27%)	Cell	2471 (30.12%)
Metabolic process	2702 (32.93%)	Cell part	2471 (30.12%)
Response to stimulus	1543 (18.81%)	Organelle	1852 (22.57%)
Developmental process	525 (6.40%)	Organelle part	1086 (13.24%)
Multicellular organismal process	497 (6.06%)	Macromolecular complex	553 (6.74%)
Biological regulation	454 (5.53%)	Cell junction	129 (1.57%)
Multi-organism process	428 (5.22%)	Symplast	129 (1.57%)
Regulation of biological process	421 (5.13%)	Extracellular region	124 (1.51%)
Localization	328 (4.00%)	Membrane-enclosed lumen	100 (1.22%)
Establishment of localization	322 (3.92%)		
Cellular component organization	302 (3.68%)	Molecular function	
Reproduction	238 (2.90%)	GO annotation	Sequences
Reproductive process	238 (2.90%)	Catalytic activity	2837 (34.58%)
Signaling	161 (1.96%)	Binding	1403 (17.1%)
Immune system process	125 (1.52%)	Transporter activity	357 (4.35%)
Death	61 (0.74%)	Structural molecule activity	219 (2.67%)
Positive regulation of biological process	59 (0.72%)	Nutrient reservoir activity	173 (2.11%)
Negative regulation of biological process	59 (0.72%)	Antioxidant activity	156 (1.9%)
Growth	50 (0.61%)	Nucleic acid binding transcription factor activity	125 (1.52%)
Rhythmic process	23 (0.28%)	Electron carrier activity	50 (0.61%)
Pigmentation	6 (0.07%)	Enzyme regulator activity	30 (0.37%)
Viral reproduction	4 (0.05%)	Molecular transducer activity	23 (0.28%)
Biological adhesion	2 (0.02%)	Receptor activity	7 (0.9%)
Cell proliferation	1 (0.01%)	Protein binding transcription factor activity	5 (0.6%)
Locomotion	1 (0.01%)	Protein tag and metallochaperone activity	5 (0.6%)

Table 3.2 Microarray analysis.

Tissues used in microarray analysis

- A) *L. x intermedia* (Grosso)- Gland from Bud B) *L. x intermedia* (Grosso)-Gland from Anthesis
- C) *L. x intermedia* (Grosso)- Gland from Flower 30% D) *L. angustifolia* (Lady)- Gland from Flower 30%
- G) *L. x intermedia* (Grosso)-Total Flower H) *L. angustifolia* (Lady)- Total Flower
- I) *L. latifolia*-Total Flower

Results of the microarray analysis

	Contig. Mem.	G Vs H Int v ang	G Vs I Int v lat	H Vs I Ang v lat	A Vs B Gr-Bud V Anth
LiCPS	53	Up-14	Down-2.28	Down -9.7	Down -29
LiLAT-1	6	Up-44	Up-1.89	Down -52	Down-4
LiLAT-4	35		Up-18	Up-18	Down-7
LiLAT-3	20	Up-1.47	Up-18.5	Up-20.79	Down-6
LiLAT-2	35	Up-2.4	Up-1.39	Down-11	Down-6

(Up) - up regulation

(Do) - down regulation

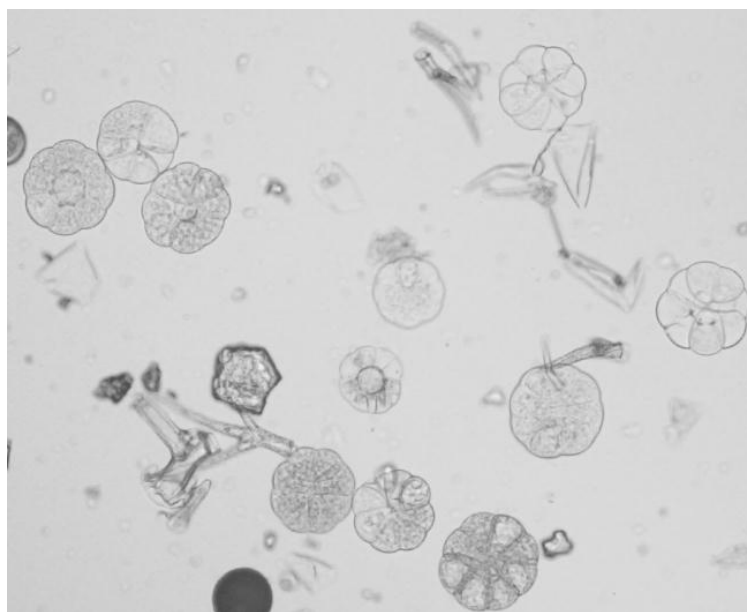


Figure 3.1 Extracted glandular trichomes from 30% flowering stage of *L. x intermedia*.
(Image was taken by Nikon dxm 1200 camera, Japan)

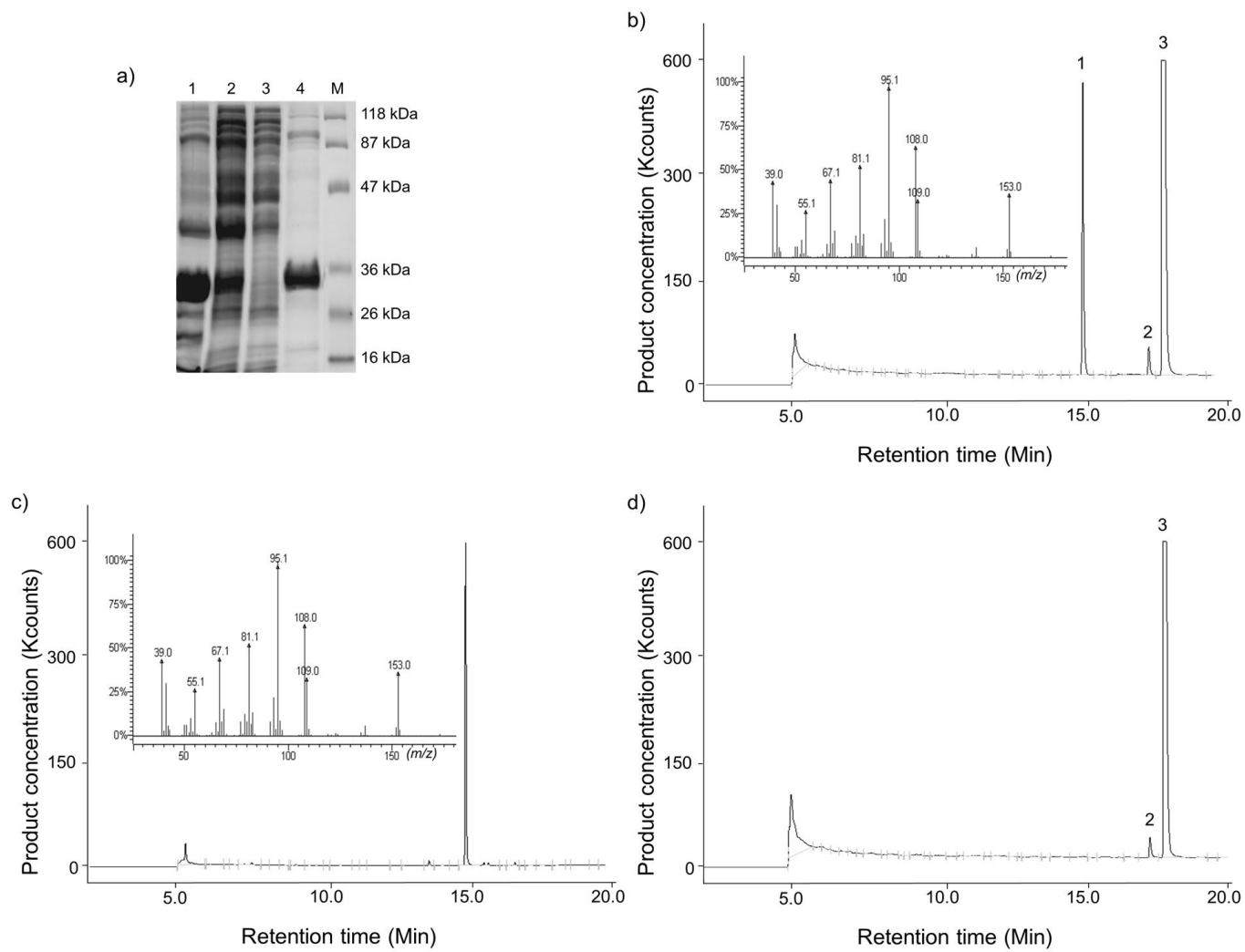


Figure 3.2 Protein purification and GC-MS analysis of LiBDH.

Figure 3.2 Protein purification and GC-MS analysis of LiBDH. a) SDS-PAGE stained with coomassie blue. M: marker, 1- pellet from induced cells, 2- soluble fraction, 3- flow through, 4- LiBDH recombinant protein purified by Ni-NTA resin column. b) GC chromatogram of LiBDH assay with mass spectrum of camphor. c) GC chromatogram and mass spectrum of authentic camphor. d) GC analysis of extract from the negative control. Peak-1 Camphor, peak-2 Iso Borneol, peak-3 Borneol. Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

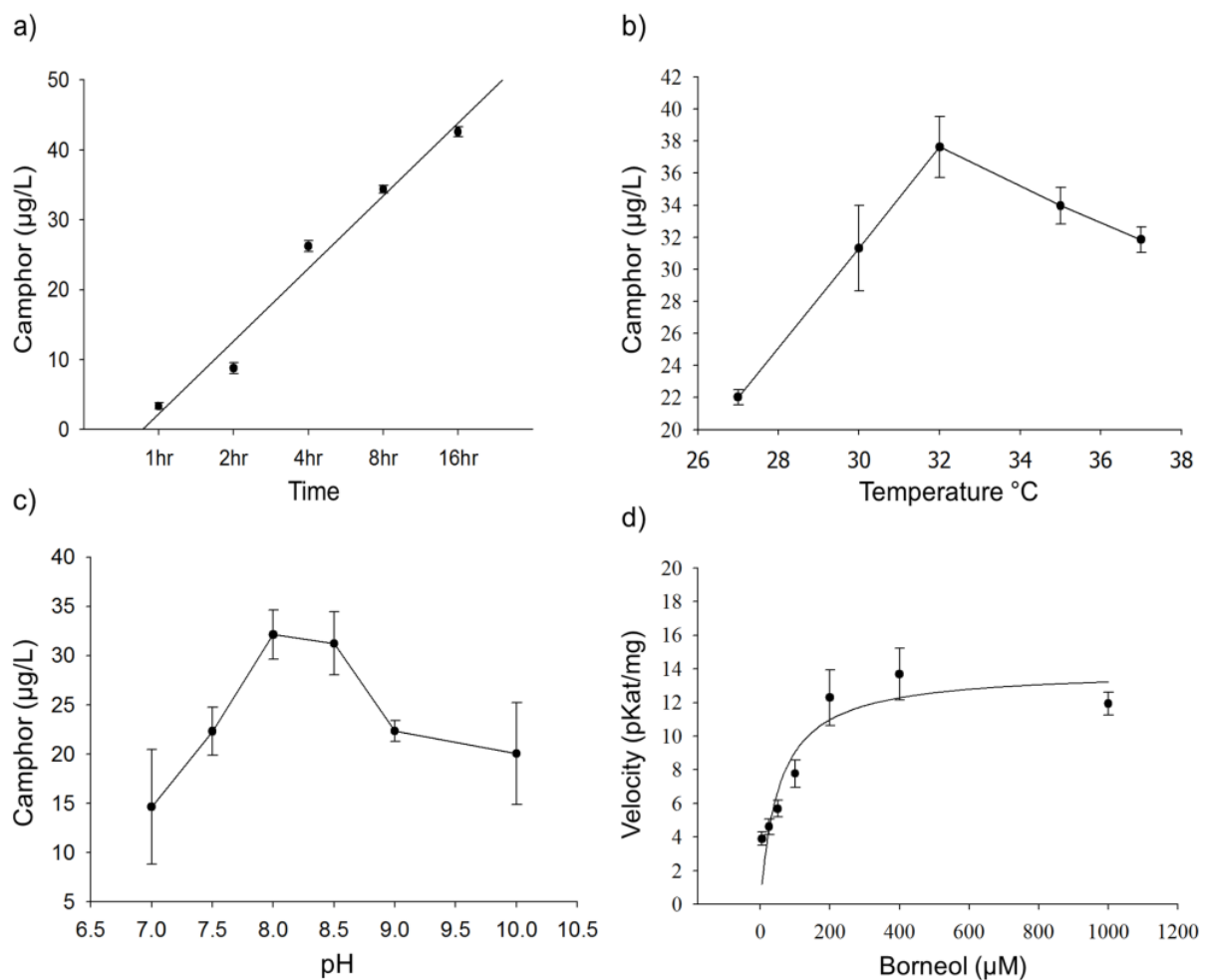


Figure 3.3 Kinetic assay of LiBDH with borneol as a substrate: a) time course assay of LiBDH activity, b) effect of pH on LiBDH activity, c) effect of temperature on LiBDH activity and d) velocity of LiBDH at increasing borneol concentrations. Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

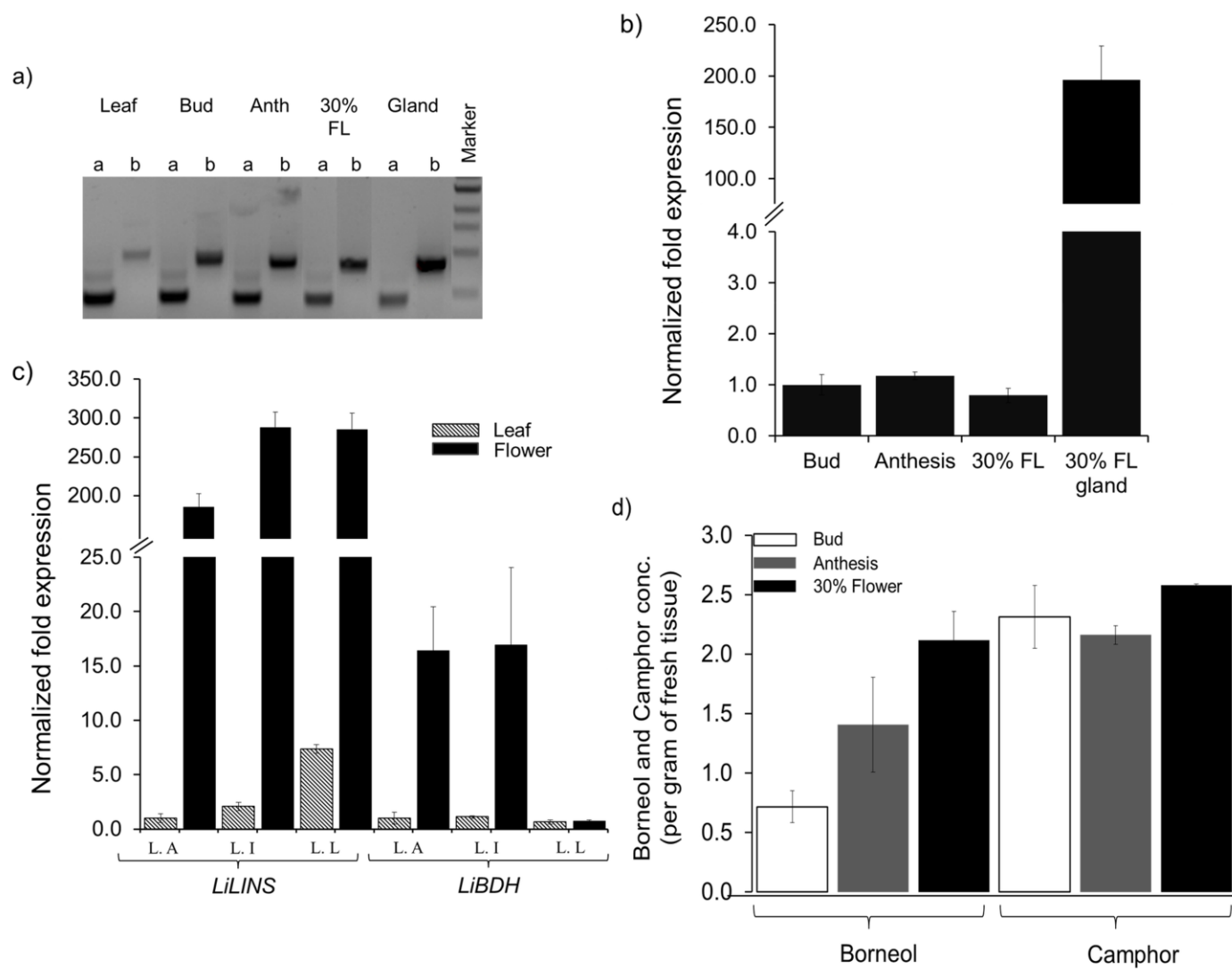


Figure 3.4 Transcriptional activity of *LiBDH* in different tissues of lavender species.

Figure, 3.4 Transcriptional activity of *LiBDH* in different tissues of lavender species. a) Standard PCR of *Actin* (a) and *LiBDH* (b) from leaf, bud, anthesis, 30% flower, and glandular trichomes of *L. x intermedia* b) Relative *LiBDH* transcript levels in developmental stages of *L. x intermedia* flowers measured by qPCR normalized to actin. c) Tissue specific transcriptional abundance of *LiBDH* and *LiLINS* in *L. angustifolia* (L.A), *L. x intermedia* (L.I), and *L. latifolia* (L.L) normalized to *actin*. d) Amount of borneol and camphor in different flowering stages of *L. x intermedia* (mg per gram of fresh tissue). Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

```

LiBDH  ---MAS-TVLRRLLEGKVALITGAASGIGESAARLFSRHGAK-VVIADIQDELALNICKDLG---STFVHCDVTKEFDVETAVNTAVSTYGKLDIMLNNAGISGAPKYISNTQLSDFKR 111
ZSD1   -----MRLEGKVALVTGGASGIGESARLFIEHGAK-ICIVDVQDELGQQVSQRLGGDPHACYFHCDDVTEDDVRRVAVDFTAKEYGTIDIMVNNAGITGDKVIDIRDADFNEFKK 109
SAD-C  MAESSSTKTGLRLAGKVAIVTGGASGIGKETARLFANQAARMVVIADIQDELGIQVAESIG-TDRCTFIHCDIRIEDDVKNLVQSTVDYGGQIDIHCNAGIISPDSQTLLELDVVSQANG 119
ISPD   MASVK-----KLAGKVAIVTGGASGIGEVITARLFAERGARAVVIADMQPEKGGTVAESIG-GRRCSYVHCDITDEQQVRSVVDWTAATYGGVDVMFCNAGTASATAQTVLDLDAQFDR 113
Adh2   ---MASLTPKARLENKVAIVTGGARGIGECIVRLFVKHGAK-VVIADVNDLGLKLLCQDLG-SKFACFVHCDVTIESDIENLINTTIAKHGQLDIMVNNAGTVDEPKLSILDNEKSDFDR 115
      : * .***:.*.***: .:** .:.*: : *.::: . . :.: * . :.***: * :.: :.: :.* :*: .*** . :.: :.: .

LiBDH  VVDVNLVGVFLGTHAARVMIP-NRSGSIISTASATAAAAGTPYPYICSKHGVLGLTRNAAVEMGGHGIRVNCVSPYYVATPMT-----RDDDWIQG-----CFSNLKGAVLTAE 216
ZSD1   VFDINVNGVFLGMKHAARIMIP-KMKGSIVSLASVSSIAGAGPHGYTGAKHAVVGLTKSVAELGRHGIRVNCVSPYAVPTRLSMPYLPESQMEDALRGFLTFFVRSNANLKGVDLMPN 228
SAD-C  VFATNAIGTALCVKHAARAMVDGKVRGSICTASISASYGVTGTGDSMSKHAVLGLMRSASVQLAKYGIRVNSVSPNGLATPLTEKLLDADAKTVEEIFS-----KFSMLKGVLRTN 233
ISPD   VMRVNARGTAACVKQAARKMVELGRGGAIICTASATVHHAGPNLTDYIMSKCGVLGLVRSASLQLGVHGIRVNSVSPALATPLTATIGLRTAADVESFYG-----QVTSKGVAITAE 227
Adh2   VVSINLAGVFLGTHAARVMIP-KCSGSIITTASICSVTGGVASHAYTSSKHGVVGLAKNAAELGKYNIRVNCVSPYFVPTKLAFKFLN---MDETSS-----FYSNLQGKTLGPQ 223
      * . * . .:*** *: .:** ** . :. :.*** :.: :.: :.***.*** :.* :.: :.* :.: .

LiBDH  DVAEAAALYLASDESKYVSGHNLVDGGVSIMNQGCNMFDLMS 259
ZSD1   DVAEAVLYLATEESKYVSGNLVIDGGFSIANHTLQVFE---- 267
SAD-C  HVADAVLFLASNESDFVTGLDLRVDG--NYITSDAVI----- 268
ISPD   HVAEAVAFLASDEAAFVTGHDLAVDGGLQCLPFVAVAK----- 265
Adh2   DIANATLFLASDESGYVSGHNLVVDGGYSVLNPAFGLFSWKP- 265
      .:*. .:**:.*: :.* :.* ** .

```

Figure 3.5 Multiple alignment of the deduced amino acid sequences of LiBDH with SDR from other plants; *Mentha x piperita* (AAU20370.1), *Pisum sativum* (AAF04253.1), *Zingiber zerumbet* (BAK09296.1), and *Artemisia annua* (ADK56099.1).(*)- identical,(;)- conserved substitution. Black bar and gray shade indicates the conserve motifs. Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

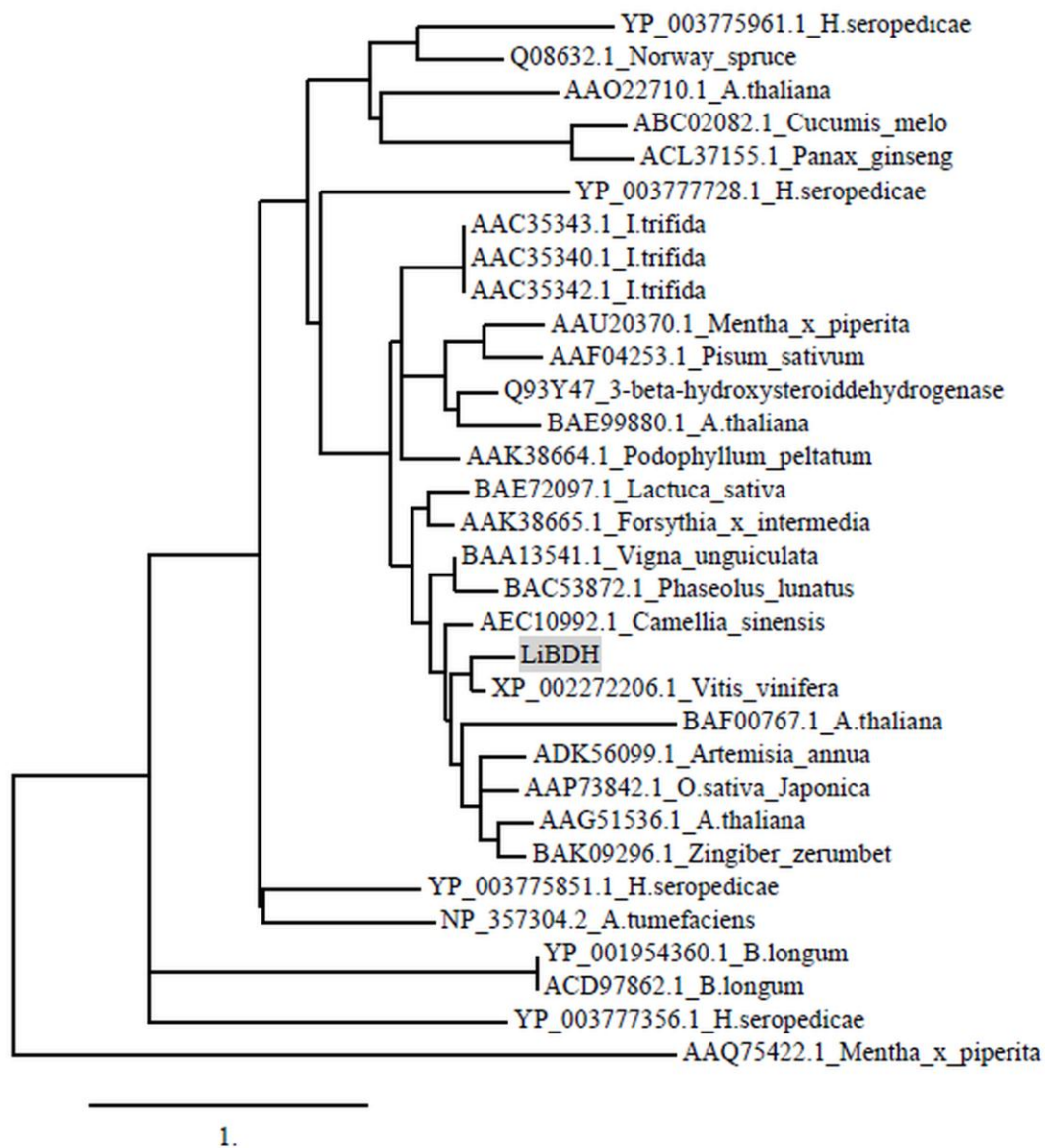


Figure 3.6 Phylogenetic tree analysis of the “Classical Group” of plant SDRs. The scale bar represents 1.0 amino acid substitutions per site. Reproduced from Sarker et al. 2012 with proper authorization from Archives of Biochemistry and Biophysics.

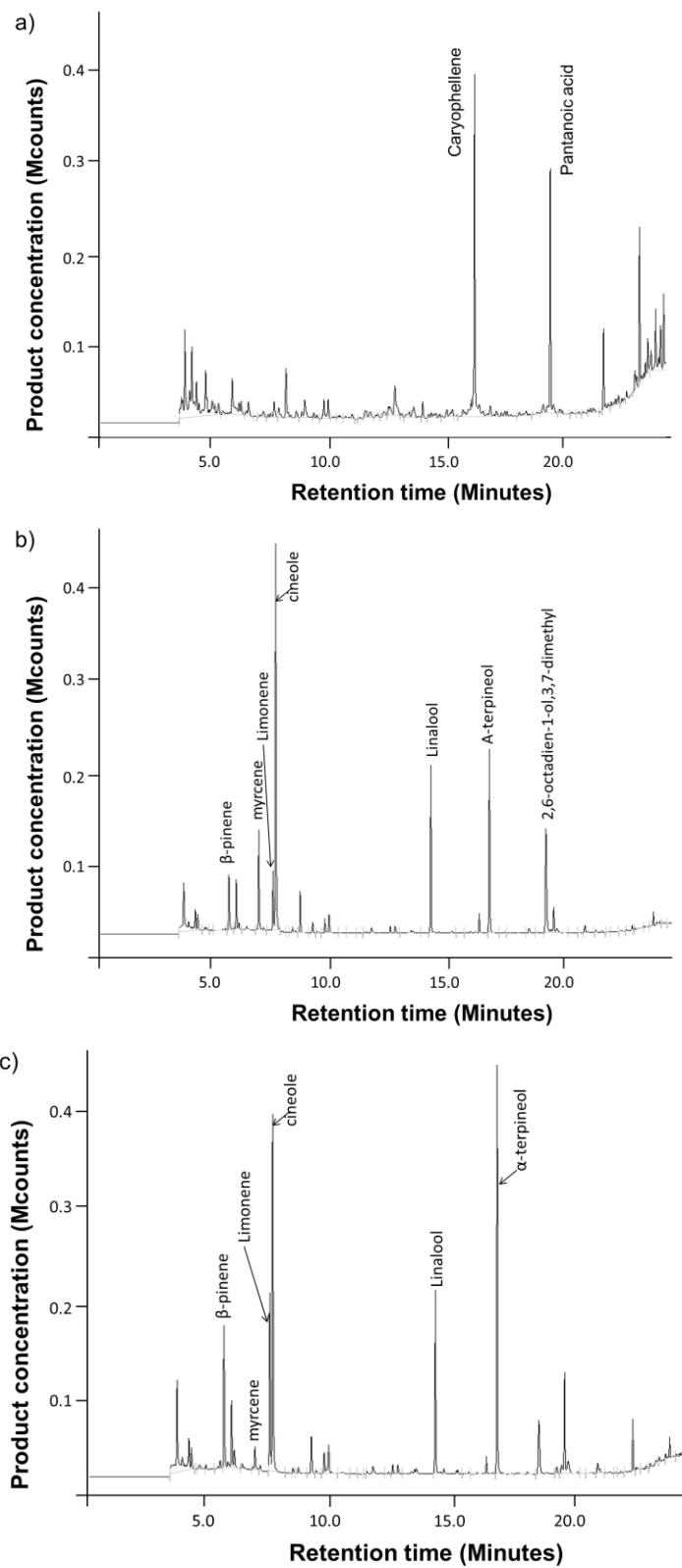


Figure 3.7 Enzymatic assays of LiCPS.

Figure 3.7 Enzymatic assays of LiCPS with a) FPP which produce caryophyllene as a major product, b) GPP which produce cineol as a major product, and c) NPP which produce α -terpineol and cineol as the major products.

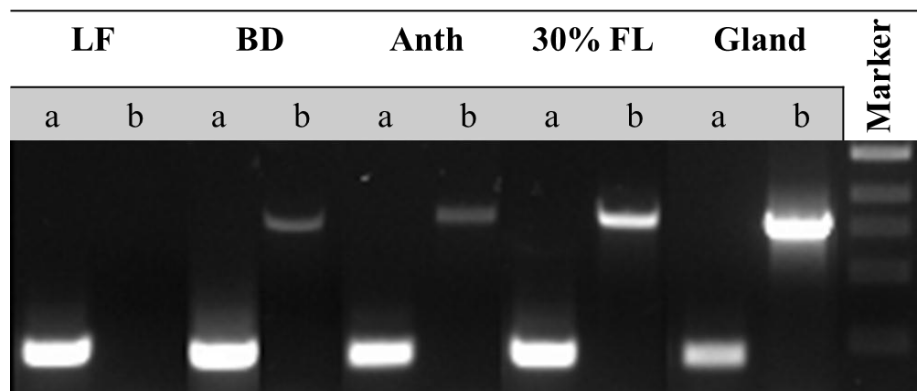


Figure 3.8 Amplification of *LiCPS* transcripts by PCR in different tissues of *L. x intermedia*.

LF-young leaf, BD-bud stage of flower, Anth-Anthesis stage of flower, 30% FL-mature flower at 30% stage, Galnd-Glandular trichome from 30% flowering stage. a- β *Actin*, b- *LiCPS*.


```

LiCPS      MEARRSANYRSAIWDHNYIQSLTSSYTGDKYVDRSQKLKFEVMMKMMVDETTDELEQLQLIHLQLRLGISYHFQDCIAKMLHNIFGSE-----NKHVEKDLHLTALKFRLLRQHGYHVPQD 115
Q5SBP4.1_ZinS MESRRSANYQASIWDDNFIQSLASPYAGEKYAEKAEKLEKTEV-KTMIDQTRDELKQLELIDNLQRLGICHHFQDLTKKILQKIYGEERNQDGHQHYKEKGLHFTALRFRILRQDGYHVPQD 119
ABB73046.1_LaBERS MEARRSGNFESSIWDDDIQSLTSSYTGKMYVDKSEKLEKIEV-KMMDEATDELEQLLELINDLQRLGISYHFQDKGIKMLNNIYKSD-----SKYMEKDLHLTALKFRLLRQHGYRVPQD 114
      *:***,*,.:***,.:***,*,*,*,*,*.:*** ** * *:***: ***:***,*. *****,:**:* *:***:*. :. :. :. **:*,***:***:***,*,***

LiCPS      VFNRTDDEGNFETWVGEDVVRGLVSLYEASYLSMEGESILDMAKDFSSHHLTEMVEQIKDES LAEEVRHALELPLHWRIERLEARWFIQAYETRPDSNPILVELAKLDFNMVQAKYQAE 235
Q5SBP4.1_ZinS VFSSFMNKAAGDFEESLSKDTKGLVSLYEASYLSMEGETILMAKDFSSHHLHKMVEDATDKRVANQIIHSLEMPHRRVQKLEAIWFIQFYECGSDANPTLVELAKLDFNMVQATYQEEL 239
ABB73046.1_LaBERS VFSSFMDDDEGNFEAWVVEDVSVLVSLYEASHISVEGESILDMAKDFSSHHLTEMVEQIGEAELAEQVKRTLELPLHWRVGRLEARWVQAYETRPNSNPTLVELAKLDFNMVQAKYQDEL 234
      **, * !. *:*** : :*, *****,:***:***** ***** :***: : :*:*** :*:*** ** * *:*** ** * *:*** *****:*** **

LiCPS      KRSSRWFKETGLPEKLSFARERPAECFFWAMGFIPEPHHGSREVMTKVGLLITVLDIYDVYGTLEELKDFNTNFERWDTSWLDRLEPEYMQICFLAIFNSVNLGYQILRDQGLNLIQN 355
Q5SBP4.1_ZinS KRLSRWYEETGLQEKLSFARHRLAEAFWLSMGIIEPGHFGYGRMHLMKIGAYITLLDDIYDVYGTLEELQVLTEIIEERWDINLLDQLPEYMQIFFLYMFNSTNELAYEILRDQGINVIS 359
ABB73046.1_LaBERS KRCSRWYEETGLPEKMSFARHRLAECFLWSLGFIPDPHHGYSREIMTKIAVLITITDDIYDIYGALEELQEFTEAFERWDINSLDLLPEYMQICFLAIFNSANELGYQILRDQGLNIIPN 354
      ** ***,:*** ***:***,*,*,*,*:***:*,*,* : :*, ** : ***:***:***: :*: :*** , ** ***** ** :***,***,*,*****:*,*

LiCPS      MRRWWTLSRVNYVEARWFHSGYVPTTEEYLNATWISITGPLLSSFGYLTTHPINNTELESLEKHPGIIYWPSMVLRLADDLGTSSDEMKGVDVSKSVQCYMNETGCSEKARHHVKNL 475
Q5SBP4.1_ZinS LKGLWVELSQCYFKEATWFHNGYTPTEEYLNACISASGPVILFSGYFTTNPINKHELQSLERHAHSL---SMILRLADDLGTSSDEMKGVDVPKAIQCFMNDTGCCCEEARQHVKRL 476
ABB73046.1_LaBERS LKRSWAELSRAYYLEARWFHNGFVPTTDQYLNATWISISGPLLSSYGYLTTHPINNKELKSLEKHPSIIRWPSMVLRLADDLGTSSDEIKRGVDVSKSIQCYMNETGCCEGDARHHVKS 474
      :. :*,***: : ** ***,*,*,***:***,*,*,* :*:*** ***:***:***:***,*. : ***:*****:*,***,*,*:***:***,*,*,*** **

LiCPS      IEAGLKRMNKEILMEKPFKDFGTAMNLARISLCFYQHGDGFGDPKSDMENKLASLFLHPFHIK- 539
Q5SBP4.1_ZinS IDAEWKKMNKDILMEKPFKNCPTAMNLRISMSFYEHGDSYGGPHSDTKKKMVSFLVQPMNITI 541
ABB73046.1_LaBERS IEVALKRMNDEILMEKPFKSFDTAMNLARISLCFYQYGDGFGKPHSDTIKNLVSLIVLPFHMP- 538
      *:*,*:***:***,*,*,*,***:***:***:*** ** ** :*:***:***:

```

Figure 3.9 Multiple alignment of the deduced amino acid sequences of LiCPS with LaBERS from *L. angustifolia* (AAB73046.1), and *Ocimum basilicum* (Q5SBP4.1). (*)- identical,(:)- conserved substitutions, and (.)-semi-conserved substitution of amino acids. Shaded sequences are conserved motifs.

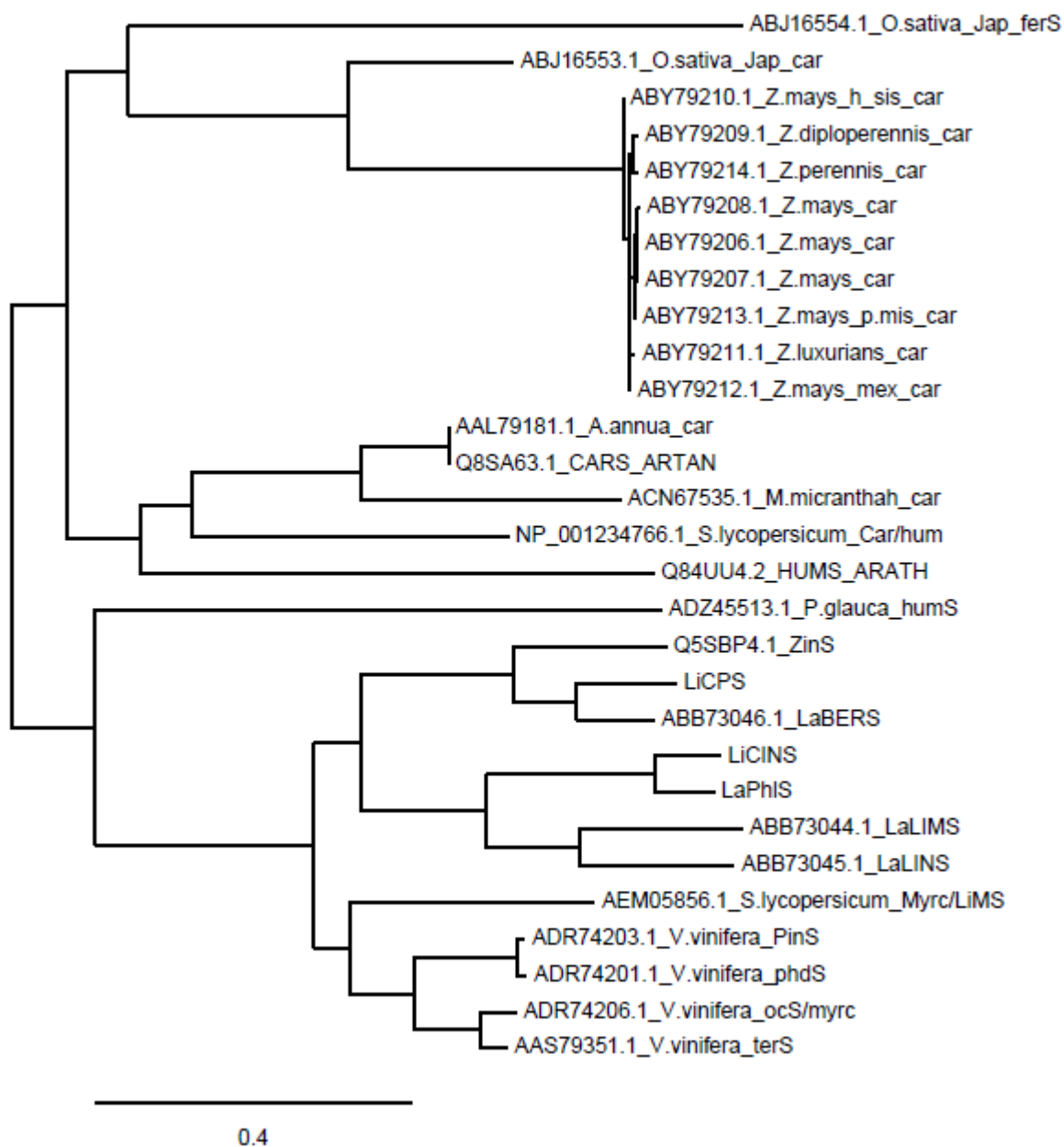


Figure 3.10 Phylogenetic tree analysis of LiCPS with other terpene synthases from different plants. The scale bar represents 0.4 amino acid substitutions per site.

```

LiLAT 1      MASTKTTLTFKVTRKDPELISPAEPTYGFKYLSDIDDQHGFRIRFSIIFYRENPSMKGK 60
LiLAT 2      MASTKTTLTFKVTRKDPELISPAEPTYGFKYLSDIDDQYGFIRFSIIFYRENPSMKGK 60
LiLAT 3      MASTKTTLTFKVTRKNRELISPAEPTYGFKYLSDIDDQDCLRFPLIFFYRENISMKGK 60
              *****: *****:*****:*****:*****
LiLAT 1      DPVKIIRDAVAKALVFYYPFAGRSRECA SRKLVVDCTGEGVIFVE----- 105
LiLAT 2      DPVKIIRDAVAKALVFYYPFAGRLRECA SRKLVVDCTGEGVIFVEADADVMMQQFGDALH 120
LiLAT 3      DPKIIRDAVAKALVFYYPFAGRLRECD SRKLVVDCTGEGVVFVEADADVMLQQFGDALH 120
              **:***** ***** *** *****:*****:*****
LiLAT 1      -----ELLLDTPGYDGIINCPLLFIQVTRLKCGGFILTYSCNHTICDAIGFSQFLSAV 158
LiLAT 2      PPFPNLEELL DTPGYDGIINCPLLFIQVTRLKCGGFILTYSCNHTICDAIGFSQFLSAV 180
LiLAT 3      PPFPNLEKLL DTPDYGTINCPILFIQVTRLKCGGFILSYSCNHTICDAAGFVQFMSAV 180
              :*****:*** *****:*****:*****:*****:***** ** **:***
LiLAT 1      GELARGATVPSIQPVWERHLLSARNPPRVSFTHREYDVLPKTNGETDKMVVRYFFFDVAD 218
LiLAT 2      GELARGATVPSIQPVWERHLLSARNPPRVSFTHREYDVLPKTNGETDKMVVRYFFFDVAD 240
LiLAT 3      GELARGATAPSIQPVWERHLLTARNPPRVSFTHREYDVVPKTNGETDKMVVRYFFFDVAD 240
              *****:*****:*****:*****:*****:*****:*****
LiLAT 1      ISALRRSLPPNLQTC SKFDVVA AFTLRCRTIAISLKPDEEVVFFSAVNIRNKITPPLPVG 278
LiLAT 2      ISALRRSLPPNLQTC SKFDVVA AFTLRCRTIAISLKPDEEVVFFSAVNIRNKITPPLPVG 300
LiLAT 3      ISALRRSLPRYLQTC SKFDIVAACAWRCRTIALSLKPDEEVVFNVTNIRNKMKPPLPVG 300
              ***** *****:*** : *****:*****:..:*****:*****
LiLAT 1      YYGNGIVSPAVVTTAEKLSKNPFHYAVELVMKAKYKATDEYVKSVDLMVMDRPSFTVA 338
LiLAT 2      YYGNGIVSPAVVTTAEKLSKNPFHYAVELVMKAKYKATDEYVKSVDLMVMDRPSFTVA 360
LiLAT 3      YYGNGIVFPVAVTTAKKLSKNPFQYAVELVMKGKYEATDDYVRSVADLMVMDRPSVGAG 360
              ***** *****:***:***:*****:***:***:***:*****:..
LiLAT 1      RNYCIVSDTTNVGF EKVDLGWGE PVYGG LAK-GIGWIPAHWYIPFKNKKGEQGTIVTVCL 397
LiLAT 2      RNYCIVSATTNVGSRKWTSGG-----ASPCM 386
LiLAT 3      MNYIIVSDTSTAGFE EVGWGK PVYGGVAKGTIDWIGVNWYIPFKNKKGEQGIIVTVCL 420
              ** *** *:..* .: * .: *:
LiLAT 1      PLNAMEEFAKQFQMMITAARTSNLSAL 424
LiLAT 2      VV----- 388
LiLAT 3      PLNAMEEFAKQFRMMITAARTNLNSAL 447
              :

```

Figure 3.11 Amino Acid sequence comparisons of three LiLAT candidates which seem to have sequence identity. Red color indicates the missing portion between the sequence alignments. Yellow color shows the conserved motifs.

```

LiLAT 1  MASTKTLTFKVTTRKDPSELISPAEPTPYGFK-YLSDIDDQHGFRIREFSIIFFYRENPSMKG 59
LiLAT 4  -----MAMIITK---QILRPSSPTPQAFKNHKLSYLDQIQAPIYIPLLFFYKNEESKYP 51
          ::: **:   ::: *:.*** .** : . **      * :.:*****: *
LiLAT 1  KDPVKIIRDAVAKALVFYYPFAGRSRE-----CASR---KLVVDCTGEGVIFVEELLLDT 111
LiLAT 4  DQISQRFKQSLSEILTIFYPLAGTMRHNSFVDCNDRGVEFVEVRVHARLAQFIQDPKMEE 111
          .: : :.:.:.: *:.***:** * . * . * : * .. . *.:.: :.
LiLAT 1  PGYDGIINCP-----LLFIQVTRLKCGGFILTYSCNHTTICDAIGFSQFLSAVGELARG 164
LiLAT 4  LKQLIPVDCISHTDDDFLLLVKISYFDCGEVVVGVCMSHKIGDGISLAAFMNAWAATCRG 171
          :.*      *.:.:.: :.** :.: . .*. * *.*: :.* . .**
LiLAT 1  -ATVPSIQPVWERHLLSARNPPRVSFTHREYDVLPKTNGETDKMVVRYFFFDVADISALR 223
LiLAT 4  ESSSEIIHPSFD---LALHFPPKD---HLSSASSFRVAIAQENIMTKRLVFDREKLEKLR 225
          :. *:* :. * : : ** : * . :. :.:.: :.* :. **
LiLAT 1  RSLPPN---LQTCSKFDVVAFTLR-----CRTIAISLKPDEEVVFFSAVNIRNKITPPL 275
LiLAT 4  KRIAASSDGVDRDPSRVEAVSVFIWKSLEAHKAESHMTETPAVSIASHAVNLRPRTVPQM 285
          : :... :. *:.***:** : : : :. : *:*:* :.* :
LiLAT 1  PVGYYGNGIVSPAVVTTAEKLSKNPFHYAVELVMKAKYKATDEYVKSVDLMVMRDRPSF 335
LiLAT 4  DQTFGNCYAPASAVVSWDEDYVHHSRLRAALREIDDDYIN--KVLKADNNYLTQDQIGD 342
          : . :.***: * . :. * . :. * * .. : :*: .
LiLAT 1  TVARNYCIVSDTTNVGFQKVDLGWGEPVYGGGLAKGIGWIPAHWYIPFKNKKGEQGTIVTV 395
LiLAT 4  LFKPENSVLSSWWRFPVYKVDFGWGKPVWVSTTT----IQYMNLIIFTSTPSEDGIEAWV 398
          . : :.:. . . *.:***:**: . :. * * *... *:* . *
LiLAT 1  CLPLNAMEEFAKQFQMMITAARTSNLSAL 424
LiLAT 4  TTTHNFFQVLQANYNKLDT----- 417
          . * :. : :. : : *

```

Figure 3.12 Amino acid sequence comparisons of LiLAT-1 and LiLAT-4. LiLAT-1 is a representative from the identicle candidates. Yellow color shows the conserve motifs.

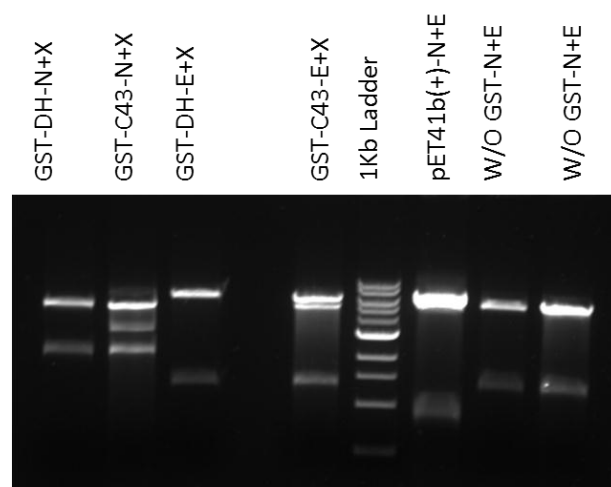


Figure 3.13 Restriction enzyme digestion of LiLAT clones after cloning into pET41b(+) vector. N= NdeI, E= EcoRI, X= XhoI, DH= *E. coli* DH10B competent cells, C43= *E. coli* C43(DE3)pLysS cells.

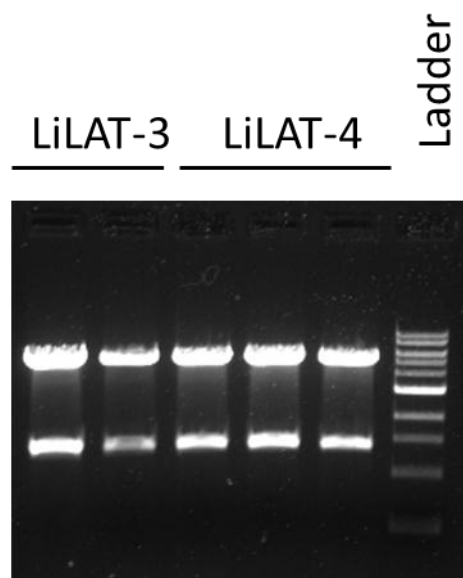


Figure 3.14 Restriction enzyme digestion of LiLAT clones in pGEX4T1 vector using EcoRI and XhoI restriction sites.

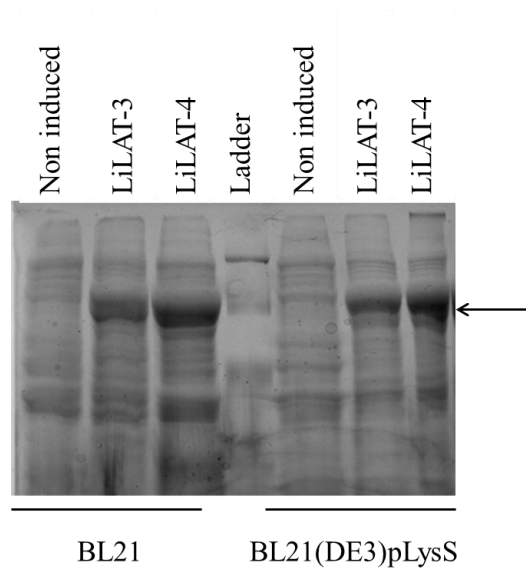


Figure 3.15 SDS-PAGE analysis of protein samples extracted from bacterial cells expressing LiLAT candidates using the pGEX4T1 vector. Non induced: total protein from non induced cells; Ladder: Protein markers; LiLAT-3 and LiLAT-4: total protein from induced bacterial cells from respective clones. BL21 and BL21(DE3)pLysS bacterial cells were used in this experiment. The band indicated by the arrow corresponds to the LiLAT-GST fusion protein.

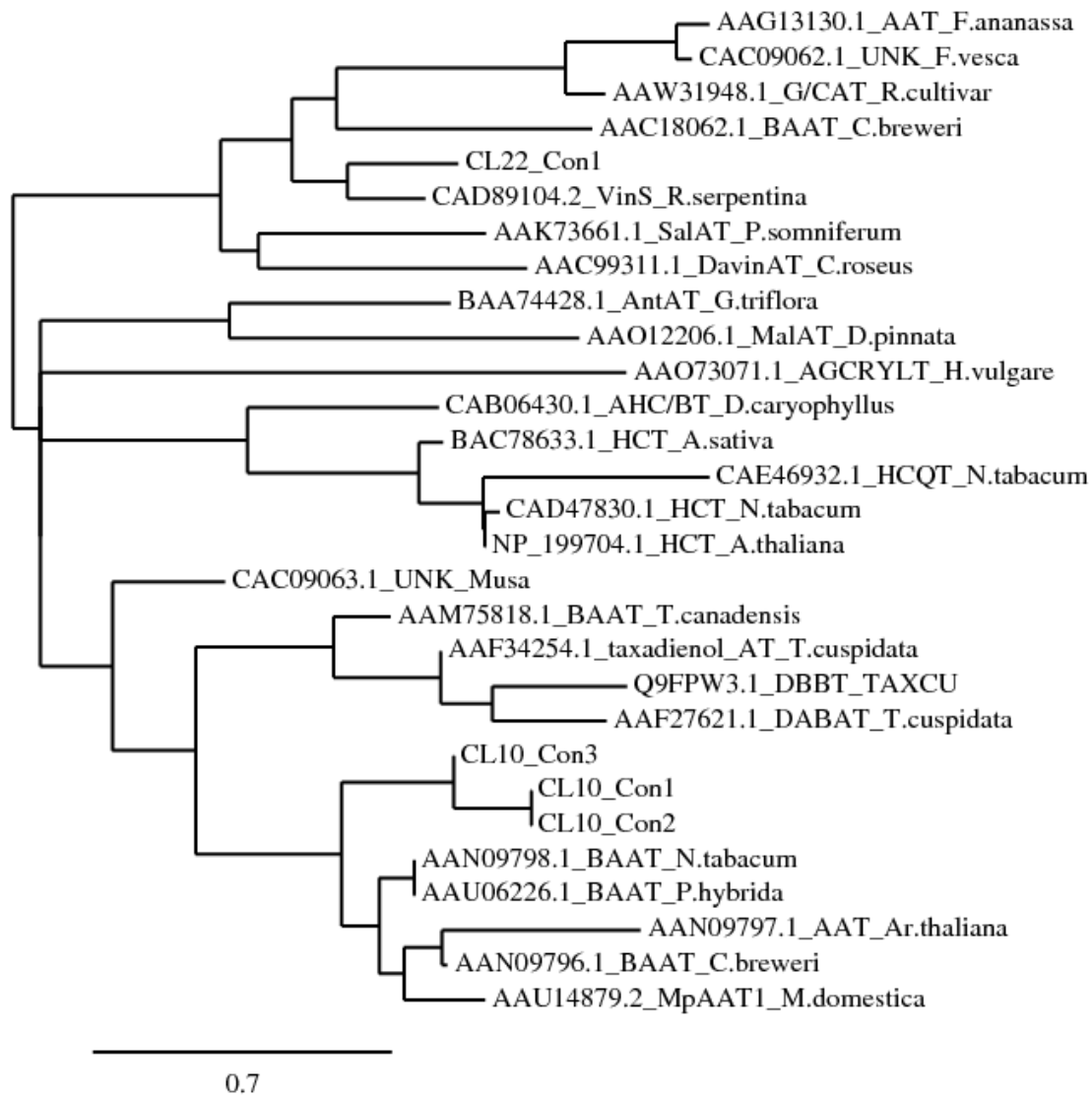


Figure 3.16 Phylogenetic tree analysis of LiLAT candidates with other acyltransferases from different plants. The scale bar represents 0.7 amino acid substitutions per site.

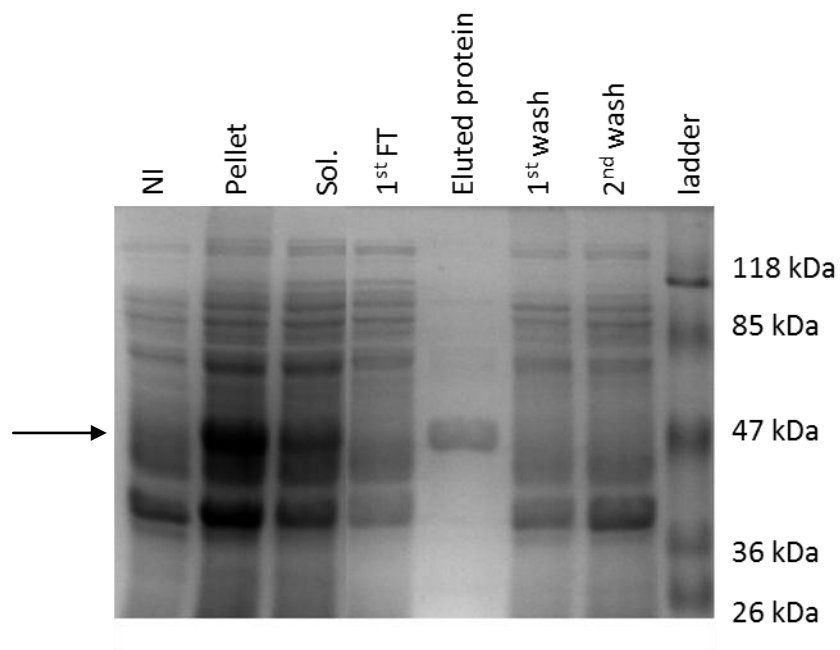


Figure 3.17 SDS-PAGE analysis of LaTPS-I. NI: Non induced; Pellet: cells after centrifugation; Sol.: Soluble fraction; 1st FT: Flow through; 1st and 2nd wash: flow through after washing the column before elution; Eluted protein: contain pure fraction of LaTPS-I protein. The arrow indicates the LaTPS-I recombinant protein.

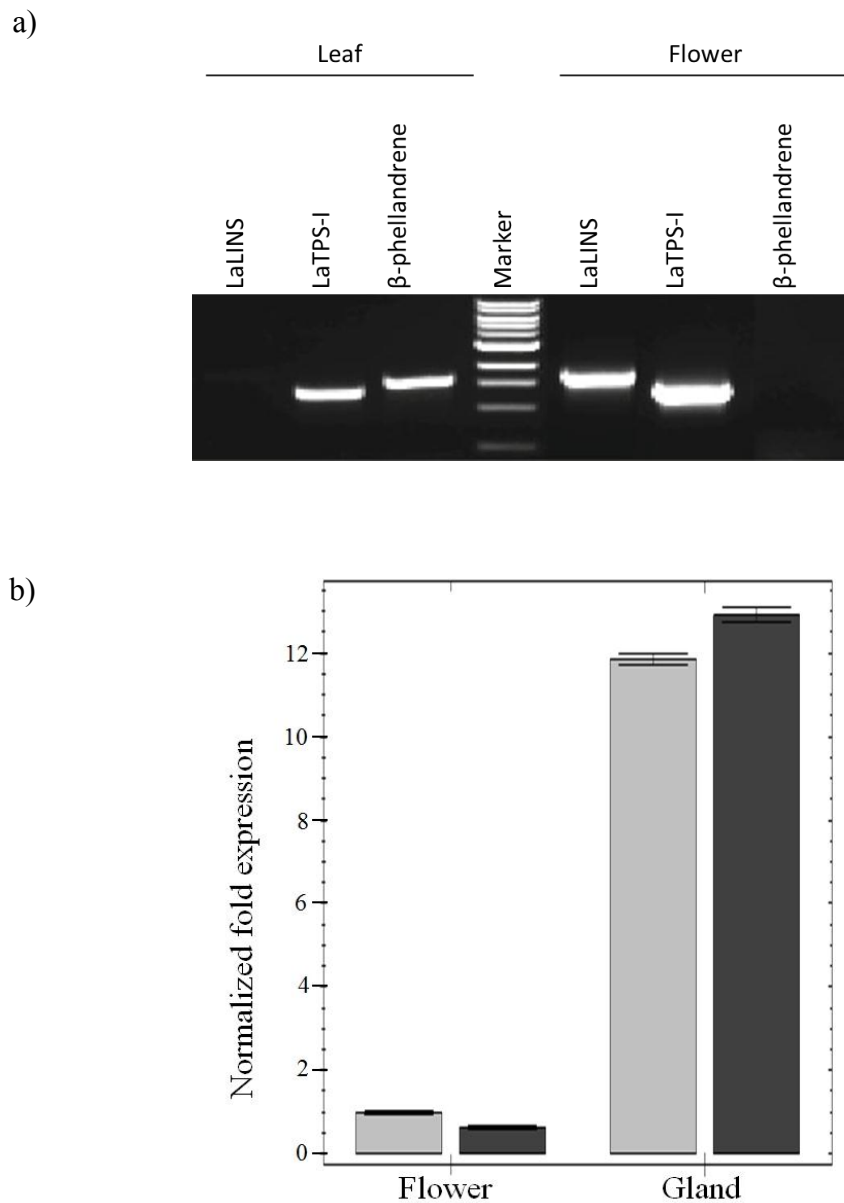


Figure 3.18 *LaTPS-I* transcript analysis. a) Standard PCR and b) qPCR. Black represents *LaLINS* and grey represents *LaTPS-I* (normalized with β -Actin, n=3). Figure 3.18 a) was reproduced from Demissie et al. 2011 with proper authorization from Planta.

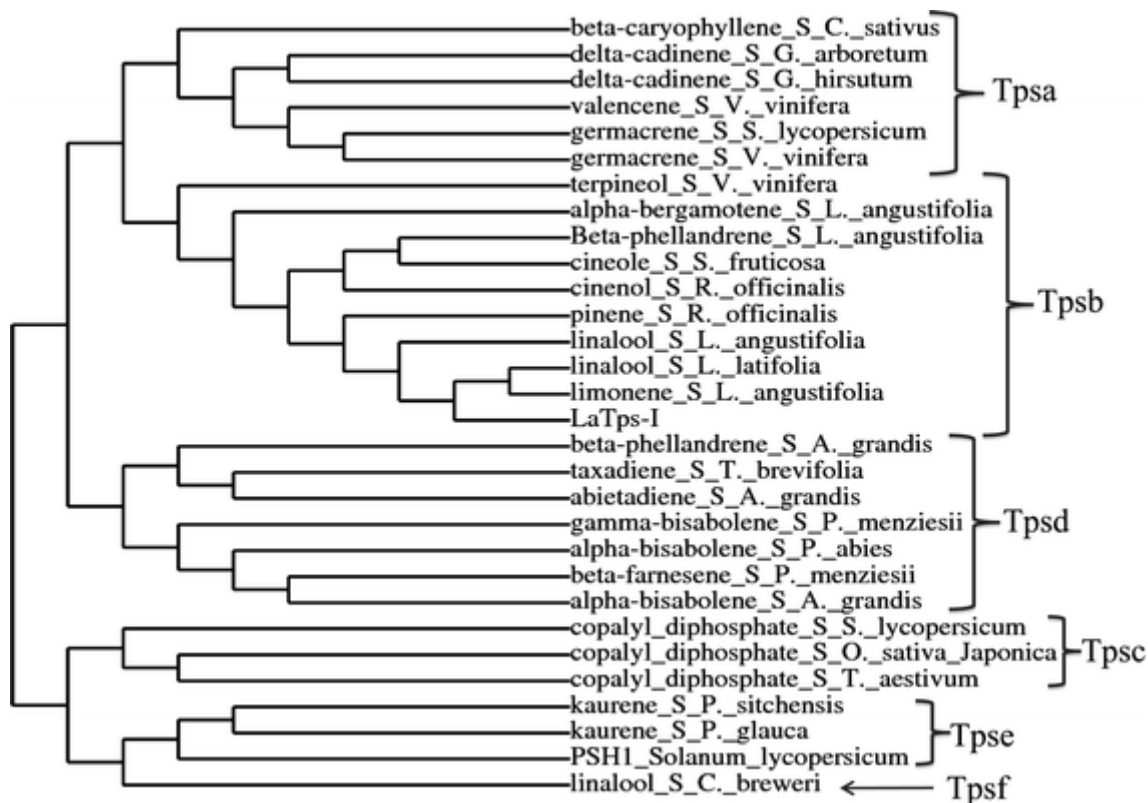


Figure 3.20 Phylogenetic tree analyses of LaTPS-I with other terpene synthases from different plants. TPSs with 50% minimum amino acid sequence identity are grouped into TPSa – TPSf subfamilies. Reproduced from Demissie et al., 2010 with permission from Planta.

4 Chapter: Discussion

4.1 Cloning and functional characterization of LiBDH

The secretory cells of glandular trichomes in *Lavandula* strongly and specifically express genes required for all stages of monoterpene metabolism. This includes those involved in the MEP pathway (e.g., DXS), which supplies precursors for monoterpene biosynthesis, and those that catalyze the formation of individual monoterpenes from GPP (e.g., linalool synthase;(Lane et al. 2010)). Also, genes encoding enzymes that catalyze the downstream modification of monoterpenes are strongly expressed in these trichomes. For example, the oxygenases, reductases and dehydrogenases that mediate the transformation of limonene to menthol in peppermint, a process that involves several biochemical reactions, are highly abundant in peppermint oil glands (Croteau et al. 2005). In order to probe the biosynthesis of EO monoterpene constituents in *Lavandula*, we have recently developed a gland-specific EST library from *L. x intermedia*. This database is highly enriched in monoterpene biosynthetic genes, and has facilitated the cloning of several terpene synthases including the *L. x intermedia* 1,8-cineole synthase (Demissie et al. 2012). We thus hypothesized that a borneol dehydrogenase is also strongly expressed in *L. x intermedia* oil glands, and proceeded to clone and functionally characterize the gene. Initially, we isolated two candidates from our *L. x intermedia* gland cDNA library, expressed them in *E. coli* cells, and assayed the dehydrogenase activities of the purified recombinant proteins using borneol and other main *Lavandula* monoterpenes as substrates. One of these candidates, LiBDH, was able to convert borneol into camphor. However, unlike many other SDRs, including a recently reported *Artemisia annua* SDR (ADH2; (Polichuk et al. 2010)) that accepts a number of

substrates, LiBDH did not produce detectable products from other monoterpenoid alcohols, indicating that this SDR is highly specific. Indeed to our knowledge, LiBDH is the first borneol specific dehydrogenase reported from plants. The recently reported *Artemisia annua* SDR (ADH2) was shown to dehydrogenate a range of substrates including: (-)-cis-carveol, (-)-artemisia alcohol, (+/-)-borneol, (-)-trans-carveol, and (-)-trans-pinocarveol. This enzyme had the highest specific activities for (-)-cis-carveol and (-)-artemisia alcohol, and the lowest specific activity for borneol (Polichuk et al. 2010), indicating that borneol is not a primary substrate for *Artemisia annua* ADH2.

LiBDH is structurally similar to other plant SDRs and contains the standard conserved motifs present in these proteins, including the structural “Rossmann fold” and the “Ser-Tyr-Lys” catalytic triad which is very important for SDR functionality (Chen et al. 1993). The protein also contains other conserved motifs present in plant SDRs including the N-terminal cofactor-binding TGXXX(AG)XG motif and the catalytic YXXXXK motif (Figure 3.5), and hence belongs to the classical SDR subfamily (Figure 3.6). In addition, several key amino acid residues were conserved, including the Ser141 residue, which helps to form the catalytic triad “Ser-Tyr-Lys” (Gani et al. 2008), and the Asp42 residue that plays a critical role in determining the coenzyme (NAD⁺ over NADP⁺) specificity of SDRs (Kallberg et al. 2002, Kallberg et al. 2010, Kavanagh et al. 2008, Ringer et al. 2003).

The recombinant LiBDH had an optimum pH of 8.5, an optimum temperature of 32 °C, a K_m of 53.6 μ M, a turnover number (k_{cat}) of 4.0×10^{-4} sec⁻¹, and a specificity constant (k_{cat}/K_m) of 7.5×10^{-6} μ M⁻¹s⁻¹. Although most of these values are in the general range of those reported for other SDRs, we noted that LiBDH is a rather slow enzyme as long incubation times were required to obtain sufficient quantities of the product for GC-MS analysis. This

could be a result of experimental conditions, and the fact that the substrate borneol is poorly soluble in aqueous buffers.

Many plant monoterpene synthases are transcriptionally regulated. For example, productions of menthofuran in peppermint leaves (Mahmoud and Croteau 2003), and linalool in *L. angustifolia* flowers increased through development and directly correlated with the transcriptional activities of the *menthofuran synthase* and *linalool synthase* genes, respectively. This was not the case with *LiBDH*, in which case transcript levels remained steady and relatively low during flower development (Figure 3.4b). The tissue camphor content also remained unchanged, although the tissue concentration of borneol (the precursor to camphor) increases with flower age (Figure 3.4d). The increases in the concentrations of the substrate implies a lack of its turnover, which in turn suggests that catalytically active LiBDH may be restricted to young tissue, where the bulk of camphor is produced, and not available during the latter stages of flower development. Consistent with these results, the activity of a number of monoterpenes synthases were very high during the early stages of leaf development (when EO synthesis is very active) and dropped rapidly in maturing leaves in peppermint (McConkey et al. 2000). Our data cannot explain whether the postulated “unavailability” of active LiBDH in older flowers is due to a lack of protein synthesis (i.e., inhibition of *LiBDH* mRNA translation), protein inactivation by inhibition, sub-cellular localization or another mechanism.

Like other plant terpene synthases (Demissie et al. 2012, Iijima et al. 2004b, Ro et al. 2002) the LiBDH transcripts were highly concentrated in floral glandular trichomes of *L.x intermedia* (Figure 3.4b). In this sense, the expression of *LiBDH* correlates with the expression of other genes involved in monoterpene metabolism in lavenders (Demissie et

al. 2012). A surprising finding of this study was that *LiBDH* transcripts were present at higher levels in *L. angustifolia*, and *L. x intermedia* flowers, compared to those of *L. latifolia* plants (Figure. 3.4c). Given that *L. latifolia* plants produce camphor as a major compound (Munoz-Bertomeu et al. 2007) and thus are expected to express a borneol dehydrogenase gene strongly, our data imply that *L. latifolia* plants may express a unique BDH, and that *L. x intermedia* (which is a natural hybrid of *L. angustifolia* and *L. latifolia*) inherited the cloned *LiBDH* from the *L. angustifolia* parent. This postulate is supported by the finding that both *L. angustifolia* and *L. x intermedia* plants express the *LiBDH* gene but *L. latifolia* does not.

4.2 Cloning and functional characterization of LiCPS

Sesquiterpenes, the C₁₅ terpenoid compounds, are one of the largest and most diverse groups of natural products isolated from plant, fungi, bacteria, and marine invertebrates. They are known for their defensive action against plant enemies, and attraction of pollinators. The thousands of different sesquiterpene compounds found in plants are derived from the precursor, farnesyl diphosphate (FPP) through the enzymatic action of sesquiterpene synthase enzymes which can yield one or more sesquiterpene products from this single substrate. Lavender EO, dominated by monoterpenes, also contains a few sesquiterpenes, including β -caryophyllene, α -bergamotene, and cadinene among others. β -caryophyllene is widely present in plants and known for its allelopathic potential and its role in plant defense. For example, this volatile compound was found to be produced abundantly in response to herbivore damage in maize. β -caryophyllene is also pharmaceutically important as it exhibits anti-inflammatory and anti-carcinogenic activities.

In this study, a β -caryophyllene synthase cDNA was cloned from the *L. x intermedia* gland cDNA library and functionally characterized *in vitro*. The encoded protein (LiCPS)

shares all conserve motifs with other known terpene synthases. These motifs include, the arginine rich N-terminal RR(X₈)W motif, required for cyclization of GPP (Williams et al. 1998a), the aspartate rich divalent metal (usually Mg²⁺) binding motif, DDXXD, involved in substrate binding and in positioning the substrate for catalysis (Cane et al. 1996a, Cane et al. 1996b, Lesburg et al. 1997, Tarshis et al. 1994). Any mutations in this region would subsequently lead to altered products formation and decreased catalytic activities (Cane et al. 1996a, Cane et al. 1996b, Seemann et al. 1999, Seemann et al. 2002a, Seemann et al. 2002b). Phylogenetic studies showed that, LiCPS belongs to the TPS-b group along with LiBERS and ZIS from sweet basil. Deduced amino acid sequence of LiCPS showed 77% conserved identity with LaBERS when aligned against known protein sequences in the NCBI data base. Amino acid sequence revealed that, like other known sesquiterpene synthases, LiCPS lacks a transit peptide.

Transcript profiling suggested that *LiCPS* is developmentally regulated during flower development in *L. x intermedia*. Further, our data indicated that this gene is more strongly expressed in flower and glandular trichomes of *L. x intermedia* compared to those of *L. angustifolia*. Standard PCR also revealed that *LiCPS* transcript is highly abundant in glandular trichomes and is developmentally regulated in *L. x intermedia* flower tissues while it is undetectable in the leaves. The purified recombinant LiCPS was able to use GPP, NPP, and FPP as substrates in *in vitro* assays. With FPP as substrate LiCPS produced β -caryophyllene as a major product in addition to some minor products such as pentanoic acid (Figure 3.7a). Like other sesquiterpenes; e.g. trans- α bergamotene synthase from *L. angustifolia*, α -zingiberene synthase from *O. basilicum*, and β -caryophyllene synthase from *A. annua*, LiCPS also produced monoterpenes (including 1,8-cineol, limonene, α -terpineol,

myrcene, and β -pinene) when assayed with GPP and NPP as substrates (Figure 3.7b, 3.7c) (Cai et al. 2002, Iijima et al. 2004a, Iijima et al. 2004b, Kollner et al. 2004, Landmann et al. 2007).

4.3 Cloning and functional characterization of LiLAT

Lavender EOs used in the perfume and cosmetic industry have a characteristically higher concentration of linalool and linalool acetate to other EO volatile compounds, in particular camphor. *L. angustifolia* plants produce the finest lavender EOs because of its higher linalool and linalool acetate to camphor ratios; however the total oil yield is very low (40 kg per hectare). On the other hand, ratios of linalool and linalool acetate to camphor are lower in *L. latifolia* and *L. x intermedia*, and the oil yield is much greater, 70 kg and 120 kg per hectare, respectively. Like other volatile esters, linalool acetate in lavender is normally generated from linalool as a result of the action of *linalool acetyltransferase* (LAT) (Aharoni et al. 2000, Harada et al. 1985). Cloning of the LAT gene would enable metabolic engineering experiments aimed at enhancing linalool acetate production in economically important lavender varieties including *L. x intermedia* plants.

A total of 117 ESTs from the gland cDNA library exhibited significant homology to known acetyltransferase genes. Based on their expression pattern (obtained from the microarray analysis experiment), four candidates, LiLAT-1, LiLAT-2, LiLAT-3 and LiLAT-4, were selected for further analysis. Although LiLAT-3 and LiLAT-4 were successfully cloned into the pET41b(+) expression vector, the recombinant protein was not expressed in bacterial cells. On the other hand, when LiLAT-3 and LiLAT-4 were cloned into pGEX4T-1 expression vector, the corresponding recombinant proteins were strongly expressed in bacteria. However, I was not able to purify the recombinant enzyme as the expressed protein

aggregated in insoluble inclusion bodies. It may be possible that the expression systems used in these studies are not suitable for the production of recombinant LAT. Thus future investigations aimed at the production of recombinant LAT should employ alternative expression systems that rely on yeast or insect cells as a host.

4.4 Cloning and functional characterization of LaTPS-I

LaTPS-I was selected due to its unusual sequence characteristics compared to other monoterpene synthases. This interesting candidate was missing a stretch of 73 amino acids in the middle, which contains the most important conserved motif found in terpene synthases, DDxxD. Further, unlike most other monoterpene synthases expressed differentially in plants, LaTPS-I transcripts were found to be present in both leaf and flower tissues. The ORF of LaTPS-I, 1482 base pairs encoding 494 amino acids, was cloned into a pET41b(+) expression vector and expressed in *E. coli* cells, and the recombinant protein was purified using Ni-NTA affinity column chromatography. The recombinant protein did not yield any product when assayed with GPP, FPP and NPP as substrates. However, small amounts of linalool and geraniol were detected. These products were also present in the negative control, recombinant protein from empty pET41b(+) vector, indicating that the linalool and geraniol were produced through solvolysis. Indeed, previous studies have confirmed that hydrolysis of GPP can yield geraniol and linalool (Tholl et al. 2001). Our results, therefore, suggested that LaTPS-I lacks *in vitro* catalytic activity. The most likely explanation for this could be the lack of the highly conserved divalent metal binding motif DDXXD in LaTPS-I. This suggestion is further supported by previous results, where substitution of any amino acid residue in this motif either reduced or completely eliminated its enzymatic activity. For example, the replacement of either the first or second aspartate (D) residue by glutamic acid

(E) in pentalenene synthase (Seemann et al. 2002b) or in aristolochene synthase (Felicetti and Cane 2004) was seen to significantly reduce the catalytic efficiency and specificity of the mutated proteins. Additionally, a few residues, C-terminal to the DDXXD motif also played a role in terpene synthase activity. Amino acid substitutions in this area, such as (N) to (A) or (L) in pentalenene synthase (Seemann et al. 2002b) and (L) in aristolochene synthase (Felicetti and Cane 2004) resulted in complete enzyme inactivation. It is therefore possible that amino acid residues, other than those in the DDXXD motif within the “missing” region (Figure 3.19), play a significant role in controlling the activity of LaTPS-I, and other related terpene synthase enzymes.

Standard PCR depicts that *LaTPS-I* transcripts are strongly expressed in both flower and leaf tissues compared to *LaLINS* and *LaβPHLS* which were present only in flower or leaf tissues, respectively (Figure 3.18a). To elucidate the tissue-specific expression of *LaTPS-I*, we compared the transcriptional activity of *LaTPS-I* with *LaLINS*, which is strongly expressed in floral oil glands. Linalool is a major component of *L. angustifolia* essential oil, often comprising over 40% of extracts, and it was recently shown that *LaLINS* expression is closely related to linalool accumulation in lavender spikes (Lane et al. 2010). Both *LaTPS-I* and *LaLINS* were present at a similar level in flower tissue. Transcript abundance of *LaTPS-I* in leaf is also comparable to its expression in flower tissue, while *LaLINS* is barely detectable in the leaves, using quantitative real-time PCR. The natural occurrence of this variant was ascertained by sequencing eight independent LaTPS-I clones amplified directly from the *L. angustifolia* flower cDNA (data not shown). However, due to lack of detectable *in vitro* activity for LaTPS-I, its biological significance in planta could not be confirmed.

4.5 Conclusion

In conclusion, the studies presented in this thesis resulted in the cloning of four essential oil related genes from *L. angustifolia* and *L. intermedia* plants. The heterologous expression of two of these genes, including LiCPS and LiBDH, in *E. coli* yielded corresponding biologically active recombinant proteins. LiCPS turned out to be a typical sesquiterpene synthase that, like many other sesquiterpene synthases, can also produce monoterpenes when assayed with the monoterpene precursors GPP and NPP as substrates *in vitro*. Although the gene is strongly expressed in glandular trichomes, LiCPS does not make a major contribution to the EO of lavender plants presumably due to the fact that it is localized to the cytosol (as it lacks a plastidial targeting sequence) where its sesquiterpene substrate (FPP) is not produced in high quantities.

LiBDH is a short chain alcohol dehydrogenase that converts borneol to camphor, and as predicted by our hypothesis, is responsible for camphor production in Lavandin plants. Based on its expression pattern, it was concluded that this gene, which is highly expressed in glandular trichomes of *L. x intermedia* plants, was likely inherited from *L. angustifolia*. Although the cloned LiBDH is highly specific, our data indicated that another *Lavandula BDH* may exist that is expressed in *L. latifolia* plants. To clone this gene, future efforts must focus on an EST database derived from oil glands of *L. latifolia* plants.

The work presented here also led to the cloning of two other genes that are strongly expressed in the glandular trichomes of lavender and lavandin plants. The recombinant form of the terpene synthase-like protein encoded by one of these genes, LaTPS-I, was not functionally active *in vitro*, presumably because it lacked the signature TPS substrate binding motif present in plant TPSs. My data does not explain why this seemingly inactive gene is

strongly expressed in lavenders. My results cannot rule out an as of yet undiscovered activity for this gene.

Based on homology to known acetyltransferases involved in plant secondary metabolism, it is very likely that the cloned putative *LiLAT* encodes the elusive linalool acetyltransferase. However, this claim could not be verified since, despite strong expression in *E. coli*, I was unable to purify the recombinant LiLAT protein and demonstrate its acetyltransferase activity.

LiBDH and LAT are two of the most important EO quality-determining genes expressed in lavenders, and their cloning represents a major milestone in lavender research. In principle, the expression of these genes may be modulated through metabolic engineering to enhance the quality of the EO (i.e., increase the production of linalool acetate and/or reduce the biosynthesis of camphor) in transgenic plants. Alternatively, the cloned genes may be used as genetic markers in targeted plant breeding programs aimed at producing lavender plants with an improved EO profile.

References

- Adersen A, Gauguin B, Gudiksen L, Jager AK (2006) Screening of plants used in Danish folk medicine to treat memory dysfunction for acetylcholinesterase inhibitory activity. *J Ethnopharmacol.* 104(3):418-22.
- Aharoni A, Keizer LCP, Bouwmeester HJ, Sun ZK, Alvarez-Huerta M, Verhoeven HA, Blaas J, van Houwelingen AMML, De Vos RCH, van der Voet H, Jansen RC, Guis M, Mol J, Davis RW, Schena M, van Tunen AJ, O'Connell AP (2000) Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. *Plant Cell.* 12(5):647-62.
- Boeckelmann A (July, 2008) Monoterpene production and regulation in Lavenders. Dissertation or Thesis, University of British Columbia.
(<https://circle.ubc.ca/handle/2429/2804>)
- Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: Molecular biology and phylogenetic analysis. *Proc Natl Acad Sci U S A.* 95(8):4126-33.
- Bohlmann J, Steele CL, Croteau R (1997) Monoterpene synthases from Grand fir (*Abies grandis*) - cDNA isolation, characterization, and functional expression of myrcene synthase, (-)(4S)-limonene synthase, and (-)-(1S,5S)-pinene synthase. *J Biol Chem.* 272(35):21784-92.
- Cai Y, Jia JW, Crock J, Lin ZX, Chen XY, Croteau R (2002) A cDNA clone for beta-caryophyllene synthase from *Artemisia annua*. *Phytochemistry.* 61(5):523-9.
- Cane DE, Xue Q, Fitzsimons BC (1996a) Trichodiene synthase. Probing the role of the highly conserved aspartate-rich region by site-directed mutagenesis. *Biochemistry (N Y).* 35(38):12369-76.
- Cane DE, Xue Q, VanEpp JE (1996b) Enzymatic formation of isochamigrene, a novel sesquiterpene, by alteration of the aspartate-rich region of trichodiene synthase. *J Am Chem Soc.* 118(35):8499-500.

- Cavanagh HMA, Wilkinson JN (2002) Biological activities of lavender essential oil. *Phytotherapy Research*. 16(4):301-8.
- Chen Z, Jiang J, Lin Z, Lee W, Baker M, Chang S (1993) Site-Specific Mutagenesis of *Drosophila* Alcohol-Dehydrogenase - Evidence for Involvement of Tyrosine-152 and Lysine-156 in Catalysis. *Biochemistry* (N Y). 32(13):3342-6.
- Christianson DW (2006) Structural biology and chemistry of the terpenoid cyclases. *Chem Rev*. 106(8):3412-42.
- Chu CJ, Kemper KJ (2001) Lavender (*Lavandula* spp.) Retrieved Aug 15, 2012 from <http://www.longwoodherbal.org/lavender/lavender.pdf>
- Clark RJ, Menary RC (1980) Environmental-Effects on Peppermint (*Mentha-Piperita* L) .I. Effect of Daylength, Photon Flux-Density, Night Temperature and Day Temperature on the Yield and Composition of Peppermint Oil. *Aust J Plant Physiol*. 7: 685-92.
- Cosentino S, Tuberose CIG, Pisano B, Satta M, Mascia V, Arzedi E, Palmas F (1999) In-vitro antimicrobial activity and chemical composition of Sardinian *Thymus* essential oils. *Lett Appl Microbiol*. 29(2):130-5.
- Croteau R, Hooper CL, Felton M (1978) Biosynthesis of Monoterpenes - Partial-Purification and Characterization of a Bicyclic Monoterpenol Dehydrogenase from Sage (*Salvia Officinalis*). *Arch Biochem Biophys*. 188(1):182-93.
- Croteau R, Davis E, Ringer K, Wildung M (2005) (-)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften*. 92(12):562-77.
- D'Auria JC (2006) Acyltransferases in plants: a good time to be BAHD. *Curr Opin Plant Biol*. 9(3):331-40.
- Degenhardt J, Koellner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry*. 70(15-16):1621-37.

- Demissie ZA, Cella MA, Sarker LS, Thompson TJ, Rheault MR, Mahmoud SS (2012) Cloning, functional characterization and genomic organization of 1,8-cineole synthases from *Lavandula*. *Plant Mol Biol*. 79(4-5):393-411.
- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-, Guindon S, Lefort V, Lescot M, Claverie J-, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res*. 36:465-9.
- Dudareva N, Martin D, Kish CM, Kolosova N, Gorenstein N, Faldt J, Miller B, Bohlmann J (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*. 15(5):1227-41.
- Dudareva N, Cseke L, Blanc VM, Pichersky E (1996) Evolution of floral scent in *Clarkia*: Novel patterns of S-linalool synthase gene expression in the *C-breweri* flower. *Plant Cell*. 8(7):1137-48.
- Fahn A (1988) Secretory-Tissues in Vascular Plants. *New Phytol*. 108 (3):229-57.
- Felicetti B, Cane DE (2004) Aristolochene synthase: Mechanistic analysis of active site residues by site-directed mutagenesis. *J Am Chem Soc*. 126(23):7212-21.
- Gani OABSM, Adekoya OA, Giurato L, Spyraakis F, Cozzini P, Guccione S, Winberg J, Sylte I (2008) Theoretical calculations of the catalytic triad in short-chain alcohol dehydrogenases/reductases. *Biophys J*. 94(4):1412-27.
- Gershenzon J, Mccaskill D, Rajaonarivony J, Mihaliak C, Karp F, Croteau R (1992) Isolation of Secretory-Cells from Plant Glandular Trichomes and their use in Biosynthetic-Studies of Monoterpenes and Other Gland Products. *Anal Biochem*. 200(1):130-8.
- Goel HC, Roa AR (1988) Radiosensitizing Effect of Camphor on Transplantable Mammary Adenocarcinoma in Mice. *Cancer Lett*. 43(1-2):21-7.
- Goff SA, Klee HJ (2006) Plant volatile compounds: Sensory cues for health and nutritional value? *Sci*. 311:815-19.

Guillandcumming D, Smith GJ (1979) Mitochondrial Respiration Depressed by Camphor - Possible Aid in Radiotherapy. *Experientia*. 35(5):659.

Harada M, Ueda Y, Iwata T (1985) Purification and some Properties of Alcohol Acetyltransferase from Banana Fruit. *Plant and Cell Physiol*. 26 (6):1067-74.

Hohl RJ (1996) Monoterpenes as regulators of malignant cell proliferation. *Adv Exp Med Biol*. 401:137-46.

Iijima Y, Davidovich-Rikanati R, Fridman E, Gang DR, Bar E, Lewinsohn E, Pichersky E (2004a) 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(3):3724-36.

Iijima Y, Gang DR, Fridman E, Lewinsohn E, Pichersky E (2004b) Characterization of geraniol synthase from the peltate glands of sweet basil. *Plant Physiol*. 134(1):370-79.

Jornvall H, Persson B, Krook M, Atrian S, Gonzalezduarte R, Jeffery J, Ghosh D (1995) Short-Chain Dehydrogenases Reductases (Sdr). *Biochem (N Y)*. 269(18):4409-17.

Kallberg Y, Oppermann U, Jornvall H, Persson B (2002) Short-chain dehydrogenase/reductase (SDR) relationships: A large family with eight clusters common to human, animal, and plant genomes. *Protein Sci*. 11(3):636-41.

Kallberg Y, Oppermann U, Persson B (2010) Classification of the short-chain dehydrogenase/reductase superfamily using hidden Markov models. *Febs J*. ;277(10):2375-86.

Kavanagh KL, Jornvall H, Persson B, Oppermann U (2008) The SDR superfamily: functional and structural diversity within a family of metabolic and regulatory enzymes. *Cellular and Molecular Life Sci*. 65(24):3895-906.

Kollner TG, Schnee C, Gershenzon J, Degenhardt J (2004) The variability of sesquiterpenes cultivars is controlled by allelic emitted from two Zea mays variation of two terpene synthase genes encoding stereoselective multiple product enzymes. *Plant Cell*. 16(5):1115-31.

- Landmann C, Fink B, Festner M, Dregus M, Engel K, Schwab W (2007) Cloning and functional characterization of three terpene synthases from lavender (*Lavandula angustifolia*). *Arch Biochem Biophys*. 465(2):417-29.
- Lane A, Boecklemann A, Woronuk GN, Sarker L, Mahmoud SS (2010) A genomics resource for investigating regulation of essential oil production in *Lavandula angustifolia*. *Planta*. ;231(4):835-45.
- Lesburg CA, Zhai GZ, Cane DE, Christianson DW (1997) Crystal structure of pentalenene synthase: Mechanistic insights on terpenoid cyclization reactions in biology. *Sci*. 277(5333):1820-4.
- Lis-Balchin M (2002) Lavender: The genus *Lavandula*. (1st ed.). Taylor and Francis Inc., New York, NY.
- Little DB, Croteau RB (2002) Alteration of product formation by directed mutagenesis and truncation of the multiple-product sesquiterpene synthases delta-selinene synthase and gamma-humulene synthase. *Arch Biochem Biophys*. 402(1):120-35.
- Liu Y, Wang H, Ye HC, Li GF (2005) Advances in the plant isoprenoid biosynthesis pathway and its metabolic engineering. *J of Integ Plant Biolog*. 47 (7):769-82.
- Mahmoud SS, Williams M, Croteau R (2004) Cosuppression of limonene-3-hydroxylase in peppermint promotes accumulation of limonene in the essential oil. *Phytochem*. 65(5):547-54.
- Mahmoud SS, Croteau RB (2002) Strategies for transgenic manipulation of monoterpene biosynthesis in plants. *Trends Plant Sci*. 7 (8):366-73.
- Mahmoud S, Croteau R (2003) Menthofuran regulates essential oil biosynthesis in peppermint by controlling a downstream monoterpene reductase. *Proc Natl Acad Sci U S A*. 100 (24):14481-86.
- McConkey M, Gershenzon J, Croteau R (2000) Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint. *Plant Physiol*. 122 (1):215-24.

- McGeady P, Croteau R (1995) Isolation and characterization of an active-site peptide from a monoterpene cyclase labeled with a mechanism-based inhibitor. *Arch Biochem Biophys.* 317(1):149-55.
- Munoz-Bertomeu J, Arrillaga I, Segura J (2007) Essential oil variation within and among natural populations of *Lavandula latifolia* and its relation to their ecological areas. *Biochem Syst Ecol.* 35(8):479-88.
- Okamoto S, Yu F, Harada H, Okajima T, Hattan J, Misawa N, Utsumi R (2011) A short-chain dehydrogenase involved in terpene metabolism from *Zingiber zerumbet*. *Febs J.* 278(16):2892-900.
- Pattnaik S, Subramanyam V, Bapaji M, Kole C (1997) Antibacterial and antifungal activity of aromatic constituents of essential oils. 89(358):39-46.
- Peffley DM, Gayen AK (2003) Plant-derived monoterpenes suppress hamster kidney cell 3-hydroxy-3-methylglutaryl coenzyme A reductase synthesis at the transcriptional level. *J Nutr.* 133(1):38-44.
- Perrucci S, Mancianti F, Cioni PL, Flamini G, Morelli I, Macchioni G (1994) In-Vitro Antifungal Activity of Essential Oils Against some Isolates of *Microsporum-Canis* and *Microsporum-Gypseum*. *Planta Med.* 60(2):184-7.
- Piel J, Donath J, Bandemer K, Boland W (1998) Mevalonate-independent biosynthesis of terpenoid volatiles in plants: Induced and constitutive emission of volatiles. *Angewandte Chemie International Edition.* 37(18):2478-81.
- Polichuk DR, Zhang Y, Reed DW, Schmidt JF, Covello PS (2010) A glandular trichome-specific monoterpene alcohol dehydrogenase from *Artemisia annua*. *Phytochem.* 71(11-12):1264-9.
- Prosser I, Altug IG, Phillips AL, Konig WA, Bouwmeester HJ, Beale MH (2004) Enantiospecific (+)- and (-)-germacrene D synthases, cloned from goldenrod, reveal a functionally active variant of the universal isoprenoid-biosynthesis aspartate-rich motif. *Arch Biochem Biophys.* 432(2):136-44.

Ringer KL, McConkey ME, Davis EM, Rushing GW, Croteau R (2003) Monoterpene double-bond reductases of the (-)-menthol biosynthetic pathway: isolation and characterization of cDNAs encoding (-)-isopiperitenone reductase and (+)-pulegone reductase of peppermint. *Arch Biochem Biophys*. 418(1):80-92.

Ro DK, Ehlting J, Douglas CJ (2002) Cloning, functional expression, and subcellular localization of multiple NADPH-cytochrome P450 reductases from hybrid poplar. *Plant Physiol*. 130(4):1837-1851.

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386.

Sakamoto R, Minoura K, Usui A, Ishizuka Y, Kanba S (2005) Effectiveness of aroma on work efficiency: Lavender aroma during recesses prevents deterioration of work performance. *Chem Senses*. 30(8):683-91.

Sarker LS, Galata M, Demissie ZD, Mahmoud SS (2012) Molecular cloning and functional characterization of borneol dehydrogenase from the glandular trichomes of *Lavandula x intermedia*. *Arch Biochem Biophys*. 528(2):163-170.

Schulz S, Buhling F, Ansorge S (1994) Prenylated Proteins and Lymphocyte-Proliferation - Inhibition by D-Limonene and Related Monoterpenes. *Eur J Immunol*. 24(2):301-7.

Seemann M, Zhai GZ, de Kraker JW, Paschall CM, Christianson DW, Cane DE (2002) Pentalenene synthase. Analysis of active site residues by site-directed mutagenesis. *J Am Chem Soc*. 124(26):7681-9.

Seemann M, Zhai GZ, Umezawa K, Cane D (1999) Pentalenene synthase. Histidine-309 is not required for catalytic activity. *J Am Chem Soc*. 121:591-92.

Shaw WV (1992) Chemical Anatomy of Antibiotic-Resistance - Chloramphenicol Acetyltransferase. *Sci Prog*. 76(301-302 Pt 3-4):565-80.

- Steele CL, Crock J, Bohlmann J, Croteau R (1998) Sesquiterpene synthases from grand fir (*Abies grandis*) - Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization and bacterial expression of delta-selinene synthase and gamma-humulene synthase. *J Biol Chem.* 273(4):2078-89.
- St-Pierre B, De Luca V (2000) Evolution of acyltransferase genes: Origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. *Evol of Metabol Pathways.* 34:285-315.
- Tanaka N, Nonaka T, Nakamura K, Hara A (2001) SDR: Structure, mechanism of action, and substrate recognition. *Cur Org Chem.* 5(1):89-111.
- Tarshis LC, Yan MJ, Poulter CD, Sacchettini JC (1994) Crystal-Structure of Recombinant Farnesyl Diphosphate Synthase at 2.6-Angstrom Resolution. *Biochem (N Y).* 33(36):10871-7.
- Theis JGW, Koren G (1995) Camphorated Oil - Still Endangering the Lives of Canadian Children. *Can Med Assoc J.* 152(11):1821-24.
- Tholl D, Croteau R, Gershenzon J (2001) Partial purification and characterization of the short-chain prenyltransferases, geranyl diphosphate synthase and farnesyl diphosphate synthase, from *Abies grandis* (grand fir). *Arch Biochem Biophys.* 386(2):233-42.
- Trapp SC, Croteau RB (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics.* 158(2):811-32.
- Turner GW, Gershenzon J, Croteau RB (2000) Distribution of peltate glandular trichomes on developing leaves of peppermint. *Plant Physiol.* 124(2):655-64.
- Upton T, Andrews S, Harriott G, King C, Langhorne J (2004) The genus *Lavandula*. (1st ed.). Timber Press, Inc., USA.
- Whittington DA, Wise ML, Urbansky M, Coates RM, Croteau RB, Christianson DW (2002) Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proc Natl Acad Sci U S A.* 99(24):15375-80.

Williams DC, McGarvey DJ, Katahira EJ, Croteau R (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. *Biochem (N Y)*. 37(35):12213-20.

Wise ML, Savage TJ, Katahira E, Croteau R (1998) Monoterpene synthases from common sage (*Salvia officinalis*) - cDNA isolation, characterization, and functional expression of (+)-sabinene synthase, 1,8-cineole synthase, and (+)-bornyl diphosphate synthase. *J Biol Chem*. 273(24):14891-9.

Wolfe N, Herzberg J (1996) Can aromatherapy oils promote sleep in severely demented patients? *Int J Geriatr Psychiatry*. 10(11):926-27.

Xu HX, Blair NT, Clapham DE (2005) Camphor activates and strongly desensitizes the transient receptor potential vanilloid subtype 1 channel in a vanilloid-independent mechanism. *J of Neurosci*. 25(39):8924-37.

Zhou K, Peters RJ (2009) Investigating the conservation pattern of a putative second terpene synthase divalent metal binding motif in plants. *Phytochem*. 70(3): 366–69.