CLONING AND FUNCTIONAL CHARACTERIZATION OF GENES INVOLVED IN THE BIOSYNTHESIS AND SECRETION OF ESSENTIAL OIL CONSTITUENTS OF *LAVANDULA*

by

Zerihun Abebe Demissie

BSc, Hawassa University, Hawassa, Ethiopia, 2001
MSc, Kenyatta University, Nairobi, Kenya, 2007

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

DOCTOR OF PHILOSOPHY

in

THE COLLEGE OF GRADUATE STUDIES

(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

April 2014

©Zerihun Abebe Demissie, 2014
Abstract

Several members of the genus *Lavandula* produce valuable essential oils (EOs) that are primarily constituted of regular and irregular monoterpenes, the C10 class of terpenes. We isolated over 22,000 ESTs from leaves, flowers and glandular trichomes of *L. angustifolia* and *L. x intermedia* species to facilitate the discovery of genes regulating the biosynthesis and secretion of EOs. In this study, we identified and studied genes involved in regular and irregular monoterpene biosynthesis and secretion. One of these genes, the *L. x intermedia* lavandulyl pyrophosphate synthase (*LiLPPS*), catalyzes the head–to–middle condensation of two DMAPP units to produce lavandulyl pyrophosphate *in vitro*. The apparent \( K_m \) and \( k_{cat} \) of *LiLPPS* were 208 \( \mu \)M and 0.1 s\(^{-1}\), respectively. *LiLPPS* is a novel cis prenyl transferase family member and its identification elucidated the biosynthetic origin of irregular monoterpenes in *Lavandula* EOs.

*L. angustifolia* \( \beta \)-phellandrene synthase (*La\( \beta \)PHLS) and *L. x intermedia* 1,8-cineole synthase (*LiCINS*) transformed geranyl pyrophosphate primarily to \( \beta \)-phellandrene and 1,8-cineole, respectively. The apparent \( K_m \) and \( k_{cat} \) of *La\( \beta \)PHLS* were 6.6 \( \mu \)M and \( 1.8 \times 10^{-2} \) s\(^{-1}\) while those of *LiCINS* were 5.8 \( \mu \)M and \( 8.8 \times 10^{-3} \) s\(^{-1}\), respectively. *La\( \beta \)PHLS* transcripts were highly abundant in young leaves where \( \beta \)-phellandrene is actively produced. *LiCINS* mRNA levels paralleled the 1,8-cineole content in flowers of three lavender species, and developmental stages of *L. x intermedia* inflorescence indicating that the production of 1,8 cineole is likely controlled through transcriptional regulation of *LiCINS*. The genomic CINS of *Lavandula* species had identical exon–intron architecture and coding sequences, but intron analysis suggests that *LiCINS* was most likely inherited from *L. latifolia*.

We also identified the *L. angustifolia* ABCB1 (*LaABCB1*) that mediates the efflux of vinblastine when expressed in *Xenopus laevis* oocytes. Treating oocytes with an oxidative uncoupler, ATPase inhibitor, classical ABC blocker and selected *Lavandula de novo* monoterpenes inhibited efflux mediated by *LaABCB1*. *LaABCB1* transcripts were constitutively expressed in the EO producing tissues of *L. angustifolia*, *L. latifolia* and *L. intermedia* plants. Both the expression pattern and the *in vitro* activity of *LaABCB1* indicate that it plays a role in monoterpene trafficking in the lavender oil glands.
Preface

Chapter 2 was originally published in the Journal of Biological Chemistry. [Zerihun A. Demissie], Lauren A. E. Erland, Mark R. Rheault, Soheil S. Mahmoud (2013) The biosynthetic origin of irregular monoterpenes in Lavandula: isolation and biochemical characterization of a novel cis prenyl diphosphate synthase gene - lavandulyl diphosphate synthase. Journal of Biological Chemistry 288: 6333 – 6341 © the American Society for Biochemistry and Molecular Biology. I identified the candidate, designed and conducted the experiments, wrote the manuscript and followed up the review process (responding to reviewers’ comments). Lauren Erland (now a graduate student under the supervision of Dr Soheil S. Mahmoud at UBC, Okanagan) was an undergraduate student working with me on the project. I conducted this work in Dr. Soheil Mahmoud’s lab under the supervision of Dr. Soheil Mahmoud and Dr. Mark R. Rheault.

Chapter 3 has been published in: [Zerihun A. Demissie], Lukman S. Sarker, Soheil S. Mahmoud (2011). Cloning and functional characterization of β-phellandrene synthase from Lavandula angustifolia. Planta 233: 685 – 696 (Used in this thesis with kind permission from Springer Science and Business Media). I designed the experiments, functionally characterized β-phellandrene synthase, wrote the manuscript and followed up the review process (responding to reviewers’ comments). Lukman S. Sarker cloned LaTPS-I. I conducted this work in Dr. Soheil Mahmoud’s lab under the supervision of Dr. Soheil Mahmoud.

Chapter 4 has been published as [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 Molecular Biology 79: 393 – 411 (Used in this thesis with kind permission from Springer Science and Business Media). I identified the candidate, designed and conducted the experiment, wrote the manuscript and followed up the review process (responding to reviewers’ comments). Monica A. Cella (now a graduate student at UBC, Okanagan school of engineering) was an undergraduate student working with me on the project. Lukamn S. Sarker cloned the genomic DNA, extracted RNA from glandular
trichomes for cDNA library construction and microarray data generation by Plant Biotechnology Institute (PBI) and the University of Toronto, respectively. Figure 4.2 was generated by Lukman S. Sarker and Travis J. Thompson. I conducted this work in Dr. Soheil Mahmoud’s lab under the supervision of Dr. Soheil Mahmoud and Dr. Mark R. Rheault.

Chapter 5 is based on a manuscript prepared for publication as [Zerihun A. Demissie], Soheil S. Mahmoud, Mark R. Rheault. I identified the candidate, tested the *Xenopus laevis* expression system using GFP, designed and conducted the experiment, and wrote the manuscript. The candidate gene used in this study was obtained from Dr. Soheil S. Mahmoud cDNA library and the work was conducted in Dr. Mark R. Rheault’s lab under the supervision of Dr. Mark R. Rheault and Dr. Soheil S. Mahmoud. Dr. Mark R. Rheault’s lab is licensed by the UBC animal care committee to use animals for experiment, application number A12-0270; and I have completed the ethics training of the Canadian Council on Animal Care (CCAC) / National Institutional Animal User Training (NIAUT) certificate number 5737 – 12. In addition, Dr. Mark R. Rheault’s lab has UBC Radioisotope Licence, permit number BIUO-3254-16, and I have completed the Laboratory Safety training.

I also co-authored the following papers during my PhD study:


Table of Contents

Abstract ................................................................. ii
Preface ........................................................................ iii
Table of Contents............................................................... v
List of Tables .................................................................. ix
List of Figures .................................................................. x
Abbreviations .................................................................. xiii
Acknowledgment ............................................................. xvii
Dedication ...................................................................... xix

Chapter 1 Introduction ..................................................... 1
  1.1 The History of Lavender (*Lavandula*) and its uses ................. 1
  1.2 Lavender EO constituents and their application ...................... 2
  1.3 Monoterpene Biosynthesis .................................................. 4
  1.4 Monoterpene Biosynthesis Regulation in *Lavandula* ............... 8
  1.5 Compartmentalization of Monoterpene Biosynthesis and Secretion 10
  1.6 ABC transporters .............................................................. 12
  1.7 Plant ABC transporters ...................................................... 13
  1.8 Research purpose ............................................................ 15

Chapter 2 The biosynthetic origin of irregular monoterpenes in *Lavandula*:
  isolation and biochemical characterization of a novel *cis* prenyl diphosphate
  synthase gene - *lavandulyl diphosphate synthase* .................. 27
  2.1 Overview ........................................................................ 27
  2.2 Materials and methods ..................................................... 30
    2.2.1 *LiLPPS* candidate selection ........................................ 30
    2.2.2 Cloning, protein expression and enrichment of *LiLPPS* .......... 30
    2.2.3 *LiLPPS* product assay and EO constituent identification .......... 31
    2.2.4 Biochemical assay ..................................................... 32
    2.2.5 Phylogenetic relationship analysis ................................... 32
  2.3 Results ............................................................................ 33
    2.3.1 *Lavandula x intermedia* cv Grosso EO composition ............. 33
    2.3.2 PDPS candidate selection and sequence analysis .................. 33
2.3.3 Cloning, expression and functional characterization of LiLPPS ..................34
2.3.4 Substrate specificity and kinetic properties of LiLPPS .............................35
2.3.5 LPP as a substrate for regular mTPSs .........................................................36
2.4 Discussion ........................................................................................................37

Chapter 3 Cloning and functional characterization of β-phellandrene synthase
from Lavandula angustifolia ..................................................................................55
3.1 Overview ...........................................................................................................55
3.2 Materials and methods ......................................................................................58
  3.2.1 Microarray analysis and clone selection .......................................................58
  3.2.2 Expression of recombinant LabPHLS and LaTPS-I in E. coli ......................58
  3.2.3 Enzyme assay ...............................................................................................59
  3.2.4 Product assay/GC–MS analysis .....................................................................60
  3.2.5 Relative expression assay .............................................................................61
3.3 Results and discussion .......................................................................................62
  3.3.1 Candidate selection and cloning .................................................................62
  3.3.2 Sequence analysis .......................................................................................63
  3.3.3 Heterologous expression .............................................................................64
  3.3.4 Functional characterization of LaβPHLS .....................................................64
  3.3.5 Enzymatic assay of LaTPS-I ......................................................................66
  3.3.6 Transcriptional regulation of LaβPHLS and LaTPS-I .................................67
  3.3.7 Phylogenetic analysis ..................................................................................68

Chapter 4 Cloning, functional characterization and genomic organization of
1,8-cineole synthases from Lavandula .....................................................................82
4.1 Overview ...........................................................................................................82
4.2 Materials and methods ......................................................................................85
  4.2.1 Glandular trichome isolation and cDNA library construction .....................85
  4.2.2 Microarray data analysis and candidate selection .......................................85
  4.2.3 Recombinant protein expression ..................................................................86
  4.2.4 Enzyme assay ..............................................................................................87
  4.2.5 Product assay/GC-MS analysis ....................................................................88
  4.2.6 Cloning of CINS cDNAs ..........................................................................88
4.2.7 Obtaining genomic CINS clones ................................................................. 89
4.2.8 Relative expression assay ........................................................................ 89
4.2.9 Phylogenetic analysis ............................................................................... 90
4.3 Results ........................................................................................................... 91
4.3.1 Construction of EST library, transcript profiling and candidate selection .... 91
4.3.2 Functional characterization of LiCINS ....................................................... 91
4.3.3 Cloning of CINS genomic DNA ............................................................... 94
4.3.4 Analysis of transcript levels for CINS by PCR ......................................... 95
4.3.5 Phylogenetic analysis .............................................................................. 95
4.4 Discussion ..................................................................................................... 97
4.4.1 Transcript profiling and EST Selection .................................................... 97
4.4.2 Recombinant production and functional assay ....................................... 98
4.4.3 Structural relation to other plant TPSs .................................................. 100
4.4.4 Phylogenic analysis and genomic organization ....................................... 101
4.4.5 Transcriptional activity and inheritance ................................................ 102

Chapter 5 Identification and functional characterization of Lavandula
angustifolia ABCB1 (LaABCB1) that is differentially inhibited by lavender
de novo monoterpenes in Xenopus laevis oocytes ................................................. 125
5.1 Overview ....................................................................................................... 125
5.2 Materials and methods ................................................................................ 128
5.2.1 Materials .................................................................................................. 128
5.2.2 Candidate selection and sequence analysis ............................................. 128
5.2.3 LaABCB1 transcriptional activity study and cloning .............................. 129
5.2.4 In vitro transcription .............................................................................. 130
5.2.5 Oocyte isolation, injection and expression ............................................ 130
5.2.6 Vinblastine accumulation in oocytes ..................................................... 131
5.2.7 Vinblastine transport inhibition by monoterpenes ................................. 132
5.2.8 Data analysis .......................................................................................... 132
5.3 Results ......................................................................................................... 133
5.3.1 LaABCB1 sequence and phylogenetic analysis ...................................... 133
5.3.2 LaABCB1 serves as an efflux pump ...................................................... 134
5.3.3 LaABCB1 mediated vinblastine efflux was differentially inhibited by monoterpene constituents of *Lavandula* oils .................................................................135
5.4 Discussion..................................................................................................................136

**Chapter 6 Conclusion and Recommendations** ..........................................................155
**References** ................................................................................................................159
**Appendices** ..............................................................................................................183
  Appendix A Vector maps used in this study. .................................................................183
List of Tables

Table 1.1 Major monoterpenes of four Lavandula species EOs (min–max %)..............26
Table 2.1 Oligonucleotides used in this study. .................................................................52
Table 2.2 Synchronized “pyrophosphate synthase” and “diphosphate synthase”
query results for prenyl diphosphate synthase homolog ESTs involved in terpene biosynthesis.................................................................53
Table 3.1 Primer sets for cloning and qRT-PCR reactions.............................................81
Table 4.1 Oligonucleotides used in this study. .................................................................120
Table 4.2 Summary of microarray analysis experiment for relative expression of terpene EO biosynthesis related genes in Lavandula using microarray technology.................................................................121
Table 4.3 Conserved amino acid level (% identity and % similarity) between LiCINS and mTPSs cloned from Lavandula and Salvia ....................122
Table 4.4 Raw data and kinetics parameters of LiCINS..................................................123
Table 4.5 Average 1,8-cineole and linalool content (mg per gm of fresh tissue)
of oils distilled from three developmental stages of L. x intermedia cv Grosso .................................................................124
Table 5.1 Oligonucleotides used this study. .................................................................154
List of Figures

Figure 1.1 Classification of the genus *Lavandula* .......................................................... 19
Figure 1.2 The biosynthesis of IPP and DMAPP ............................................................... 20
Figure 1.3 Schematic representations of biosynthetic pathways leading to regular (left) and irregular (right) monoterpenes in *Lavandula*. ......................................................... 22
Figure 1.4 Structure based classification of common monoterpenes derived from GPP (geranyl diphosphate) in lavenders ................................................................. 24
Figure 1.5 Distribution of glandular trichomes on *L. angustifolia* leaf. ............................ 25
Figure 2.1 Schematic representations of IPP and DMAPP coupling reactions catalyzed by typical “trans” and “cis” PDPS family members ................................................. 40
Figure 2.2 GC chromatogram of *L. x intermedia* cv Grosso essential oil ............................ 41
Figure 2.3 Multiple alignment of LiLPPS with NPPS and cisFPPS of tomato .......................... 42
Figure 2.4 GC chromatogram and mass spectrum of LiLPPS product and authentic lavandulol standard ................................................................. 44
Figure 2.5 GC chromatogram of negative controls ............................................................. 45
Figure 2.6 LPP accumulation vs DMAPP and IPP consumption ....................................... 47
Figure 2.7 Effects of temperature and pH on the activity of LiLPPS ................................... 48
Figure 2.8 Kinetic properties of LiLPPS at increasing DMAPP concentrations ...................... 49
Figure 2.9 Phylogenetic relationship and classification of prenyltransferases ....................... 50
Figure 3.1 *L. angustifolia* young leaf and young flower EO chromatogram ........................ 69
Figure 3.2 Multiple alignment of mTPSs ............................................................................. 71
Figure 3.3 SDS-PAGE analysis of protein samples from bacterial cells expressing LaβPHLS. .............................................................................................................. 73
Figure 3.4 Analysis of LaβPHLS assay products and authentic standard by GC–MS .......... 74
Figure 3.5 Biochemical characterization of LaβPHLS ......................................................... 76
Figure 3.6 Transcriptional activity of *L. angustifolia* terpene synthases .............................. 78
Figure 3.7 Phylogenetic tree of plant TPSs ....................................................................... 80
Figure 4.1 Alignment of *Lavandula* 1,8-cineole synthases with mTPSs of *L. angustifolia*, and 1,8-cineole synthases of Salvia ................................................................. 104
Figure 4.2 Transcriptional activity of MEP-pathway genes in leaves and floral tissues of *L. angustifolia*, *L. x intermedia* and *L. latifolia* measure by PCR ........................... 106
Figure 4.3  SDS-PAGE analysis of protein samples from bacterial cells expressing LiCINS, and those transformed with the empty expression vector.                          107
Figure 4.4  GC chromatograms and mass spectra of products produced by the recombinant LiCINS............................................................................................................. 108
Figure 4.5  GC chromatograms of Ni-NTA affinity chromatography purified soluble fraction from induced cells transformed with empty pET41(b+) product from GPP.       110
Figure 4.6  Kinetic assay of LiCINS............................................................................................................................................................................................................. 111
Figure 4.7  Multiple alignments of 1,8-cineole synthase cDNAs of *L. latifolia*,  
*L. angustifolia* and *L. x intermedia*................................................................................................................................. 113
Figure 4.8  Schematic representation of LiCINS genomic DNA......................................................................................................................... 114
Figure 4.9  The transcriptional activity of *LiCINS, LaβPHLS, LaLINS* and the reference gene (*β-actin*) in lavenders measured by PCR. ................................................................. 115
Figure 4.10 Phylogenetic relationship and classification of TPSs. ................................................................................................................................. 117
Figure 5.1  Schematic representation of HMMTOP predicted LaABCB1 membrane topology model. ..................................................................................................................... 140
Figure 5.2  Multiple alignment of LaABCB1 with selected functionally characterized ATP binding cassette subfamily B members from *Arabidopsis thaliana* and *Coptis japonica*.................................................................................................................. 143
Figure 5.3  Phylogenetic relationship and classification of ABCBs.................................................................................................................................. 144
Figure 5.4  The transcriptional activity of *LaABCB1* and the reference gene (*β-actin*) in *L. angustifolia* and *L. latifolia* leaf and floral tissues, and *L. x intermedia* leaf, floral and secretory cells of glandular trichomes after 30 cycles of PCR. 146
Figure 5.5  LaABCB1 and GFP mRNA resolved on a 1.5 % agarose gel using TBE buffer along with ssRNA ladder (NEB). .................................................................................. 147
Figure 5.6  Olympus FluoView FV1000 laser scanning confocal microscope images of water and GFP mRNA injected *Xenopus laevis* oocytes (10x). .................................................. 148
Figure 5.7  LaABCB1 mediated vinblastine efflux in *Xenopus laevis* oocytes after 60 minutes.......................................................................................................................... 149
Figure 5.8  Effect of oxidative uncoupler (KCN) and ATPase blocker (Vanadate) treatment on the LaABCB1 mediated vinblastine efflux. .............................................. 150
Figure 5.9 LaABCB1 mediated vinblastine efflux in *Xenopus laevis* oocytes over time course. .....................................................................................................................................................151
Figure 5.10 LaABCB1 mediated vinblastine efflux kinetic assay. ..................................................152
Figure 5.11 Inhibition of LaABCB1 mediated vinblastine efflux by *Lavandula de novo* monoterpenes ..................................................................................................................................................153
Figure A.1 Map of pET41b(+) vector (Novagen, EMD chemicals, Germany).................................183
Figure A.2 Schematic representation of pXOON vector map ...........................................................184
Figure A.3 Map of RSET-EmGFP vector (Invitrogen) ....................................................................185
Abbreviations

ABC    ATP-binding cassette
AACT   acetoacetyl-coenzyme A thiolase
ACT    acetyl CoA transferase
BDH    borneol dehydrogenase
BornS  borneol synthase
bp     Base pair
BSA    Bovine serum albumin
CinS   1,8-cineole synthase
cisPDPS cis prenyl pyrophosphate synthases
Contig A set of overlapping DNA segments that together represent a
        consensus region of DNA
cDNA   Complementary DNA
CPP    chrysanthemyl diphosphate
CMK    4-(cytidine-5'-diphospho)- 2-C-methyl-D-erythritol kinase
CPPS   chrysanthemyl diphosphate synthase
cv.    Cultivar
dDPPS  dehydrodolichol diphosphate synthase
DMAPP  dimethylallyl diphosphate
DNA    Deoxyribonucleic acid
dNTPs  Deoxynucleotide triphosphate mix
dTPS(s) diterpene synthase(s)
DTT    Dithiothreitol
DXS    1-deoxy-D-xylulose-5-phosphate synthase
DXR    1-deoxy-D-xylulose-5-phosphate reductoisomerase
EDTA   Ethylenediaminetetraacetic acid
EO(s)  essential oil(s)
EST(s) expressed sequence tag(s)
FPP    farnesyl diphosphate
FPPS(s) farnesyl diphosphate synthase(s)
GC-MS  gas chromatography-mass spectrometry
GGPPS(s)  geranylgeranyl diphosphate synthase(s)
GO      Gene Ontology
GOI     Gene of interest
GPP     geranyl diphosphate
GPPS(s)  geranyl diphosphate synthases(s)
GST     Glutathione S-transferase
HDS     1-hydroxy-2-methyl-E-butenyl-4-diphosphate synthase
HDR     1-hydroxy-2-methyl-E-butenyl-4-diphosphate reductase
HMGS    3-hydroxy-3-methylglutaryl-CoA synthase
HMGR    3-hydroxy-3-methylglutaryl-CoA reductase
IDT     Integrated DNA technologies
IPP     isopentenyl diphosphate
IPPI    IPP isomerase
IPTG    Isopropyl-b-D-thiogalactopyranoside
ISO     International organization for standardization
LaABCB1  L. angustifolia ATP-binding cassette subfamily B1
LaBERS   L. angustifolia bergamotene synthase
LaCinS(s) L. angustifolia 1,8-cineole synthase(s)
LaβPhlS(s) L. angustifolia β-phellandrene synthase(s)
LaLINS(s) linalool synthase(s)
LaLIMS(s) limonene synthase(s)
LaTPS-I  L. angustifolia terpene synthas-like protein
LavS     lavandulol synthase
LiCinS(s) L. x intermedia 1,8-cineole synthase(s)
LiLPPS(s) L. x intermedia lavandulyl diphosphate synthase(s)
LiCinS(s) L. latifolia 1,8-cineole synthase(s)
LinS     linalool synthase
LimS     limonene synthase
LPP      lavandulyl diphosphate
LPPS(s)  lavandulyl diphosphate synthase(s)
NPP      neryl diphosphate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPPS(s)</td>
<td>neryl diphosphate synthase(s)</td>
</tr>
<tr>
<td>MCT</td>
<td>2-C- methyl-D-erythritol-4-phosphate cytidyltransferase;</td>
</tr>
<tr>
<td>MDS</td>
<td>2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase;</td>
</tr>
<tr>
<td>MDR(s)</td>
<td>multidrug resistance protein(s)</td>
</tr>
<tr>
<td>MEP</td>
<td>2-C-methyl-D-erythritol 4-phosphate</td>
</tr>
<tr>
<td>mg/g</td>
<td>mg per gram</td>
</tr>
<tr>
<td>MK</td>
<td>mevalonate kinase;</td>
</tr>
<tr>
<td>MOPSO</td>
<td>3-morpholino-2-hydroxypropanesulfonic acid</td>
</tr>
<tr>
<td>MPK</td>
<td>mevalonate 5-phosphate kinase</td>
</tr>
<tr>
<td>MPDS</td>
<td>mevalonate-5- diphosphate decarboxylase</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>MRP(s)</td>
<td>multidrug resistance associated protein(s)</td>
</tr>
<tr>
<td>mTPS(s)</td>
<td>monoterpenene synthase(s)</td>
</tr>
<tr>
<td>MVA</td>
<td>mevalonate</td>
</tr>
<tr>
<td>Ni-NTA agarose</td>
<td>Nickel-charged nitrilotriacetic acid affinity resin column</td>
</tr>
<tr>
<td>NpPDR1/NpABCG1</td>
<td><em>Nicotiana plumbaginifolia</em> pleiotropic drug protein 1</td>
</tr>
<tr>
<td>NtPDR1</td>
<td><em>Nicotiana tabacum</em> pleiotropic drug protein 1</td>
</tr>
<tr>
<td>NBD(s)</td>
<td>nucleotide-binding domain(s)</td>
</tr>
<tr>
<td>ORF</td>
<td>Open reading frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDR(s)</td>
<td>pleiotropic drug protein(s)</td>
</tr>
<tr>
<td>Pgp(s)</td>
<td>p-glycoprotein(s)</td>
</tr>
<tr>
<td>PhlS</td>
<td>β-phellandrene synthase</td>
</tr>
<tr>
<td>PSH1</td>
<td>tomato β-phellandrene synthase</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethanesulfonyl fluoride</td>
</tr>
<tr>
<td>PTV</td>
<td>Programmable Temperature Vaporizing</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>reverse transcriptase PCR</td>
</tr>
<tr>
<td>RE</td>
<td>Restriction enzyme</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SpTUR2</td>
<td><em>Spirodela polyrrhiza</em> PDR2</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>sTPS(s)</td>
<td>sesquiterpene synthase(s)</td>
</tr>
<tr>
<td>Tm</td>
<td>Melting temperature</td>
</tr>
<tr>
<td>TMD(s)</td>
<td>trans-membrane domain(s)</td>
</tr>
<tr>
<td>TPS(s)</td>
<td>terpene synthase(s)</td>
</tr>
<tr>
<td>transFPPS</td>
<td>farnesyl diphosphate synthase</td>
</tr>
<tr>
<td>transGGPPS</td>
<td>trans geranylgeranyl diphosphate synthase</td>
</tr>
<tr>
<td>transPDPS(s)</td>
<td>trans prenyl pyrophosphate synthase(s)</td>
</tr>
<tr>
<td>U</td>
<td>Unit- amount of the enzyme that catalyzes the conversion of 1 micromole of substrate per minute</td>
</tr>
</tbody>
</table>
Acknowledgment

I am highly indebted to all UBC Okanagan campus faculty, staff and graduate students who have helped me in different ways during my study. I am particularly grateful to my supervisors Dr. Soheil S. Mahmoud and Dr. Mark R. Rheault for entrusting in my abilities and providing me the opportunity to work under their supervision. Their guidance and unreserved support both in academia and life in general has allowed me to flourish and succeed. I also would like to extend my heartfelt gratitude to Dr. Mary Forrest, Dr. Joyce Boon, Dr. Cedric Saucier, Dr. Kirsten Wolthers and Dr. Paul Shipley who served as my advisory committee members. I am very fortunate to get the chance to be mentored by all of you. I particularly would like to thank Dr. Kirsten Wolthers and Dr. Mary Forrest for their technical advice and research inputs used in this study.

I also would like to thank Dr. Soheil S. Mahmoud’s and Dr. Mark R. Rheault’s past and present graduate students whom I worked with during my PhD study for their assistance and friendship. This includes, but not limited to, Dr. Grant Woronuk, Lukman S. Sarker, Carly Hofmann, Lauren A. E. Erland, Mariana Galata, Monica Cella, Travis Thompson, Laura B. Finney, Breanne Cadham, Sierra Padmoroff (from Dr. Mahmoud lab), Brieanne J. Matier, Kimberly Lemky, Lily Kotzev, Matt Glover, Melissa Cruz (from Dr. Rheault’s lab), Carla Meints and Caitlyn Makins (from Dr. Kirsten Wolthers’ lab) for sharing an unbelievable times during my time here in Kelowna. I particularly would like to thank Lukman S. Sarker and Brieanne J. Matier for making my early transition bearable and for being there when I needed them at all times. I also would like to thank Lauren A. E. Erland for taking time to edit my thesis.

I would like to acknowledge the financial assistance through the UBC Okanagan PhD Tuition fee Scholarships, UBC Okanagan International Partial Tuition fee Scholarship, UBC Okanagan University Graduate Fellowship, UBC Okanagan Special Awards and the Pacific Century Graduate Scholarships (PCGS) award from the province of British Columbia through the Ministry of Advanced Education. I would also like to extend my heart felt appreciation to UBC Okanagan campus, Investment Agriculture Foundation of British Columbia, NRC Plant Biotechnology Institute through the NAPGEN program, and Genome British Columbia for supporting my research through grants and/or in-kind contributions to
Dr. Soheil S. Mahmoud, and to Natural Sciences and Engineering Research Council of Canada through grants to Dr. Soheil S. Mahmoud and Dr. Mark R. Rheault.

I finally would like to extend my sincere gratitude to my mom Genete Burka Telila, my lovely sister Mitiqua Teressa, my brothers Tesfahun Meberehatu and Misikir Meberehatu and my nephew Kidus Ephrem for being the source of my inspirations. I roll up my sleeves and get to work everyday thinking about you guys. Special thanks to my beautiful wife Salelsh M. Woldeyes for sticking with me through the ups and downs, and enduring all the lonely moments and late nights during my study. I am lucky to have you in my life, and you deserve all the accolades and the honor that comes with this PhD as much as I do.
Dedication

To my Grandma (Gete Bisa),
I am where I am because of you.
There was never a time I did not think about you, rest in peace!!!
Chapter 1 Introduction

1.1 The History of Lavender (Lavandula) and its uses

The genus Lavandula consists of diverse aromatic flowering shrubs that belong to the family Lamiaceae or the mint plant family. The genus is believed to have originated from the Mediterranean region and is now grown worldwide with Australia, Bulgaria, China, England, France, Italy, Russia, Spain and Ukraine as the leading producers (Adam 2006). Over 39 species are currently recognized in the genus Lavandula, which is further classified into three subgenera known as Lavandula, Fabricia and Sabaudia based on their morphological and molecular similarities. Subgenus Lavandula consists of three sections (Lavandula, Dentatae and Stoechas), subgenus Fabricia contains four sections (Pterostoechas, Subnudae, Chaetostachys and Hasikenenses) and subgenus Sabaudia containing the only section Sabaudia. Each section contains one or more species and any intra-section natural hybrids (Figure 1.1). For example, the section Lavandula contains three species: L. angustifolia, L. latifolia, L. lanata, and four natural intra-section hybrids: L. x intermedia also called lavandin (L. angustifolia subspecies angustifolia x L. latifolia), L. x aurigerana (L. angustifolia subspecies pyrenaica x L. latifolia), L. x losae (L. latifolia x L. lanata) and L. x chaytorae (L. angustifolia subspecies angustifolia x L. lanata) (Upson and Andrews 2004).

The name Lavandula is believed to be derived from the Latin word ‘lavare’ meaning ‘to wash”, which was in reference to its ancient use by Greeks to scent their baths, beds, clothes and even hair. An alternative theory for the origin of the name lavender suggests that it is a derivative of the Latin “livere”, meaning ‘to be livid or bluish’ in reference to the flower color (Lis-Balchin 2002; Upson and Andrews 2004). Lavenders have been cultivated for over 2500 years for their floral essential oil (EO), which possess medicinal properties and have been used in hygiene products for over 2,500 years (Boeckelmann 2008). Lavender EOs have been used as topical ointment to treat and alleviate insect bites, small cuts, burns and inflammatory conditions, and in medications prescribed for treating indigestion and heartburn. Dried flowers, leaves and stems of lavender were included in pillows and sachets by the ancient Greeks to promote sleep and relaxation (Castle and Lis-Balchin 2002; Woronuk et al. 2011). The first written account of lavender use dates back to 65 AD to the Greek physician Dioscorides who wrote about the medicinal value of L. stoechas. Most
written herbal materials from the fifteenth and sixteenth centuries also describe the medicinal and food uses of *L. angustifolia* and *L. latifolia* (Lis-Balchin 2002; Upson and Andrews 2004). Today lavenders are used industrially in the production of cosmetics, perfumes, soaps, disinfectants, antiseptic and anti-inflammatory agents among others. Certain lavender cultivars are also prominent in alternative medicine industry and specialty foods. A few species are also valued as ornamental plants (Bakkali et al. 2008; Brunke et al. 1993; Castle and Lis-Balchin 2002; Festing 1982). The economic and social importance of lavenders was further confirmed when the Herb Growing and Marketing Network of United States of America named lavender as ‘Herb of The Year” in 1999 (Cavanagh and Wilkinson 2002).

Although 39 species are recognized in the genus *Lavandula*, only two species belonging to the section *Lavandula* and their natural hybrid (*L. x intermedia*) are widely cultivated. The preference for *L. angustifolia*, *L. latifolia* and *L. x intermedia* species is due to their distinct commercially and medicinally desirable oil profile. Oils distilled from *L. angustifolia* species are regarded as the finest of all oils and are mainly used in the perfumery and aromatherapy industries, while those of *L. latifolia* are mainly used in the manufacturing of antiseptics, and personal care and medicinal products (Lis-Balchin 2002; Upson and Andrews 2004). Lavandin oils have a characteristics of the oils of *L. angustifolia* and *L. latifolia* and are often used in both industries. However, due to their superior oil yield and similarity to *L. angustifolia* oil, they are predominantly used in the perfume industry (Boeckelmann 2008; Lis-Balchin 2002).

### 1.2 Lavender EO constituents and their application

Monoterpenes, the C<sub>10</sub> class of terpenes, are the major constituents of lavender EOs. Fifty to sixty different monoterpenes can often be identified in a given lavender EO, but the dominant ones are linalool, linalyl acetate, 1,8-cineole, β-ocimene (usually both cis- and trans-), terpinen-4-ol, lavandulol and camphor (Table 1.1) (Cavanagh and Wilkinson 2002; Flores et al. 2005; Harborne and Williams 2002; Kreis and Mosandl 1992). The relative abundance of these constituents is the primary determinant of the characteristic aroma, bioactivity, and market value of a given lavender EO, and varies significantly depending on the developmental stage genotype, and growth environment of the source plant (Boeckelmann 2008). The perfume and cosmetic industry requires high quality lavender EO, as defined by a high percentage of linalool and linalyl acetate and low camphor content.
whereas the non-perfumery applications like medicinal and insect repellency require relatively high camphor and 1,8-cineole content (Cavanagh and Wilkinson 2002). Oils distilled from *L. angustifolia* are characterized by high levels of the monoterpenes linalool (25 – 38%) and linalyl acetate (25 – 45%), and negligible quantities of 1,8-cineole (0 – 1.5%) and camphor (0 – 0.5%), and are considered as the finest oil for the perfumery, alternative medicine and aromatherapy industries (Boeckelmann 2008). *L. latifolia* oils, on the other hand, are characterized by high linalool (28 – 38%), relatively lower linalool acetate (0 – 28% depending on cultivar), and high 1,8-cineole (6 – 11%) and camphor content (7 – 11%). The relatively high level of 1,8-cineole and camphor makes them the preferred oils in antiseptic and antimicrobial products (Munõz-Bertomeu et al. 2007). The EOs of *L. x intermedia* is a blend of *L. angustifolia* and *L. latifolia* components, including linalool (25 – 35%), linalool acetate (25 – 35%), 1,8-cineole (4 – 7%) and camphor (6 – 8%) and are popular in both perfume and antiseptic products (Boeckelmann 2008; Desautels et al. 2009; Lis-Balchin 2002).

Some monoterpenic constituents of lavenders also have medicinally attractive bioactive properties when used in isolation. Linalyl acetate and linalool have a sedative and local anesthetic effect whereas 1,8-cineole acts as a spasmylytic, local anesthetic and antibacterial agent. Camphor, α-terpineol and terpen-4-ol, α- and β-pinene and p-cymene have all demonstrated strong antimicrobial activity (Chu and Kemper 2001; Loutrari et al. 2004; Perrucci et al. 1997; Woronuk et al. 2011). However, the anticancer and antimutagenic properties of some lavender EO constituents, like perillyl alcohols, are the most exciting therapeutic prospects of lavender EOs (Loutrari et al 2004; Woronuk et al. 2011). For example, perillyl alcohol and terpen-4-ol have shown strong antimutagenic properties against TA98 bacterial cells, and development of topological ointments from lavender oils has been suggested as a skin cancer prevention strategy.

There is a great deal of interest in improving the quality and quantity of the EO in lavenders. In particular, enhancing the perfumery quality of *L. x intermedia* species - while maintaining their yield - is a top priority for lavender growers. This could be achieved by altering the relative proportion of monoterpenes in the oil using targeted breeding and/or metabolic engineering tools, which require a full understanding of the monoterpenic biosynthetic machinery. Research presented in this thesis focused on the isolation and
functional characterization of structural genes involved in the biosynthesis and secretion of EO constituents in *Lavandula*. The transcriptional activity of the isolated genes in relation to their product concentration in the oil was also studied to confirm the suitability of these genes in breeding and/or metabolic engineering experiments.

1.3 Monoterpenic Biosynthesis

Monoterpenes are derived from isoprene units, hence their biosynthesis begins with the synthesis of the universal terpene building blocks isopentenyl diphosphate (IPP) and its allylic isomer dimethylallyl diphosphate (DMAPP). Two independent but interactive pathways, called the mevalonate (MVA) or cytosolic and the 2-C-methyl-D-erythritol 4-phosphate (MEP) or plastidial pathways, produce IPP and DMAPP in lavenders and other plants (Arigoni et al. 1997; Bick and Lange 2003; Gershenzon et al. 2000; Laule et al. 2003) (figure 1.2). Though cross talk between the two pathways has been reported, monoterpenes of lavenders and other Lamaiceae family members are predominantly derived from the MEP pathway products (Boeckelmann 2008; Lane et al. 2008; Mahmoud and Croteau 2003). This pathway commences by condensation of pyruvate and D-glyceraldehyde-3-phosphate into 1-deoxy-D-xylulose 5-phosphate (DXP), catalyzed by 1-deoxy-D-xylulose-5-phosphate synthase (DXS). DXP is subsequently transformed into IPP and DMAPP through the sequential action of 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), 2-C-methyl-D-erythritol-4-phosphate cytidylyltransferase (MCT), 4-(cytidine-5'-diphospho)-2-C-methyl-D-erythritol kinase (CMK), 2-C-methyl-D-erythritol-2,4-cyclophosphate synthase (MDS), 1-hydroxy-2-methyl-E-butene-4-diphosphate synthase (HDS) and 1-hydroxy-2-methyl-E-butene-4-diphosphate reductase (HDR) (Phillips et al. 2008). DXS and DXR catalyze the rate limiting steps of MEP pathway, and as such are commonly targeted to alter the supply of monoterpenic precursor molecules (Mahmoud and Croteau, 2001; 2002; 2003).

Synthesis of IPP and DMAPP through the MVA pathway involves the synthesis of mevalonic acid from acetyl-CoA through the sequential activity of acetoacetyl-CoA thiolase, HMG-CoA synthase (HMGS) and HMG-CoA reductase (HMGR), respectively. The synthesis of mevalonic acid is the rate-limiting step, and HMGS and HMGR are the rate-limiting enzymes of this pathway. Mevalonic acid is then phosphorylated and carboxylated to generate IPP through the action of mevalonate kinase, phosphomevalonate kinase and mevalonate diphosphate decarboxylase (Vranová et al. 2013). Figure 1.2 illustrates stages of
the two pathways and their key products in lavender EO. As IPP is relatively unreactive in the downstream processes, it is isomerized to DMAPP by isopentenyl diphosphate isomerase (Vranová et al. 2013).

In the second stage of monoterpene biosynthesis, IPP and DMAPP are condensed in a “head-to-tail fashion” or “non-head-to-tail fashion” to form the C_{10} monoterpene immediate precursors called geranyl diphosphate (GPP), neryl diphosphate (NPP, a cisoid isomer of GPP), lavandulyl diphosphate (LPP) or chrysanthemyl diphosphate (CPP) by a family of enzymes collectively known as prenyl pyrophosphate synthases (PDPSs). While GPP and NPP are derived from head-to-tail (c’1-4) condensation of isoprenes or through chain elongation reaction, LPP and CPP are derived from head-to-middle (c’1-2) condensation or branching reaction and cyclopropanation (c’1-2-3) reaction, respectively (Thulasiram et al. 2008). Lavender EOs are mainly constituted of monoterpenes derived from GPP, a condensation reaction product catalyzed by geranyl diphosphate synthase (GPPS) (Figure 1.3). GPPSs cloned from different species often share high sequence similarity and are distinguished by two aspartate rich conserved motifs, DD_{2-4}D and (N/D)DxxD. These motifs serve as substrate and divalent metal ion cofactor, often Mg^{2+}, binding sites for the carbocation rearrangement mediated condensation reactions (Liang et al. 2002; Walsh 2007). A third conserved motif, the CxxxC motif (where “x” is any hydrophobic residue), is present in heterodimeric GPPSs (Wang and Dixon 2009). GPPSs cloned from Lavandula close relatives Mentha x piperita (Burke et al. 1999) and Salvia miltiorrhiza (Ma et al. 2012) comprise a large and a small subunit, which interact through the CxxxC conserved motif to be catalytically active.

Lavender EOs also contain small quantities of a structurally and functionally unique monoterpene called lavandulol and its derivative lavandulyl acetate (Lis-Balchin 2002; Woronuk et al. 2011). These monoterpenes were identified in pheromones of major insect pests and are subsequently used in artificial pheromone preparations to disrupt the mating behaviors of economically important pests (Franco et al. 2009; Walton et al. 2006). Based on chemical synthesis studies, the immediate precursor for these monoterpenes is believed to be the head-to-middle condensation reaction product LPP. However, a cDNA that encodes for lavandulyl diphosphate synthase (LPPS) has not yet been reported and the biosynthetic origin of these monoterpenes in lavenders was unknown before this study (Figure 1.3). In general,
cDNAs encoding for enzymes involved in non-head-to-tail isoprene condensation are scarce. Chrysanthemyl diphosphate synthase (CPPS) that produces the pyrethrin branch point precursor CPP in *Chrysanthemum cinerariaefolium* (Rivera et al. 2001) and *Artemisia tridentate ssp. spiciformis* (Hemmerlin et al. 2003) are the only enzyme with confirmed non-head-to-tail condensation catalysis roles.

The third stage of monoterpene synthesis involves an ionization dependent transformation of GPP or LPP into their respective monoterpene backbone types. The transformation process is catalyzed by a group of enzymes collectively referred to as monoterpene synthases (mTPSs). For example, Landmann et al (2007) cloned two mTPSs called linalool synthase and limonene synthase, which are responsible for the synthesis of linalool and limonene in *L. angustifolia*, respectively (Figure 1.3). However, several genes that control the biosynthesis of other key EO constituents in commercially important species of *Lavandula* are expected because 1) a large array of structurally diverse monoterpenes are identified in EOs derived from a given lavender plant (Lis-Balchin 2002), and 2) *in silico* analysis of their transcriptional profile shows that they encode a large number of mTPS homologs (Lane et al. 2010).

In recent years, numerous mTPS cDNAs have been cloned and functionally characterized from a wide range of plants (Degenhardt et al. 2009 for detailed review). mTPSs share certain unique features including conserved motifs and a 5’ end transit peptide for localization into plastids. The first conserved motif is the arginine-rich N-terminal RR(x8)W signature motif required for cyclization of GPP (Williams et al. 1998), it may, however, be absent in mTPSs that produce acyclic products (e.g., linalool synthase). The arginine rich motif is often preceded by a transit peptide to direct them to plastids where monoterpene synthesis is concentrated (Bohlmann et al. 1998). The second motif is the highly conserved aspartate-rich DDxxD motif required for divalent metal, usually Mg$^{2+}$, assisted substrate binding and ionization (Christianson 2006; Nieuwenhuizen et al. 2009). The other two are partially conserved amino acid sequences, including LQLYEASFLL (Wise et al. 1998) and (N,D)D(L,I,V)x(S,T)xxxE (Degenhardt et al. 2009; Roeder et al. 2007), that play roles in catalysis and second metal ion binding, respectively.

The final stage of monoterpene biosynthesis involves the decoration of the mTPS output monoterpene’s backbone and is largely responsible for the structural diversity
observed in monoterpenes. The common functional groups attached to monoterpenes backbones are hydroxyl, carbonyl, aldehyde and carboxyl groups. For example, addition of acetyl CoA to the monoterpenes linalool and lavandulol produces linalool/linalyl acetate and lavandulyl acetate, respectively while removal of hydrogen from borneol generates camphor (Sarker et al. 2013) (Figure 1.3).

Over 5,000 structurally different monoterpenes are presently known, of which 50–60 are commonly identified in oils derived from a given lavender species. They are classified based on their structural similarity as acyclic or linear (e.g. linalool, lavandulol, geraniol, etc.), monocyclic (e.g. limonene, β-phellandrene, α-phellandrene, β-pinene, ocimene, etc.) and bicyclic (e.g. 1,8-cineole, camphor, etc.) (Figure 1.4), or based on the functional group they carry such as alcohols, ketones, aldehydes and carboxylic acids. Alternatively, monoterpenes may also be categorized as regular or irregular monoterpenes based on the condensation orientation of their C_{10} precursors. Monoterpenes derived from head-to-tail condensed isoprenes are called regular monoterpenes and are the most common in nature while those derived from non-head-to-tail condensed isoprenes are referred to as irregular monoterpenes and are observed less frequently in nature (Thulasiram et al. 2008). For example, linalool is a monoterpene derived from head-to-tail condensed C_{10} precursor molecule called GPP and classified into the regular monoterpene class, while lavandulol is derived from head-to-middle condensed C_{10} precursor molecule, called LPP, and classified into the irregular monoterpene class (Figure 1.3).

Regular monoterpenes, like linalool, are the most abundant in nature and have significant economic interest. Subsequently, their biosynthetic route is better understood, and cDNAs encoding for GPPSs (Bouvier et al. 2000; Burke and Croteau 2002; Burke et al. 1999; Schmidt and Gershenzon 2008; Tholl et al. 2004) and mTPSs capable of transforming GPP and/or NPP into core monoterpenes have been described from diverse plant species. When the work presented in this thesis started only two cDNAs encoding enzymes catalyzing the synthesis of regular monoterpenes had been cloned and functionally characterized in lavenders (Landmann et al. 2007). At that time cDNAs encoding for enzymes involved in the biosynthesis of lavender irregular monoterpene constituents had not been identified yet. Identifying cDNAs encoding for mTPSs catalyzing the synthesis of other oil constituents – regular and irregular – is important to employ targeted breeding and/or metabolic engineering
to improve oil quality and quantity. In line with this, our group has developed three cDNA libraries to aid the isolation and characterization of cDNAs involved in regular and irregular monoterpene biosynthesis in lavenders (Lane et al. 2010; Sarker et al. 2013).

1.4 Monoterpene Biosynthesis Regulation in *Lavandula*

The biosynthesis of monoterpenes is regulated by physiological, developmental and environmental cues (Gershenzon et al. 2000). For instance, tomato infestation by the fungus *Botrytis cinerea* is accompanied by elevated emission levels of various monoterpenes (Jansen et al. 2009), while conifers deploy their monoterpene rich resin to defend themselves against the bark beetle and other pest attacks (Bohlmann 2012). *Nicotiana attenuata* plants and cotton (*Gossypium hirsutum L.*) seedlings emit volatile blends dominated by the monoterpenes linalool and (E)-β-ocimene in response to herbivore attack (Kessler and Baldwin 2002). Similarly, forest tent caterpillars have been shown to induce the emission of (E)-β-ocimene and linalool, and other terpenes from the leaves of hybrid poplar (Arimura et al. 2004). Monoterpene biosynthesis is also regulated through allelopathic means by monoterpenes and other volatile compounds released from neighbouring plants. Lima bean plants exposed to (E)-β-ocimene emitted from neighbouring transgenic tobacco plants were induced to release monoterpenes and were more successful in deterring two-spotted spider mites (*Tetranychus urticae*) as compared to non-exposed plants (Arimura et al. 2012; Muroi et al. 2011; Sugimoto et al. 2013).

The monoterpene profile of lavenders and related plants is also regulated by the developmental stage of the plant. In peppermint, monoterpane content is low in newly emerging leaves, peaks rapidly between 12 and 20 days post leaf emergence, levels off as full expansion is reached and remains stable for the remainder of the leaf life (Turner et al. 1999). The monoterpene profile and accumulation level in lavender flowers is also determined by the flowering stages of the plant. These flowering stages may be defined as unopened bud stage, 30% of the flowers in bloom, 70% of the flowers in bloom and full bloom or 100% of the flowers in bloom. EO (monoterpene) content in lavender flowers is low in unopened bud stage flowers, reaches its maximum level when 30% of the flowers are in bloom and decreases slightly at later stages. For example, the amount of total oil distilled from *L. x intermedia* cv. Grosso was 11.6 – 13 mg/g flower fresh weight at unopened bud stage, reached 27.2 – 35.5 mg/g flower fresh weight at 30% boom, and decreased slightly to 20.8 –
25.4 mg/g flower fresh weight at full bloom out stage (Boeckelmann 2008; Lane et al. 2010).

In some instances, the effect of environmental and developmental cues on monoterpene accumulation is closely associated with changes in gene expression level of biosynthetic enzymes. In line with this, the abundance of mRNA, protein and enzymatic activity of β-ocimene synthase in snapdragon, S-linalool synthase expression in Clarkia breweri flowers and menthofuran synthase expression in peppermint leaves was highly correlated with the emission levels of the corresponding monoterpenes (Dudareva et al. 1996; 2003; Mahmoud and Croteau 2003; Turner et al. 2000a). These findings suggest that monoterpene production is partly regulated at the level of gene transcription (Dudareva et al. 1996, McConkey et al. 2000). Additionally, suppression of menthofuran synthase expression in peppermint leaves through metabolic engineering reduced the production of menthofuran, an undesired constituent of peppermint EO (Lang et al. 2011; Mahmoud and Croteau 2003).

In Lavandula, the expression of the linalool synthase gene and its product (linalool) concentration in lavender oil increased steadily as the flower matured (Boeckelmann 2008; Lane et al. 2010). In contrast, some monoterpenes are predominantly produced at early stages. For example, the abundances of camphor and its biosynthetic enzymes (e.g., borneol dehydrogenase, which converts borneol to camphor) in lavender reached their maximum levels at early stages of flower development, slightly decreased when flowers were at 30% bloom, and remained constantly low for the remainder of the flower life (Boeckelmann 2008; Sarker et al. 2013).

Transcriptional level regulation of monoterpene biosynthesis is not limited to monoterpene synthases. Changes in the expression of enzymes involved in the MEP pathway can also affect monoterpene synthesis by limiting or increasing the precursor supply level. In peppermint, the monoterpene biosynthesis level and EO profile was significantly affected by modulating the transcriptional activity of DXS, DXR and menthofuran synthase (Mahmoud and Croteau 2001; 2002; 2003; Lang et al. 2011). Similarly, the abundance of DXR, not HMGR, in lavender glandular trichomes was found to correlate well with the synthesis level of monoterpenes indicating the transcriptional level regulation of monoterpene biosynthesis in this plant (Boeckelmann 2008; Lane et al. 2010). Further, over-expression of Arabidopsis thaliana DXS in L. latifolia species through metabolic engineering also enhanced the production of monoterpenes (Muñoz-Bertomeu et al. 2006). These findings suggest that
monoterpene biosynthesis in lavenders is partly regulated at the transcription level of genes encoding for enzymes involved in IPP and DMAPP biosynthesis, those catalyzing the condensation of IPP and/or DMAPP to monoterpene intermediate precursors (GPP and LPP), and monoterpene synthase themselves. It also implies that identifying cDNAs encoding for these enzymes is necessary to employ targeted breeding and/or metabolic engineering tools to improve the quality of lavender EO.

1.5 Compartmentalization of Monoterpene Biosynthesis and Secretion

Cell and tissue specialization as well as cellular compartmentalization are key features of the evolutionary development of eukaryotes that demarcates biosynthetic, catabolic and storage areas. This has enabled eukaryotic cells and tissues to concentrate substrates and catalysts, isolate incompatible substrates or processes, and allowed for the safe accumulation or disposal of potentially harmful products into specialized areas (Sanders and Bethke 2000). When produced at high concentration monoterpenes are among these potentially toxic natural products, (Caissard et al. 2004; Goossens et al. 2003a & b) hence, their biosynthesis, as well as accumulation is localized to specialized structures called glandular trichomes, or oil glands, which are found on floral parts and leaves of lavenders (Gershenzon et al. 2000; McCaskill and Croteau 1995; McConkey et al. 2000) (Figure 1.5a&b). Two types of glandular trichomes are found in lavenders: peltate and capitate. Floral peltate glands are the major sources of monoterpenes, while capitate glands may also produce them in small quantities (Biswas et al. 2009; Lane et al. 2010). Peltate glands contain up to eight secretory cells (disc cells), a stalk cell and a basal cell, which anchor the trichomes in the epidermis, and an epicuticular storage cavity, which is separated by a plasma membrane from the cytoplasm of the secretory cells (Boeckelmann 2008; Fahn 1988; Turner and Croteau 2004) (Figure 1.5c).

The cytosol and leucoplast of the secretory cells are of particular importance with respect to monoterpene biosynthesis and secretion in Lavandula and other Lamiaceae family members. This is because monoterpenes are synthesized in leucoplasts, where the MEP pathway and mTPSs are localized, and then transported to cytosols for downstream transformation and/or secretion (Fahn 1988; Turner and Croteau 2004; Turner et al. 2000; Wagner 1991). This has been shown by Turner et al (1999) who found that both the small and large subunits of GPPS, and limonene synthase of peppermint, and therefore limonene
synthesis, are localized to the leucoplast. Limonene produced in the leucoplasts was then transported to the cytosol for enzymatic transformation to menthol, which involves a series of coordinated function of enzymes residing in three different compartments: limonene-6-hydroxylase (L6OH) in the endoplasmic reticulum (ER), isopiperitenol dehydrogenase (IPD) in the mitochondrial matrix, and the apparently soluble pulegone reductase (PR) in the cytosol. PR catalyzes the final reaction in menthol synthesis in the cytosol, which is then pumped/secreted into the storage cavity of the peppermint glandular trichomes (Turner and Croteau 2004). These findings imply the existence of a highly coordinated intra-cellular trafficking and secretion mechanism for monoterpenes and/or their precursors.

Though the exact mechanisms are not yet known, active transport mechanisms, possibly mediated by ATP-binding cassette (ABC) transporter proteins, are predicted to play a role in accumulation and release of menthol and other monoterpenes. This is based on a trafficking study of menthol in peppermint from its synthesis to storage area that was unidirectional and showed selective retention towards certain monoterpenes (McCaskill et al. 1992; Turner and Croteau 2004), a typical feature of ABC mediated transport. Another study found that approximately 63% of the menthofuran produced in peppermint glands is selectively retained within the glandular secretory cells for further processing (Turner and Croteau 2004). In addition, the transcriptional activities of ABC transporters in N. tabacum, N. plumbaginifolia and the aquaphyte Spirodela polyrhiza have been found to correlate well with isoprenoid synthesis and/or emission. To further support this hypothesis, putative ABC transporter transcripts were found to be well represented in expressed sequence tag (EST) libraries derived from gland secretory cells of Lavandula (Lane et al. 2010; Sarker et al. 2013), peppermint (Lange et al. 2001), and hop (Wang et al. 2008). Hence, current efforts to understand the mechanisms responsible for monoterpene (and isoprenoid) secretion put emphasis on ABC transporter homologs expressed in monoterpene biosynthetic tissues. This may be true of lavenders as both lavenders and peppermints have structurally and functionally similar glandular trichomes (Lane et al. 2010). For example, Lane et al (2010) showed that both DXS and linalool synthase transcripts were predominantly produced in the glandular trichomes (oil glands) of lavenders where the bulk of monoterpenes are produced.
1.6 ABC transporters

ABC transporter proteins, the largest super-family in the ABC protein family, are membrane proteins involved in the transport of mineral ions, lipids, and peptides across membranes (Rea 2007) and regulation of other transporters (Carvajal et al. 1997). ABC transporters can function as importers, which are involved in metabolite influx, and exporters, which are involved in metabolite efflux. Prokaryotes employ both importers and exporters while eukaryotes are mainly enriched with exporters (Davidson et al. 2008). Both influx and efflux functions of ABC transporters are aided by energy derived from hydrolysis of ATP, a key feature distinguishing ABC transporters from other transport mechanisms. The efflux activity of ABC transporters is particularly involved in drug and antibiotic tolerance of human cancerous and bacterial cells, respectively, and functional defect of ABC transporters causes certain diseases like cystic fibrosis and Tangier disease (Dean et al. 2001).

Functional ABC transporters are made up of four domains: two transmembrane domains (TMDs) and two nucleotide-binding domains (NBDs). These four domains are usually coded as a single polypeptide in eukaryotes and referred to as ‘full–length transporters’, but also as separate polypeptides consisting of two domains and referred to as ‘half transporters’, that perform their function as a complex. When similar domains (TMD-TMD or NBD-NBD) are encoded together they are called homodimers, while a TMD and an NBD encoded as one unit are called heterodimers (Higgins 2001; Hollenstein et al. 2007). The TMD segments of ABC transporters span the membrane and share low sequence similarity, enabling ABC transporters to create a gateway for diverse substrates. In contrast, NBDs (each about 200 amino acid long), generally share high sequence similarity and retain three conserved motifs: Walker A box (GxxGxGK[S/T]), where x is any amino acid) and Walker B box [øøøøDE, where ø is any hydrophobic residue] and an ABC signature (or alias C or linker peptide, commonly referred to as LSGGQ) that is situated upstream of Walker B box (Davidson et al. 2008; Kos and Ford 2009; Rea 2007). Biochemical studies and structural predictions of NBDs implicated the role of these motifs in ATP binding and hydrolysis (Kos and Ford 2009) to generate motional energy to aide in the translocation of the TMD bound substrate.
1.7 Plant ABC transporters

Ubiquitous presence of ABC transporters in humans, animals and prokaryotes is well established (Davidson et al. 2008; Higgins 2001; Hollenstein et al. 2007; Kos and Ford 2009). Recent studies have revealed that plants are also rich in ABC proteins, the majority of which are membrane bound. The genome of the model plant *Arabidopsis thaliana* encodes more than 120 ABC proteins compared to about 52 in human (Dean et al. 2001) and 50–70 in *Escherichia coli* depending on strain (Linton and Higgins 1998). Of the total ABC proteins identified in *Arabidopsis*, at least 103 likely encode membrane proteins (Garcia et al. 2004; Sanchez-Fernandez et al. 2001; Verrier et al. 2008). The abundance of ABC transporters in plant genomes is believed to stem from their role in intracellular trafficking of structurally unrelated metabolites including chlorophyll catabolites (Lu et al. 1998; Tommasini et al. 1998), the phytohormone auxin (Geisler and Murphy 2006), lipids (Buda et al. 2013; Chen et al. 2011; Hu et al. 2009), virulence resistance factors (Shang et al. 2009), and other functions like transcriptional regulation of other ABC transporters (Carvajal et al. 1997) and ion guard channel control (Klein et al. 2003). Nonetheless, their primary role is maintaining the cellular homeostasis by actively pumping potentially harmful endogenous and exogenous compounds into specialized structures like the epicuticular cavity of glandular trichomes and vacuoles (Rea 2007).

Plant ABC transporters are assigned to nine subfamilies referred to as ABCA through ABCI based on their orientation (forward, TMD–NBD arrangement or reverse, NBD–TMD arrangement), the presence or absence of idiotypic transmembrane and/or linker domains and overall sequence similarity to prototype subfamily members from humans and yeasts (Verrier et al. 2008). Of the nine plant ABC transporter subfamilies, however, only the full–length members of the ABCB, ABCC and ABCG subfamilies are of particular interest in terms of secondary metabolite transport. These three subfamily members include both “half” and “full–length” transporters and are distinguished from each other based mainly on their TMD-NBD arrangements. Both ABCB and ABCC members have a forward orientation, but an additional ABCC-specific N-terminal transmembrane characterizes the later and ABCG members have a reverse orientation (Klein et al. 2006). The full–length members of ABCB (also called multidrug resistance proteins, MDRs in short, or p-glycoproteins) and ABCC (also known as multidrug resistance associated proteins abbreviated as MRPs) subfamilies

13
are encoded by all eukaryotes while ABCGs, also called pleiotropic drug proteins (PDRs), are encoded only in plants and fungi (Rae 2007; Kang et al. 2011).

The first plant ABC transporter identified was the ATP-dependent glutathione S-conjugate (GS-X) export/efflux pump involved in intracellular sequestration of glutathionylated xenobiotics into vacuoles (Martinoia et al. 1993). Lu et al. (1997) cloned the cDNA encoding for the *Arabidopsis thaliana* multiple drug resistance associated protein1, abbreviated as AtMRP1/ABCC1, and determined that it functions as a GS-X pump.

Following the identification of AtMRP1, several plant ABC transporter genes have been studied for their role in specialized metabolite transport including the *Coptis japonica* MDRs, also known as CjABCB1/CjMDR1 (Shitani et al. 2003) and CjABCB2/CjMDR2 (Shitani et al. 2013), and their homologs from *Arabidopsis thaliana* (Geisler et al. 2003; 2005; Jasinski et al. 2001), maize (*Zea mays; brachytic2*) and sorghum (Geisler et al. 2003; 2005). The expressions CjMDR1 and CjMDR2 were localized to the alkaloid-accumulating rhizomes and *in vitro* transcribed CjMDR1 mRNA expressed in *Xenopus laevis* oocytes and CjMDR2 expressed in yeast were found to mediate the alkaloid berberine influx (Sakai et al. 2002; Shitan et al. 2003; 2013). The first plant PDR to be cloned and studied *in vitro* was SpTUR2 from *Spirodela polyrrhiza* (van den Brûle et al. 2002). The isolation of SpTUR2 led to the identification of its homologs from different plants including NpPDR1/NpABCG1 from *Nicotiana plumbaginifolia* and NtPDRs from *Nicotiana tabacum* trichomes (Crouzet et al. 2013).

*In vitro* functionally characterized plant PDRs (SpTUR2, NpPDR1 and NtPDR1) were identified based on the observation that, 1) they were the major ABC transporters expressed in tissues specialized for terpene biosynthesis and release (Bienert et al. 2012; Bultreys et al. 2009; Ruocco et al. 2011), and 2) their transcriptional activity was found to correlate well with terpene synthesis and release (Smart and Fleming 1996; Jasinski et al. 2001; Sasabe et al. 2002; Stukkens et al. 2005). In addition, silencing NpPDR1 and NtPDR1 increased host susceptibility to some diterpenes like sclareol (Bienert et al. 2012; Bultreys et al. 2009; Ruocco et al. 2011; van den Brule et al. 2002; Crouzet et al. 2013), while heterologous expression of SpTUR2 in *A. thaliana* conferred resistance to antimicrobial compounds including the antifungal diterpene sclareol (Jasinski et al. 2001) and kanamaycin (Mentewab & Stewart 2005). Therefore, PDR homologs expressed in tissues specialized for
terpene (and monoterpane) biosynthesis and release have been considered the de facto transporter candidates. Interestingly, although NtPDR1, a PDR extensively studied in this regard, was isolated from tissues known to accumulate and release monoterpenes in response to fungal attack, its over-expression in *N. tabacum* BY2 cells did not confer resistance against tobacco de novo terpenes like the monoterpane 1,8-cineole and the diterpene geranylgeraniol (Crouzet et al. 2013). In fact, the only PDR that conferred resistance to monoterpenes *in vitro* so far was isolated from the mountain pine beetle-fungal symbiont *Grosmannia clavigera* (Wang et al. 2013).

Although plant MDRs are predominantly known for their role in cellular auxin transport (Geisler et al. 2003; 2005; Geisler and Murphy 2006) and influx of the isoquinoline alkaloid berberine (Sakai et al. 2002; Shitan et al. 2003; 2013), recent competitive inhibition studies showed that they have potential to transport monoterpenes. The *in vitro* accumulation of [³H]digoxin mediated by the human MDR1 was reversed by *Zanthoxyl* Fructus ethyl acetate extracts and its dominant monoterpane constituents. Of the 25 monoterpenes tested, (R)-(+)citronellal, (S)-(−)β-citronellol, δ-terpinene, terpinolene and (−)-β-pinene inhibited digoxin transport (Yoshida et al. 2005; 2006; 2008). This selective inhibition of MDRs by monoterpenes implies that MDRs expressed in tissues specialized for monoterpane synthesis and accumulation could play a role in their transport. This also holds true for MDRs expressed in glandular trichomes of *Lavandula*.

1.8 Research purpose

Despite the economic importance of lavender EOs, which consist primarily of monoterpenes, very little is known about their biosynthesis and secretion. First, the biosynthetic pathways leading to irregular monoterpenes in lavenders has not been investigated. The irregular monoterpenes found in lavenders, lavandulol and its ester derivative lavandulyl acetate, have been found to be pheromones of some major insect pests such as *Planococcus ficus* and are subsequently used in artificial pheromone preparations to disrupt the mating behavior of these economically important pests (Franco et al. 2009; Walton et al. 2006). Therefore, further investigation and elucidation of this pathway may allow for optimization of these processes, which is of significant economic value.

Second, although 50–60 structurally diverse monoterpenes have been identified in EOs derived from any given lavender plant, only three mTPSs catalyzing the synthesis of
linalool, limonene and camphor have been identified so far (Landmann et al. 2007; Sarker et al. 2013). Identifying mTPSs catalyzing the synthesis of other oil constituents is a prerequisite to improve oil quality and quantity through targeted breeding and/or metabolic engineering. For example, limonene synthase isolated from *L. angustifolia* species was used to modify the production and composition of EO of transgenic lavandin (Tsuru and Assada 2013).

Third, lavender EO accumulation is the cumulative effect of the amount synthesized and amount translocated from its synthesis area to storage cavity (Fahn 1988; Gershenzon et al. 2000; McCaskill and Croteau 1995; McConkey et al. 2000). However, very little is known about the secretion mechanism. Identifying cDNAs encoding for proteins involved in the secretion process and studying them *in vitro* is the first step to understand this mechanism and will allow for the design of effective strategies to improve lavender EO quality and quantity.

Although environment and flower development stage plays a role, monoterpene biosynthesis is mainly regulated at the level of transcription. Numerous studies have shown correlation between biosynthetic gene transcript abundance and monoterpene accumulation level (Boeckelmann 2008; Bohlmann et al. 1997; Dudareva et al. 1996, 2003; Mahmoud and Croteau 2003). Furthermore, transcriptional profile databases have been successfully used to isolate and functionally characterize cDNAs involved in monoterpene biosynthesis from different plants including *Lavandula* (Sarker et al. 2013). The main objective of the work presented in this thesis was to address the above research gaps by cloning and functionally characterizing genes responsible for the production and secretion of regular and irregular monoterpenes, and thus I hypothesized:

1) Irregular monoterpenes of lavenders are derived from the head-to-middle condensation of IPP and/or DMAPP via the C\(_{10}\) precursor LPP, that a PDPS enzyme (i.e., LPPS) catalyzes its synthesis; and that like other monoterpenes related biosynthetic genes, *LPPS* would be highly expressed in secretory cells of *Lavandula peltate* glandular trichomes.

To examine this hypothesis, I predicted that an EST database derived from a *Lavandula* secretory cell cDNA library would contain PDPSs capable of catalyzing both head-to-tail and non-head-to-tail reactions. As sequence similarity among PDPSs favors functional similarity,
I anticipated that the two types of PDPSs would share low sequence similarity but follow similar developmental and tissue specific regulation pattern that is consistent with monoterpane biosynthesis. Therefore, I screened our library for a unique PDPS homolog that shared low sequence homology with PDPSs known to catalyze head-to-tail condensation reactions (for example GPPS) to identify an LPPS candidate. The selected cDNA was cloned into the pET41b+ cloning/expression vector (Figure A.1), the recombinant protein was expressed in *E. coli*, purified using Ni-NTA affinity chromatography, and assayed for LPP synthetic activity using IPP and/or DMAPP as a substrate.

2) β–Phellandrene and 1,8-cineole are synthesized from GPP through the function of β–phellandrene (βPHLS) and 1,8-cineole (CINS) synthases. βPHLS is highly expressed in *L. angustifolia cv*. Lady leaves where its product is the major monoterpane. 1,8-Cineole is the major lavender EO constituents and CINS would be highly expressed in *Lavandula* secretory cells. Heterologously produced βPHLS and CINS would transform GPP to β–phellandrene and 1,8-cineole, respectively *in vitro*.

To test the above two hypotheses, putative β–phellandrene and 1,8-cineole synthase candidates fulfilling the above outlined candidate selection criteria were obtained from our cDNA libraries and cloned into pET41b+ cloning/expression vector (Figure A.1). The recombinant protein expressed in *E. coli* was enriched using the Ni-NTA affinity chromatography and assayed *in vitro* for β–phellandrene and 1,8-cineole synthesis catalytic activity, respectively from GPP and its cisoid isomer NPP.

3) ABC transporters expressed in secretory cells of glandular trichomes are involved in monoterpane trafficking and secretion of EO constituents, and can be functionally expressed in *Xenopus laevis* oocytes.

Lavenders, like many other plants (Rea 2007; Kang et al. 2011), encode a large number of ABC transporter homologs. Therefore, we limited our search to full–length PDR and MDR homologs expressed in glandular trichome secretory cells to identify a possible candidate. An MDR homolog that fulfilled the candidate selection criteria was obtained from our cDNA library and cloned into the eukaryotic expression vector pXOON (Figure A.2) (Jespersen et al. 2002) and used to produce mRNA *in vitro* The *in vitro* transcribed mRNA was then injected into *Xenopus laevis* oocytes for expression and transport assays. The transport activity mediated by the candidate ABC transporter was determined by using the
MDR substrate [³H]vinblastine sulphate. The possible role of the ABC transporter in mediating monoterpenes transport *in vitro* was derived from inhibition transport assays using commercial standards of selected lavender monoterpenes constituents, linalool and 1,8-cineole.

Results generated in this thesis will enhance our understanding of monoterpenes biosynthesis and secretion in lavenders, and may also be used to improve the quality and quantity of lavender EOs. They may also be used for mass production of these compounds using synthetic biology. To date, β-phellandrene synthase cDNA reported in this thesis has already been used in “a proof of concept” study by researchers from University of California, Davis. The codon optimized β-phellandrene synthase cDNA was expressed in cyanobacteria to demonstrate that monoterpenes based renewable biofuel can be produced through a “photosynthesis-to-fuel” approach (Bentley et al. 2013).
Figure 1.1 Classification of the genus *Lavandula*.
Subgenera names are given in oval shape and their respective sections names are given in a rectangle. Only species belonging to the section *Lavandula* are given because they are the focus of this thesis.
Figure 1.2 The biosynthesis of IPP and DMAPP.
Figure 1.2 The biosynthesis of IPP and DMAPP via the mevalonate (MVA) pathway (left) and the 2-C- methylerythritol-4-phosphate (MEP) pathway (right). Arrows indicate reactions catalyzed by AACT, acetoacetyl-coenzyme A thiolase; CMK, 4-(cytidine-5'-diphospho)- 2-C-methyl-D-erythritol kinase; DMAPP, dimethylallyl diphosphate; DXS, 1-deoxy-D- xylulose-5-phosphate synthase; DXR, 1- deoxy-D-xylulose-5-phosphate reductoisomerase; HDS, 1-hydroxy-2-methyl-E-butenyl-4-diphosphate synthase; HDR, 1-hydroxy-2-methyl-E- butenyl-4-diphosphate reductase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGCR, 3-hydroxy-3-methylglutaryl-CoA reductase; IPP, isopentenyldiphosphate; IPPI, IPP isomerase; MCT, 2-C- methyl-D-erythritol-4-phosphate cytidyltransferase; MDS, 2-C- methyl-D-erythritol-2,4-cyclodiphosphate synthase; MK, mevalonate kinase; MPK, mevalonate 5-phosphate kinase and MPDS, mevalonate-5- diphosphate decarboxylase. Adapted from Lane et al (2010) with permission from Planta.
Figure 1.3 Schematic representations of biosynthetic pathways leading to regular (left) and irregular (right) monoterpenes in *Lavandula.*
Figure 1.3 Schematic representations of biosynthetic pathways leading to regular (left) and irregular (right) monoterpenes in Lavandula. Arrows indicate reactions catalyzed by ACT, acetyl CoA transferase; BDH, borneol dehydrogenase; BornS, borneol synthase; CinS, 1,8-cineole synthase; DMAPP, dimethylallyl diphosphate; GPP, geranyl diphosphate; GPPS, geranyl diphosphate synthase; IPP, isopentenyl diphosphate; MEP, 2-C- methylerythritol-4-phosphate; MVA, mevalonate; LinS, linalool synthase; LimS, limonene synthase; LPP, GPPS: lavandulyl diphosphate; LPPS, lavandulyl diphosphate synthase; LavS, lavandulol synthase and PhlS, β-phellandrene synthase. * denote enzymes identified and characterized in this study, ** denotes enzyme unidentified as of yet and carbons of IPP and/or DMAPP joined during condensation reaction are circled.
Figure 1.4 Structure based classification of common monoterpenes derived from GPP (geranyl diphosphate) in lavenders.
Figure 1.5 Distribution of glandular trichomes on *L. angustifolia* leaf.

a) Leaf and glandular trichomes viewed under light microscopy (*scale bar* ~0.25 cm); b) Scanning electron microscopy image of glandular trichomes (*scale bar* ~50 µm) (from Lane et al. 2010) with permission from Planta; c) Schematic representation of peltate glandular trichomes (from Turner and Croteau 2004 with permission from American Society of Plant Biologists).
Table 1.1 Major monoterpenes of four Lavandula species EOs (min–max %).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Linalool</td>
<td>25 – 38</td>
<td>28 – 38</td>
<td>25 – 35</td>
<td>–</td>
</tr>
<tr>
<td>Linalyl acetate</td>
<td>25 – 45</td>
<td>20 – 28</td>
<td>28 – 38</td>
<td>–</td>
</tr>
<tr>
<td>Camphor</td>
<td>0 – 0.5</td>
<td>7 – 11</td>
<td>6 – 8</td>
<td>–</td>
</tr>
<tr>
<td>1,8-cineole</td>
<td>0 – 1.5</td>
<td>6 – 11</td>
<td>4 – 7</td>
<td>–</td>
</tr>
<tr>
<td>Terpinene-4-ol</td>
<td>2 – 6</td>
<td>0 – 1</td>
<td>2 – 4</td>
<td>–</td>
</tr>
<tr>
<td>β-Phellandrene</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>12 – 32</td>
</tr>
<tr>
<td>α-Phellandrene</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>6 – 16</td>
</tr>
<tr>
<td>Lavandulol and</td>
<td>1.0 – 14</td>
<td>0.5 – 0.8</td>
<td>0.2 – 1.5</td>
<td>–</td>
</tr>
<tr>
<td>Lavandulyl acetate</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Limonene</td>
<td>0.2 – 0.4</td>
<td>0.9 – 1.5</td>
<td>1.0 – 2.2</td>
<td>–</td>
</tr>
</tbody>
</table>
Chapter 2 The biosynthetic origin of irregular monoterpenes in *Lavandula*: isolation and biochemical characterization of a novel *cis* prenyl diphosphate synthase gene - lavandulyl diphosphate synthase

2.1 Overview

Monoterpenes, the C\textsubscript{10} class of the isoprenoids, are derived from the universal terpene building blocks IPP and DMAPP predominantly synthesized through the MEP metabolic pathway in plants (Croteau et al. 2000). *En route* to monoterpane synthesis IPP and DMAPP, or two DMAPP units, can be condensed to form structurally diverse branch point C\textsubscript{10} precursor molecules (Walsh 2007). The diversity of these molecules arises from the capability of PDPSs to condense isoprene in head–to–tail or non–head–to–tail orientations (Thulasiram et al. 2007; 2008), and outcomes with “*cis*” or “*trans*” geometric configurations (Liang et al. 2002). For example, the head–to–tail coupling of two isoprene units to GPP or its *cisoid* isomer NPP is catalyzed by GPPS and neryl diphosphate synthase (NPPS), respectively. The non–head–to–tail condensation of isoprene units to LPP and CPP, on the other hand, is catalyzed by LPPS and CPPS, respectively (Figure 1.3 & 2.1). The final stage of monoterpane synthesis – transforming and elaborating the different precursor molecules to monoterpenes – is catalyzed by another group of enzymes called mTPSs.

The head–to–tail condensation of IPP and DMAPP, and the enzymes catalyzing these reactions, are the most common in nature. Thus, monoterpenes derived from GPP and NPP like linalool, 1,8-cineole, limonene, and so forth, are widely distributed and are referred to as “regular monoterpenes”. Subsequently, their biosynthetic route is well established, and cDNAs encoding for GPPS, (Bouvier et al. 2000; Burke et al. 1999; 2002; Schmidt and Gershenzon 2008; Tholl et al. 2004) NPPS (Schilmiller et al. 2009) and mTPSs capable of transforming GPP and/or NPP into core monoterpenes have been described from diverse plant species (Chen et al. 2011; Degenhardt et al. 2009). GPPS, a *trans*PDPS, and NPPS, a *cis*PDPS, elongate the linear prenyl chain by coupling DMAPP with IPP through a chain elongation (c1ˊ- 4) reaction to generate GPP and NPP, respectively (Thulasiram et al. 2008). Despite catalyzing the same reaction, however, GPPSs and NPPS are more distantly related to one another than they are to PDPSs of similar geometric outcomes. GPPSs share high sequence similarity with *trans*PDPSs catalyzing the higher order terpene precursors and are distinguished by two aspartate rich conserved motifs, DD\textsubscript{2-4}D and (N/D)DxxD. These
motifs serve as a substrate and divalent metal ion cofactor, often Mg$^{2+}$, binding sites for the carbocation rearrangement mediated condensation reactions (Liang et al. 2002; Walsh 2007). A third conserved motif, the CxXXC motif (where “x” is any hydrophobic residue), is present in heterodimeric PDPSs (Wang et al. 2009). Heterodimeric PDPSs, like the GPPSs cloned from *Mentha x piperita* (Burke et al. 1999), *Salvia miltiorrhiza* (Ma et al. 2012), etc., are enzymes constituted of two subunits (large and small) which interact through the CxXXC conserved motif to be catalytically active. NPPS and other cisPDPSs, on the other hand, do not necessarily retain the above motifs but share five conserved regions designated as Region I – V. In particular the aspartate residue in Region IV and the glutamate residue of Region V are catalytically essential (Kharel and Koyama 2003; Sallaud et al. 2009).

PDPSs catalyzing non–head–to–tail coupling reactions, and monoterpenes derived from LPP and CPP such as lavandulol and pyrethrins, respectively are encountered less frequently in nature (Dewick 2009; Epstein et al. 1991; Rivera et al. 2001). Irregular monoterpenes, like pyrethrins, are the major ingredients in leading botanical and EPA certified insecticides (Matsuda et al. 2005; Matsuda 2012). Lavandulol and its ester derivative lavandulyl acetate were identified in pheromones of major insect pests and are subsequently used in artificial pheromone preparations to disrupt the mating behaviors of economically important pests (Franco et al. 2009; Walton et al. 2006). Yet, little is known about the biosynthetic pathways leading to these monoterpenes and their derivatives. Only two cDNAs encoding for CPPSs, both transPDPS family members, have been isolated and functionally characterized from *Chrysanthemum cinerariaefolium* (Rivera et al. 2001) and *Artemisia tridentate* ssp. *Spiciformis* (Hemmerlin et al. 2003) so far. CPPSs predominantly catalyze the cyclopropanation (c1´-2-3) reaction in which the non–head–to–tail condensation of two DMAPP molecules generates the pyrethrin branch point intermediate, CPP. In addition to their major product, the *A. tridentate* CPPS and its chimeric derivatives were found to catalyze the non–head–to–tail coupling of two DMAPP molecules through branching (c1´-2) to generate LPP (Thulasiram et al. 2007; 2008). LPP is the branch point precursor of monoterpenes with head–to–middle condensed PDP backbones, such as lavandulol and lavandulyl acetate (Epstein et al. 1991). It is also the source of the lavandulyl side group of sophoraflavanone G in *Sophora flavescens* Ait (Zhao et al. 2003) that determines their anti-tumour (Ko et al. 2000) and phospholipase-Cγ1-inhibition properties.
(Lee et al. 2012). To our knowledge a wild type LPPS gene has not been described from plants or other organisms.

EOs of the genus Lavandula are primarily constituted of a few “regular” monoterpenes and their derivatives. For example, the economically important EOs derived from L. angustifolia and L. x intermedia species, contain large amounts of linalool, linalool acetate and 1,8-cineole. Consequently, these oils are industrially utilized as ingredients of various cosmetic and antiseptic products (Warnock et al. 2011). The biosynthetic pathway leading to these monoterpenes from IPP and DMAPP have been defined both experimentally (Demissie et al. 2011; 2012; Landmann et al. 2007) and through in silico analysis of Lavandula EST databases (Demissie et al. 2011; 2012; Lane et al. 2010). In this regard, four cDNAs encoding for mTPSs that transform GPP and NPP into linalool, 1,8-cineole, limonene and β-phellandrene in vitro have been cloned from L. angustifolia and L. x intermedia species (Demissie et al. 2011; 2012; Landmann et al. 2007). However, lavender EOs – particularly L. x intermedia species – also contain the irregular monoterpane lavandulol and its ester derivative lavandulyl acetate (Lis-Balchin 2002) whose biosynthetic origin has not yet been investigated. Here, we report the isolation and functional characterization of a novel cisPDPS cDNA encoding for LPPS, an enzyme that condenses two DMAPP molecules to generate the lavandulol branch point precursor LPP in vitro, from a L. x intermedia oil gland cDNA library.
2.2 Materials and methods

2.2.1 LiLPPS candidate selection

Our group recently reported the construction of a cDNA library and its corresponding annotated EST database from *L. x intermedia cv* Grosso secretory cells of oil glands, tissues specialized for EO biosynthesis and secretion (Demissie et al. 2012). All PDPS homologs in the database were retrieved by searching the EST database using the strings “diphosphate synthase” and “pyrophosphate synthase”. We then synchronized the search results and manually excluded EST homologs known to be involved in regular terpene biosynthesis including GPPS, *trans* farnesyl diphosphate synthase (*transFPPS*) and *trans* geranylgeranyl diphosphate synthase (*transGGPPS*), and PDPS homologs that are not involved in terpene biosynthesis. This led us to acquire a novel *cis*PDPS homolog contig that was later determined to be *L. x intermedia* lavandulyl diphosphate synthase (*LiLPPS*). The transcriptional expression pattern of this contig and previously described mTPSs throughout *L. x intermedia cv* Grosso flower developmental stages was assessed by microarray analysis using the Agilent oligo-based microarray technology, through services provided by the University Health Network Microarray Centre (Toronto, Canada). The three floral developmental stages were as follows: unopened buds, anthesis and mature flowers in which 30% of the buds were in bloom (for photographic descriptions refer to Lane et al. 2010). After validating the microarray data using standard PCR, we selected the contig for further detailed analysis.

2.2.2 Cloning, protein expression and enrichment of LiLPPS

Glandular trichome secretory cells were isolated from mature flowers of *L. x intermedia cv* Grosso plants grown at the University of British Columbia, Okanagan campus lavender field following a previously described modified glass bead abrasion method (Demissie et al. 2012). Total RNA was extracted from 100 mg of the tissue using an RNA extraction kit (OMEGA bio-tek, USA), and reverse transcribed in a reaction containing the oligo d(T) primer (Fisher Scientific, Canada) and *M-MuLV* Reverse Transcriptase enzyme (New England Biolabs, USA) following the manufacturer’s directions. The putative *LiLPPS* ORF was amplified with the full length cloning primer sets (Table 2.1) and *iProof™* High-
Fidelity DNA Polymerase (Bio-Rad, USA). The PCR program used was 95 °C for 5 min, followed by 37 cycles of 95 °C for 1 min, 58 °C for 30 sec and 72 °C for 1 min, and a 5 min final extension at 72 °C. The amplified fragments were cloned into the NdeI/EcoRI sites of the pET41b(+) expression vector (Figure A.1) and expressed in *E. coli* BL21(DE3) strain (EMD Chemicals, Darmstadt, Germany) following previously described procedures (Demissie et al. 2011; 2012). Except for the bind buffer that was slightly modified (0.5 M NaCl, 50 mM KH₂PO₄, pH 8.0) the procedure described in the aforementioned papers was followed to enrich the expressed protein. Sequence information for *LiLPPS* was deposited in the NCBI public databank with the accession number JX985358.

2.2.3 *LiLPPS* product assay and EO constituent identification

Initial *in vitro* enzyme activity assays were performed in a 500 µl reaction volume containing the assay buffer (50 mM Tris/HCl, 5% glycerol, 1 mM MnCl₂, 1 mM MgCl₂, pH 8.0), 1 mM DTT, 40 µM IPP and 40 µM DMAPP (Echelon, Salt Lake City, UT, USA), and 2.5 - 50 µg LiLPPS. After two hour incubation at 30 °C, the reaction mix was heated at >80 °C for 10 min and kept on ice for 2 min. Then 30 units of calf intestinal alkaline phosphatase (CIP; New England Biolabs, USA) were added to the reaction mix, overlaid by 500 µl of pentane and incubated overnight at 32.5 °C to hydrolyze the prenyl product and unreacted substrates. GPP and NPP standards (Echelon, Salt Lake City, UT, USA) were also hydrolyzed as a control of the hydrolysis reaction. The reaction was stopped by vigorous vortexing followed by flash freezing in liquid nitrogen, and stored in a -80 °C freezer until analyzed. Assay product identification and quantification, and EO extraction and quantification were performed by gas chromatography-mass spectrometry (GC-MS) following procedures described previously (Demissie et al. 2011; 2012; Falk et al. 2009). Purified protein extract of *E. coli* BL21(DE3) strain transformed with empty pET41b(+) expression vector (Figure A.1), i.e. without insert, were also assayed under the same conditions as a negative control.
2.2.4 Biochemical assay

Biochemical characterization of LiLPPS was performed following the procedure described above with slight modifications: the reaction volume was reduced to 125 µl, 2.5 µg LiLPPS was used and assays were incubated at 37 °C for 1 hour after adding CIP. Six temperature levels (25, 27.5, 30, 32.5, 35 and 37.5 °C) were tested to determine LiLPPS optimum temperature, and the optimum pH was determined using MES and MOPS buffers at pH 5.5, 6.0, 6.5, 7.0, 7.5, 8.0 and 8.5. The kinetic properties of LiLPPS were deduced from LPP accumulation data in assays with increasing DMAPP concentrations (6.25, 12.5, 25, 50, 100, 150, 200, 300, 450 and 600 µM) at the optimal conditions. The GraphPad Prism software version 5.0d (GraphPad Software, Inc., La Jolla, CA) was used to calculate the \( V_{\text{max}} \), \( K_m \) and the degree of cooperativity (h) values and fit the data into a sigmoidal substrate concentration dependent enzyme response curve using the equation

\[
V_o = \frac{V_{\text{max}} [S]^h}{K_m + [S]^h} \quad \text{equation 2.1}
\]

where \( V_o \) is initial reaction rate, \( V_{\text{max}} \) is maximum reaction rate, \( K_m \) is the substrate concentration at which the reaction rate is half of \( V_{\text{max}} \), [S] is substrate concentration and h is the degree of cooperativity or hill coefficient. Substrate specificity of the enzyme was determined from the type and amount of reactants consumed versus product accumulated in assays containing a constant IPP concentration (150 µM) combined with various DMAPP levels (6.25 - 200 µM) under the optimized conditions. We also tested LPP as a potential substrate for \( L. x \) intermedia 1,8-cineole synthase (LiCINS), an enzyme that primarily produced 1,8-cineole from both GPP and NPP (Demissie et al. 2012).

2.2.5 Phylogenetic relationship analysis

LiLPPS was aligned with cisPDPSs isolated from tomato using the default parameters of the ClustalW alignment tool available at the EBI platform (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The phylogenetic relation of LiLPPS with other PDPSs was constructed using the Jukes-Cantor genetic distance model and UPGMA tree building method of Geneious Tree Builder module (Geneious 5.0.3 software, Auckland, New Zealand).
2.3 Results

2.3.1 *Lavandula x intermedia cv* Grosso EO composition

GC-MS analysis of EO distilled from *L. x intermedia cv* Grosso floral tissues profiled the following terpenes in their respective abundance order: linalool, linalool acetate, 1,8-cineole, camphor, iso-borneol, lavandulyl acetate, α-terpineol, α-bisabolol, α-cadinol, β-ocimen, β-caryophyllene, lavandulol, limonene, δ-carene, myrcene, geraniol, nerol, neryl-acetate and a few other minor products (Figure 2.2). Except for the irregular monoterpenes lavandulol and lavandulyl acetate, which are derived from LPP, the remaining EO mono- and sesquiterpene constituents arise from the corresponding linear precursors GPP and *trans*FPP, respectively.

2.3.2 PDPS candidate selection and sequence analysis

Searching our EST library using the string “pyrophosphate synthase” as a query identified 11 different prenyltrasferase homolog unigenes while the string “diphosphate synthase” identified 31. Duplicate search results and prenyltrasferases with no known involvement in terpene biosynthesis (e.g., ribose-phosphate pyrophosphokinase and cytidine diphosphate diacylglycerol synthase) were disregarded to consolidate the search results. The final synchronized search result identified 25 unigenes that were homologous to known PDPSs involved in terpene biosynthesis (Table 2.2). From these sequences 12 were homologous to PDPSs involved in the MEP pathway, eight were homologous to *trans*PDPSs (GPPSs, FPPSs, GGPPSs) and one was homologous to bornyl diphosphate synthase. The remaining four candidates were homologous to *cis*PDPSs. Two of these homologs corresponded to dehydrodolichol diphosphate synthase (dDPPS) genes, which catalyzes the prenyl chain elongation reaction to produce the polyprenyl backbone of dolichol. The other unigenes, one of which belonged to a contig with 18 EST members and the other was a singleton, exhibited significant homology to NPPS [FJ797956] and *cis*FPPS [ACJ38408] from cultivated (*Solanum lycopersicum*) (Schilmiller et al. 2009) and wild tomato (*Solanum habrochaites*) (Sallaud et al. 2009), respectively. The singleton was later determined to be a splice variant of the contig and was disregarded.
The transPDPS and dDPPS homologs were anticipated to be present in a database derived from *L. x intermedia* secretory cells because they catalyze reactions that are consistent with EO composition in this tissue. However, given the fact that GPP is the preferred substrate for regular monoterpenes (Croteau and Karp 1979; Schilmiller et al. 2009) and typical sesquiterpenes synthesized from cisFPP (Sallaud et al. 2009) were not detected in the *L. x intermedia* EO (Figure 2.2), the identification of NPPS and cisFPPS homologs were unforeseen. In addition, the results of our transcript profiling experiment indicated that the transcriptional expression of this EST paralleled those of other lavender mTPSs (Demissie et al. 2011; 2012; Lane et al. 2010), and was developmentally regulated in *L. x intermedia* flowers (data not shown). Therefore, we decided to further investigate this contig which was later determined to be *LiLPPS*. The ORF of LiLPPS was expressed in *E. coli* BL21(DE3) cells for further analysis.

The ORF of *LiLPPS* is 918 nucleotides long encoding for a 305 amino acid protein that includes a short N-terminal plastidial transit peptide (30 amino acids), and shares 66.4% and 66% sequence similarity with NPPS and cisFPPS of tomato, respectively. Like many other cisPDPSs, LiLPPS lacks the two aspartate rich transPDPS signature motifs DDx2–4D and (N/D)DxxD, but maintains all five semi-conserved regions (I – V) that define the cisPDPSs (Figure 2.3). In particular the catalytically important aspartate and glutamate residues of Region IV and V, were conserved in LiLPPS at 201 and 276 positions from the N-terminal, respectively (Figure 2.3). Multiple alignment of LiLPPS and its splice variant, both at nucleotide and amino acid level, revealed that the splice variant lacks a stretch of 22 amino acids in the middle of its ORF that included a portion of the conserved region III (Figure 2.3). Repeated attempts to amplify the splice variant from reverse transcribed total mRNA isolated from *L. x intermedia* glandular secretory cells failed possibly due to the low abundance of the corresponding cDNA. It is also possible that the splice variant singleton was an artefact that resulted from sequence assembly errors during database construction. Therefore, only the full-length contig was considered for further investigation.

2.3.3 Cloning, expression and functional characterization of LiLPPS

The ca. 34.5 kDa recombinant LiLPPS was successfully expressed in bacterial cells and enriched using the Ni-NTA agarose affinity chromatography system. Incubation of the
purified recombinant protein with IPP and DMAPP followed by the hydrolysis of the assay mix by CIP, resulted in the production of lavandulol as the only product (Figure 2.4a). The identity of the product was determined by comparing its mass spectrum and retention time to those of an authentic lavandulol standard (Sigma, Canada) (Figure 2.4b). Given that the alkaline CIP-mediated hydrolysis of LPP is known to yield lavandulol (Epstein et al. 1991; Thulasiram et al. 2007; 2008), we concluded that LiLPPS catalyzes the synthesis of LPP from isoprene units. In this respect, the hydrolysis of GPP and NPP with CIP under the same conditions produced geraniol and nerol, respectively, as anticipated. Hydrolysis of the control assay mixes contained only the prenyl alcohol derivatives of IPP (buten-1-ol, <3-methyl-3) and DMAPP (buten-1-ol, <3-methyl-2) (Figure 2.5a&b).

2.3.4 Substrate specificity and kinetic properties of LiLPPS

When LiLPPS was incubated with IPP alone, detectable amounts of lavandulol or any other terpene prenyl alcohols were not identified after the alkaline hydrolysis step. We also observed that assay reactions containing 5 μM of each of IPP and DMAPP, or 5 μM DMAPP alone produced equivalent amounts of LPP. Furthermore, lowering the concentration of DMAPP in the assay resulted in a parallel reduction in LPP production, while reducing the concentration or excluding IPP from the assay had no effect. When LiLPPS was incubated with constant IPP concentration (150 μM) combined with an increasing DMAPP amount (6.25 – 200 μM), the amount of LPP accumulated correlated (R² = 0.99; P < 0.0001) with the amount of DMAPP consumed while the IPP consumption remained close to zero at all combinations (Figure 2.6).

The optimum temperature and pH of LiLPPS were found to be 30 °C and 8.0 (Figure 2.7a &b), respectively. Unlike other PDPSs cloned from related plant species, the substrate concentration dependent saturation curve of LiLPPS did not follow the standard Michaelis-Menten kinetics. Instead the amount of LPP accumulated in response to increasing DMAPP concentration levels (6.25 – 600 μM) fitted a sigmoidal saturation curve (Figure 2.8), typical of allosterically related enzymes. The equation (equation 2.1) was used to calculate the $K_m$ and $V_{max}$ of LiLPPS by non-linear least square analysis, which were determined to be 208 ± 12 μM, and 448 ± 22 pmole/min, respectively. The hill coefficient (h) value of LiLPPS was
2.7 ± 0.3 while the $k_{cat} (V_{max}/[E])$ and catalytic efficiency ($k_{cat}/K_m$) of the enzyme were calculated as 0.1 s$^{-1}$ and 5 x $10^{-10}$ µM$^{-1}$ s$^{-1}$, respectively.

2.3.5 LPP as a substrate for regular mTPSs

mTPSs involved in the biosynthesis of regular monoterpenes in lavenders accept both GPP and NPP as *in vitro* substrates. We thus examined the ability of the recombinant LiCINS to utilize LPP as a substrate in standard assay reactions. As anticipated, LiCINS, which produced 1,8-cineole as a major product upon incubation with GPP and NPP (Demissie et al. 2012), did not produce detectable quantities of a product from LPP.
2.4 Discussion

Like those of many other plants, the EOs of lavenders are dominated by regular monoterpenes synthesized through the head–to–tail condensation of isoprene units. It was therefore not surprising that the EST database derived from *L. x intermedia* secretory cells – tissues specialized for EO biosynthesis – contained a substantial number of ESTs corresponding to enzymes involved in the biosynthesis of regular monoterpenes (Table 2.2). Given that *L. x intermedia* plants also accumulate irregular monoterpenes, particularly the LPP derivatives lavandulol and lavandulyl acetate (Epstein et al. 1991) (Figure 2.2), it was also not surprising to find ESTs related to irregular monoterpene biosynthesis – including the cloned LiLPPS – in our database. However, it was somewhat unexpected to find a cisPDPS in *L. x intermedia* glandular trichomes because: 1) all known Lamiaceae PDPSs (Bouvier et al. 2000; Burke et al. 1999; 2002; Schmidt and Gershenzon 2008; Tholl et al. 2004), including those of *Lavandula* (Demissie et al. 2011; 2012; Landmann et al. 2007), are of the transPDPS type. 2) The only reported PDPSs with confirmed role in irregular monoterpene biosynthesis, CPPSs (Thulasiram et al. 2007; 2008; Rivera et al. 2001), also belong to the transPDPSs family. The previously reported cisPDPSs, NPPS and cisFPPS, catalyzed the biosynthesis of regular (and not irregular) mono- and sesquiterpene precursors, respectively, in tomato glandular trichomes.

The “cis” and “trans” PDPS family members often share higher sequence similarity with each other irrespective of the genetic relatedness among donor organisms (Kharel and Koyama 2003; Liang et al. 2002). This is contrary to some other enzymes involved in terpene biosynthesis, particularly terpene synthases (TPSs). Typically, TPSs of a given species are more related to one another than to those of distantly related species, even if they catalyze the same reaction. For example β-phellandrene synthase of *L. angustifolia* shares 48 and 47% identity with linalool and limonene synthases of the same species, respectively. However, the gene only shares 31 and 25% similarity with β-phellandrene synthases of grand fir and tomato, respectively (Demissie et al. 2011). In this study, we observed that *L. x intermedia* transPDPS homologs (GPPS, FPPS and GGPPS) share a higher sequence similarity among each other and with other transPDPSs cloned from genetically unrelated species than they do with LiLPPS. LiLPPS is closely related to NPPS and cisFPPS of cultivated (*Solanum lycopersicum*) (Schilmiller et al. 2009) and wild (*Solanum habrochaites*) (Sallaud et al. 2009).
tomato, respectively. The phylogenetic tree presented in figure 2.9 clearly suggests that
LiLPPS together with NPPS and cisFPPS diverged very early from transPDPSs.

LiLPPS catalyzes the head–to–middle condensation of two DMAPP molecules to
synthesize the linear lavandulol branch point precursor, LPP (C₁₀). As demonstrated by
Thulasiram et al (2007; 2008), LPP synthesis proceeds via rearrangement of the double bond
position to create the highly reactive carbocation intermediate lavandulyl cation (L⁺). L⁺ will
eventually be transformed to LPP following proton loss. LiLPPS catalyzes these reactions
with an apparent $K_m$ and $k_{cat}$ of 208 ± 12 µM and 0.1 s⁻¹, respectively. These catalytic
properties are within the range of previously reported cisPDPSs including NPPS that has a
$K_m$ and $k_{cat}$ of 177 µM and 0.2 s⁻¹, respectively (Schilmiller et al. 2009). The $K_m$ and $k_{cat}$
values of the aromatic prenyltrasferase cloned from fungi were 325 µM and 0.03 s⁻¹,
respectively (Haug-Schifferdecker et al. 2010). Unlike other PDPSs, LiLPPS displayed
a sigmoidal saturation curve (Figure 2.8). This is a typical feature of allosterically related
enzymes, in which the binding of the first substrate affects the affinity of the enzyme for the
second substrate through conformational changes or stabilization of the active pocket
environment (Birrell et al. 2003). LiLPPS, like any other typical PDPS enzyme, has binding
sites for a divalent metal ion cofactor and two substrates (Kharel and Koyama 2003; Liang et
al. 2002). PDPSs accepting IPP and DMAPP (e.g., GPPS or NPPS) can be saturated with one
of the substrates and forced to follow Michaelis-Menten kinetics typical of single substrate
enzymes for the other substrate. Since DMAPP was the only substrate for LiLPPS, saturating
one site was not an option. Thus, plotting LiLPPS product (LPP) accumulation against
increasing substrate (DMAPP) concentrations assumed a sigmoidal saturation curve with a
hill coefficient (h) value of 2.7. The positive hill coefficient value (2.7) indicates a positive
cooperativity among the binding pockets.

If LPP was a condensation result of IPP and DMAPP, an equimolar consumption of
the two isoprenes would be expected in our assays. However, LPP production was
independent of IPP, and required only DMAPP. In our assays, a reduction in DMAPP supply
resulted in an equivalent reduction in LPP accumulation. However, reducing the
concentration or excluding IPP from the reaction did not alter LPP production. This result
was also confirmed when IPP was provided at a constant high concentration level (150 µM)
while the amount of DMAPP was increased progressively (6.25 - 200 µM) and the reaction
was allowed to run until LPP synthesis seized or DMAPP molecules were nearly consumed. At all IPP and DMAPP concentration combinations assayed, IPP consumption remained very close to zero, while that of DMAPP increased in parallel to the amount of LPP synthesized (Figure 2.6). This outcome indicated that, like the CPPS enzymes reported by Thulasiram et al (2007) and Rivera et al (2001), LiLPPS utilizes DMAPP as the only substrate to synthesize LPP. The structural features of LiLPPS underlying this catalytic property or residues involved in DMAPP recognition and c1’-c2 coupling reaction catalysis are yet to be determined. Our results, however, suggest that the active sites of LiLPPS selectively binds two DMAPP (and not IPP) units and positions them in such a way that the first carbon atom of one unit is in close proximity to the second carbon of the other to facilitate the c1’-c2 bond formation. One possibility is that LiLPPS preferentially recognizes DMAPP by identifying its double bond position. It is also possible that the position of the double bond in IPP is not favorable for head–to–middle condensation.

With over 55,000 members, isoprenoids are the most structurally and stereochemically diverse biochemical compounds known to mankind (Christianson 2008). Much of this diversity has been attributed to the astounding mechanistic heterogeneity of TPSs. PDPSs – “cis” or “trans” – also play a major role in the structural diversity of isoprenoids by providing intermediate precursor molecules of varying chain length (C₁₀, C₁₅, C₂₀, C₃₀, etc) destined for different isoprenoid groups. In addition, PDPSs generate the structurally distinct C₁₀ precursor molecules GPP/NPP, LPP, CPP, maconelliyl diphosphate (C₁₀; MPP) or planococcyl diphosphate (C₁₀; PPP) by simply changing the position of the carbon-carbon bond (Thulasiram et al. 2007; 2008; Rivera et al. 2001). These linear precursors are then elaborated upon by mTPSs to create various monoterpenes (Christianson 2008; Gao et al. 2012). In Lavandula, GPP and LPP are the linear precursors for the biosynthesis of regular and irregular monoterpenes, respectively (Demissie et al. 2011; 2012; Landmann et al. 2007).
Figure 2.1 Schematic representations of IPP and DMAPP coupling reactions catalyzed by typical “trans” and “cis” PDPS family members. The newly identified enzyme and the reaction it catalyzed are circled. CPP: chrysanthemyl diphosphate; CPPS: chrysanthemyl diphosphate synthase; DMAPP: dimethylallyl diphosphate; IPP: isopentenyl diphosphate; GPP: geranyl diphosphate; GPPS: geranyl diphosphate synthase; LPP: lavandulyl diphosphate; LiLPPS: L. x intermedia lavandulyl diphosphate synthase; LinS: linalool synthase; NPP: neryl diphosphate and NPPS: neryl diphosphate synthase.
Figure 2.2 GC chromatogram of *L. x intermedia* cv Grosso essential oil. Peaks correspond to 1) 1,8-Cineole, 2) ß-Ocimen, 3) Camphor, 4) Linalool, 5) Linalool acetate, 6) Terpinen-4-ol, 7) Lavandulyl acetate 8) Menthol – internal standard, 9) iso-Borneol, 10) ß-Bisabolol, 11) Geraniol, 12) α-Cadinol, 13) α-Bisabolol, and *) lavandulol.
Figure 2.3 Multiple alignment of LiLPPS with NPPS and cisFPPS of tomato.
Figure 2.3 Multiple alignment of LiLPPS with NPPS and cisFPPS of tomato. Bars indicate the five conserved Regions (I-V), and the aspartate and glutamate residues in Region IV and V, respectively are boxed. Identical amino acid residues are represented by asterisks, conserved amino acid substitutions are represented by a semicolon while semi-conserved amino acid substitutions are represented by a period. LiLPPS [JX985358]: *L. x intermedia* lavandulyl diphosphate synthase; zFPS_SOLHA [B8XA40.1]: *z,z*-farnesyl diphosphate synthase from (*Solanum habrochaites*) and NPPS [NP_001234633.1]: neryl diphosphate synthase from (*Solanum lycopersicum*).
Figure 2.4 GC chromatogram and mass spectrum of LiLPPS product and authentic lavandulol standard.
(a) product from DMAPP after calf intestinal alkaline phosphatase hydrolysis and (b) authentic lavandulol standard. Peaks correspond to: 1) IPP prenyl alcohol, 2) DMAPP prenyl alcohol, 3) lavandulol.
Figure 2.5 GC chromatogram of negative controls.
Figure 2.5 GC chromatogram of negative controls. GC chromatogram of assay products when IPP and DMAPP were incubated overnight with (a) soluble protein extracted from E. coli BL21(DE3) strains transformed with the empty pET41(b+) expression vector (Figure A.1), and (b) in assay buffers without protein. Peaks correspond to: 1) IPP prenyl alcohol and 2) DMAPP prenyl alcohol.
Figure 2.6 LPP accumulation vs DMAPP and IPP consumption.

\( R^2 = 0.99, \ P < 0.0001, \ n = 2 \) and error bars indicate standard deviation.
Figure 2.7 Effects of temperature and pH on the activity of LiLPPS.
(a) Effect of temperature (b) Effect of pH (n = 3; Error bars indicate standard deviation).
Figure 2.8 Kinetic properties of LiLPPS at increasing DMAPP concentrations.

The non-linear regression equation used to fit LPP accumulation against DMAPP concentrations was 

\[ V = V_{\text{max}} \frac{[S]^b}{K_{\text{m}}^b+[S]^b} \]

\( n = 3 \) and error bars indicate standard deviation.

\( V_{\text{max}} = 448 \pm 22 \text{ pmol/min} \)
\( K_{\text{m}} = 207.86 \pm 12.33 \mu M \)
Figure 2.9 Phylogenetic relationship and classification of prenyltransferases.
Figure 2.9 Phylogenetic relationship and classification of prenyltransferases.

Prenyltransferases within the same class share a minimum of 50% amino acid identity. The \textit{trans}PDPS family members are boxed in broken line, \textit{cis}PDPS family members are circled in solid line and the prenyltransferase described in this paper is boxed with solid line. The scale bar represents amino acid substitutions per site and numbers represent the branch support values in percent. Accession numbers of prenyltransferases used to generate the phylogenetic tree are: AsCPPS: AY308478.1; AtdDPPS1: NP_565551.1; AtdDPPS4: NP_568883.1; AtuDPPS: AAM67372.1; \textit{cis}SlFPPS: B8XA40.1; \textit{cis}SlPDPS6: AFW98430.1; LiLPPS: JX985358; lsu-MpGPPS: AAF08793.1; lsu-SmGPPS: AEZ55681.1; MpFPPS: AAK63847.1; MtdDPPS: XP_003615624.1; PkGPPS: AAW66658.1; SmFPPS: ABV08819.1; SmGPPS: ACR19637.1; SlGPPS: NP_001234302.1; SINPps: NP_001234633.1; ssu-AmGPPS: AAS82859.1; ssu-LiGPPS: JX985359; ssu-MpGPPS: AAF08792.1; ssu-SmGPPS: AEZ55678.1 and TcCPPS: HQ235057.1
Table 2.1 Oligonucleotides used in this study.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiLPPS set I</td>
<td>F: 5′–CACTCATATGGCATTCCTGCAGTTACC-3’</td>
</tr>
<tr>
<td></td>
<td>R1: 5′–CTCTTCATGAATTCTTTTTCGTTCAACCG-3’</td>
</tr>
<tr>
<td>LiLPPS set II</td>
<td>F: 5′–CACTCATATGGCATTCCTGCAGTTACC-3’</td>
</tr>
<tr>
<td></td>
<td>R2: 5′–AATTCCTCTTCATGAATTCTTTTGACCCGG-3’</td>
</tr>
</tbody>
</table>
Table 2.2 Synchronized “pyrophosphate synthase” and “diphosphate synthase” query results for prenyl diphosphate synthase homolog ESTs involved in terpene biosynthesis.

<table>
<thead>
<tr>
<th>Sequence ID</th>
<th>E-value</th>
<th>Homology description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL15Contig3</td>
<td>0</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase EC(s): 1.17.4.3</td>
</tr>
<tr>
<td>CL15Contig2</td>
<td>0</td>
<td>4-Hydroxy-3-methylbut-2-en-1-yl diphosphate synthase EC(s): 1.17.4.3</td>
</tr>
<tr>
<td>CL15Contig4</td>
<td>0</td>
<td>4-Hydroxy-3-methylbut-2-en-1-yl diphosphate synthase EC(s): 1.17.4.3</td>
</tr>
<tr>
<td>CL15Contig5</td>
<td>5e-153</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>CL2Contig6</td>
<td>2e-90</td>
<td>(+)-Bornyl diphosphate synthase EC(s): 5.5.1.8</td>
</tr>
<tr>
<td>CL15Contig1</td>
<td>7e-49</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>CL39Contig1</td>
<td>1e-28</td>
<td>Dehydrodolichyl diphosphate synthase 2 EC(s): 2.5.1</td>
</tr>
<tr>
<td>LINFO1NG-RP-013_K24_6MAY2010_086</td>
<td>2e-09</td>
<td>E,E-Geranylgeranyl pyrophosphate synthase 3</td>
</tr>
<tr>
<td>LINFO2NG-RP-008_D17_07MAY2010_077</td>
<td>2e-94</td>
<td>E,E-Farnesyl pyrophosphate synthase EC(s): 2.5.1.10</td>
</tr>
<tr>
<td>Code</td>
<td>Accession</td>
<td>Description</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>LINFO1NG-RP-008_P10_21APR2010_033</td>
<td>2e-14</td>
<td>E,E-Geranyl pyrophosphate synthase small subunit</td>
</tr>
<tr>
<td>LINFO2NG-RP-012_E15_07MAY2010_060</td>
<td>1e-46</td>
<td>E,E-Geranyl<a href="E">geranyl pyrophosphate synthase</a>: 2.5.1.30</td>
</tr>
<tr>
<td>LINFO1NG-RP-013_L03_6MAY2010_006</td>
<td>3e-15</td>
<td>Dehydrodolichyl diphosphate synthase 2 (E): 2.5.1</td>
</tr>
<tr>
<td>LINFO2NG-RP-010_L06_07MAY2010_021</td>
<td>1e-107</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>LINFO1NG-RP-008_O14_21APR2010_049</td>
<td>6e-43</td>
<td>(E)-4-Hydroxy-3-methylbut-2-enyl diphosphate synthase (E): 1.17.4.3</td>
</tr>
<tr>
<td>LINFO1NG-RP-013_L23_6MAY2010_086</td>
<td>5e-83</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>LINFO1NG_RPC_103_101_20APR2010_007</td>
<td>2e-20</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>LINFO2NG-RP-014_J11_07MAY2010_040</td>
<td>1e-59</td>
<td>E,E-Farnesyl diphosphate synthase 2</td>
</tr>
<tr>
<td>LINFO1NG_RPC_006_J21_21APR2010_087</td>
<td>3e-40</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>LINFO2NG_UP_1001_L06_18FEB2010_021</td>
<td>1e-32</td>
<td>4-Hydroxy-3-methylbut-2-enyl diphosphate synthase</td>
</tr>
<tr>
<td>LINFO1NG_RPC_005_F08_20APR2010_028</td>
<td>4e-43</td>
<td>Z,Z-Neryl diphosphate synthase 1 (E): 2.5.1.28</td>
</tr>
<tr>
<td>LINFO2NG_RPC_002_F16_21APR2010_060</td>
<td>2e-37</td>
<td>Copalyl diphosphate synthase</td>
</tr>
</tbody>
</table>

Rows shaded in grey contain the ‘cis prenyl diphosphate synthase’ homologous candidates identified in the library.
Chapter 3 Cloning and functional characterization of β-phellandrene synthase from *Lavandula angustifolia*

3.1 Overview

Lavender (*Lavandula*) EOs, commonly used in personal care products, household and industrial cleaners, and traditional and modern medicines, are mainly composed of monoterpenes, the C10 class of the isoprenoids or terpenoids. Though 50–60 monoterpenes have been identified in lavenders, typically a few predominate in the oil and determine the characteristic scent of the species in which they occur (Table 1.1). For example, the EO from *L. angustifolia* is dominated by linalool and linalool acetate, and the EO from *L. latifolia* contains linalool, camphor and 1,8-cineole as major constituents. The EO of *L. pinnata* is dominated by α- and β-phellandrene, and the EO of *L. intermedia* (a hybrid plant resulting from a cross between *L. angustifolia* and *L. latifolia*) contains high quantities of linalool, linalool acetate, camphor and 1,8-cineole (Cavanagh and Wilkinson 2002; Figueiredo et al. 1995; Flores et al. 2005; Harborne and Williams 2002; Kreis and Mosandl 1992; Lis-Balchin 2002). Although EO composition is determined primarily by plant genotype, environmental cues (e.g., temperature and length of the day) and agricultural practices (e.g., irrigation and fertilization) can also influence oil profile in these plants (Boira and Blanquer 1998; Muníoz-Bertome et al. 2007).

Lavender essential oil is produced in specialized glandular trichomes found on the surface of leaves and flowers, although oil yield and composition are different between the two tissues. For example, oil yield in *L. angustifolia* plants grown in British Columbia, Canada, ranged 0.7–2.9 mg/g fresh weight in leaves, and 13.9–15.3 mg/g fresh weight in flower. Further, floral oil was dominated by linalool and linalool acetate, while leaf oil contained primarily phellandrenes and a few minor products such as camphor and borneol (Falk et al. 2009).

Monoterpane biosynthesis begins with the condensation of IPP and DMAPP to form the basic intermediary product GPP (Figure 1.2). mTPSs enzymes then convert GPP to respective monoterpenes (Figure 1.3). To date, mTPS cDNAs have been cloned and functionally characterized from a wide range of plants including linalool, limonene and bergamotene synthases from *L. angustifolia* (Landmann et al. 2007); betaocimene and 1,8-
cineole synthases from *Citrus unshiu* (Shimada et al. 2005), 1,8-cineole synthase from *Nicotiana suaveolens* (Roeder et al. 2007); linalool synthase from *Clarkia breweri* (Pichersky et al. 1995), *Artemisia anua* (Jia et al. 1999) and *Mentha* (Crowell et al. 2002); camphene, β-phellandrene, terpinolene and α-pinene synthases from grand fir (Bohlmann et al. 1999) and many others (for detail review of Terpene synthases consult Degenhardt et al. 2009). In some instances, monoterpane production has been directly correlated with the transcriptional activity of the corresponding mTPS. For example, linalool accumulation in *L. angustifolia* directly correlates with transcript levels for linalool synthase gene (Lane et al. 2010). Similarly, in peppermint the production of menthofuran (an undesired constituent of peppermint EO) is primarily regulated at the level of transcription (Mahmoud and Croteau 2003).

TPSs cloned from different plant species display distinct phylogenetic relationship, and have been classified into six subfamilies assigned TPSa through TPSf. Those producing primary metabolites are grouped in subfamilies TPSc and TPSe while TPSs synthesizing secondary metabolites are assigned to the subfamilies TPSa, TPSb, TPSd and TPSf. Angiosperm sesquiterpene synthases (sTPSs) are classified into TPSa subfamily while their monoterpane counterparts, including mTPSs of Lamiaceae, are classified into subfamily TPSb. mTPSs of gymnosperms are categorized under subfamily TPSd, and TPSf containing only linalool synthase of *Clarkia breweri* stands alone (Bohlmann et al. 1998). Despite the intriguing differences among the subfamilies, TPSs share several conserved motifs. The first is the mTPSs signature arginine-rich N-terminal RR(x8)W motif (Williams et al. 1998) required for cyclization of GPP. The second is the highly conserved aspartate-rich DDxxD motif required for divalent metal, usually Mg$^{2+}$, assisted substrate binding (Nieuwenhuizen et al. 2009). The other two are partially conserved amino acid sequences, including LQLYEASFLL (Wise et al. 1998) and (N,D)D(L,I,V)x(S,T)xxxE (Degenhardt et al. 2009; Roeder et al. 2007), that play roles in catalysis and second metal ion binding, respectively. In mTPSs the arginine rich motif required for cyclization of GPP is preceded by a transit peptide, another characteristic feature of mTPSs, to direct them to plastids where monoterpane synthesis is concentrated (Bohlmann et al. 1998).
Despite the economic importance of lavender EOs, to date only three cDNAs, encoding two mTPSs, linalool synthase (LaLINS) and limonene synthase (LaLIMS), and one sTPS, bergamotene synthase (LaBERS), have been cloned and functionally characterized in these plants (Landmann et al. 2007). Further, regulation of expression of the essential oil biosynthetic genes is poorly understood. We have recently developed substantive genomics resources, including leaf and flower cDNA libraries, to facilitate gene discovery in lavenders and to gain insight into the regulation of essential oil metabolism in these plants. Here we report the isolation, functional characterization in *E. coli* and transcriptional regulation of β-phellandrene synthase and a monoterpene synthase-like cDNA from *L. angustifolia*. 
3.2 Materials and methods

3.2.1 Microarray analysis and clone selection

The construction of *L. angustifolia* leaf and flower cDNA libraries was reported earlier by our group (Lane et al. 2010). Transcriptional activity for all leaf and flower ESTs were evaluated in developing leaves and flowers using the Agilent oligo-based microarray system. 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). The microarray analysis results highlighted two previously uncharacterized TPS-like sequences including *LaβPHLS* and *LaTPS-I* (identified using BLASTX algorithm of the National Center for Biotechnology Information—http://www.ncbi.nlm.nih.gov/) that were strongly expressed in young leaves or flowers, where essential oil metabolism actively takes place. These cDNAs were selected for further analysis.

3.2.2 Expression of recombinant LabPHLS and LaTPS-I in *E. coli*

Full length clones for *LaβPHLS* (accession number HQ404305) and *LaTPS-I* (accession number HQ404306), which were present in our cDNA libraries, were fully sequenced. The ORF excluding the N-terminal transit peptides for both clones were amplified by PCR using Deep Vent DNA polymerase (New England Biolabs, Beverly, MA, USA) to generate restriction sites for cloning in the bacterial expression vector pET41 b(+) (Figure A.1) (EMD Chemicals, Darmstadt, Germany). The PCR program used was 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 58 °C for 30 s and 72 °C for 2 min, and 5 min final extension at 72 °C. PCR products were purified using QIAquick Kit (Qiagen, Valencia, CA, USA), and fused to sequences encoding eight C-terminus Histidine residues in the expression vector to facilitate protein purification. The resulting constructs were electroporated into *E. coli* Rosetta (DE3) cells. The transformed cells were spread on LB media supplemented with 30 mg/l kanamycin and grown overnight at 37 °C. Cells from selected single colonies were inoculated in 5 ml LB media containing kanamycin (30 mg/l) and grown overnight at 37 °C. The next morning 500 µl of the overnight grown culture was added to 100 ml LB media supplemented with 30 mg/l kanamycin and incubated at 37 °C.
until OD₆₀₀ of 0.6–0.7 was achieved. Isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 0.1 mM was added and the cultures were incubated at 20 °C for 14–16 h to induce heterologous protein expression. After protein expression, culture flasks were kept on ice for 15–20 min and cells were harvested by centrifugation at 3,220g and 4 °C for 20 min. The pellet was resuspended in half the initial volume of ice-cold wash buffer (20 mM Tris, 10 mM EDTA, 10% Triton X-100, pH 7.6) and collected twice by centrifugation at 3,220g and 4 °C for 20 min. The washed cells, 0.6–0.85 mg fresh weight, were resuspended in 5–6 ml Novagen bind buffer (0.5 M NaCl, 20 mM Tris/HCl, 5 mM imidazole, pH 7.9; EMD Chemicals, Germany) containing 0.5 mg/ml of lysozyme (Sigma, Canada) and 1 mM phenylmethanesulfonyl fluoride (PMSF). The lysozyme digestion was performed on ice for 30 min with brief vortexing at 5 min intervals. Cells were then sonicated on ice using a Sonic Dismembrator Model 100 (Fisher Scientific, Ottawa, ON, Canada) to complete bacterial membrane disruption. The clear lysate containing the soluble protein fraction was collected after centrifugation for 15 min at 15,000g and 4 °C. The His-tagged protein was purified from the soluble fraction by Ni–NTA agarose affinity chromatography (EMD Chemicals, Germany) following the manufacturer’s procedure. Purified proteins, and proteins extracted from non-induced and induced cells, were resolved in 10% SDS–PAGE and visualized by staining with Coomassie Brilliant Blue.

3.2.3 Enzyme assay

Enzyme activity was assayed as previously reported (Mahmoud et al. 2004). Typical assays were performed in 500–800 µl reaction volume, containing the assay buffer (50 mM Tris/HCl, 5% glycerol, 1 mM MnCl₂, 1mM MgCl₂, 1 mg/ml Bovine Serum Albumin [BSA], pH 7.0), 1 mM DTT, 25 µM substrate (GPP, NPP or farnesyl diphosphate [FPP]; Echelon, Salt Lake City, UT, USA), and 5–50 µg purified protein. The mixture was overlaid by 400 µl of pentane and incubated at 30 °C for 15 min. Purified protein extracted from E. coli Rosetta (DE3) cells transformed with empty expression vector was also assayed under the same conditions as a control. The reaction was stopped by freezing in liquid nitrogen and storing in -80 °C freezer until analyzed. Internal standard, 100 ng of camphor, was added to the reaction mixture prior to transferring the liquid phase – which contained the assay products –
to an ice-cold glass tube. The assay products were concentrated by evaporating ~90% of the pentane using a gentle stream of highly purified helium gas.

Seven reaction time points (5, 10, 15, 20, 30, 45 and 60 min) and five temperature levels (25, 27.5, 30, 32.5 and 35 °C) were selected to evaluate kinetics of LaβPHLS, and to determine optimum temperature, respectively. The optimum pH of the enzyme was determined using MES and MOPS buffers with pH 6.0, 6.5, 7.0, 7.5 and 8.0. The saturation curve of the enzyme kinetics was analyzed from assays containing seven different substrate concentrations (5, 10, 25, 50, 75, 100 and 200 µM) under optimum conditions. SigmaPlot v.10.00 (Systat Software, Erkrath, Germany) software was used to fit the data in the Michaelis–Menten equation, and calculate the Vmax and Km values. Substrate specificity of the enzyme was determined using the three substrates (GPP, NPP or FPP) under optimized conditions.

3.2.4 Product assay/GC–MS analysis

Identification and quantification of the assay product were conducted by GC–MS. All GC–MS analyses were performed on Varian GC 3800 Gas Chromatographer (Varian Inc., Palo Alto, CA, USA) coupled with a Saturn 2200 Ion Trap mass detector. 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 CO2 cooled 1079 Programmable Temperature Vaporizing (PTV) injector (Varian Inc., 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 product identity was confirmed by comparing its retention time and mass spectrum to those for an authentic β-phellandrene sample (Chemos GmbH, Regenstauf, Germany). Essential oils of L. angustifolia leaves and flowers were distilled as previously reported (Falk et al. 2009), and analyzed as described above.
3.2.5 Relative expression assay

*L. angustifolia* plants (cv. lady) grown at the University of British Columbia Okanagan lavender field were used in this study. The transcriptional activity of *LaβPHLS* and *LaTPS-I* was evaluated in developing leaves and floral tissues. Total RNA was isolated from 100 mg fresh tissue using the RNeasy Plant Mini Kit (Qiagen; Valencia, CA, USA), and treated with DNase using the Qiagen on-column DNase digestion kit to remove genomic DNA. Total RNA was reverse transcribed with oligo d(T) primer and M-MuLV Reverse Transcriptase (New England Biolabs, Beverly, MA, USA) following the manufacturer’s directions. The reverse transcribed cDNA was amplified with cloning primers using Taq DNA Polymerase (New England Biolabs, Beverly, MA, USA). The standard PCR program used was 95 °C for 5 min, followed by 30 cycles of 95 °C for 1 min, 58 °C for 30 s and 72 °C for 2 min, and final extension of 72 °C for 5 min. Transcript abundance of *LaβPHLS* in young, medium and old leaves, and young floral tissues, and *LaTPS-I* in leaves and 30% flower were quantified by CFX96™ real-time PCR detection system (Bio-Rad; Hercules, CA, USA) using the SsoFast™ Eva-Green® Supermix (Bio-Rad; Hercules, CA, USA) along with approximately 150 ng of cDNA template and 500 nM of each primer in a total volume of 20 µl. Gene-specific primers (see Table 3.1) used in quantitative real-time PCR experiments were designed using IDT primer quest (http://www.idtdna.com/Scitools/Applications/Primerquest/) to amplify 120–150 bp products. The following program was used for all of these PCR reactions: 95 °C for 30 s followed by 40 cycles of 5 s at 95 °C and 30 s at 58 °C. Normalized expression values (ΔΔCT) of *LaβPHLS*, *LaTPS-I* and *LaLINS* were calculated by CFX96™ manager software (Bio-Rad; Hercules, CA, USA) using b-actin and 18S rRNA as a reference gene.
3.3 Results and discussion

3.3.1 Candidate selection and cloning

We have recently reported the isolation and sequencing of over 14,000 ESTs for *L. angustifolia* (Lane et al. 2010). This collection includes several ESTs homolog to known terpene synthase genes from *Lavandula* and other plants. In order to gain insight into their regulation, we examined the expression pattern of our ESTs by microarray analysis in developing *L. angustifolia* flowers and leaves. Both tissues produce essential oil constituents in glandular trichomes although the composition of the essential oil is different. While young leaves accumulate mainly β-phellandrene (Figure 3.1a), young flowers – which do not produce this monoterpane – accumulate large quantities of linalool (Figure 3.1b).

The microarray analysis experiment revealed that many of the terpene synthase transcripts are differentially produced, following an anticipated expression pattern. For example, transcripts for linalool synthase gene were strongly expressed in flowers where linalool is produced, but were not detected in leaves (data not shown). However, the expression pattern for two ESTs appeared unusual.

One of the ESTs (*LaTPS-I*) was initially identified as *L. angustifolia* limonene synthase based on homology to the top hit during a Genebank search, although the clone lacked an internal stretch of 73 amino acids, including the DDxxD terpene synthase signature motif (Figure 3.2). Given that limonene is produced mainly in lavender leaves, we expected to see a strong transcriptional activity for *LaTPS-I* in leaves, but not flowers. Instead, our microarray data did not indicate a differential expression pattern for this gene, as presumably the same amounts of the transcript were present in both tissues. The other sequence, *LaβPHLS*, was highly homologous to the 1,8-cineole synthase gene from *Salvia* spp. Again, given that 1,8-cineole is a major constituent of floral essential oil in some lavender species (e.g., *L. x intermedia* (Lis-Balchin 2002)), higher *LaβPHLS* transcript levels was anticipated in flowers. However, our microarray data suggested that LaβPHLS is strongly expressed in leaves of *L. angustifolia*. These unexpected expression results, which were later confirmed by standard and real-time PCR (see below), prompted further investigation of these clones. We thus expressed both ESTs in bacteria, and functionally characterized LaβPHLS recombinant protein *in vitro*. We also examined the expression pattern for these genes in more detail.
3.3.2 Sequence analysis

The ORF for \textit{LaβPHLS} was 1,791 bp long, encoding a 597 amino acid protein with a predicted molecular weight of 66.3 kDa. The ORF for \textit{LaTPS-I} contained 1,482 nucleotides, and encoded a 494 amino acid protein with a predicted molecular weight of 54.3 kDa. Both \textit{LaβPHLS} and \textit{LaTPS-I} contained predicted N-terminal transit peptides (Signal-3L, http://www.csbio.sjtu.edu.cn/bioinf/Signal-3L/) composed of 52 and 49 amino acids, respectively, which resembled other plant TPS transit peptides (Cai et al. 2010; Keegstra et al. 1989; Von Heijne et al. 1989). Further, \textit{LaβPHLS} contained all conserved TPS domains including the RRx8W (Williams et al. 1998), the DDxxD, the (N,D)D(L,I,V)x(S,T)xxxE, and the LQLYEASFLL motifs (Degenhardt et al. 2009; Roeder et al. 2007) (Figure 3.2). \textit{LaTPS-I} also contained all signature motifs except for the highly conserved DDxxD sequence located within the 73 amino acid long missing region towards the C-terminus of the protein (Figure 3.2). This motif is highly conserved in all functionally characterized TPSs across species (Bohlmann et al. 1998; Degenhardt et al. 2009; Nagegowda 2010), and serves as divalent metal binding site (Seemann et al. 2002). The LQLYEASFLL motif is often partially conserved, as different variations have been reported for linalool synthase (Landmann et al. 2007), 1,8-cineole synthases (Roeder et al. 2007; Shimada et al. 2005; Wise et al. 1998) and β-phellandrene synthase of grand fir (Bohlmann et al. 1999).

Alignment of the amino acid sequences revealed that \textit{LaβPHLS} shares 48 and 47% identity with other lavender TPSs \textit{LaLINS} and \textit{LaLIMS}, respectively, while \textit{LaTPS-I} exhibits 66 and 70% identity with \textit{LaLINS} and \textit{LaLIMS}, respectively. The sequence identity between \textit{LaβPHLS} and β-phellandrene synthases from grand fir (AAF61453.1) and tomato (ACO56896.1) was only 31 and 25%, respectively. The identity between β-phellandrene synthases from grand fir (AAF61453.1) and tomato (ACO56896.1) was also only 26%. However, the β-phellandrene synthase cDNA cloned from grand fir retained 64 and 74% identity with limonene and myrcene synthases, respectively of the same species. As noted by Bohlmann et al. (1998), Degenhardt et al. (2009) and Nagegowda (2010), TPSs sequence conservation tends to favor evolutionary relationship despite functional divergence. Accordingly, the genetic distance between grand fir (\textit{Abies grandis}), tomato (\textit{Solanum lycopersicum}), and lavenders (\textit{L. angustifolia}) could explain the low sequence identity among the three β-phellandrene synthases.
3.3.3 Heterologous expression

For functional expression in *E. coli*, N-terminal transit peptides were removed from both TPSs. It has previously been demonstrated that exclusion of transit peptide enhances solubility of recombinant proteins in bacterial hosts (Williams et al. 1998). The coding regions excluding the transit peptide for LaβPHLS and LaTPS-I with a C-terminal histidine tag (eight His residues) was expressed in *E. coli* Rosetta (DE3) cells, in which most of the recombinant proteins were produced in soluble forms. The recombinant proteins were purified by affinity chromatography using Ni–NTA resin (EMD Chemicals, Germany), which yielded highly purified proteins with anticipated masses of approximately 62.3 kDa for LaβPHLS (Figure 3.3). Purified proteins were subsequently assayed with GPP, NPP or FPP as substrates.

3.3.4 Functional characterization of LaβPHLS

The purified recombinant LaβPHLS produced β-phellandrene as its major product (86.1%) in vitro when assayed with GPP or NPP as a substrate (Figure 3.4a). However, the enzyme did not produce a detectable product with FPP as a substrate (data not shown). Further, bacterial extracts harvested from control cells (those transformed with the empty expression vector, or cells expressing LaTPS-I) did not produce detectable amounts of β-phellandrene when assayed with either GPP or NPP, although in all cases negligible quantities of linalool and/or geraniol were detected in assays that were incubated for over two hours (data not shown). These products correspond to the hydrolysis outputs of GPP/NPP, which occurs spontaneously in long incubations (Schilmiller et al. 2009; Tholl et al. 2001). The identity of the assay product was confirmed by comparing its retention time and mass spectrum to those of authentic β-phellandrene (Figure 3.4b).

Despite a clear preference for GPP, LaβPHLS efficiently synthesized β-phellandrene from NPP, a cisoid isomer of GPP. The preferential affinity toward GPP was evidenced by the amount of β-phellandrene synthesized from the two substrates under identical conditions (data not shown). This contradicts a previous report demonstrating that a tomato β-phellandrene synthase (PSH1) preferentially utilized NPP as a substrate (Schilmiller et al. 2009). This difference in substrate specificity is most likely due to the structural differences between PSH1 and LaβPHLS. PSH1 (ACO56896.1) is distantly related to most other plant
TPSs, exhibiting relatively low homology to them. In addition, PSH1 lacks three of the conserved TPS signature motifs, in particular the RRx8W motif required for isomerization of GPP (Figure 3.2). LaβPHLS, on the other hand, is closely related to known monoterpene synthases and includes all key signature motifs present in these proteins.

Although LaβPHLS produced β-phellandrene as its major product, assays with both GPP and NPP consistently yielded small quantities of limonene as well. This is not unusual, as mTPSs often convert substrate/s to a multitude of products. In this context, β-phellandrene synthases from grand fir and tomato produced multiple products. The grand fir β-phellandrene synthase produced 52% β-phellandrene, 34% β-pinene, 8.5% α-pinene and 6% 4S-limonene (Bohlmann et al. 1999). The tomato PSH1 produced mainly β-phellandrene and small amounts of carene, α-phellandrene, γ-terpinene, and limonene when NPP was used as substrate (Schilmiller et al. 2009). The enzyme, however, produced equivalent amounts of β-phellandrene, myrcene, and ocimene together with trace amounts of linalool when GPP was used as substrate (Schilmiller et al. 2009).

Under our assay conditions, LaβPHLS was active from 5 to over 180 min (Figure 3.5a), with a linear activity range 5–60 min. Like many other mTPSs, the optimum temperature for LaβPHLS was found to be 30 °C (Figure 3.5b). However, its optimum pH of 6.5 (Figure 3.5c) was slightly lower than pH 7.0 reported for mTPSs cloned from L. angustifolia (Landmann et al. 2007) and salvia officinalis (Wise et al. 1998). This is not unusual as there are other terpene synthases with slightly acidic pH optima, including linalool synthase of bergamot mint (Mentha citrata Ehrh) that had optimum pH of 6.5 (Crowell et al. 2002).

Kinetics assays of the LaβPHLS were conducted under optimum conditions. The initial velocity data showed a hyperbolic dependence on substrate concentration when fit to the Michaelis–Menten equation by non-linear square analysis using SigmaPlot software v.10.00 (Systat Software, Erkrath, Germany) (Figure 3.5d). The $K_m$ of the enzyme was calculated to be 6.55 ± 1.01 µM, while the $V_{max}$, $k_{cat}$ ($V_{max} / [E]$), and catalytic efficiency ($k_{cat} / K_m$) were 125.7 ± 3.86 pKat/mg, 1.7 ± 10^{-2} s^{-1} and 2.83 ± 10^{-3} µM^{-1} s^{-1}, respectively. The $K_m$ of LaβPHLS enzyme for GPP appears to be in close range with the most mTPSs reported. For example, the $K_m$ of the sesquiterpene synthase LaBERS for FPP cloned from lavender was 4.7 µM (Landmann et al. 2007) while two mTPSs cloned from snapdragon flower
exhibited 7.57 and 7.68 µM $K_m$ for GPP (Nagegowda et al. 2008), and the $K_m$ of sabinene and 1,8-cineole synthase cloned from Common Sage (*Salvia officinalis*) were determined to be 7.0 and 7.4 µM, respectively (Wise et al. 1998). LaβPHLS has nine to ten fold lower $K_m$ values compared to other mTPSs cloned from *L. angustifolia* and some other plants. LaLINS and LaLIMS exhibited a $K_m$ of 42.7 ± 4.6 and 47.4 ± 3.8 µM (Landmann et al. 2007) and the $K_m$ of linalool synthase cloned from bergamot mint was 56.5 ± 5 µM (Crowell et al. 2002). Furthermore, the catalytic activity of the enzyme is within the range reported for other mTPSs, 0.01–0.1 s$^{-1}$ (Wise and Croteau 1999).

3.3.5 Enzymatic assay of LaTPS-I

Incubation of the affinity purified LaTPS-I with GPP, NPP or FPP as substrates for 2 h did not produce any products. The only products detected were very small quantities of linalool and geraniol, which were also formed in control assays devoid of LaTPS-I, indicating that they most likely arose from the hydrolysis of the substrate after long incubation (data not shown). 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 proposition is supported by the results of previous investigations, where substitution of the first aspartate by asparagine in the DDxxD motif abolished the enzymatic activity of aristolochene synthase (Felicetti and Cane 2004). Furthermore, replacement of either the first or second aspartates (D) residue in the DDxxD motif by glutamic acid (E) in pentalenene synthase (Seemann et al. 2002), and aristolochene synthase (Felicetti and Cane 2004) significantly reduced the catalytic efficiency and specificity of the mutated protein. In addition, substitution of asparagine (N) residue located further downstream of the DDxxD motif by either alanine (A) or leucine (L) in pentalenene synthase (Seemann et al. 2002) and by leucine (L) in aristolochene synthase (Felicetti and Cane 2004) completely inactivated the enzymes. It is therefore possible that amino acid residues other than those in the DDxxD motif within the “missing” region (Figure 3.2) play a significant role in controlling the activity of LaTPS-I, and perhaps other related enzymes.
3.3.6 Transcriptional regulation of LaβPHLS and LaTPS-I

We compared transcription activity for LaβPHLS and LaTPS-I to that of LaLINS (which is strongly expressed in flowers) in young flowers and leaves using standard and quantitative real time PCR (see Table 3.1 for PCR primers used in this study). As anticipated (based on our microarray data) LaβPHLS transcripts were barely detectable in flowers, while they were strongly expressed in young leaves (Figure 3.6a&b). This transcript production pattern correlated with β-phellandrene content in L. angustifolia tissue, as β-phellandrene is the major component of young leaf EO, but it is barely detected in EO distilled from flowers.

The developmental regulation of LaβPHLS was determined by measuring transcript levels for LaβPHLS in young, medium and old leaf tissues. The results indicated that this gene is strongly expressed in young leaves, while its transcriptional activity rapidly decreases with leaf age (Figure 3.6a&c). This expression pattern resembles those of some mTPSs (for example, limonene synthase expression in peppermint reached its peak within 21 days after the emergence of leaves. See McConkey et al. 2000), and is opposite to others. For example, the expression of linalool synthase gene in English Lavender increases with age until flower maturity (Boeckelmann 2008; Lane et al. 2010). Finally, transcript levels for LaβPHLS correlate with β-phellandrene content in leaves with young leaves producing large quantities of both, indicating that like many other TPSs (Lane et al. 2010) LaβPHLS is regulated mainly at the level of transcription.

Although LaTPS-I transcripts were slightly more abundant in flowers than leaves, they were strongly expressed in both tissues, compared to LaLINS and LaβPHLS (Figure 3.6a). To elucidate the tissue-specific expression of LaTPS-I, we compared the transcriptional activity of this gene to that of a gene coding for linalool synthase, 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). 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 leaves, using quantitative real-time PCR. The natural occurrence of this variant was ascertained by sequencing eight independent LaTPS-I clones amplified directly from 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 using the assay procedures described above.

3.3.7 Phylogenetic analysis

A phylogenetic tree was constructed using the default parameters of PhyML phylogenetic tree building software (Dereeper et al. 2008) available at “phylogeny.fr” platform (http://www.phylogeny.fr). PhyML software employs MUSCLE software to generate multiple alignments and the maximum likelihood computational method to construct the phylogenetic tree. Both LaβPHLS and *LaTPS-I* were clustered into subfamily TPSb that contains mTPSs from angiosperms including *Lamiacea*, while b-phellandrene and other monoterpenes cloned from grand fir (*Abies grandis*) were distinctly clustered in another subfamily TPSd (Figure 3.7). This is in agreement with TPSs classification reported by Bohlmann et al. (1998), Degenhardt et al. (2009) and Nagegowda (2010). PHS1 (β-phellandrene synthase cloned from tomato) was rooted in subfamily TPSe together with kuarene synthase (Schilmiller et al. 2009). Within TPSb LabPHLS was rooted with 1,8-cineole syntheses cloned from *Salvia fruticosa* and *Rosmarinus officinalis*, while LaTPS-I was rooted to LaLIMS and LaLINS, and linalool synthase from *L. angustifolia*. 
Figure 3.1 *L. angustifolia* young leaf and young flower EO chromatogram.
Figure 3.1 *L. angustifolia* young leaf and young flower EO chromatogram: (a) leaf and (b) young flower. Peaks designated by numbers represent individual components (1) careen, (2) α-phellandrene, (3) limonene, (4) β-phellandrene, (5) linalool, (6) linalool acetate, and (7) terpene-4-ol. Note that the oils analyzed here were obtained from young tissue. In older flowers, linalool acetate is present at much higher levels (Falk et al. 2009).
Figure 3.2 Multiple alignment of mTPSs.
Figure 3.2 Multiple alignment of mTPSs. LaβPHLS: *L. angustifolia* β-phellandrene synthase, AAF61453.1: grand fir β-phellandrene synthase, ACO56896.1: tomato β-phellandrene synthase, LaLINS: *L. angustifolia* linalool synthase, LaLIMS: *L. angustifolia* limonene synthase, LaTPS-I: *L. angustifolia* terpene synthas-like protein. Symbols below the sequences indicate identical (asterisk), conserved (colon), and semi-conserved (period) amino acids. The missing region of LaTPS-I is blocked, and conserved motifs are shown on top of the sequences.
Figure 3.3 SDS-PAGE analysis of protein samples from bacterial cells expressing LaβPHLS. (1) total protein from cells expressing LaβPHLS, (2) soluble proteins from cells expressing LaβPHLS, (3) purified LaβPHLS and, (4) protein marker.
Figure 3.4 Analysis of LaβPHLS assay products and authentic standard by GC–MS.
Figure 3.4 Analysis of LaβPHLS assay products and authentic standard by GC–MS. a) Chromatogram and mass spectrum of product of LaβPHLS with GPP as a substrate: (1) limonene, and (2) β-phellandrene. b) Chromatogram and mass spectra of authentic β-phellandrene: (1) limonene, (2) β-phellandrene, and (3) cymene. Limonene and cymene are impurities present in the commercially obtained standard.
Figure 3.5 Biochemical characterization of LaβPHLS.
Figure 3.5 Biochemical characterization of LaβPHLS. a) Progress curve determined by plotting incubation time in minutes against β-phellandrene concentration, b) effect of temperature on LaβPHLS activity, c) effect of pH on LaβPHLS activity and d) effect of substrate (GPP) concentration on LaβPHLS activity. (n = 3 and error bars indicate standard deviation)
Figure 3.6 Transcriptional activity of *L. angustifolia* terpene synthases.
Figure 3.6 Transcriptional activities of *L. angustifolia* terpene synthases. Transcript levels were normalized to β-actin and 18S rRNA. a) Amplification of full length (1) *LaLINS*, (2) *LaTPS-I*, (3) *LabβPHLS* clones from leaf and flower cDNA, and L: 1 kb DNA ladder. b) Transcriptional activity of *LaβPHLS* (black) and *LaLINS* (gray). c) Developmental expression of *LaβPHLS* mRNA. Error bars indicate standard deviation (n = 3).
Figure 3.7 Phylogenetic tree of plant TPSs.

Phylogenetic tree of plant TPSs reconstructed using maximum likelihood computational method with PhyML phylogenetic tree building software (Dereeper et al. 2008). TPSs with 50% minimum amino acid sequence identity are grouped into TPSa to TPSf subfamilies.
Table 3.1 Primer sets for cloning and qRT-PCR reactions.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Target gene</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><em>LaβPHLS</em> ORF without signal peptide</td>
<td>F: 5'- TTTTTTTTCATATGACGGCGCGCCGATCCCGGA -3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'- GCACTGGAATTCTCGATCAACGTCTCTT -3'</td>
</tr>
<tr>
<td>2</td>
<td><em>LaTPS-I</em> ORF without signal peptide</td>
<td>F: 5'-GGCCATATGGATTAGCAGTCATCAATG-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'-ATACTCGAGCTCTGCAAGATTCGCTGC-3'</td>
</tr>
<tr>
<td>3</td>
<td><em>LaTPS-I</em> ORF with signal peptide</td>
<td>F: 5'-AGCGGCCCATATGTCTAGCTTGTACATG-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'-ATACTCGAGCTCTGCAAGATTCGCTGC-3'</td>
</tr>
<tr>
<td>4</td>
<td><em>LaLINS</em></td>
<td>F: 5'-CGCATATGCGATCAATATCAACAT-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'-ATAGAATTCTGCGATCGCTAACA-3'</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Actin</td>
<td>F: 5'- TGTGGATTGCAAGCCAGAGT-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'- ATGAGCAAGGCGAACAGCA-3'</td>
</tr>
<tr>
<td>2</td>
<td>18s rRNA</td>
<td>F: 5'- ACGACTGCTCGGCAACGGCT-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'- GCGTTCAAGACATCGAGTTG-3'</td>
</tr>
<tr>
<td>3</td>
<td><em>LaβPHLS</em></td>
<td>F: 5'-TGCTCGTTGCAAGTGAGTGATGAT-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'-CCCTTCTAGTGCTCTTCTCC-3'</td>
</tr>
<tr>
<td>4</td>
<td><em>LaTPS-I</em></td>
<td>F: 5'- ACTACACTGAGGGTGCAAAAGA-3'</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5'- AATCTGGGAATCGCATTTGGCG-3'</td>
</tr>
</tbody>
</table>
Chapter 4 Cloning, functional characterization and genomic organization of 1,8-cineole synthases from *Lavandula*

4.1 Overview

The genus *Lavandula* (lavenders), a member of the *Lamiaceae* (mint) family of plants, is composed of over 32 morphologically distinct species including *L. angustifolia*, *L. latifolia*, and their natural hybrid *L. x intermedia* (Upson 2002). These plants are widely grown for their EOs, which are extensively used in the manufacturing of perfumes, food flavours, antiseptics, and personal care and medicinal products (Upson and Andrew 2004). *Lavandula* EOs are enriched in a few monoterpenes - the C10 class of the isoprenoids - in a species-specific manner. For example, *L. angustifolia* oils are dominated by linalool and linalool acetate (Boeckelmann 2008), while *L. latifolia* oils are characterized by high levels of linalool, 1,8-cineole and camphor (Munõz-Bertomeu et al. 2007). The EOs of *L. x intermedia* plants contain a blend of *L. angustifolia* and *L. latifolia* components, including linalool, linalool acetate, 1,8-cineole and camphor as major oil constituents (Desautels et al. 2009; Lis-Balchin 2002). However, the relative composition of each of these monoterpenes in EOs distilled from the three species varies considerably. In particular, 1,8-cineole accounts for 20.5 - 42.4% and 7 – 11% of the EOs distilled from *L. latifolia* and *L. x intermedia* species, respectively. However, *L. angustifolia* species accumulate only small amounts (0 – 1.5% of the oil) of this monoterpene (Lis-Balchin 2002) (Table 1.1).

In lavenders and other EO-producing plants (e.g., mints) the biosynthesis of EO constituents takes place in specialized structures known as glandular trichomes or oil glands. Clusters of six to eight secretory cells situated in these tissues are specialized to produce and secrete large quantities of EO constituents into a subcuticular storage cavity of the oil gland (McCaskill and Croteau 1995, McCaskill et al. 1992, Turner et al. 2000a & b). The production of EO constituents begins with the synthesis of the universal terpene precursors IPP and its allylic isomer DMAPP, mainly through the MEP or plastidial pathway of isoprenoid metabolism (Dudareva et al. 2005, Lane et al. 2010, McCaskill and Croteau 1995, McCaskill et al. 1992, Rodriguez-Concepcion et al. 2001). The MEP pathway commences by condensation of pyruvate and DXP, catalyzed by DXS and DXP is subsequently transformed into IPP and DMAPP through the sequential action of the following enzymes: DXR, MCT,
CMK, MDS, HDS and HDR (Phillips et al. 2008). IPP and DMAPP are then condensed head-to-tail by GPPS to produce the linear monoterpane precursor GPP, which is subsequently transformed into various monoterpenes by specific enzymes collectively known as mTPSs. For example, the *L. angustifolia* LaLINS and LaLIMS transform GPP mainly to linalool and limonene, respectively, (Landmann et al. 2007) while LaßPHLS produces β-phellandrene (Demissie et al. 2011) from GPP or its cisoid isomer NPP* in vitro* (Figure 1.2 and 1.3).

Over the last three decades, numerous mTPSs have been described from gymnosperms and angiosperms and have been reviewed by Degenhardt et al (2009). Bohlmann et al (1998) and Chen et al (2011) exploited the evolutionary relationship among TPSs isolated from different species, as explained by their amino acid similarity level, to classify them into six subfamilies - TPSa through TPSf. According to this criterion, angiosperm and gymnosperm mTPSs were classified into TPSb and TPSd subfamilies, respectively. An alternate classification system groups TPSs into three classes based on the architecture of their genomic DNA. With six introns and seven exons, *Lamiaceae* mTPSs and their angiosperm counterparts are classified into the Class III clade while gymnosperm mTPSs are classified into Class II clade with nine introns and ten exons (Lee and Chappell 2008; Trapp and Croteau 2001). Despite their evolutionary, structural, and functional heterogeneity, gymnosperm and angiosperm mTPSs retain four conserved functional motifs. These include the divalent metal binding aspartate-rich DDxxD and (N,D)D(L,I,V)x(S,T)xxxE motifs, the catalytic LQLYEASFL motif, and the RR(x8)W motif (Figure 4.1) (Bohlmann et al. 1998; Christianson 2006; Degenhardt et al. 2009; Roeder et al. 2007; Trapp and Croteau 2001; Williams et al. 1998; Wise et al. 1998). This last motif is involved in the cyclization of the linear GPP into cyclic products (Williams et al. 1998), and can be absent in mTPSs that produce acyclic products (e.g., linalool synthase) (Bohlmann et al. 1998).

The EOs of most *Lavandula* species contain 50-60 monoterpenes including 1,8-cineole, also known as eucalyptol, named after the *Eucalyptus* species from which it was first isolated. This monoterpane occurs widely in plants, where it performs important ecological functions, for example to repel insects, deter herbivores, and repress germination and growth of competing plants (Franks et al. 2012; Khan et al. 2008; Gershenzone and Croteau 1991;
Southwell et al. 2003). Industrially, 1,8-cineole is widely used in hygiene products, food flavors, and pharmaceutical preparations. These include prescribed topical ointments for inflammation and pain relief (Juergens et al. 2003; 2004; Santos et al. 2000), nasal sprays, medication for the treatment of bronchial asthma and non-purulent rhinosinusitis (Juergens et al. 2003; Kehrl et al. 2004; Tesche et al. 2008), mouthwashes and cough suppressants (Lahora et al. 2002), disinfectants (Gilles et al. 2010), and insect repellents (Klocke et al. 1987; Maciel et al. 2010; Sfara et al. 2009), among others.

There is great interest in improving the quality and yield of the EO in lavenders. These objectives may be met through metabolic engineering, as has been demonstrated for peppermint (Mahmoud and Croteau, 2001; Mahmoud and Croteau, 2004), once the genes that control the biosynthesis of key EO constituents in commercially important species of _Lavandula_ are identified. To date, only three monoterpene synthases (limonene synthase, linalool synthase, and ß-phellandrene synthase) and a single sesquiterpene synthase (bergamotene synthase) have been reported from _L. angustifolia_ (Demissie et al. 2011; Landmann et al. 2007). In this study, we obtained over 8,200 ESTs from floral oil glands of _L. x intermedia_ plants, and examined their transcriptional activity using microarrays in the flowers of three lavender species, _L. x intermedia_ and its parents _L. angustifolia_ and _L. latifolia_. Here, we report the cloning, heterologous protein expression in _E. coli_, purification, and functional characterization of 1,8-cineole synthase from _L. x intermedia_. We also analyzed the genomic architecture/organization of CINS genes in the aforementioned lavender species.
4.2 Materials and methods

4.2.1 Glandular trichome isolation and cDNA library construction

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 hr in ice-cold extraction buffer (200 mM sorbitol, 10 mM sucrose, 25 mM MOPS, 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 aurintricarboxylic acid, 5 mM thiourea, and 2 mM DTT at pH 6.6. Cells were then isolated, washed by a wash buffer (10% glycerol, 25 mM PO₄ buffer, 1 mM EDTA, 2 mM aurintricarboxylic acid, 5 mM thiourea, and 2 mM DTT), flash frozen in liquid N₂ and stored in a -80 °C freezer until used. Total RNA was extracted from the secretory cells by the Qiagen RNAeasy mini Kit, and used to construct a cDNA library using the Zap-cDNA® Library Construction Kit (Agilent Technologies, Palo Alto, CA, USA). A total of 10,000 ESTs were isolated and partially sequenced from the 5’ end.

4.2.2 Microarray data analysis and candidate selection

In order to evaluate the expression pattern of lavender genes in relation to the biosynthesis of EO constituents, we evaluated the relative transcriptional activity of genes corresponding to our ESTs in secretory cells isolated from three stages of developing L. x intermedia flowers using the Agilent oligo-based microarray technology. The three floral developmental stages were: unopened buds or bud I (A), anthesis (B) and mature flowers in which 30% of the buds were in blooms (C) (Photographic description of lavender flower ontologies is available in Boeckelamnn 2008). Further, in order to trace the origin of L. x intermedia EO biosynthetic genes, the abundance of mRNAs corresponding to all ESTs was evaluated in mature flowers of L. x intermedia (D) and its parents L. angustifolia (E) and L. latifolia (F). The following comparisons were made: A vs B, B vs C, D vs E, D vs F, and E vs F. 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). The expression profile data were used to select a putative 1,8-cineole synthase EST.
4.2.3 Recombinant protein expression

The putative LiCINS full-length sequence was obtained from our *L. x intermedia* gland cDNA library. The ORF - excluding the N-terminal transit peptide predicted using ChloroP1.1 software ([http://www.cbs.dtu.dk/services/ChloroP/](http://www.cbs.dtu.dk/services/ChloroP/)), and stop codon - was cloned into the Ndel/EcoRI sites of pET41b(+) expression vector (Figure A.1) using Sticky-End PCR (Zeng et al. 1998). This cloning replaced the vector sequences that code for glutathione S-transferase (GST) with the LiCINS ORF. The coding region of LiCINS was amplified by PCR using set I & II cloning primers (Table 4.1) and *Deep Vent* DNA polymerase (New England Biolabs, Beverly, MA, USA) in separate tubes. The PCR program used was 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 60 °C for 30 sec and 72 °C for 2 min, and a 5 min final extension at 72 °C. The PCR products were purified using a Gel extraction/PCR purification kit (OMEGA bio-tek, USA). To generate sticky ends the purified PCR products were combined and denatured at 95 °C for 5 min followed by renaturation at room temperature for 30 min. The coding region of LiCINS was fused to sequences encoding eight C-terminus Histidines during ligation in the pET41b(+) vector (Figure A.1) in order to facilitate its purification by Ni-NTA agarose affinity chromatography (EMD Chemicals, Darmstadt, Germany). The recombinant sequence was expressed in *E. coli* Rosetta™(DE3)pLysS cells (EMD Chemicals, Darmstadt, Germany) at 20 °C for 14 -16 h in LB media supplemented with 30 mg/l kanamycin and IPTG at 0.1 mM final concentration. Following expression, cells were kept on ice for 15 - 20 min and harvested by centrifugation at 3,220 g and 4 °C for 20 min. The pellet was resuspended in half the initial volume of ice-cold wash buffer (20 mM Tris/HCl, 10 mM EDTA, 10% Triton X-100, pH 7.6) and collected twice by centrifugation at 3,220 g and 4 °C for 20 min. The washed cells, 0.7 - 0.8 mg fresh weight, were resuspended in 5 - 6 ml Novagen bind buffer (0.5 M NaCl, 20 mM Tris/HCl, 5 mM imidazole, pH 7.9; EMD Chemicals, Germany) that contained 0.5 mg/ml lysozyme (Sigma, Canada) and 1mM PMSF. The lysozyme digestion was performed on ice for 30 min with brief vortexing at 5 min intervals. Cells were then sonicated on ice using a Sonic Dismembrator Model 100 (Fisher Scientific, Ottawa, ON, Canada) to complete bacterial membrane disruption. The soluble fraction containing proteins was separated from cell debris by centrifugation at 15,000 g and 4 °C for 15 min. The His-tagged protein was then harvested from the soluble cellular content by Ni-NTA agarose affinity chromatography (EMD...
Chemicals, Germany) following the manufacturer’s procedure. Purified proteins, and total proteins extracted from non-induced and IPTG-induced cells, were resolved using 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by staining with Coomassie Brilliant Blue.

4.2.4 Enzyme assay

*In vitro* enzyme activity was assayed as previously described (Mahmoud et al. 2004; Demissie et al. 2010). Typical assays were performed in 500 µl reaction volume, containing the assay buffer (50 mM Tris/HCl, 5% glycerol, 1 mM MnCl₂, 1 mM MgCl₂, 1 mg/ml Bovine Serum Albumin [BSA], pH 7.0), 1 mM DTT, 25 µM substrate (GPP, NPP or farnesyl diphosphate [FPP]; Echelon, Salt Lake City, UT, USA), and 5 - 50 µg purified protein. The mixture was overlaid by 400 µl of pentane and incubated at 30 °C for 30 min. Purified protein extracted from *E. coli* Rosetta™ (DE3) plysS cells transformed with empty expression vector was also assayed under the same conditions as a control. The reaction was stopped by vigorous vortexing followed by flash freezing in liquid nitrogen, and stored in a -80 °C freezer until analyzed. An internal standard, 100 ng of camphor, was added to the reaction mixture prior to transferring the liquid phase, which contained the assay products, to an ice-cold glass tube. The assay products were concentrated by evaporating ≈90% of the pentane using a gentle stream of highly purified helium gas.

Seven reaction time points (5, 10, 20, 30, 60, 90 and 120 min) and five temperature levels (25, 27.5, 30, 32.5 and 35 °C) were selected to analyze the linear kinetic properties of LiCINS, and to determine its optimum temperature, respectively. The optimum pH of the enzyme was determined using MES and MOPS buffers at pH 5.5, 6.0, 6.5, 7.0, 7.5 and 8.0. A saturation curve was constructed using the data obtained from assays performed with seven different substrate (GPP) concentrations (5, 10, 25, 50, 75, 100 and 200 µM) at the optimal time, temperature and pH. The SigmaPlot software (Systat Software, Germany) was used to produce a Michaelis-Menten saturation curve, and to calculate the $V_{max}$ and $K_m$ values. Substrate specificity of the enzyme was determined by assaying the enzyme with GPP, NPP and FPP under the optimized conditions.
4.2.5 Product assay/GC-MS analysis

Assay product identification and quantification was performed by gas chromatography-mass spectrometry (GC-MS) on a Varian GC 3800 Gas Chromatographer coupled to a Saturn 2200 Ion Trap mass detector. 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 CO₂ cooled 1079 Programmable Temperature Vaporizing (PTV) injector (Varian Inc., USA). Samples were injected at 40 °C. 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 identities of products were confirmed by comparing their retention times and mass spectra to those of authentic standards (Sigma, Canada) analyzed under the same conditions. EOs of L. angustifolia, L. x intermedia and L. latifolia flowers were distilled and analyzed as previously reported (Falk et al. 2009). The components were identified by comparison of obtained mass spectra to those in the NIST library and authentic standards, and quantified using menthol (1mg/ml) as internal standard.

4.2.6 Cloning of CINS cDNAs

L. angustifolia and L. x intermedia plants grown at the University of British Columbia, Okanagan campus lavender field, and L. latifolia leaf and floral tissues generously provided by Dr. Tim Upson (Cambridge University, UK) were used to clone CINS cDNA from each species. Briefly, total RNA was extracted from 100 mg floral tissues, collected at 30% flowering stage, using an RNA extraction kit (OMEGA bio-tek, USA), and treated with the on-column DNaseI digestion kit (Qiagen, USA) to degrade genomic DNAs. The total RNA was then reverse transcribed in a reaction containing the oligo d(T) primer (Fisher Scientific, Canada) and M-MuLV Reverse Transcriptase enzyme (New England Biolabs, USA) following the manufacturer’s directions. The full-length cDNAs corresponding to CINSs were amplified with set-I cloning primers (Table 4.1) and iProof™ High-Fidelity DNA Polymerase (Bio-Rad, USA). The PCR program used was 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 60 °C for 30 sec and 72 °C for 2 min, and a 5 min final extension at 72 °C. The amplified fragments were cloned into pGEM-T Easy Vector System.
following the manufacturer’s procedure (Promega, USA). Ten independent clones from each species were sequenced, and contigs were constructed using the ClustalX module of the Geneious 5.0.3 software (Auckland, New Zealand) (Drummond et al. 2009). The three cDNA sequences are available at the NCBI database with the following accession numbers: JN701459 (LiCINS), JN701460 (L. latifolia 1,8-cineole synthase; LICINS) and JN701461 (L. angustifolia 1,8-cineole synthase; LaCINS).

4.2.7 Obtaining genomic CINS clones

Genomic DNAs of L. angustifolia, L. x intermedia and L. latifolia were extracted from young bud tissues using DNeasy Plant Mini Kit (Qiagen, USA). Set I cloning primers (Table 4.1) and iProof™ High-Fidelity DNA Polymerase (Bio-Rad, USA) were used to amplify CINSs from corresponding genomic DNAs. The PCR program included an initial heating of the reaction mixture at 95 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 60 °C for 30 sec and 72 °C for 4 min, and a 5 min final extension at 72 °C. Approximately 2.8 kb amplified fragments were cloned into pGEM-T Easy Vector System following the manufacturer’s procedure (Promega, USA), and fully sequenced. Intron/exon number, placement, phase and sizes were predicted using the NCBI Spidey genomic DNA-mRNA aligner (http://www.ncbi.nlm.nih.gov/IEB/Research/Ostell/Spidey/spideyweb.cgi), and further analyzed manually by aligning the CINS cDNA sequences of the three Lavandula species and other CINS cDNAs available in public database against the genomic DNA sequence results.

4.2.8 Relative expression assay

The transcriptional activity of LiCINS in floral - at 30% flowering stage - and young leaf tissues of L. angustifolia, L. x intermedia and L. latifolia were assessed by standard PCR based on the intensity of CINS bands amplified with set I cloning primers (Table 4.1) and Taq DNA Polymerase (New England Biolabs, USA). The standard PCR program used was 95 °C for 5 min, followed by 30 cycles of 95 °C for 1 min, 60 °C for 30 sec and 72 °C for 2 min, and a final extension at 72 °C for 5 min. Further, the relative transcript abundance of LiCINS in secretory cells isolated from bud I, anthesis and 30% flowering stages of L. x intermedia developing flowers was also assessed by CFX96™ real-time PCR detection
system (Bio-Rad, USA) using the SsoFast™ Eva- Green® Supermix (Bio-Rad, USA) along with approximately 150 ng of cDNA template and 500 nM of each of the primers in 20 µl reaction volume. Gene specific primers (see Table 4.1) used in quantitative real-time PCR experiments were designed using the IDT primer quest software (http://www.idtdna.com/Scitools/Applications/Primerquest/) targeting 180 - 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 (ΔΔC_T) of LiCINS and LaLINS were calculated by CFX96™ data manager (Bio-Rad, USA) using β-actin as a reference gene.

4.2.9 Phylogenetic analysis

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. TPSs that shared a minimum of 50% amino acid identity were clustered into distinct subfamilies of TPSa through TPSf (Bohlmann et al. 1998; Chen et al. 2011).
4.3 Results

4.3.1 Construction of EST library, transcript profiling and candidate selection

In an attempt to obtain genes involved in the biosynthesis and storage of lavender EO, we isolated and partially sequenced approximately 10,000 ESTs from secretory cells of *L. x intermedia* floral oil glands. The experiment yielded 8,205 high quality reads, which were clustered into 4,116 unigenes. The unigene library contained 3,075 singletons and 1,041 contigs corresponding to 5,130 sequences, representing over 62.53% of the reads. Several contigs – most of which corresponded to TPS-like genes – contained numerous EST members, suggesting that the corresponding genes were transcriptionally strongly active in secretory cells. For example the contig corresponding to linalool synthase contained over 278 members. Linalool is one of the most abundant oil constituents, and linalool synthase is strongly expressed in oil glands. 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. We examined the expression of mRNA species corresponding to our EST collection (focusing on TPS-like ESTs) in the oil glands of developing *L. x intermedia* flowers. Probes corresponding to our ESTs were loaded onto microarrays, which were then hybridized with labeled RNA obtained from secretory cells isolated from *L. x intermedia* flowers in three different developmental stages including bud-I (A), anthesis (B) and mature (30% in bloom) flowers (C). Using the same strategy, we also examined the relative abundance of EO-related transcripts in *L. x intermedia* (D) and its parents, *L. angustifolia* (E) and *L. latifolia* (F), mature flowers. The results of this study revealed that in general EO-biosynthetic genes were expressed in developing flowers in a predicted manner. For example, all ESTs homologous to the MEP pathway genes of isoprenoid biosynthesis and functionally characterized mTPSs of *Lavandula* (Figure 1.2 & 1.3) were down-regulated in secretory cells isolated from flowers at the bud-I stage, where only small amounts of EOs are produced, compared to open flowers (anthesis and 30% bloom), where EO synthesis actively takes place. Further, all the MEP pathway genes were expressed at similar levels in flowers of the three species examined, although slight variations were observed in some cases (Table 4.2). We were not able to
validate these expression variations by PCR (Figure 4.2), and thus concluded that these differences are due to experimental error. In contrast, a number of TPS homologs were differentially expressed both in developing flower tissues of *L. x intermedia* and flowers of the three species. In particular, transcript levels for a previously uncharacterized EST were substantially lower (9 and 35 folds) in the flowers of *L. x intermedia* and *L. angustifolia* compared to *L. latifolia*, respectively (Table 4.2, column D vs F and E vs F). However, they were considerably more abundant (12 fold) in *L. x intermedia* flowers compared to *L. angustifolia* (Table 4.2, column D vs E). This EST displayed significant homology to CINSs of *Salvia fruticosa* and *Salvia officinalis* (Sage) (Kampranis et al. 2007; Wise et al. 1998), and hence designated as LiCINS. Interestingly, the transcriptional activity for this EST (obtained from the microarray data) paralleled the 1,8-cineole content in the floral tissue of *L. x intermedia*, *L. angustifolia* and *L. latifolia* EOs (Lis Balchin 2002). Once the results of the microarray experiment were confirmed by semi-quantitative PCR, this EST was selected for sub-cloning and subsequent functional characterization.

4.3.2 Functional characterization of LiCINS

The complete ORF of LiCINS was 1,749 bp long encoding for 582 amino acids with a predicted mass of ca. 68.5 kDa. The encoded protein retained all conserved motifs of mTPSs with slight modifications observed in the second divalent metal binding site (N,D)D(L,I,V)x(S,T)xxxE and the catalytic LQLYEASFLL motifs (Figure 4.1). The highly conserved DDxxD motif - the main divalent metal binding site - and RR(x₈)W motif - signature sequences of mTPSs synthesizing cyclic monoterpenes - were fully retained in LiCINS. Further, LiCINS retained a 90% sequence similarity with LaβPHLS of *L. angustifolia* (ADQ73631.1), 77% similarity with cineole synthases of *Salvia fruticosa* (ABH0767.1) and *Salvia officinalis* (AAC26016.1), and a 65% and 64% sequence similarity to LaLINS (ABB73045.1) and LaLIMS (ABB73044.1) of *L. angustifolia*, respectively (Table 4.3). The N-terminal 52 amino acids were predicted to code for a transit peptide using ChloroP1.1 peptide prediction tool (http://www.cbs.dtu.dk/services/ChloroP/) and Signal 3L software (http://www.csbio.sjtu.edu.cn/bioinf/Signal-3L/), which was ultimately excluded during cloning, resulting in a 1,599 bp long ORF that encoded for 532 amino acids with a theoretical mass of ca. 63.2 kDa.
Recombinant LiCINS was produced in *E. coli* Rosetta™(DE3)plysS strain using the pET41(b+) expression system (Figure A.1) (EMD Biosciences, USA), and highly enriched using Ni-NTA agarose affinity chromatography (Figure 4.3). Upon incubation with GPP as a substrate the purified recombinant protein produced 1,8-cineole as its major product (80%) and a few other minor products including sabinene (7.9%), α-phellandrene (6.6%), limonene (2.3%), and α-terpineol (1.7%) (Figure 4.4a). When incubated with NPP, the cisoid isomer of GPP, LiCINS still produced 1,8-cineole as its major product (61%), but the proportions of sabinene (18.5%), α-phellandrene (8%), α-terpineol (5.5%), and limonene (3.3%) in the product mix were increased (Figure 4.4b). Trace quantities of linalool were also detected in assays containing either of the substrates. The major product of the recombinant LiCINS had identical retention time and mass spectrum to those of a 1,8-cineole analytical standard (Sigma, Canada) (Figure 4.4c), thus confirming the identity of the product. We also confirmed the authenticity of limonene and α-terpineol by comparing their retention times and mass spectra with those of authentic standards (Sigma, Canada). The identity of sabinene and α-phellandrene were determined by comparing their mass spectra to those of sabinene and α-phellandrene, respectively, in the National Institute of Standards and Technology (NIST) library. As a negative control, we also performed identical assays with protein extracts of *E. coli* Rosetta™(DE3)plysS cells transformed with empty vector. These assays did not produce detectable amounts of any of the products, except for trace amounts of the typical GPP hydrolysis products linalool (Figure 4.5). The linear chemical kinetics of LiCINS extended from 5 to 120 min (Figure 4.6a), while its optimum pH (Figure 4.6b) and temperature (Figure 4.6c) were found to be 6.5 and 30 °C, respectively. $K_m$ and $V_{max}$ were determined by fitting the Michaelis–Menten equation to the initial velocity data, which showed a hyperbolic dependence on substrates concentrations. The equation was fit by nonlinear square analysis using SigmaPlot software v.10.00 (Systat Software, Erkrath, Germany) (Figure 4.6d). The $K_m$ of LiCINS was calculated to be 5.75 ± 0.91 µM, while the $V_{max}$, $k_{cat}$ ($V_{max}/[E]$), and catalytic efficiency ($k_{cat}/K_m$) were calculated as 1.05 x $10^{-6}$ µmole s$^{-1}$ or 138.73 ± 3.96 pKat/mg, 8.8 x $10^{-3}$ s$^{-1}$ and 1.53 x $10^{-3}$ µM$^{-1}$ s$^{-1}$, respectively (raw data and detail of enzyme kinetics analysis is provided in Table 4.4). The enzyme was inactive upon incubation with FPP (data not shown).
4.3.3 Cloning of CINS genomic DNA

We cloned and sequenced several copies of the full length complementary and genomic DNA for CINS from each of *L. x intermedia* and its parents, *L. angustifolia*, and *L. latifolia*. The ORFs of the cDNAs from the three species had exactly the same nucleotide sequences except for a polymorphic nucleotide in *L. angustifolia* cDNA at the 1,468\textsuperscript{th} position of the bacterially produced recombinant protein devoid of the transit peptide (at the 1615\textsuperscript{th} position when the transit peptide is included) (Figure 4.7). At this position, the substitution of thymine (T) by cytosine (C) in *L. angustifolia* cDNA altered the encoded amino acid at the 490\textsuperscript{th} (539\textsuperscript{th} when the transit peptide is included) position from tyrosine (Y) to histidine (H) (Figure 4.1). We expressed and assayed (with GPP) the recombinant *L. angustifolia* ortholog as before under the optimal conditions. The results showed that the amino acid change did not alter the product profile of the enzyme (data not shown).

The genomic DNA of CINS contained six introns and seven exons, placing this gene in the Class III TPS clade (Figure 4.8). Exons and introns were numbered based on their proximity to the 5’ terminus with “exon1” being the closest to the 5’ end and “intron1” being the first non-coding sequences interrupting “exon1” and “exon2”. “Exon1” was 234 bp long encoding for 78 amino acids, “exon2” was 255 bp long encoding for 85 amino acids and the longest exon, “exon3”, contained 359 bp and encoded for 119 amino acids. These three exons were interrupted by “intron1” and “intron2” that were 264 and 80 bp long, respectively. The mTPSs signature motif RR(\(x_8\))W was placed on “exon1” while the catalytic LQLYEASFLLL motif was placed on the 3\textsuperscript{rd} exon. The aspartate-rich divalent metal binding DDxxD motif was placed on “exon4” that was 219 bp long and encoded for 73 amino acids. The other divalent metal binding motif [(N,D)D(L,I,V)x(S,T)xxxE] was shared between “exon6”, which was 234 bp and 78 amino acids long, and “exon7”, which was 309 bp and 103 amino acids long. “Exon5” was the shortest of all exons with 139 bp long encoding for 46 amino acids. “Intron3” was 155 bp long, “intron5” contained 73 nucleotides while the shortest and longest introns, “intron4” and “intron6”, had 58 and 417 nucleotides, respectively. Intron phase describes the placement of an intron on the proximate codon nucleotide. Introns placed on the first nucleotide of the proximate codon are described as “0”, while “1” and “2” describe the placement of the intron on the second and third nucleotide of the codon, respectively. “Exon1”, “exon2”, “exon5”, “exon6” and “exon7” were placed on the first
nucleotide of their proximate codon and had phase 0, while “exon 3” and “exon4” were placed on the second nucleotide of their proximate codon and thus had intron phase 1.

4.3.4 Analysis of transcript levels for CINS by PCR

The transcriptional activity of CINS gene in young leaves and floral tissues (30% flowering) of *L. angustifolia*, *L. x intermedia* and *L. latifolia* plants was determined by PCR using set I full-length primers. In agreement with the microarray results, the end-point PCR analysis showed that CINS mRNA was present in flower tissues of all plants (Figure 4.9a). Further, the mRNA corresponding to this gene was not detected in leaf tissues of these species (Figure 4.9b). We also evaluated transcriptional activity of the β-phellandrene synthase gene in the same tissues as a control. As anticipated (Demissie et al. 2011), β-phellandrene synthase mRNA was detected in the flowers and leaves of *L. angustifolia* plants only (Figure 4.9a&b).

The transcriptional activity of LiCINS paralleled the EO 1,8-cineole content through *L. x intermedia* flower development. The 1,8-cineole concentrations in bud-I, anthesis and 30% blooming flowers were virtually the same, amounting to 2.9, 2.7 & 2.3 mg/gm of fresh tissue, respectively (Table 4.5). In these tissues the LiCINS mRNA abundance followed a similar pattern and remained relatively constant (Figure 4.9c). On the other hand, the transcript levels for *LaLINS* (measured as a control) followed a previously reported trend (Lane et al. 2010), and were 9 and 12 folds higher in anthesis and 30% flowering stages, respectively, compared to bud-I (Figure 4.9c). In these tissues linalool contents also increased by age, and were at 0.7, 5.9 & 11.3 mg/gm of fresh tissue weight in bud-I, anthesis and 30% blooming flowers, respectively (Table 4.5). In other words, both 1,8-cineole and LiCINS mRNA levels remained at low levels, while linalool content and *LaLINS* mRNA abundance increased during flower development as previously reported (Lane et al. 2010).

4.3.5 Phylogenetic analysis

Based on its amino acid sequence, LiCINS was clustered into the TPSb subfamily of TPSs (Figure 4.10), which contains all angiosperm mTPSs including the CINS from *Salvia* and other mTPSs of *Lavandula* (Bohlmann et al. 1998; Chen et al. 2011; Demissie et al. 2011; Landmann et al. 2007). As expected mTPSs, sTPSs and dTPSs of conifers were
clustered together in subfamily TPSd while sTPSs of angiosperms were grouped in subfamily TPSa.
4.4 Discussion

4.4.1 Transcript profiling and EST Selection

To facilitate the discovery of genes that control EO formation in lavenders, we used microarray-assisted transcript profiling to study the relative expression of ESTs isolated from oil glands of *L. x intermedia* flowers in various *Lavandula* species. In particular, we examined ESTs corresponding to genes involved in the MEP pathway, the predominant route producing precursors (IPP and DMAPP) for terpene production in lavender oil glands, and those catalyzing the formation of EO monoterpenes (i.e., mTPSs) in two experimental sets. The first experiment was designed to evaluate the expression of these genes in *L. x intermedia* oil glands isolated from flowers of three developmental stages - bud, anthesis and 30% bloom. In agreement with previous findings (Boeckelmann 2008; Lane et al. 2010), transcripts corresponding to the MEP pathway and mTPS genes (Figure 1.2 & 1.3) were less abundant in oil glands isolated from unopened flower buds compared to flowers at anthesis (Table 4.2, column A vs B). However, they were evenly abundant in oil glands of the later two stages (Table 4.2, column B vs C). Among the MEP pathway genes, however, transcripts corresponding to *DXS* and *DXR* were more abundant than the others. Although transcripts corresponding to most of the known lavender mTPSs were generally more abundant in developing flowers compared to unopened buds, the mRNA corresponding to *LaLINS* was substantially more abundant (over 47 fold) than those of other mTPSs in maturing flowers. These results are not surprising given that linalool is the most abundant essential oil constituent in lavender flowers, and that the expression of the linalool synthase gene is primarily regulated at the level of transcription (Lane et al. 2010).

In the second set of experiments, the abundances of EO-related transcripts were compared in mature flowers of *L. x intermedia*, *L. latifolia* and *L. angustifolia* in order to identify differentially expressed genes in these species. There was little variation in the abundance of the transcripts corresponding to the MEP pathway genes among the three species. However, mRNA levels for the three functionally characterized lavender mTPSs were much more variable (Table 4.2, column D vs E, column D vs F and column E vs F). For example, *LaLINS* and *LaLIMS* transcripts were much more abundant in *L. x intermedia* than *L. angustifolia* and *L. latifolia* flowers. This enhanced expression level could be a result
of “hybrid vigor”. Typically, *L. x intermedia* varieties (e.g., Grosso lavender) have a much better EO yield than those of either parents, a phenomenon attributed to hybrid vigor (Harborne and Baxter 2001). In this regard, the estimated overall oil content for *L. x intermedia* (cv Grosso) was 182.9 mg, while that of *L. angustifolia* (cv Lady) was only 68.1 mg per gram of fresh tissue (Boeckelmann 2008). The transcript profiling experiment revealed that the mRNA corresponding to one EST (later established as *LiCINS*) was more abundant in *L. latifolia* than both *L. x intermedia* and *L. angustifolia*. This result was confirmed by end-point PCR using the full-length primer set for CINS (Table 4.1). The amplified fragments were sequenced to confirm their identity. Given this unique expression pattern, the EST was selected for functional analysis.

4.4.2 Recombinant production and functional assay

The coding region of the putative *LiCINS* was expressed in bacterial cells in order to obtain the recombinant protein. The predicted N-terminal transit peptide, which resembled those found in other plants (Cai et al. 2010; Keegstra et al. 1989; Von Heijne et al. 1989), was excluded during cloning to enhance the solubility of the recombinant protein. A transit peptide targets heterologously expressed proteins in *E. coli* into the periplasmic cavity where proteins often aggregate and form insoluble inclusion bodies (Hannig and Makrides 1998). Its exclusion, therefore, enhances the deposition of the recombinant proteins in the bacterial cytoplasm in a soluble form (Williams et al. 1998). The estimated molecular weight of the recombinant LiCINS, according to its resolution on SDS-PAGE (Figure 4.3), was slightly greater than the theoretical molecular weight. Such inconsistencies between predicted and SDS-PAGE based molecular weight estimates are routinely observed (Fischer et al. 2004), and could result from post-translational modification of the heterologously expressed proteins. For example, the frequently observed gluconoylation of recombinant proteins in *E. coli* BL21(DE3) (Aon et al. 2008) increases the size of the expressed protein (Kim et al. 2001).

Like many other plant TPSs the recombinant LiCINS proved to be a multiproduct enzyme and produced mainly 1,8-cineole, along with sabinene, α-phellandrene, α-terpineol, and limonene as minor products. With the exception of α-phellandrene, these monoterpenes are common in the product mixes of other “cineole cassette”-type mTPSs (Fähnrich et al. 2001).
For example, the recombinant common sage (*Salvia officinalis*) CINS produced 1,8-cineole (the major product), α-pinene, β-pinene, myrcene and sabinene (Wise et al. 1998), and the recombinant *Nicotiana suaveolens* CINS produced the same products as well as (E)-β-ocimene and α-terpineol (Roeder et al. 2007). Further, the bacterially produced α-terpineol synthases from *N. alata* and *N. langsdorfii* produced α-terpineol as the major product, and smaller amounts of 1,8-cineole, α-pinene, myrcene and sabinene (Fähnrich et al. 2011). Given that the above TPSs produce products represented in the source plant, and that α-phellandrene is a minor constituent of the EO of *Lavandula*, it is not surprising that LiCINS can produce small amounts of this monoterpen. In this respect, all reported *Lavandula* mTPSs can produce low levels of α-phellandrene *in vitro* (Demissie et al. 2011; Landmann et al. 2007).

Some plant mTPSs utilize both GPP and NPP as a substrate. In this context, NPP was shown to be an effective substrate for some plant TPSs, including those recently reported from glandular trichomes of tomato (Schilmiller et al. 2009), and LaßPHLS (Demissie et al. 2011). LiCINS also utilized both GPP and NPP as substrates to produce 1,8-cineole as its major product *in vitro*. However, GPP was the preferred substrate for the enzyme, which produced more of the minor products when fed with NPP. In particular, the production of sabinene, α-phellandrene, α-terpineol and limonene were increased from 7.9%, 6.6%, 1.7% and 2.3%, respectively (for GPP), to 18.5%, 8%, 5.5% and 3.3% (for NPP) of the product mix, respectively. These increases were accompanied by a reduction in the production of 1,8-cineole (Figure 4.4a & b). The enhanced production of minor products could be due to the geometric configuration of NPP, which may be a better substrate for their synthesis through direct cyclization (Chayet et al. 1984). The ability of TPSs (including LiCINS) to produce multiple products may have resulted from incomplete evolution of the active sites of the ancestral protein to achieve precision in its catalytic activity (Christianson 2006).

The optimum temperature and pH of LiCINS were in the range of those reported for other related mTPSs. The optimum temperature of 30 °C was similar to LaLiNS, LaLiMS and LaßPHLS, while the optimal pH of 6.5 was similar to LaßPHLS (Demissie et al. 2011) and linalool synthase of bergamot mint (Crowell et al. 2002). The $K_m$ of LiCINS for GPP (5.75 µM) was close to those of most mTPSs reported. For example, the $K_m$ of the LaßPHLS was 6.75 µM (Demissie et al. 2011), and that of LaBERS was 4.7 µM for FPP (Landmann et
al. 2007). Further, two snapdragon flower mTPSs exhibited a $K_m$ of 7.57 and 7.68 µM for GPP, (Nagegowda et al. 2008), while the $K_m$ of sabinene and CINSs cloned from Common Sage (*Salvia officinalis*) were determined to be 7.0 and 7.4 µM, respectively (Wise et al. 1998). Finally, the catalytic activity of the enzyme $(8.8 \times 10^{-3} \text{s}^{-1})$ was close to the $0.01 – 0.1 \text{s}^{-1}$ ranges reported for other plant mTPSs (Wise and Croteau 1999).

4.4.3 Structural relation to other plant TPSs

The biosynthesis of cyclic terpenes from GPP occurs in three stages. These include the rearrangement and isomerization of GPP to the highly reactive intermediate linalyl diphosphate (LPP); re-ionization of LPP to produce a cyclic $\alpha$-terpinyl cation; and finally, conversion of the cation to the end product (for example, 1,8-cineole). Four conserved structural motifs of mTPSs have experimentally confirmed roles in these processes (Figure 4.1). Two of these motifs, RR(x8)W and DDxxD, are fully conserved in most mTPSs across species. The DDxxD motif serves as a binding site for a divalent metal ion cofactor, often Mg$^{2+}$, required for the ionization and isomerization of GPP into LPP by the RR(x8)W motif (Christianson 2006; Degenhardt et al. 2009; Dewick 2009; Roeder et al. 2007; Wendt and Schulz 1998). Like many other previously reported mTPSs, such as LaLINS and LaLIMS (Landmann et al. 2007), and LaßPHLS (Demissie et al. 2011) of *L. angustifolia*, LiCINS fully retained these motifs. However, slight alterations were observed in the catalytic motif LQLYEASFLL and the second divalent metal ion binding site (N,D)D(L,I,V)x(S,T)xxxE. For example, the C-terminus leucine residue of the catalytic motif was replaced by serine residue in LiCINS (Figure 4.1). Similar changes were reported for LaßPHLS (Demissie et al. 2011) and LaLINS (Landmann et al. 2007) of *L. angustifolia*, CINSs of *Nicotiana suaveolens* (Roeder et al. 2007), *Citrus unishu* (Shimada et al. 2005) and *Salvia officinalis* (Wise et al. 1998), and $\beta$-phellandrene synthase of grand fir (Bohlmann et al. 1999).

Although TPSs of a given species are often more related to one another than they are to TPSs of similar function in genetically distant plants, functionally similar TPSs of closely related species (e.g., *Lavandula* and *Salvia*) display higher sequence similarity to each other than to those with different functions (Bohlmann et al. 1998; Chen et al. 2011; Trapp and Croteau 2001). In this context, the CINSs reported here were closer in amino acid sequence to the CINSs from *Salvia* than to most of the *Lavnadula* TPSs. For example, LiCINS
exhibited ~77% homology to the CINSs of *Salvia fruticosa* (ABH0767.1 and ACM89961.1) and *Salvia officinalis* (AAC26016.1) (Figure 4.1), while it displayed a ~65% sequence similarity to LaLINS (ABB73045.1) and 64% to LaLIMS (ABB73044.1) of *L. angustifolia* (Figure 4.1). An exception was that LiCINS was highly homologous (over 90% at the amino acid level) to the *L. angustifolia* β-phellandrene synthases, indicating that these two enzymes are very closely related and likely evolved from one another, or from the same (relatively recent) parent.

4.4.4 Phylogenic analysis and genomic organization

Based on the protein primary structure plant TPSs are grouped into six subfamilies, where TPSs cloned from the same or closely related species are rooted together within the same subfamily regardless of their catalytic properties (Bohmann et al. 1998; Chen et al. 2011; Trapp and Croteau 2001). In this regard we have previously reported that LaβPHLS is placed in subfamily TPSb and closely rooted with mTPSs of *Lamiaceae*, including *Lavandula* and *Salvia* (Demissie et al. 2011). Similarly, the three CINSs of *Lavandula* were closely rooted with LaβPHLS and CINSs of *Salvia*, while CINSs cloned from *C. unshiu*, *N. suaveolens* and *A. thaliana* (that are genetically more distant from *Lamiaceae*) were rooted in a separate clade within TPSb (Figure 4.10).

Angiosperm, including *Lamiaceae*, mTPSs contain seven exons and six introns, and are classified in class III clade (Landmann et al. 2007; Lee and Chappell 2008; Trapp and Croteau 2001). The genomic DNA of the CINSs cloned from *L. angustifolia*, *L. latifolia* and *L. x intermedia* included seven exons and six introns, and fell in this category. Consistent with the organization of the limonene synthase gene from *P. frutescens*, also classified in class III clade, exon3 is the longest, while exon5 is the shortest exon in these genes (Trapp and Croteau 2001). In both *P. frutescens* limonene synthase and CINS genes, the RR(x8)W and DDxxD motifs are placed on the first and third exon, respectively. However, contrary to the phase 0 placement for all *P. frutescens* limonene synthase introns, the 3rd and 4th introns of the CINS genes were placed on the second nucleotide of their proximate codon. This could have been resulted either from intron sliding (Mathews and Trotman 1998) due to insertion or deletion of a single nucleotide in these introns, or from base calling errors during sequencing.
4.4.5 Transcriptional activity and inheritance

It has previously been shown that production of certain monoterpenes is regulated through transcriptional control of the corresponding genes. For example, the production of menthofuran in peppermint directly correlated with the abundance of menthofuran synthase mRNA (Mahmoud and Croteau 2003). Also, Lane et al. (2010) established a direct relation between LaLINS transcript level and quantity of linalool in the EO of L. angustifolia flower. Further, Boeckelmann (2008) reported a concerted increase in accumulation of LaLINS transcript and linalool (product of LaLINS) in L. angustifolia and L. x intermedia plants during flower development. Similar results were also reported in other plants (Dudareva et al. 2005; Turner et al. 1999; Turner et al. 2000b). In the present study, LaLINS transcript levels paralleled the tissue linalool content and increased with flower age. On the other hand, the abundances of 1,8-cineole and LiCINS transcripts did not change during flower development (Figure 4.9c). Consistent with the present result, Boeckelamnn (2008) reported that the concentration of some of L. x intermedia EO components, including borneol and camphor, did not significantly change during flower development. Our data indicate that LiCINS might be transcriptionally regulated, although detailed experiments must be performed to examine the possible involvement of other regulatory mechanisms.

To establish the parental origin of the expressed LiCINS, we obtained several cDNA and genomic clones from each of the three lavender species studied. The nucleotide and amino acid sequences, and the genomic organization of the gene were highly conserved among L. x intermedia and its parents, except for a single polymorphic nucleotide in the coding region of the L. angustifolia CINS ortholog. LaCINS coded for a histidine residue at this position instead of a tyrosine residue found in those of the other two species. This substitution did not detectably alter the product profile of the enzyme in vitro. Conservation of exon/intron structure between genes with similar function is a common phenomenon in plants (Hardison 1996). For example, the plant CYCD gene, that encodes a D-type cyclin, has identical genomic architecture across angiosperms (Menges et al. 2007). The non-coding sequences of the three genomic CINS clones were also highly conserved, although a few nucleotide substitutions were observed in L. angustifolia introns. This phenomenon (i.e., nucleotide substitutions in non-coding regions) is particularly common in genes involved in secondary metabolism (Kulheim et al. 2009). The presence of essentially the same CINS
gene in *L. latifolia* and *L. angustifolia* implies that these plants are closely related and most likely share a close common ancestor.
Figure 4.1 Alignment of *Lavandula* 1,8-cineole synthases with mTPSs of *L. angustifolia*, and 1,8-cineole synthases of *Salvia*.
Figure 4.1 Alignment of *Lavandula* 1,8-cineole synthases with mTPSs of *L. angustifolia*, and 1,8-cineole synthases of *Salvia*. Bold letters indicate conserved motifs. Identical amino acid residues are marked by asterisks, conserved amino acids by semicolons, and semi-conserved amino acid by period. LaβPHLS (ADQ73631.1): β-phellandrene synthase from *L. angustifolia*, LaLIMS (ABB73044.1): limonene synthase from *L. angustifolia*, LaLINS (ABB73045.1): linalool synthase from *L. angustifolia*, LaCINS (JN701461): 1,8-cineole synthase from *L. angustifolia*, LiCINS (JN701459): 1,8-cineole synthase from *L. x intermedia*, LiCINS (JN701460): 1,8-cineole synthase from *L. latifolia*, SfCINS (ABH07677.1): 1,8-cineole synthase from *S. fruticosa*, and SoCINS (AAC26016.1): 1,8-cineole synthase from *S. officinalis*. 
Figure 4.2 Transcriptional activity of MEP-pathway genes in leaves and floral tissues of *L. angustifolia, L. x intermedia* and *L. latifolia* measure by PCR.

Abbreviations stand for: CMK, 4-(cytidine-5'-diphospho)-2-C-methylerythritol kinase; DXR, 1-deoxyxylulose-5-phosphate reductoisomerase; DXS, 1-deoxyxylulose-5-phosphate synthase; FL, 30% flower; GPPS, geranyl pyrophosphate synthase; HDR, 1-hydroxy-2-methyl-E-butene-4-diphosphate reductase; HDS, 1-hydroxy-2-methyl-E-butene-4-diphosphate synthase; IPP, isopentenyl diphosphate; IPPI, IPP isomerase; LF, leaf and MDS, 2-C-methylerythritol-2,4-cyclodiphosphate synthase.
Figure 4.3 SDS-PAGE analysis of protein samples from bacterial cells expressing LiCINS, and those transformed with the empty expression vector. (a) protein marker, (b) total protein from cells expressing LiCINS, (c) soluble proteins from cells expressing LiCINS, (d) purified LiCINS and, (e) purified protein (GST) from cells transformed with the empty pET41(b+) vector (Figure A.1).
Figure 4.4 GC chromatograms and mass spectra of products produced by the recombinant LiCINS.
Figure 4.4 GC chromatograms and mass spectra of products produced by the recombinant LiCINS from GPP (a) and NPP (b), and for authentic 1,8-cineole standard (c). Peaks correspond to: 1) 1,8-cineole, 2) sabinene, 3) α-terpineol, 4) α-phellandrene, 5) limonene and 6) linalool.
Figure 4.5 GC chromatograms of Ni-NTA affinity chromatography purified soluble fraction from induced cells transformed with empty pET41(b+) product from GPP. Peak (1) is linalool and asterisks represent peaks without hit in the National Institute of Standards and Technology (NIST) library.
Figure 4.6 Kinetic assay of LiCINS.
Figure 4.6 Kinetic assay of LiCINS with GPP: a) time course assay of LiCINS activity, b) effect of pH on LiCINS activity, c) effect of temperature on LiCINS activity and d) velocity of LiCINS at increasing GPP concentrations.
Figure 4.7 Multiple alignments of 1,8-cineole synthase cDNAs of *L. latifolia*, *L. angustifolia* and *L. x intermedia*.

Asterisks indicate conserved nucleotides in the three cDNAs while the polymorphic nucleotides are in bold and bigger font size.
Figure 4.8 Schematic representation of LiCINS genomic DNA.

Exons are denoted by rectangular boxes (Exon1 through Exon7), and introns are denoted by lines connecting adjacent exons (I1 through I6). Numbers inside the box indicate the number of amino acids encoded by that exon. The four conserved motifs are given below the exons, and the black arrows indicate their approximate position. Note: the conserved motif (N,D)D(L,I,V)x(S,T)xxxE amino acids are partly encoded by “Exon 6” and partly by “Exon 7”.
Figure 4.9 The transcriptional activity of LiCINS, LaβPHLS, LaLINS and the reference gene (β-actin) in lavenders measured by PCR.
Figure 4.9 The transcriptional activity of LiCINS, LaβPHLS, LaLINS and the reference gene (β-actin) in lavenders measured by PCR. PCR amplified fragments corresponding to LiCINS, LaβPHLS, LaLINS and the reference gene (β-actin) in L. angustifolia, L. x intermedia and L. latifolia (a) in floral tissue, (b) in leaf tissue and (c) the relative transcriptional activity of LaLINS and LiCINS in secretory cells isolated from developing floral tissues of L. x intermedia relative to bud-I floral stage (see Boeckelmann 2008 or Lane et al. (2010) for detailed description of floral developmental stages). In figures a & b letters in upper-case denote: A) 1 kb DNA ladder (NEB, Canada), B) LaβPHLS in L. angustifolia, C) LaβPHLS in L. x intermedia, D) LaβPHLS in L. latifolia, E) LiCINS in L. angustifolia, F) LiCINS in L. x intermedia, G) LiCINS in L. latifolia, H) LaLINS in L. angustifolia, I) LaLINS in L. x intermedia, J) LaLINS in L. latifolia, K) β-actin in L. angustifolia, L) β-actin in L. x intermedia and M) β-actin in L. latifolia. Relative expression (c) was normalized to β-actin and error bars indicate standard deviation (n = 3).
Figure 4.10 Phylogenetic relationship and classification of TPSs.
Figure 4.10 Phylogenetic relationship and classification of TPSs. TPSs within the same class share a minimum of 50% amino acid identity. The scale bar represents 0.5 amino acid substitutions per site. All angiosperm mTPSs, including LiCINS, were grouped under the TPSb class of TPSs and LiCINS, LiCINS and LaCINS were closely rooted with LaβPHLS. Accession numbers of terpene synthases used to generate the phylogenetic tree are: LiCINS: JN701459; LiCINS: JN701460; LaCINS: JN701461; beta-Phellandrene_S_L._angustifolia: HQ404305; Sabinene_S_S._officinalis: AAC26018.1; Bornyldiphosphate_S_S._officinalis: AAC26017.1; Cineole_S_S._officinalis: AAC26016.1; beta-Pinene_S_C._limon: AAM53945.1|AF514288_1; beta-Ocimene_S_C._unshiu: BAD91046.1; Valencene_S_V._vinifera: AAS66358.1; beta-Phellandrene_S_A._grandis: AAF61453.1|AF139205_1; beta-Phellandrene_S_P._abies: AAK39127.2; Limonene-alpha-pinene_S_A._grandis: AAF61455.1|AF139207_1; Pinene_S_Q._ilex: CAK55186.1; GermacreneA_S_V._vinifera: ADR66821.1; Germacrene_S_S._lycopersicum: AEM05858.1; delta-Cadinene_S_G._hirsutum: AAC12784.1; beta-Caryophyllene_S_C._sativus: AAU05952.1; delta-Cadinene_S_G._arboreum: AAA93064.1; trans-alpha-Bergamotene_S_L._angustifolia: ABB73046.1; Linalool_S_L._angustifolia: ABB73045.1; Limonene_S_L._angustifolia: ABB73044.1; Cineole_S_R._officinalis: ABI20515.1; Cineole_S_S._fruticosa: ABH07677.1; Cineole_S_A._thaliana: AAU01970.1; Linalool_S_L._latifolia: ABD77417.1; Carene_S_S._stenophylla: AAM89254.1|AF527416_1; Limonene_S_M._longifolia: AAD50304.1|AF175323_1; Limonene_S_P._citridora: AAF65545.1; Myrcene_S_P._frutescens: AAF76186.1; Pinene_S_R._officinalis: ABP01684.1; Sabinene_S_S._pomifera: ABH07678.1; Terpinolene_S_O._basilicum: AAV63792.1; beta-Myrcene_S_O._basilicum: AAV63791.1; Fenchol_S_O._basilicum: AAV63790.1; Geraniol_S_P._frutescens: ABB30218.1; Linalool_S_P._setoyensis: ACN42009.1; Copalyldiphosphate_S_S._trichocarpa: XP_002306777.1; Copalyldiphosphate_S_S._lycopersicum: BAA84918.1; ent-Copalyldiphosphate_S_T._aestivum: BAH56558.1; Taxadiene_S_T._brevifolia: AAC49310.1; Abietadiene_S_A._grandis: AAK83563.1; beta-Farnesene_S_P._menziesii: AAX07265.1; gamma-Bisabolene_S_P._menziesii: AAX07266.1; alpha-Bisabolene_S_A._grandis: AAC24192.1; alpha-Bisabolene_S_P._abies: AAS47689.1; Linalool_S_C._breweri: AAD19840.1; Linalool_S_C._concinna: AAD19839.1; ent-
Kaurene_S_P._glaucu: ACY25275.1; ent-Kaurene_S_S._lycopersicum: AEP82778.1; Cineole_S_C._unshiu: BAD91045.1; Cineole_S_N._suaveolens: ABP88782.1; alpha-terpineol_S_V._vinifera: AAS79352.1 and beta-ocimene/myrcene_S_V._vinifera: ADR74206.1
<table>
<thead>
<tr>
<th>Primers</th>
<th>Target gene</th>
<th>Primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><em>LiCinS</em> set I (without transit peptide)</td>
<td>F1: 5’ - TATGATCCAAACGGGCGACGAT-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R1: 5’ - CGATTCGTAAGCGCTCGAACAAC-3’</td>
</tr>
<tr>
<td>2</td>
<td><em>LiCinS</em> set II (without transit peptide)</td>
<td>F2: 5’ - TGATCCAAACGGGCGACGAT-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R2: 5’ - AATTCGATTCGTAAGCGCTCGAACAAC-3’</td>
</tr>
<tr>
<td>3</td>
<td><em>LaLinS</em></td>
<td>F: 5’ - CCGCATATGTCGATCAATATCAACAT-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5’ - ATAGAATTCTGCGTACGGCTCGAACA-3’</td>
</tr>
<tr>
<td>4</td>
<td><em>β-Actin</em></td>
<td>F: 5’ - GCACGGAATTGTGAGCAATTGGGA-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5’ - TTATGTCCCTCAGATTTCCCGCT-3’</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td><em>LiCinS</em></td>
<td>F1: 5’ - CCAAGCCTCAGCCATGATAGA-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R1: 5’ - TTGACATCGATGGCTATCGTA-3’</td>
</tr>
<tr>
<td>2</td>
<td><em>LaLinS</em></td>
<td>F: 5’ - ACACGCACGACAATTTGCCA-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5’ - AGGCCCTCAAATGAAGTGAGAT-3’</td>
</tr>
<tr>
<td>3</td>
<td><em>β-Actin</em></td>
<td>F: 5’ - TGTGGATTGCCAAGGCAGAGT-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: 5’ - AATGAGCAGGCCAGCAACAGCA-3’</td>
</tr>
</tbody>
</table>
Table 4.2 Summary of microarray analysis experiment for relative expression of terpene EO biosynthesis related genes in *Lavandula* using microarray technology.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>A vs B</th>
<th>B vs C</th>
<th>D vs E</th>
<th>D vs F</th>
<th>E vs F</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>DXS</em></td>
<td>down (12.14)</td>
<td>ns</td>
<td>up (2.14)</td>
<td>down (2.39)</td>
<td>down (5.79)</td>
</tr>
<tr>
<td><em>DXR</em></td>
<td>down (12.24)</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>down (1.5)*</td>
</tr>
<tr>
<td><em>CMK</em></td>
<td>down (3.11)</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>down (1.39)</td>
</tr>
<tr>
<td><em>MDS</em></td>
<td>down (2.47)</td>
<td>ns</td>
<td>up (2.04)</td>
<td>ns</td>
<td>down (1.5)*</td>
</tr>
<tr>
<td><em>HDS</em></td>
<td>down (5.35)</td>
<td>ns</td>
<td>down (2.05)</td>
<td>ns</td>
<td>up (1.84)</td>
</tr>
<tr>
<td><em>HDR</em></td>
<td>down (5.56)</td>
<td>ns</td>
<td>ns</td>
<td>up (1.75)</td>
<td>up (2.21)</td>
</tr>
<tr>
<td><em>GPPS</em></td>
<td>down (3.57)</td>
<td>ns</td>
<td>ns</td>
<td>up (5.66)*</td>
<td>up (4.01)</td>
</tr>
<tr>
<td><em>IPPi</em></td>
<td>down (7.95)</td>
<td>ns</td>
<td>down (1.76)</td>
<td>ns</td>
<td>up (1.93)</td>
</tr>
<tr>
<td><em>LaLINS</em></td>
<td>down (47.31)</td>
<td>ns</td>
<td>up (11.23)</td>
<td>ns</td>
<td>down (7.72)</td>
</tr>
<tr>
<td><em>LaLIMS</em></td>
<td>ns</td>
<td>ns</td>
<td>up (34.13)</td>
<td>up (4.76)</td>
<td>down (58.95)</td>
</tr>
<tr>
<td><em>LiCINS</em></td>
<td>down (3.05)</td>
<td>ns</td>
<td>up (11.76)*</td>
<td>down (9.02)</td>
<td>down (34.79)</td>
</tr>
</tbody>
</table>

**A**: *L. x intermedia* glands isolated from floral tissues at bud stage, **B**: *L. x intermedia* glands isolated from floral tissues at anthesis stage, **C**: *L. x intermedia* glands isolated from floral tissues at 30% flowering stage, **D**: *L. x intermedia* floral tissues at 30% flowering stage, **E**: *L. angustifolia* floral tissues at 30% flowering stage and **F**: *L. latifolia* floral tissues at 30% flowering stage. For each of the samples below “down” means the gene is down-regulated relative to the comparison sample and “up” is vice versa. Numbers in bracket indicate average expression fold changes, n = 4. For example: *DXS* was 12.6 times down-regulated in sample “A” relative to sample “B”. Asterisk represents values derived from n=1 and ns means non-significant difference.
Table 4.3 Conserved amino acid level (% identity and % similarity) between LiCINS and mTPSs cloned from *Lavandula* and *Salvia*.

<table>
<thead>
<tr>
<th></th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
<th>% Identity</th>
<th>% Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>LiCINS (JN701459)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>LaβPHLS (ADQ73631.1)</td>
<td>84</td>
<td>90</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>LaLINS (ABB73045.1)</td>
<td>50</td>
<td>65</td>
<td>49</td>
<td>66</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>LaLIMS (ABB73044.)</td>
<td>47</td>
<td>64</td>
<td>47</td>
<td>66</td>
<td>63</td>
<td>77</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>SfCINS (ABH07677.1)</td>
<td>66</td>
<td>77</td>
<td>61</td>
<td>75</td>
<td>52</td>
<td>69</td>
<td>53</td>
<td>70</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>SoCINS (AAC26016.1)</td>
<td>66</td>
<td>77</td>
<td>61</td>
<td>75</td>
<td>52</td>
<td>70</td>
<td>52</td>
<td>69</td>
<td>95</td>
<td>97</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 4.4 Raw data and kinetics parameters of LiCINS.

<table>
<thead>
<tr>
<th>GPP (µM)</th>
<th>µmole sec(^{-1})</th>
<th>pmole sec(^{-1}) mg(^{-1}) of enzyme (pKat mg(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rep1</td>
<td>Rep2</td>
</tr>
<tr>
<td>5</td>
<td>4.89E-07</td>
<td>5.54E-07</td>
</tr>
<tr>
<td>10</td>
<td>6.13E-07</td>
<td>6.23E-07</td>
</tr>
<tr>
<td>25</td>
<td>8.33E-07</td>
<td>9.05E-07</td>
</tr>
<tr>
<td>50</td>
<td>9.48E-07</td>
<td>9.69E-07</td>
</tr>
<tr>
<td>75</td>
<td>1.02E-06</td>
<td>1.05E-06</td>
</tr>
<tr>
<td>100</td>
<td>9.25E-07</td>
<td>1.14E-06</td>
</tr>
<tr>
<td>200</td>
<td>9.74E-07</td>
<td>1.01E-06</td>
</tr>
</tbody>
</table>

Enzyme kinetic parameters

<table>
<thead>
<tr>
<th>K(_m) (µM)</th>
<th>V(_{max}) (µM sec(^{-1}))</th>
<th>V(_{max}) (pKat mg(^{-1}))</th>
<th>k(<em>{cat}) (sec(^{-1})) = V(</em>{max})/[E]</th>
<th>k(_{cat})/K(_m) (µM(^{-1}) sec(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.75 ± 0.91</td>
<td>1.05E-06</td>
<td>105.43 ± 3.01</td>
<td>8.8E-03 ± 4.9E-8</td>
<td>1.53E-03</td>
</tr>
</tbody>
</table>

\(k_{cat}\) (sec\(^{-1}\)) = V\(_{max}\)/[E]

Since the molecular weight of our protein is 63.2 kDa or 63,200 µg/µmole and we used 7.6 µg of LiCINS or 10 µg of Ni-NTA purified protein of which ≈76% was our protein of interest

\[ [E] = 7.6 \text{ µg}/63,000 \text{ µg}/\mu\text{mole} \]
\[ = 0.00012 \text{ µmole} \]

Then, \(k_{cat}\)

\[ = 1.05E-06 \text{ µmole sec}^{-1}/0.00012 \text{ µmole} \]
\[ = 8.8E-03 \text{ sec}^{-1} \]

\(k_{cat}/K_m\)
\[ = 8.8E-03 \text{ sec}^{-1}/5.75 \text{ µM} \]
\[ = 1.53E-03 \text{ µM}^{-1} \text{ sec}^{-1} \]
Table 4.5 Average 1,8-cineole and linalool content (mg per gm of fresh tissue) of oils distilled from three developmental stages of *L. x intermedia* cv Grosso.

<table>
<thead>
<tr>
<th>Flower tissue</th>
<th>Average mg 1,8-cineole per gm of fresh tissue</th>
<th>Average mg linalool per gm of fresh tissue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bud-I</td>
<td>2.9 ± 0.35</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td>Anthesis</td>
<td>2.8 ± 1.01</td>
<td>5.9 ± 0.14</td>
</tr>
<tr>
<td>30% flower</td>
<td>2.3 ± 0.35</td>
<td>11.25 ± 0.71</td>
</tr>
</tbody>
</table>
Chapter 5 Identification and functional characterization of *Lavandula angustifolia* 
ABCB1 (*LaABCB1*) that is differentially inhibited by lavender *de novo* monoterpenes in *Xenopus laevis* oocytes

5.1 Overview

Lavenders (*Lavandula*) are members of the *Lamiaceae* family widely grown for their essential oils (EOs), which are mainly constituted of a few monoterpenes, the C\textsubscript{10} class of the isoprenoids or terpenoids. For example, the EOs of *L. angustifolia* are dominated by the monoterpenes linalool (25 − 38%), linalool acetate (25 − 45%), 1,8-cineole (0 − 1.5%) and camphor (0 − 0.5%) (Demissie et al. 2013; Lis-Balchin 2002). These monoterpenes play significant biological, physiological and ecological roles such as pollinator attractants, and allelopathic and defense compounds (Cavanagh and Wilkinson 2002; Upson and Andrews 2004; Woronuk et al. 2011); however, they are also potentially toxic to cells when produced at higher concentrations (Caissard et al. 2004; Goossens et al. 2003; Goossens et al. 2003). Hence, their biosynthesis as well as accumulation is limited to specialized tissues called glandular trichomes or oil glands (Gershenzon et al. 2000; McCaskill and Croteau 1995; McConkey et al. 2000). In these glands a group of six to eight secretory cells actively produce and secrete the monoterpenes into an epicuticular storage cavity that is demarcated by a plasma membrane (Gershenzon et al. 2000; McCaskill et al. 1992; McCaskill and Croteau 1995; McConkey et al. 2000; Turner and Croteau 2004) (Figure 1.5c). Although the exact mechanism and transport proteins involved in this process have yet to be identified, the secretion of monoterpenes into the storage cavity, as observed in menthol trafficking in peppermint glandular trichomes, does display characteristic features of ATP binding cassette (ABC) transporter mediated transport (McCaskill et al. 1992; Turner and Croteau 2004).

ABC transporters are a large family of membrane proteins that utilize energy derived from direct ATP hydrolysis to actively transport structurally diverse compounds. Plant genomes encode more ABC transporter proteins than any other organism (Kanga et al. 2011; Rea 2007) and putative ABC transporter transcripts are well represented in expressed sequence tag (EST) libraries derived from gland secretory cells of *Lavandula* (Demissie et al. 2012; Lane et al. 2010), peppermint (Lange et al. 2001), and hop (Wang et al. 2008). Thus, identification of ABC transporters involved in monoterpane transport has been under intensive research focus recently. In spite of this, a plant ABC transporter capable of
transporting monoterpenes has yet to be reported (Kanga et al. 2011; Rea 2007, Theodoulou 2000).

Current efforts aimed at identifying potential monoterpene (and isoprenoid) transporter proteins have placed emphasis on the full–length members of the ABCG subfamily, also known as pleiotropic drug resistance transporters (PDRs). This is due to reports that the transcriptional activity of PDRs in *N. tabacum* and *N. plumbaginifolia* (Jasinski et al. 2001; Sasabe et al. 2002; Smart and Fleming 1996; Stukkens et al. 2005) have been found to correlate with isoprenoid synthesis and also that PDRs are the major ABC transporters expressed in tissues specialized for isoprenoid accumulation and/or release (Bienert et al. 2012; Bultreys et al. 2009; Ruocco et al. 2011). In addition, heterologous over-expression of PDR, conferred resistance to some diterpenes like sclareol, while silencing them increased host susceptibility (Crouzet et al. 2013; van den Brule et al. 2002). Interestingly, although NtPDR1, a PDR extensively studied in this regard, was isolated from tissues known to accumulate and/or release monoterpenes in response to fungal attack, its over-expression in *N. tabacum* BY2 cells did not confer resistance against tobacco *de novo* isoprenoids like geranylgeraniol and 1,8-cineole (Crouzet et al. 2013). In fact, the only PDR cDNA that conferred resistance to monoterpenes *in vitro* was isolated from the mountain pine beetle-fungal symbiont *Grosmannia clavigera* (Wang et al. 2013). While important in the study of plant-pathogen interaction mechanisms and the evolutionary relationships of ABC transporters, the isolation of a fungal transporter involved in monoterpene resistance does not provide significant insight into their *in planta* transport role. In contrast, recent competitive inhibition studies have shown that the full–length members of the ABCB subfamily, also known as multidrug resistance transporters (MDRs) have the potential to transport monoterpenes. The *in vitro* accumulation of [3H]digoxin mediated by human MDR1 was inhibited by *Zanthoxyli Fructus* ethyl acetate extracts and its dominant monoterpene constituents. Of the 25 monoterpenes tested, (R)-(+)−citronellal, (S)(−)−β-citronellol, δ-terpinene, terpinolene and (−)−β-pinene inhibited digoxin transport (Yoshida et al. 2005; Yoshida et al. 2006; Yoshida et al. 2008). This selective inhibition of MDRs by monoterpenes implies that MDRs expressed in tissues specialized for monoterpene synthesis and accumulation could play a role in their transport. This also holds true for MDRs expressed in *Lavandula* glandular trichomes.
We have previously identified and functionally characterized several terpene synthases (Demissie et al. 2011; 2012; Sarker et al. 2012; 2013) and a novel cis-prenyl transferase (Demissie et al. 2013) from *L. angustifolia* and *L. x intermedia* leaf, flower and glandular trichome cDNA libraries (Demissie et al. 2013; Lane et al. 2010). With cDNAs encoding for their biosynthetic enzymes identified, the biosynthesis of major *Lavandula* monoterpenes is fairly well understood (Demissie et al. 2011; 2012, Landmann et al. 2007; Sarker et al. 2012; 2013), however, little is presently known regarding the secretion of these monoterpenes into the storage cavity. In an attempt to isolate proteins involved in *Lavandula* monoterpene secretion, we studied the above databases to identify ABCB (MDR) and/or ABCG (PDR) full–length homologs expressed in the glandular trichomes of *Lavandula*. Here, we report the isolation and functional characterization of the *L. angustifolia* ABCB1 cDNA (a multidrug resistance transporter homolog), termed *LaABCB1* in *Xenopus laevis* oocytes. Transcripts corresponding to *LaABCB1* were expressed in *L. angustifolia*, *L. x intermedia* and *L. latifolia* leaf and flower tissues, and in the secretory cells of *L. x intermedia* glandular trichomes. *Xenopus* oocytes injected with *in vitro* transcribed *LaABCB1* mRNA showed significantly higher efflux of the MDR generic substrate vinblastine compared to non-injected oocytes or those injected with nuclease free water. Vinblastine efflux in oocytes expressing *LaABCB1* was inhibited by treatment with an oxidative uncoupler (KCN) and ATPase inhibitor (Vanadate). Of the eight most common monoterpenes identified in *Lavandula* oils five of them inhibited the *LaABCB1* mediated vinblastine efflux better than verapamil, a known MDR transporter competitive inhibitor. Monoterpenes that displayed significant inhibition property were linalool, geraniol, sabinene, borneol and camphor while linalool acetate, 1,8-cineole and limonene showed no effect. Few plant ABC transporters have been cloned from tissues specialized for monoterpene biosynthesis and secretion so far. However, this is the first plant ABC transporter to be selectively inhibited by monoterpenes produced by the plant from which the cDNA is isolated. This result paves the way for the identification of other genes encoding for transporter proteins involved in monoterpene transport in lavenders and other plants.
5.2 Materials and methods

5.2.1 Materials

[3H]vinblastine sulphate (specific activity = 1.85 MBq) and vinblastine sulphate were purchased from American Radiolabelled Chemicals (St. Louis, MO), and (±)-verapamil hydrochloride, potassium cyanide, sodium orthovanadate, (±)-Linalool, linalool acetate, geraniol, sabinene hydrate, borneol, camphor, limonene and 1,8-cineole were purchased from Sigma, Canada. Female African clawed frog, *Xenopus laevis*, were purchased from VWR (Mississauga, ON, Canada) and maintained as per the UBC Animal Care SOPs and guidelines. The pXOON vector (Figure A.2) was kindly donated by Dr. Dmitri Y. Boudko, Department of Physiology and Biophysics, Rosalind Franklin University and pRSET::EmGFP (Figure A.3) was kindly donated by Dr. Mary Forrest, Department of Biology, The University of British Columbia (Okanagan).

5.2.2 Candidate selection and sequence analysis

Full–length ABC transporter homologs present in our EST databases were identified using the string “ABC transporter” as a search key, and/or using NtPDR1 and CjMDR1 nucleotides as a query for the “BLASTx” module of local blast search engine. A cDNA was designated “full–length” based on its size (~ 4 kb for MDRs and ~ 4.5 – 5 kb for PDRs) and alignment position to known plant ABC transporters, which led to the identification of a 4,027 bp long MDR homolog termed *LaABCB1* (KJ135790) from *L. angustifolia* cDNA library. Multiple alignment of *LaABCB1* with functionally characterized plant MDRs was generated using the default parameters of the TM-Coffee multiple aligning module available at http://tcoffee.crg.cat/apps/tcoffee/do:tmcoffee (Chang et al. 2012) and its phylogenetic relationship with other plant MDRs was deduced from a phylogenetic tree developed using default parameters of Tree module in Geneious® version 7.0.6 (Biomatters, San Francisco, CA). The trans-membrane regions of *LaABCB1* and its orientation were determined from the TOPO2 graphic output of the HMMTOP trans-membrane region prediction software available at http://www.sacs.ucsf.edu/cgi-bin/hmmtop.py (Tusnday and Simon 2001).
5.2.3 LaABCB1 transcriptional activity study and cloning

The LaABCB1 transcript abundance in leaf, flower and glandular trichome secretory cells of L. angustifolia and L. x intermedia – a hybrid result of L. angustifolia and L. latifolia – were assessed by standard PCR based on the intensity of LaABCB1 fragment amplified with set I primers (Table S1) and Taq DNA Polymerase (New England Biolabs, USA). The PCR program used was 95 °C for 1 min, followed by 30 cycles of 95 °C for 30 sec, 60 °C for 30 sec and 72 °C for 90 sec, and a 3 min final extension at 72 °C. The LaABCB1 open reading frame was then cloned in–frame into the BamHI/HindIII restriction sites of the Xenopus expression vector pXOON (Figure A.2) (Jespersen et al. 2002) using Sticky-End PCR cloning strategy as described by Zeng (1998). Briefly, the coding regions of LaABCB1 were amplified by PCR using set II & III cloning primers (Table 5.1) and KAPA HiFi DNA polymerase (Wilmington, MA, USA) in separate tubes. The PCR program used was 95 °C for 3 min, followed by 7 cycles of 95 °C for 1 min, 65 °C for 30 sec and 72 °C for 4 min, additional 35 cycles of 95 °C for 1 min, 60 °C for 30 sec and 72 °C for 4 min, and a 5 min final extension at 72 °C. The PCR products were purified using a Gel extraction/PCR purification kit (OMEGA bio-tek, USA). To generate sticky ends equal amounts of the purified PCR products were combined and denatured at 95 °C for 5 min followed by renaturation at room temperature for 30 min. The coding regions of EmGFP were amplified by PCR using cloning primers (Table 5.1) and KAPA HiFi DNA polymerase (Wilmington, MA, USA) using the following PCR program used was 95 °C for 3 min, followed by 40 cycles of 95 °C for 30 sec, 60 °C for 30 sec and 72 °C for 1 min, and a 5 min final extension at 72 °C. Restriction enzyme digested PCR product and pXOON plasmid were purified using a Gel extraction/PCR purification kit (OMEGA bio-tek, USA) and ligated using T4 DNA following manufacturers direction (NEB, NE). The forward cloning primer contains the Kozak consensus sequences (GCCGCC) upstream of the translation initiation codon while the HA-tag (YPYDVPDYA) encoding sequences were included in the reverse primer before the stop codon. The pXOON (Figure A.2) vector contains a T7 promoter, 5’ and 3’ end non-coding Xenopus β-globin sequences, a multiple cloning site and additional unique restriction sites downstream of the poly(A) tail for linearization. The in–frame cloning of LaABCB1 and EmGFP were re-confirmed by sequencing the ligated plasmids.
5.2.4 *In vitro* transcription

pXOON::LaABCB1 and pXOON::EmGFP plasmids were linearized at the XbaI restriction site situated downstream of the poly(A) tail and capped mRNAs were synthesized using the T7 mMESSAGE mMACHINE kit as per the manufacturers guidelines (Ambion, Austin, TX). The *in vitro* synthesized *LaABCB1* mRNA was purified using MEGAClear™ purification kit (Ambion, Austin, TX) following manufacturers procedure after the template DNA was degraded by DNase I enzyme (Ambion, Austin, TX). A sample of the mRNAs were resolved on a 1.5% agarose gel alongside a single stranded RNA ladder (NEB, NE) to confirm its purity and size.

5.2.5 Oocyte isolation, injection and expression

A female African clawed frog, *Xenopus laevis*, was anesthetised in RO water (pH 7.5) containing 0.75 mg/L MS-222 (Tricaine Methanesulfonate). Part of one ovary was removed, divided into small portions, and placed in Ca$^{++}$ free ND96 medium (96.0 mM NaCl, 2.0 mM KCl, 1.0 mM MgCl$_2$, 10.0 mM HEPES, pH 7.4) containing 2.0 mg/ml collagenase (Sigma, Canada). Oocytes were deflocculated in the collagenase solution for 1–1½ hour on a rocking platform and incubated overnight at 18 °C in collagenase free oocyte maintenance media (Ca$^{++}$ free ND-96 medium containing 2.5 mM sodium pyruvate, 2.5% horse serum, 1.8 mM CaCl, 50 mg/L gentamycin and 30 mg/L kanamycin) after preliminary sorting. Mature stage VI oocytes were injected with the *in vitro* synthesized *LaABCB1* mRNAs (50.6 nl, ~0.7 ng/nl), or nuclease-free water for control oocytes, by using a glass micropipette mounted to an automatic Drummond “NANOJECT II” nanoliter injector (Nanoject, Drummond Scientific, Broomall, PA). The glass micropipettes were prepared by pulling 3.5” Drummond borosilicate capillaries on a P-97 Flaming/Brown Micropipette Puller (Sutter instrument Co., Novato, CA, USA) using the following program: Heat = 570, pull = 30, velocity = 120 and time 200 then broke the tip under microscope. Injected oocytes were kept at 18 °C for 3–4 days in oocyte maintenance medium, which ensured over 97.5% survival. Media was replaced and damaged oocytes were separated from the healthy ones every day until used.
5.2.6 Vinblastine accumulation in oocytes

Initial vinblastine efflux studies were conducted by incubating groups of 15
LaABCB1 and water injected oocytes, three days after injection, in 100 µl efflux assay
solution at room temperature for 60 minutes. The efflux assay solution is oocyte
maintenance medium containing 2 µCi of [³H]vinblastine sulphate combined with different
concentrations of unlabeled vinblastine to make 0.5 – 100 µM vinblastine solutions. Uptake
was terminated by washing the oocytes three times with ice-cold oocyte maintenance
medium containing 500 µM unlabelled vinblastine and lysed in 150 µl of Ca⁺⁺ free ND96
medium containing 50 mM Tris HCl (pH 8.0), 5 mM EDTA and 2% SDS. Lysed cells were
added into 3 ml Ultima Gold™ scintillation cocktail (PerkinElmer, MA), and the
radioactivity level accumulated in LaABCB1 and nuclease free water injected oocytes were
determined using the Beckman LS 6000SC Liquid Scintillation Counter (Beckman Coulter,
Mississauga, ON). The amount of vinblastine accumulated (nmol) per oocytes was
calculated based on the disintegration per minute (DPM) value of 1 µl of the 10 µmol
[³H]vinblastine sulphate stock solution (Kelly et al. 1995; Strauss et al. 2013). The
LaABCB1 mediated vinblastine efflux amount was determined by calculating the difference
in the amount of vinblastine accumulated in water injected oocytes to that of LaABCB1
injected oocytes. The ATP-dependence of vinblastine efflux in LaABCB1 injected oocytes
was determined by including a metabolic inhibitor that depletes ATP (1 mM potassium
cyanide) and an ATPase blocker that competes for the ATP binding site of the transporter (1
mM vanadate) in the efflux assay solution. Time course – 15, 30, 45, 60 and 120 minutes –
accumulation levels were determined after exposing a batch of 15 oocytes to 10 µM
vinblastine containing 2 µCi [³H]vinblastine sulphate while the transport kinetic properties
were determined from accumulation data collected from oocytes incubated for 30 minutes in
oocyte maintenance medium containing 2 µCi of [³H]vinblastine sulphate combined with
different concentrations of unlabeled vinblastine to prepare 2.5, 10, 25, 50, 100 and 250 µM
solutions. Oocytes were lysed and the accumulated radioactivity levels were determined as
described above. The LaABCB1 mediated % efflux was calculated using the formula
\[
\% \text{ efflux} = \left( \frac{[\text{Vinblastine}]_{\text{water}} - [\text{Vinblastine}]_{\text{LaABCB1}}}{[\text{Vinblastine}]_{\text{water}}} \right) \times 100 \quad (\text{equation 5.1})
\]

while efflux amount was calculated as

\[
\text{efflux amount} = [\text{Vinblastine}]_{\text{water}} - [\text{Vinblastine}]_{\text{LaABCB1}} \quad (\text{equation 5.2})
\]

where \([\text{Vinblastine}]_{\text{water}}\) is Vinblastine amount in water injected oocytes and 
\([\text{Vinblastine}]_{\text{LaABCB1}}\) is Vinblastine amount in LaABCB1 mRNA injected oocytes.

5.2.7 Vinblastine transport inhibition by monoterpenes

The role of LaABCB1 in monoterpene transport was determined by measuring the inhibitory effect of monoterpenes on LaABCB1 mediated vinblastine efflux in injected Xenopus oocytes. A batch of 10 – 15 oocytes, three days after injection, were incubated at room temperature for 30 minutes in 100 µl efflux solution containing 10 µM vinblastine (including 2 µCi \([^3\text{H}]\text{vinblastine}\) sulphate and 500 µM of verapamil (classical MDR blocker) or lavender de novo monoterpane pure standards. Oocytes were lysed and accumulated radioactivity levels were determined as previously described. LaABCB1 mediated vinblastine efflux percent inhibition by verapamil, linalool or 1,8-cineole pure standards were determined from changes in efflux amount compared to that of uninhibited efflux assay.

5.2.8 Data analysis

One-way ANOVA and multiple mean separation modules of the GraphPad prism (version 7.0, Graphpad Software, San Diego, CA) were used to calculate p values and compare the means, respectively. The same software was used to calculate the \(J_{\text{max}}\) and \(K_t\) values by fitting the Michaelis–Menten equation to the efflux data (nmol/min/oocyte) by non-linear least square analysis.
5.3 Results

5.3.1 LaABCB1 sequence and phylogenetic analysis

Through homology-based searches we identified 13 MDR and 14 PDR homolog ESTs in our Lavandula EST databases (data not shown). Only one of the MDR candidates (now identified as LaABCB1), was found to be of full-length. The 4,027 bp LaABCB1 cDNA included a 3,726 bp ORF that encoded for a 1,241 amino acid protein with an apparent molecular mass of 136 kDa. The encoded protein contained two trans-membrane (TMD) and two nucleotide binding domains (NBD) arranged in TMD1-NBD1-TMD2-NBD2 order, a typical feature of MDR transporters. Each TMD contain included six predicted membrane-spanning regions, and the NBDs retained the ABC protein signature motifs Walker A (GX4GK[ST]) and Walker B (DEATSALD) boxes, as well as the LSGG motif, also referred as Walker C box, between Walker A and B (Figure 5.1 and 5.2).

Plant ABCB transporter subfamily members show a unique phylogenetic relationship among each other and are classified into clades I – III based on their sequence homology level (Takanashi et al. 2012). LaABCB1 was closely rooted with putative ABCBs from Solanum lycopersicum, Theobroma cacao, Medicago truncatula, and the A. thaliana homologs AtABCB2 and AtABCB10. LaABCB1 shares 89.2% amino acid similarity with its homolog from Solanum lycopersicum (XP_004245909.1), 88.9% with EOY25142.1 (the Theobroma cacao ABCB homolog), 88.2% with Medicago truncatula ABCB homolog (XP_003612850.1) and 76% with the A. thaliana AtABCB2 (NP_194326.2), but shares only 49% with the clade I prototype AtABCB1 (NP_181228.1). LaABCB1 and AtABCB1, an auxin efflux pump isolated from A. thaliana, are the only in vitro characterized members of clade I. ABCBs with the highest similarity with LaABCB1 have not been functionally characterized and though sequence similarity does not necessarily translate to functional similarity, it is worth noting that the volatile emissions of Solanum lycopersicum, Theobroma cacao and Medicago truncatula contain either linalool or its structural analogs myrcene and ocimene as their characteristic aroma feature. Clade II contains the most intra-related members and more have been functionally characterized than the other clades. The influx transport mediators AtABCB4 (the proto-type of the clade), AtABCB21, CjMDR1 and CjABCB2, and the alkaloid efflux mediator LjABCB1 belong to the clade II class of
plant ABCB sub-families. To date there are no functionally characterized transporters from clade III members (Figure 5.3).

Transcripts corresponding to LaABCB1 were detected in leaves, flowers and glands, but were most prominent in flower and gland tissues. The internal control β-actin was constitutively expressed in all tissues and the gene is conserved in L. x intermedia and both its parental lines L. angustifolia and L. latifolia (Figure 5.4). Due to the fact that monoterpenes have been shown to inhibit digoxin transport by human and mouse MDRs, and LaABCB1 was highly expressed in secretory cells of glandular trichomes, we decided to investigate this cDNA further.

5.3.2 LaABCB1 serves as an efflux pump

In vitro transcription of pXOON::LaABCB1 produced a single mRNA product with the correct size (Figure 5.5). We expressed this mRNA in Xenopus oocytes to determine its function. In vitro transcribed GFP mRNA was injected (in separate oocytes) alongside LaABCB1 as a positive expression control, and 95-100% of oocytes showed GFP expression when observed using laser scanning confocal microscopy (Figure 5.6). For initial efflux assays, LaABCB1 mRNA and water injected oocytes were incubated in assay solutions containing 0.5, 1.0, 10 or 100 µM vinblastine (a standard substrate for MDR transporters) for 60 minutes. LaABCB1 mRNA-injected oocytes incubated in 10 and 100 µM vinblastine assay solution pumped significantly lower amounts of vinblastine compared to oocytes incubated in 0.5 and 1 µM assay solution (Figure 5.7). Hence, subsequent experiments leading to the dose response assays were performed in assay solutions containing 10 µM vinblastine. Including 1mM KCN (an oxidative uncoupler) or Vanadate (ATPase inhibitor) in efflux solutions significantly increased (p < 0.05) the vinblastine amount retained in LaABCB1 mRNA injected oocytes compared to those incubated in efflux solution containing vinblastine alone (Figure 5.8).

Time course incubation of LaABCB1 mRNA and water injected oocytes in efflux assay solutions containing 10 µM vinblastine showed that 30 minutes was required for the LaABCB1 mediated vinblastine efflux to reach to its maximum level, 14.6 nmol per oocyte. The efflux amount gradually decreased as the incubation time increased from 30 to 60 minutes and approached the diffusion rate when oocytes were incubated for 120 minutes.
There was a significant difference in the efflux amount at 30 and 45 minutes (p < 0.05) as compared to water injected oocytes, but the difference was not significant when compared to the efflux amount attained at 15, 60 and 120 minutes (Figure 5.9). The substrate concentration dependent LaABCB1 mediated vinblastine efflux has a maximum efflux rate ($J_{\text{max}}$) of $1.34 \pm 0.34 \text{ nmol/min/oocyte}$ and a $K_t$ value of $20.3 \pm 18.7 \mu\text{M}$ (Figure 5.10). These values are within the range of previously characterized plant ABCB homologs (Shitan et al. 2003).

5.3.3 LaABCB1 mediated vinblastine efflux was differentially inhibited by monoterpenic constituents of Lavandula oils

To avoid monoterpenic toxicity/stress effects on oocytes during the inhibition assay, three replicates of 15 oocytes were incubated in oocyte maintenance media containing 100 – 500 µM of the selected monoterpenes at room T°C and assessed periodically. Oocytes were monitored under a stereo microscope for visible damage to the cell, color separation (healthy oocytes have a clearly separated dark and light brown side) and death at 15, 30, 45, 60, 90 and 120 minutes at room temperature, followed by overnight incubation at 18 °C. No oocytes showed the above stress signs either during the 2 hour room temperature incubation period or after the overnight incubation at 18 °C (data not shown). As a result, 500 µM of monoterpenes or verapamil (MDR competitive inhibitor) were included in the efflux assay solution to determine their inhibitory effects on LaABCB1 mediated vinblastine efflux. The inclusion of 500 µM verapamil, linalool, geraniol, sabinene, borneol and camphor in the efflux solution inhibited the vinblastine efflux capacity of LaABCB1 injected oocytes significantly (p < 0.05) compared to those incubated in assay solutions containing vinblastine alone. Verapamil reduced the LaABCB1 mediated vinblastine efflux by 76%, while camphor, borneol, linalool, geraniol and sabinene reduced the efflux amount by 100%, 99.3%, 93%, 92% and 88%, respectively (Figure 5.11). In contrast, linalool acetate, 1,8-cineole and limonene did not affect LaABCB1 mediated vinblastine efflux. There was no significant difference in the vinblastine efflux capacity of LaABCB1 injected oocytes incubated in solutions containing vinblastine alone or those containing 500 µM linalool acetate, 1,8-cineole and limonene (Figure 5.11).
5.4 Discussion

The trafficking of monoterpenes from their site of synthesis to their storage area in peppermint, a close relative of lavenders, is unidirectional and shows selective retention towards certain monoterpenes (McCaskill et al. 1992; Turner and Croteau 2004). For example, approximately 63% of the menthofuran produced in peppermint glands is selectively retained within the glandular secretory cells for further processing (McCaskill et al. 1992). Given that heterologously expressed PDR and MDR-type transporters have demonstrated roles in isoprenoid trafficking (Crouzet et al. 2013; Wang et al. 2013; Yoshida et al. 2005; 2006; 2008), and that the secretory cells of lavender oil glands accumulate ABC transporter transcripts (Demissie et al. 2012; Lane et al. 2010), we decided to investigate the potential role of these proteins in monoterpene transport. Since lavenders, like other plants (Kanga et al. 2011; Rea 2007), encode a large number of ABC transporter homologs, two criterions were used to limit the number of candidates: first, like all other isoprenoid biosynthesis related cDNAs, the candidate had to be expressed in secretory cells; and second, because previous findings have implicated the possible role of PDRs and MDRs in isoprenoid secretion, the candidate had to be a full–length PDR and/or MDR homolog. Our search resulted in 14 PDR and 13 MDR homolog ESTs, but only one MDR homolog cDNA was found to be full–length. The low abundance of full–length candidates is associated with the inherent limitations of a cDNA library construction technique to clone full–length mRNAs of large genes like PDRs and MDRs (Demissie et al. 2012; Lane et al. 2010).

Putative cDNAs involved in monoterpene biosynthesis are often selected for in vitro functional analysis based on their stronger transcriptional activity in secretory cells relative to other flower and leaf cell types (Demissie et al. 2011; 2012; 2013; Sarker et al. 2012; 2013). Similarly, a cDNA encoding for a transporter protein involved in monoterpene secretion is expected to follow a paralell transcriptional pattern. In this regard, previously characterized plant ABC transporters were selected based on their localized expression to specialized tissues and/or their over-expression in response to elicitor treatments (Bultreys et al. 2009; Sakai et al. 2002; Shitan et al. 2003; 2013, Smart and Fleming 1996; van den Brule et al. 2002; Yazaki et al. 2001). We also selected LaABCB1 for monoterpene transport assays because its corresponding cDNAs were abundant in the glandular trichomes of L. x intermedia. However, while the LaABCB1 fragment amplified from leaf tissues was weaker,
the band amplified from flower tissues was as intense as the one from secretory cells (Figure 5.4). This ubiquitous expression of MDRs could be due to their ability to transport many structurally diverse metabolites, which may be synthesized and/or accumulated in different plant cells. Plant MDRs have shown transport substrate promiscuity towards structurally unrelated endogenous/exogenous secondary metabolites (Goossens et al. 2003; Rea 2007; Yazaki 2006) including chlorophyll catabolites (Lu et al. 1998), auxins (Buda et al. 2013), lipids (Chen et al. 2011) and virulence resistance factors (Shang et al. 2009) among others. They also play a role in the transcriptional regulation of other ABC transporters (Carvajal et al. 1997) and ion guard channels (Klein et al. 2003). It is therefore expected, that MDRs be expressed in all flower cell types, including gland secretory cells, at a similar level. We expressed LaABCB1 in Xenopus laevis oocytes to determine its transport activity because, firstly, transcripts for this gene were present in cDNA libraries derived from secretory cells. Secondly, distant homologs of this gene (mouse and human MDRs) were shown to be inhibited by monoterpenes derived from Zanthoxyli fructus, which suggests their role in monoterpene trafficking (Yoshida et al. 2005; 2006; 2008).

Although Xenopus laevis oocytes are the preferred heterologous hosts for membrane protein expression, using monoterpenes as a substrate to measure efflux through LaABCB1 would be extremely challenging for various reasons. Firstly, monoterpenes are considerably volatile, making direct injection of known monoterpene concentrations into oocytes difficult, and secondly their volatility makes assaying their efflux equally challenging. Therefore, we opted to characterize LaABCB1 mediated transport using [3H]vinblastine, a known MDR substrate and followed the standard inhibition assay method to determine the inhibitory properties of Lavandula monoterpene constituents. LaABCB1 injected oocytes incubated for 30 minutes in 10 μM efflux solution accumulated 24% less vinblastine than water-injected oocytes and the efflux rate gradually decreased as incubation time increased (Figure 5.9). The efflux difference was inhibited by including the oxidative uncoupler KCN or ATPase inhibitor vanadate (Figure 5.8), indicating that LaABCB1 mediated efflux is energy dependent and an active transport process. In addition, inhibition of LaABCB1 mediated efflux by verapamil, a classical MDR competitive inhibitor (Figure 5.11); established LaABCB1 as a plant ABCB full–length (MDR) efflux mediator. The J_max (1.34 nmol/min/oocyte) and K_i (20.3 μM) values of LaABCB1 (Figure 5.10) were also within the
range of previously reported plant ABCBs. For example the $K_t$ and $J_{max}$ values of Cjmdr1-injected oocytes for berberine influx were $54.62 \pm 5.46 \, \mu\text{M}$ and $0.75 \pm 0.015 \, \text{nmol/10 min/mg of protein}$, respectively (Shitan et al. 2003). This is in contrast to the efflux rate mediated by the human MDR1 under the same expression system for the same substrate. Oocytes expressing human MDR1 accumulated only 5-7\% of the vinblastine amount accumulated in water-injected oocytes after 60 minutes incubation in 1 \mu M efflux solution (Castillo et al. 1990). It could be argued that the efflux capacity difference between LaABCB1 and MDR1 could have been partly attributed to the 10-fold higher vinblastine concentration we used in our assay solution. However, this is unlikely as LaABCB1 mRNA and water-injected oocytes incubated in 1 \mu M efflux solution for 60 minutes during our initial assay did not show a significance difference in vinblastine accumulation (Figure 5.7). This clearly implies a lower vinblastine extrusion capacity of LaABCB1 compared to the human MDR1.

One probable cause for the lower efflux capacity of LaABCB1 compared to human MDR1 could be a sub-optimal expression of the former due to source-host incompatibility. We could not quantitate LaABCB1 expression levels in oocytes using western blotting due to cross-reactivity of our secondary antibody with endogenous proteins (data not shown). However, lack of an appropriate heterologous host is one of the main reasons why very few plant ABC transporters have been functionally characterized, and low expression or complete failure has been previously reported (Yazaki et al. 2006). One such example was the unsuccessful attempt by the authors who expressed CjMDR1 (BAB62040.1) in Xenopus laevis oocytes to express CjABCB2 (BAM11098.1), despite the two proteins sharing 91\% amino acid similarity (Shitan et al. 2013). The efflux capacity difference may also be due to low LaABCB1 affinity towards vinblastine compared to MDR1. In fact the two homologs share only 37\% amino acid similarity (data not shown) and vinblastine (an indole alkaloid) is not produced in lavender de novo. LaABCB1 showed selective inhibition even towards monoterpenes present in L. angustifolia oil. In the inhibition study, linalool, geraniol, sabinene, borneol and camphor reduced the vinblastine efflux mediated by LaABCB1 to background levels while linalool acetate, limonene and 1,8-cineole did not (Figure 5.11). To our surprise, linalool acetate showed efflux stimulation capacity, though it did not result in significantly higher accumulation compared to oocytes incubated in efflux assay solutions.
containing 10 µM vinblastine alone (Figure 5.11). Contrary to LaABCB1, none of the monoterpenes tested in this study modulated [³H]digoxin-transport activity of the human MDR1 (Yoshida et al. 2006). This demonstrates that despite the ability to transport structurally diverse compounds, MDRs show preferential selectivity towards some compounds and LaABCB1 could have low affinity for vinblastine compared to the human MDR1. In addition, as noted before, possible differential transport of monoterpenes (menthofuran and menthol) has been reported before (McCaskill et al. 1992; Turner and Croteau 2004). Our findings parallel this report, and suggest that multiple transport pathways for monoterpenes may exist in plant oil glands.

Finally, we investigated the phylogenetic relationship of LaABCB1 to other ABCB homologs. LaABCB1 was closely rooted (78 – 81%) with putative ABCBs isolated from Solanum lycopersicum, Theobroma cacao and Medicago truncatula (Figure 5.3). These species are known to accumulate and/or release linalool and/or its structural analogs myrcene and ocimene. While linalool is the characteristic feature of volatile emissions of Solanum lycopersicum (van Schie et al. 2007) and Theobroma cacao (Kadow et al. 2013), the volatile emissions of Medicago truncatula are dominated by myrcene and ocimene (Gomez et al. 2005; Navia-Gine et al. 2009). Furthermore, LaABCB1 was rooted in clade I with the only other in vitro characterized efflux pump from A. thaliana, AtABCB1 (Lin and Wang 2005), while the influx mediator ABCB homologs AtABCB4 (Terasaka et al. 2005), AtABCB21 (Kamimoto et al. 2012), CjMDR1 (Shitan et al. 2003) and CjABCB2 (Shitan et al. 2013) are rooted together in clade II. The only exception in this fascinating efflux-influx transporters clade distribution is the efflux mediator LjABCB1 (Takanashi et al. 2012) that was closely rooted with CjMDR1. Very few plant MDR homologs have been functionally characterized, limiting the ability to draw strong conclusion about the sequence similarity level association with efflux vs influx transporters evolution (Figure 5.3). The functional characterization of additional ABCB homologs will help to determine the structural and evolutionary relationship between efflux and influx transport systems. Developing a Lavandula global expression profile library using high throughput sequencing technologies could help to identify more full–length MDR and PDR homologs, and their functional characterization would help clarifying their role in essential oil production in plants.
Figure 5.1 Schematic representation of HMMTOP predicted LaABCB1 membrane topology model.
Figure 5.1 Schematic representation of HMMTOP predicted LaABCB1 membrane topology model. LaABCB1 is composed of two trans-membrane (TMD) and two nucleotide-binding (NBD) domains. Each TMD is constituted of six trans-membrane helixes and the two NBDs contain the conserved motifs Walker A, Walker B and LSGGQ. Brown cylinders represent trans-membrane helixes of TMD-1 and purple cylinders represent those from TMD-2, NBD-1 region was shaded by light pink background and NBD-2 with light blue background with the conserved motifs Walker A, Walker B and LSGGQ denoted by blue oval shape, purple oval shape and green rectangle, respectively. The light blue background represents the extracellular area while the light yellow background represents the intra-cellular (cytoplasmic) area. HMMTOP default parameters were used.
Figure 5.2 Multiple alignment of LaABCB1 with selected functionally characterized ATP binding cassette subfamily B members from *Arabidopsis thaliana* and *Coptis japonica*.

The conserved motifs, Walker A, Walker B and LSGGQ, are indicated in bold letters with their respective label on top of the sequences. Amino acids shaded in yellow are inward facing (cytoplasmic) sequences, red are trans-membrane helixes (labeled as TM helix 1 through 12) and blue shade indicates extracellular amino acids. Identical amino acid residues are marked by asterisks, conserved amino acids by semicolons, and semi-conserved amino acid by period. LaABCB1 (KJ135790): *L. Angustifolia* ATP binding cassette B1; AtABCB1 (NP_181228.1): *A. thaliana* ATP binding cassette B1; AtABCB2 (NP_194326.2): *A. thaliana* ATP binding cassette B2; and CjMDR1 (BAB62040.1): *Coptis japonica* multidrug resistance1.
Figure 5.3 Phylogenetic relationship and classification of ABCBs.
Figure 5.3 Phylogenetic relationship and classification of ABCBs. ABCBs within the same class share a minimum of 50% amino acid identity and numbers represent branch length. Different clades are shaded in different colors: Clade I – purple, Clade II – green and Clade III – yellow. Asterisks indicated in vitro functionally characterized ABCB members and LaABCB1 is boxed. Accession numbers of ABCB subfamily members used to generate the phylogenetic tree are: AtABCB1: NP_181228.1; AtABCB2: NP_194326.2; AtABCB3: NP_192091.1; AtABCB4: NP_182223.1; AtABCB5: NP_192092.1; AtABCB6: NP_181480.1; AtABCB7: NP_199466.1; AtABCB8: Q9LHK4.1; AtABCB9: NP_193539.6; AtABCB10: NP_172538.1; AtABCB11: NP_171753.1; AtABCB12: NP_171754.1; AtABCB13: NP_174115.1; AtABCB14: NP_174122.1; AtABCB15: NP_189475.1; AtABCB16: NP_189477.4; AtABCB17: NP_189479.1; AtABCB18: NP_189480.1; AtABCB19: NP_189528.1; AtABCB20: NP_191092.1; AtABCB21: NP_191774.2; AtABCB22: NP_683599.1; CjABCB1: BAM11098.1; CjABCB2: BAM11099.1; CjABCB3: ; CjMDR1: BAB62040.1; CrABCB2: ABG56413.1; EOY25142; XP_003612850; LjABCB1: BAM33973.1; LaABCB1: KJ135790 and XP_004245909
Figure 5.4 The transcriptional activity of LaABCB1 and the reference gene (β-actin) in L. angustifolia and L. latifolia leaf and floral tissues, and L. x intermedia leaf, floral and secretory cells of glandular trichomes after 30 cycles of PCR.
Figure 5.5 LaABCB1 and GFP mRNA resolved on a 1.5 % agarose gel using TBE buffer along with ssRNA ladder (NEB).
Figure 5.6 Olympus FluoView FV1000 laser scanning confocal microscope images of water and GFP mRNA injected *Xenopus laevis* oocytes (10x).

a) Image of GFP expression measure using 488 nm laser in water injected oocyte, b) transmission light microscopy image of water injected oocyte observed using a 635 nm laser, c) image of GFP expression measure using 488 nm laser in GFP mRNA injected oocyte and d) transmission light microscopy image of GFP mRNA injected oocyte observed using a 635 nm laser.
Figure 5.7 LaABCB1 mediated vinblastine efflux in *Xenopus laevis* oocytes after 60 minutes.

Asterisks indicate the significant level difference between the two points spanned by the straight line while ns indicates non-significant differences. Error bars indicate standard means of error (n = 3).
Figure 5.8 Effect of oxidative uncoupler (KCN) and ATPase blocker (Vanadate) treatment on the LaABCB1 mediated vinblastine efflux. Treatment with 1 mM KCN and 1 mM Vanadate significantly increased (p < 0.05) the amount of vinblastine retained in *Xenopus laevis* oocytes injected with *LaABCB1* mRNA compared to non-treated oocytes. Error bars indicate standard means of error (n = 3).
Figure 5.9 LaABCB1 mediated vinblastine efflux in *Xenopus laevis* oocytes over time course.

The broken line represents percent vinblastine efflux in water-injected oocytes (which was 0) while points connected by the solid line denote percent vinblastine efflux mediated by LaABCB1 after 15, 30, 45, 60 and 120 minutes incubation in efflux assay solution. Error bars indicate standard means of error (n = 3) and asterisks indicate a significance level efflux difference between LaABCB1 mRNA-injected oocytes compared to their water-injected counter parts (p < 0.05).
Figure 5.10 LaABCB1 mediated vinblastine efflux kinetic assay.
Error bars indicate standard means of error (n = 3) and the curve was fit using the Michaelis–Menten equation ($J_{\text{max}} = 1.34 \pm 0.34$ nmol/min/oocyte per 30 minutes and a $K_i = 20.3 \pm 18.7$ µM).
Figure 5.11 Inhibition of LaABCB1 mediated vinblastine efflux by *Lavandula de novo* monoterpenes.
Asterisks indicate the significant level difference between the two points spanned by the straight line while treatments with non-significant difference were not indicated (p < 0.05; n = 3).
Table 5.1 Oligonucleotides used this study.

<table>
<thead>
<tr>
<th>Target gene</th>
<th>Primer sequence</th>
</tr>
</thead>
</table>
| LaABCB1 set I | F1: 5’ - TCGCCCGGATGTGTTGATCTT -3’  
R1: 5’ - TCCCCGAAGACACCTTCACTGGCT -3’ |
| LaABCB1 set II | F2: 5’ - GATCCGCCACATGAGTCATCAATCTGCA -3’  
R2: 5’ - TTTAAGCGTAAATCTGGAACCATCGTATGGGTAAGTTTCTGT -3’ |
| LaABCB1 set III | F3: 5’ - GCCACCATGAGTCATCAATCTGCA -3’  
R3: 5’ - AAGCTTTACGCTATCTGGAACCATCGTATGGGTAAGTTTCTGT -3’ |
| EmGFP | F: 5’ - ATATAAGGATCCGCCACATGGGGATCCGAATTCGC -3’  
R: 5’ - TACGTAAGCTTTACGCTATCTGGAACCATCGTATGGGTAAGCTTCTGAGTTAC -3’ |
| β-Actin | F: 5’ - GCACGGAATTGTGAGCAATTGGGA -3’  
R: 5’ - TTATGTCCTCACGATTTCACCCTGT -3’ |

Sequences shaded in yellow – are BamHI restriction sites, pink and underlined – are KOZAK consensus sequences, dark green – are HindIII restriction sites, blue – are HA tag coding sequences, red are stop codons and gene specific sequences are underlined.
Chapter 6 Conclusion and Recommendations

Lavenders are cultivated for their EO’s applied in perfumery, personal care and medicinal products, to name a few. The EOs are predominantly constituted of a few monoterpenes whose relative proportion - which varies considerably among species depending on plant genotype, environmental conditions and developmental stages of the flower - determines the quality, application range, and hence market value of the oil. Oil yield in lavenders also depends on plant genotype, and varies notably among various species. For example, oil yield in L. x intermedia cv. Grosso at 30% bloom stage was 27.2 – 35.5 mg/g flower fresh weight, while that of L. angustifolia cv. Lady grown in a similar environmental condition was 12.9-20.6 mg/g flower fresh weight (Boeckelamnn 2008). The amount of total oil accumulated in a given species is attributed to the number of peltate glandular trichomes, as well as the plant’s capacity to synthesize and translocate the oil from its synthesis area to storage cavity (Fahn 1988; Gershenzon et al. 2000; McCaskill and Croteau 1995; McConkey et al. 2000). In turn, the EO synthesis capacity of a given plant is regulated by the amount of precursor supplied for monoterpenes production, and likely by the efficiency of the transport mechanism employed to secrete the EO constituents into the storage cavity. As shown previously by several researchers, transcriptional regulation of genes encoding for proteins involved in precursor supply, synthesis of monoterpenes and transport plays significant role in determining the quality and quantity of lavender essential oils. Hence, the study presented in this thesis proposed to identify and functionally characterize cDNAs corresponding to genes involved in these processes using genomic resources developed in our laboratory. The isolation and in vitro characterization of genes involved in the production of the irregular monoterpenes precursor lavandulyl diphosphate (LPP), synthesis of monoterpenes from geranyl diphosphate (GPP) and secretion of monoterpenes into storage area were particularly targeted. Candidates were identified from our cDNA library based on their transcriptional level regulation in relation to oil synthesis, and studied in vitro for their respective roles.

The first objective in this thesis was to elucidate the biosynthetic origin of irregular monoterpenes constituents of Lavandula EOs. This was achieved through the identification and functional characterization of LiLPPS, a gene that encodes for a novel cisPDPS family
member enzyme. LiLPPS catalyzes the head-to-middle condensation of two DMAPP units to generate LPP \textit{in vitro}, a C\textsubscript{10} precursor of the monoterpenes lavandulol and derivative lavandulyl acetate. To our knowledge, this is the first report of a \textit{cis}PDPS involved in the biosynthesis of irregular monoterpenes. In addition, \textit{LiLPPS} is the first wild type gene that catalyzes the unusual head–to–middle condensation of two DMAPP molecules to synthesize LPP (C\textsubscript{10}) as its primary function. The irregular monoterpenes of lavenders, lavandulol and its ester derivative lavandulyl acetate, were also identified in pheromones of major insect pests and are subsequently used in artificial pheromone preparations to disrupt the mating behaviors of economically important pests (Franco et al. 2009; Walton et al. 2006). Therefore, the elucidation of this pathway enables researchers to further investigate their biosynthesis in lavenders and other plants to maximize their economic benefit. Further, the cloned gene could be used to modulate the accumulation of lavandulol, lavandulyl acetate and prenylated metabolites with a lavandulyl side group to a desired level through metabolic engineering (Christianson 2008). As described by Zhao et al. (2003), Ko et al. (2000) and Lee et al. (1997) LPP is the source of the lavandulyl side chain of sophorafлавonone G in \textit{Sophora flavescescens} Ait that determines their anti-tumour and phospholipase-C\gamma\textsubscript{1}-inhibition properties.

As part of this thesis two mTPSs, \textit{LaβPHLS} – a gene responsible for β-phellandrene production – and \textit{LiCINS} – a gene encoding for an enzyme that catalyzes the synthesis of 1,8-cineole – were also cloned and functionally characterized \textit{in vitro}. \textit{LaβPHLS} cDNA was isolated from \textit{L. angustifolia} leaves, where it is catalyzes the synthesis of the major leaf essential oil constituent. In addition to \textit{L. angustifolia} leaves, β-phellandrene is the major constituents of \textit{L. pinnata} oil (Figueiredo and Pais 1994; Huang et al. 2008) and the identified gene can be used to enhance the concentration of β-phellandrene in these plants. \textit{LiCINS} was cloned from the secretory cells of \textit{L. x intermedia} floral glandular trichomes. The \textit{LiCINS} mRNA levels paralleled the 1,8-cineole content in mature flowers of \textit{L. x intermedia}, \textit{L. angustifolia} and \textit{L. latifolia} species, and in developmental stages of \textit{L. x intermedia} inflorescence indicating that the production of 1,8 cineole in \textit{Lavandula} is most likely controlled through transcriptional regulation of \textit{LiCINS}. We also elucidated \textit{LiCINS} genomic architecture and expression in three lavender species. Although the gene was well conserved among the studied plants, our data indicated that \textit{L. x intermedia} most likely
inherited its expressed LiCINS from its other parent, *L. latifolia*. This can be further verified when sequence information for the genome of these lavender species becomes available. Genome wide sequencing can also reveal the exact contribution of each parent to the genome of *L. x intermedia* plants, and lead to the discovery of other key EO biosynthetic genes.

The thesis also describes the identification of the first ABC transporter, termed LaABCB1, that is strongly expressed in secretory cells of *Lavandula* and displays selective *in vitro* inhibition by some *Lavandula de novo* monoterpenes. This *in vitro* property of LaABCB1 and its strong expression in secretory cells of *L. x intermedia* glandular trichomes are indicative of LaABCB1’s *in planta* monoterpene secretion role. However, this needs to be confirmed, for example by over-expressing or silencing LaABCB1 through metabolic engineering in lavenders or related plants. Lack of LaABCB1 inhibition by some monoterpenes also suggests the presence of another substrate selective transporter protein in lavenders and that oil secretion into storage cavity is a concerted action result of several membrane transporters. LaABCB1 could be used in metabolic engineering efforts to enhance monoterpene accumulation in lavenders, related plants or heterologous isoprenoid production systems like microorganisms engineered for monoterpene production (Jespersen et al. 2002).

Understanding the expression of *Lavandula* EO biosynthetic genes and those involved in oil secretion in relation to EO metabolism could generate critical information regarding the regulation of EO biosynthesis in higher plants. Genes described in this thesis enhance our understanding of lavender essential oil biosynthesis and secretion, and can be used for future applications. As demonstrated by Tsuro and Assada (2013), the expression of these genes and their resultant monoterpene concentration in lavenders can be modified through metabolic engineering to enhance the quality and quantity of essential oils in transgenic plants. Alternatively, the cloned genes may be used as genetic markers in targeted plant breeding programs or in synthetic biology. As an example, LaβPHLS was expressed in cyanobacteria to demonstrate that monoterpene based renewable biofuel can be produced through “photosynthesis-to-fuel” approach (Bentley et al. 2013).
Through the work reported in this thesis and other collaborations (see preface), most structural genes regulating the production of lavender essential oil key constituents are now identified. However, some aspects of lavender essential oil biosynthesis and secretion requires further investigation. These includes, but not limited to, 1) using the cloned genes to enhance the quality and quantity of lavender essential oils through targeted breeding and/or metabolic engineering, 2) developing molecular markers linked with essential oil biosynthesis to fingerprint lavender species with desirable oil profiles, 3) identifying regulatory genes involved in monoterpene biosynthesis, 4) studying the post-transcriptional regulation mechanisms involved in oil biosynthesis regulation, 5) studying the development of glandular trichomes and 6) identifying more membrane transporter genes to study their *in vitro* and *in vivo* role. In addition, one of the major challenges I encountered during the course of this study was identifying full-length MDR and PDR homologs from our cDNA library. This is likely associated with the inherent limitation of cDNA libraries in including very long transcripts such as those that encode ABC transporters (~4 – 5 kb). Hence, an alternative strategy (for example sequencing the lavender transcriptom by Illumina technology) must be used to obtain sequence information for additional full length MDR and PDR candidates. A direct *in vitro* monoterpene transport quantification method should also be designed to obtain direct transport data and determine the kinetic properties of the transport, instead of relying on indirect inhibition studies.
References


Bohm J (2012) Pine terpenoid defences in the mountain pine beetle epidemic and in other conifer pest interactions: specialized enemies are eating holes into a diverse, dynamic and durable defence system. Tree Physiology 32 (8):943–945


Enantioenriched Odor Active Compounds. Chem Rev 111:4036–4072


Chen F, Tholl D, Bohlmann J, Pichersky E (2011) The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. Plant J 66:212–229


Demissie Z, Sarker L, Mahmoud S (2011) Cloning and functional characterization of β-
phellandrene synthase from *Lavandula angustifolia*. Planta 233:685–696


Figueiredo AC, Barroso JG, Pedro LG, Sevinate-Pinto I, Antunes T, Fontinha SS, Looman


Kelly SM, Butler JP, Macklem PT (1995) Control of cell volume in oocytes and eggs from


among four *Eucalyptus* species for genes from secondary metabolite biosynthetic pathways. BMC Genom 10:452–463


Lee HS, Ko HR, Ryu SY, Oh WK, Kim BY, Ahn SC, Mheen TI, Ahn JS (2012) Inhibition of phospholipase C gamma 1 by the prenylated flavonoids from *Sophora flavescens*. Planta Med 63:266–268


Eur J Biochem 269:3339–3354


insecticidal effects on *Lutzomyia longipalpis*. Vet Parasit 167:1–7


reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. Plant J 27:213–222


Thulasiram HV, Erickson HK, Poulter CD (2007) Chimeras of Two Isoprenoid Synthases Catalyze All Four Coupling Reactions in Isoprenoid Biosynthesis. Science 316:73–76


Tuesday GE, Simon I (2001) Bioinformatics 17: 849–850


Appendices

Appendix A Vector maps used in this study.

Figure A.1 Map of pET41b(+) vector (Novagen, EMD chemicals, Germany).
BamHI: restriction site of BamHI enzyme, CamV: cytomegalovirus immediately early enhancer/promoter, EmGFP: GFP gene ORF, Ha tag: sequences encoding for HA tag, HindIII: restriction site of HindIII enzyme, Kam R: a gene encoding for kanamycin resistance protein, KOZAK: eukaryotic mRNA consensus sequence, LaABCB1: LaABCB1 gene ORF, linearization site: unique restriction sites to linearize the vector (e.g. restriction site of XbaI enzyme), MCS: multiple cloning sites, PolyA: polyadenylation consensus sequences, T7: T7 RNA polymerase promoter, 3’ UTR: 3’ end β-globin untranslated region, 5’ UTR: 5’ end β-globin untranslated region.

Figure A.2 Schematic representation of pXOON vector map.
Figure A.3 Map of RSET-EmGFP vector (Invitrogen).