Cloning, functional characterization and genomic organization of 1,8-cineole synthases from *Lavandula*

Zerihun A. Demissie • Monica A. Cella • Lukman S. Sarker • Travis J. Thompson • Mark R. Rheault • Soheil S. Mahmoud

Department of Biology, University of British Columbia, 1177 Research Road, Kelowna, British Columbia V1V 1V7, Canada

Corresponding author e-mail: soheil.mahmoud@ubc.ca

Abstract

Several members of the genus *Lavandula* produce valuable essential oils (EOs) that are primarily constituted of the low molecular weight isoprenoids, particularly monoterpenes. We isolated over 8,000 ESTs from the glandular trichomes of *L. x intermedia* flowers (where bulk of the EO is synthesized) to facilitate the discovery of genes that control the biosynthesis of EO constituents. The expression profile of these ESTs in *L. x intermedia* and its parents *L. angustifolia* and *L. latifolia* was established using microarrays. The resulting data highlighted a differentially expressed, previously uncharacterized cDNA with strong homology to known 1,8-cineole synthase (CINS) genes. The ORF, excluding the transit peptide, of this cDNA was expressed in *E. coli*, purified by Ni-NTA agarose affinity chromatography and functionally characterized *in vitro*. The ca. 63 kDa bacterially produced recombinant protein, designated *L. x intermedia* CINS (LiCINS), converted geranyl diphosphate (the linear monoterpene precursor) primarily to 1,8-cineole with K_m and k_{cat} values of 5.75 µM and 8.8 x 10^{-3} s^{-1}, respectively.

The genomic DNA of CINS in the studied *Lavandula* species had identical exon-intron architecture and coding sequences, except for a single polymorphic nucleotide in the *L. angustifolia* ortholog which did not alter protein function. Additional nucleotide variations restricted to *L. angustifolia* introns were also observed, suggesting that *LiCINS* was most likely inherited from *L. latifolia*.

The *LiCINS* mRNA levels paralleled the 1,8-cineole content in mature flowers of the three lavender species, and in developmental stages of *L. x intermedia* inflorescence indicating that the production of 1,8 cineole in *Lavandula* is most likely controlled through transcriptional regulation of *LiCINS*.

Keywords: *L. x intermedia; L. angustifolia; L. latifolia; essential oil; isoprenoids; monoterpene synthases; 1,8-cineole synthase; intron/exon*

Abbreviations- 1,8-Cineole synthase(s): CINS(s); Diterpene synthase(s): dTPS(s); Essential oil(s): EO(s); Expressed Sequence Tag(s): EST(s); Geranyl diphosphate: GPP; Monoterpene synthase(s): mTPS(s); Sesquiterpene synthase(s): sTPS(s); Terpene synthase(s): TPS(s); *L. angustifolia* 1,8-cineole synthase: LaCINS; *L. angustifolia* linalool synthase: LaLINS; *L. angustifolia* limonene synthase: LaLIMS; *L. angustifolia* ß-phellandrene synthase: LaßPHLS; *L. x intermedia* 1,8-cineole synthase: LiCINS; *L. latifolia* 1,8-cineole synthase: LlCINS; Neryl diphosphate: NPP

Accession numbers *LiCINS*: JN701459; *LlCINS*: JN701460; *LaCINS*: JN701461

Introduction

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 essential oils (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 C_{10} 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).

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 isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP), mainly through the 2-C-methyl-D-eryhtritol 4-phosphate (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 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 the following enzymes: 1-Deoxy-D-xylylose 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-cyclodiphosphate synthase (MDS), 4-Hydroxy-3-methylbut-2-enyl diphosphate synthase (HDS) and 4-Hydroxy-3-methylbut-2-enyl diphosphate reductase (HDR) (Phillips et al. 2008). IPP and DMAPP are then condensed head-to-tail by geranyl diphosphate synthase (GPPS) to produce the linear monoterpene precursor geranyl diphosphate (GPP; C_{10}), which is subsequently transformed into various monoterpenes by specific enzymes collectively known as monoterpene synthases (mTPSs). For example, the *L. angustifolia* linalool synthase (LaLINS) and limonene synthase (LaLIMS) transform GPP mainly to linalool and limonene, respectively, (Landmann et al. 2007) while ß-phellandrene synthase (LaßPHLS) produces ß-phellandrene (Demissie et al. 2011) from GPP or neryl diphosphate (NPP) *in vitro* (Fig. 1).

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 terpene synthases (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 LQLYEASFLL motif, and the RR(x_8)W motif (Fig. 2) (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 monoterpene 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* (Landmann et al. 2007; Demissie et al. 2011). 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.

Materials and methods

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 h in ice-cold extraction buffer (200 mM sorbitol, 10 mM sucrose, 25 mM MOPSO, 0.5 mM PO_4 buffer, 10 mM sodium bisulfate, 10 mM ascorbic acid, 1 mM EDTA, 1% PVP-40, and 0.6% methylcellulose) containing 2 mM aurinticarboxylic acid, 5 mM thiourea, and 2 mM DTT at pH 6.6. Cells were then isolated, washed by a wash buffer (10% glycerol, 25 mM PO4 buffer, 1 mM EDTA, 2 mM aurinticarboxylic acid, 5 mM thiourea, and 2 mM DTT), flash frozen in liquid N_2 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.

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.

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/), and stop codon - was cloned into the NdeI/EcoRI sites of pET41b(+) expression vector 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 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 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^{TM}(DE3)plysS cells (EMD Chemicals, Darmstadt, Germany) at 20 °C for 14 -16 h in LB media supplemented with 30 mg/l Kanamycin and Isopropyl-ß-D-thiogalactopyranoside (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 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 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.

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 $MnCl_2$, 1 mM $MgCl_2$, 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^{TM} (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.

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 CO_2 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.

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 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; LlCINS) and JN701461 (*L. angustifolia* 1,8-cineole synthase; LaCINS).

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 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.

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 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 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 ($\Delta\Delta C_T$) of *LiCINS* and *LaLINS* were calculated by CFX96™ data manager (Bio-Rad, USA) using ß-actin as a reference gene.

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).

Results

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* (Fig. 1) 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 2; raw microarray expression data is given as supplementary Table 1). We were not able to validate these expression variations by PCR (Supplementary Fig. 1), and thus concluded that these slight 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 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 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.

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 (Fig. 2). The highly conserved DDxxD motif - the main divalent metal binding site - and RR(x_8)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 (Supplementary Table 2). 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^{TM}(DE3)plysS strain using the pET41(b+) expression system (EMD Biosciences, USA), and highly enriched using Ni-NTA agarose affinity chromatography (Supplementary Fig. 2). 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%) (Fig. 3a). 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 (Fig. 3b). Trace quantities of linalool were also detected in LiCINS assays (Fig 3a & b), and in negative control assays (supplementary Fig. 3) containing either of the substrates. The negative control contained protein extracts of *E. coli* Rosetta^{TM}(DE3)plysS cells transformed with empty vector instead of LiCINS. 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) (Fig. 3c), 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. The linear chemical kinetics of LiCINS extended from 5 to 120 min (Fig. 4a), while its optimum pH (Fig. 4b) and temperature (Fig. 4c) were found to be 6.5 and 30 °C, respectively. The Michaelis–Menten enzyme saturation curve was generated using the hyperbolic enzyme kinetics analysis module of the SigmaPlot software v.10.00 (Systat Software, Erkrath, Germany) (Fig. 4d). 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 supplementary Table 3). The enzyme was inactive upon incubation with FPP (data not shown).

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 1,468th position of the bacterially produced recombinant protein devoid of the transit peptide (at the 1615th position when the transit peptide is included) (see supplementary Fig. 4). At this position, the substitution of thymine (T) by cytosine (C) in *L. angustifolia* cDNA altered the encoded amino acid at the 490th (539th when the transit peptide is included) position from tyrosine (Y) to histidine (H) (Fig. 2). 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 (Supplementary Fig. 5).

The genomic DNA of CINS contained six introns and seven exons, placing this gene in the Class III TPS clade (Fig. 5). 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 LQLYEASFLL motif was placed on the 3rd 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.

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 (Fig. 6a). Further, the mRNA corresponding to this gene was not detected in leaf tissues of these species (Fig. 6b). We also evaluated transcriptional activity of the ß-phellandrene synthase gene in the same tissues as a control. As anticipated (Demissie et al. 2011), ß-phellandrene synthase mRNA was detected in the flowers and leaves of *L. angustifolia* plants only (Fig. 6a & 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, 2.8 & 2.3 mg/gm of fresh tissue, respectively (Table 3). In these tissues the *LiCINS* mRNA abundance followed a similar pattern and remained relatively constant (Fig. 6c). 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 (Fig. 6c). 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 3). 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).

Phylogenetic analysis

Based on its amino acid sequence, LiCINS was clustered into the TPSb subfamily of TPSs (Fig. 7), 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.

Discussion

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 monoterpene constituents (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-I, anthesis and 30% bloom. In agreement with previous findings (Boeckelmann 2008; Lane et al. 2010), transcripts corresponding to the MEP pathway and mTPS genes (Fig. 1) were less abundant in oil glands isolated from unopened flower buds compared to flowers at anthesis (Table 2, column A vs B). However, they were evenly abundant in oil glands of the later two stages (Table 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 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 1). The amplified fragments were sequenced to confirm their identity. Given this unique expression pattern, the EST was selected for functional analysis.

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 heterlogously 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 (Supplementary Fig. 2), 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 1,8-cineole as the main product, along with sabinene, α-phellandrene, α-terpineol, and limonene as minor products. Trace quantities of linalool were also detected in all assays, including the negative control (supplementary Fig. 3), indicating that this monoterpene was most likely a product of spontaneous GPP/NPP hydrolysis as previously reported (Schilmiller et al. 2009; Tholl et al. 2001). With the exception of α-phellandrene, monoterpenes produced by LiCINS are common in the product mixes of several other mTPSs, in particular the "cineole cassette" enzymes that produce both 1,8-cineole and α-terpineol along with other minor products (Fähnrich et al. 2011; Raguso et al. 2006). 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 their respective source plants, and that α-phellandrene is a minor constituent of the EO of *Lavandua*, it is not surprising that LiCINS can produce small amounts of this monoterpene. In this respect, all reported *Lavandula* mTPSs 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 (Fig. 3a & 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 Tº and pH of LiCINS were in the range of those reported for other related mTPSs. The optimum Tº of 30 ^{o}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 (8.8×10^{-3} s^{-1}) was close to the 0.01 – 0.1 s^{-1} ranges reported for other plant mTPSs (Wise and Croteau 1999).

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 α-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 (Fig. 2). Two of these motifs, RR(x_8)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(x_8)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 "L" residue of the catalytic motif was replaced by an "S" residue in LiCINS (Fig. 2). 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 ß-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) (Fig. 2), while it displayed a ~65% sequence similarity to LaLINS (ABB73045.1) and 64% to LaLIMS (ABB73044.1) of *L. angustifolia* (Fig. 2). An exception was that LiCINS was highly homologous (over 90% at the amino acid level) to the *L. angustifolia* ß-phellandrene synthases, indicating that these two enzymes are very closely related and likely evolved from one another, or from the same (relatively recent) parent.

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 (Bohlmann 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 (Fig 7).

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. frutesceus*, 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. frutesceus* limonene synthase and CINS genes, the RR(x_8)W and DDxxD motifs are placed on the first and third exon, respectively. However, contrary to the phase 0 placement for all *P. frutesceus* 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.

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 (Fig. 6c). 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 indicates 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* (Supplementary Fig. 5). 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.

In summary, the present study resulted in the cloning of CINS, and elucidation of its genomic architecture and expression in three lavender species. Although the gene was well conserved among the studied plants, our data indicated that *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. Understanding the expression of *Lavandula* EO biosynthetic genes in relation to EO metabolism could generate critical information regarding the regulation of EO biosynthesis in higher plants.

Acknowledgments: This work was supported through grants or in-kind contributions to SSM by UBC Okanagan campus, Investment Agriculture Foundation of British Columbia, NRC Plant Biotechnology Institute through the NAPGEN program, and Genome British Columbia, and to SSM and MRR by Natural Sciences and Engineering Research Council of Canada. ZAD would like to acknowledge the financial support through the Pacific Century Graduate Scholarships (PCGS) award from the province of British Columbia through the Ministry of Advanced Education. We would also like to thank Dr Tim Upson (Cambridge University, UK) for providing the *L. latifolia* leaf and flower tissues used in this study.

References

Aon JC, Caimi RJ, Taylor AH, Lu Q, Oluboyede F, Dally J, Kessler MD, Kerrigan JJ, Lewis TS, Wysocki LA, Patel PS (2008) Suppressing posttranslational gluconoylation of heterologous proteins by metabolic engineering of *Escherichia coli*. Appl Environ Microbiol 74(4): 950–958

Boeckelmann A (2008) Monoterpene production and regulation in lavenders (*Lavandula angustifolia* and *Lavandula x intermedia*). MSc Thesis, University of British Columbia Okanagan, pp 70–71, 154-156

Bohlmann J, Meyer-Gauen G, Croteau R (1998) Plant terpenoid synthases: molecular biology and phylogenetic analysis. Proc Natl Acad Sci USA 95(8): 4126–4133

Bohlmann J, Phillips M, Ramachandiran V, Katoh S, Croteau R (1999) cDNA cloning, characterization, and functional expression of four new monoterpene synthase members of the Tpsd gene family from grand fir (*Abies grandis*). Arch Biochem Biophys 368(2): 232–243

Cai Y, He J, Li X, Feng K, Lu L, Feng K, Kong X, Lu W (2010) Prediction of protein subcellular locations with feature selection and analysis. Protein Pept Lett 17(4): 464–472

Carretero-Paulet L, CairòA, Botella-Pavìa P, Besumbes O, Campos N, Boronat A, Rodrìguez-Concepciòn M (2006) Enhanced flux through the methylerythritol 4-phosphate pathway in *Arabidopsis* plants overexpressing deoxy-D-xylulose 5-phosphate reductoisomerase. Plant Mol Biol 62: 683–695

Cavanagh HMA, Wilkinson JM (2002) Biological activities of lavender essential oil. Phytother Res 16(4): 301–308

Chayet L, Rojas MC, Cori O (1984) Complexes of bivalent cations with neryl and geranyl pyrophosphate: their role in terpene biosynthesis. Bioorg Chem 12: 329–338

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

Christianson WD (2006) Structural biology and chemistry of the terpenoid cyclases. Chem Rev 106: 3412–3442

Crowell AL, Williams DC, Davis EM, Wildung MR, Croteau R (2002) Molecular cloning and characterization of a new linalool synthase. Arch Biochem Biophys 405(1): 112–121

Cseke L, Dudareva N, Pichersky E (1998) Structure and evolution of linalool synthase. Mol Biol Evol 15(11): 1491–1498

Degenhardt J, Köllner TG, Gershenzon J (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. Phytochemistry 70(15–16): 1621–1637

Demissie ZA, Sarker LS, Mahmoud SS (2011) Cloning and functional charaterization of ß-phellandrene synthase from *Lavandula angustifolia*. Planta 233: 685–696

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J, Guindon S, Lefort V, Lescot M, Claverie J, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 1(36): W465–W469

Desautels A, Biswas K, Lane A, Boeckelmann A, Mahmoud SS (2009) Suppression of linalool acetate production in *Lavandula x intermedia*. Natural Product Communications 4(11): 1533–1536

Dewick PM (2009) Medicinal Natural products: A biosynthetic approach. John Wiley and Sons Ltd, Chichester, West Sussex, United Kingdom pp 167–186

Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J, Kearse M, Markowitz S, Moir R, Stones-Havas S, Sturrock S, Thierer T, Wilson A (2009) Geneious v5.0.3, Available from http://www.geneious.com/

Dudareva N, Pichersky E (2000) Biochemical and molecular genetic aspects of floral scents. Plant Physiol 122(3): 627–633

Dudareva N, Andersson S, Orlova I, Gatto N, Reichelt M, Rhodes D, Boland W, Gershenzon J (2005) The nonmevalonate pathway supports both monoterpene and sesquiterpene formation in snapdragon flowers. Proc Natl Acad Sci USA 102: 933–938

Estévez JM, Cantero A, Reindl A, Reichler S, Leoñ P (2001) 1- Deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. J Biol Chem 276: 22901–22909

Fähnrich A, Krause K, Piechulla B (2011) Product variability of the 'Cineole Cassette' monoterpene synthases of related *Nicotiana* species. Mol Plant 49: 1–20

Falk L, Biswas K, Boeckelmann A, Lane A, Mahmoud SS (2009) An efficient method for the micropropagation of lavenders: regeneration of a unique mutant. J Essent Oil Res 21(3): 225–228

Fischer H, Polikarpov I, Craievich AF (2004) Average protein density is a molecular-weight-dependent function. Protein Sci 13(10): 2825–2828

Franks SJ, Wheeler GS, Goodnight C (2012) Genetic variation and evolution of secondary compounds in native and introduced populations of the invasive plant *Melaleuca qunquenervia*. doi:10.1111/j.1558-5646.2011.01524.x

Gershenzone J, Croteau R (1991) Herbivores: Their Interactions with Secondary Plant Metabolites. In: Rosenthal GA, Berenbaum M(eds) Terpenoids: The Chemical Participants. Academic Press, San Diego, California, Vol I (2nd ed), pp 165–219

Gershenzon J, McCaskill D, Rajaonarivony JIM, Mihiliak C, Karp F, Croteau R (1992) Isolation of secretory cells from plant glandular trichomes and their use in biosynthetic studies of monoterpenes and other gland products. Anal Biochem 200(1): 130–138

Gershenzon J, McConkey ME, Croteau RB (2000) Regulation of monoterpene accumulation in leaves of peppermint. Plant physiol 122(1): 205–214

Gilles M, Zhao J, An M, Agboola S (2010) Chemical composition and antimicrobial properties of essential oils of three Australian Eucalyptus species. Food Chem 119: 731–737

Hannig G, Makrides SC (1998) Strategies for optimizing heterologous protein expression in *Escherichia coli*. Trends Biotechnol 16(2): 54–60

Harborne JB, Baxter H (2001) Chemical dictionary of economic plants. John Wiley and Sons Ltd, Chichester, West Sussex, United Kingdom pp 80–81

Hardison RS (1996) A brief history of hemoglobins: Plant, animal, protist, and bacteria. Proc Natl Acad Sci USA 93: 5675–5679

Juergens UR, Dethlefsenw U, Steinkampz G, Gillissen A, Repgesz R, Vetter H (2003) Anti-inflammatory activity of 1,8 -cineole (*eucalyptol*) in bronchial asthma: a double-blind placebo-controlled trial. Respiratory Med 97: 250–256

Juergens UR, Engelen T, Racke K, Stöber M, Gillissen A, Vetter H (2004) Inhibitory activity of 1,8-cineole (eucalyptol) on cytokine production in cultured human lymphocytes and monocytes. Pulmonary Pharmacol and Therap 17: 281–287

Kampranis SC, Ioannidis D, Purvis A, Mahrez W, Ninga E, Katerelos NA, Anssour S, Dunwell JM, Degenhardt J, Makris AM, Goodenough PW, Johnson CB (2007) Rational conversion of

substrate and product specificity in a Salvia monoterpene synthase: structural insights into the evolution of terpene synthase function. Plant Cell 19(6): 1994–2005

Keegstra K, Olsen LJ, Theg SM (1989) Chloroplastic precursors and their transport across the envelope membranes. Annu Rev Plant Phys 40(1): 471–501

Kehrl W, Sonnemann U, Dethlefsen U (2004) Therapy for acute nonpurulent rhinosinusitis with cineole: results of a double-blind, randomized, placebo-controlled trial. Laryngoscope 114: 738–742

Kim KM, Yi EC, Baker D, Zhang KY (2001) Post-translational modification of the N-terminal His tag interferes with the crystallization of the wild-type and mutant SH3 domains from chicken src tyrosine kinase. Acta Crystallogr D Biol Crystallogr 57(P5): 759–762

Khan MA, Hussain I, Khan EA (2008) Allelopathic effects of Eucalyptus (*Eucalyptus camaldulensis* L.) on germination and seedling growth of wheat (*Triticum aestivum* L.). Pak J Weed Sci Res 14(1-2): 9–18

Klocke JA, Darlington MV, Balandrin MF (1987) 1,8-Cineole (Eucalyptol), a mosquito feeding and ovipositional repellent from volatile oil of *Hemizonia fitchii* (Asteraceae). J Chem Ecol 13(12): 2131–2141

Külheim C, Yeoh SH, Maintz J, Foley WJ, Moran GF (2009) Comparative SNP diversity among four *Eucalyptus* species for genes from secondary metabolite biosynthetic pathways. BMC Genomics 10: 452–463

Lahora S, Figueiredo AF, Magalhães PJC, Leal-Cardoso JH (2002) Cardiovascular effects of 1,8-cineole, a terpenoid oxide present in many plant essential oils, in normotensive rats. Can J Physiol Pharmacol 80: 1125–1131

Landmann C, Fink B, Festner M, Dregus M, Engel K, Schwab W (2007) Cloning and functional characterization of three terpene synthases from lavender (*Lavandula angustifolia*). Arch Biochem Biophys 465(2): 417–429

Lane A, Boecklemann A, Woronuk G, Sarker L, Mahmoud S (2010) A genomics resource for investigating regulation of essential oil production in *Lavandula angustifolia*. Planta 231(4): 835–845

Lee S, Chappell J (2008) Biochemical and genomic characterization of terpene synthases in *Magnolia grandiflora*. Plant Phyiol 147: 1017–1033

Lis-Balchin M (2002) Chemical composition of essential oils from different species, hybrids and cultivars of *Lavandula*. In: Lis-Balchin M (ed) Lavender: the genus Lavandula. Taylor & Francis, London, pp 251–262

Lois LM, Rodrı´guez-Concepciòn M, Gallego F, Campos N, Boronat A (2000) Carotenoid biosynthesis during tomato fruit development: regulatory role of 1-deoxy-D-xylulose 5-phosphate synthase. Plant J 22: 503–513

Maciel MV, Morais SM, Bevilaqua CML, Silva RA, Barros RS, Sousa RN, Sousa LC, Brito ES, Souza-Neto MA (2010) Chemical composition of *Eucalyptus* spp. essential oils and their insecticidal effects on *Lutzomyia longipalpis*. Vet Parasit 167: 1–7

Mahmoud SS, Croteau RB (2001) Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. Proc Natl Acad Sci USA 98(15): 8915–8920

Mahmoud SS, Croteau RB (2003) Menthofuran regulates essential oil biosynthesis in peppermint by controlling a downstream monoterpene reductase. Proc Natl Acad Sci USA 100(24): 14481–14486

Mahmoud SS, Croteau RB (2004) Cosuppression of limonene-3-hydroxylase in peppermint promotes accumulation of limonene in the essential oil. Phytochemistry 65(5): 547–554

Mathews CM, Trotman CAN (1998) Ancient and recent intron stability in Artemia hemoglobin gene. J Mol Evol 47: 763–771

McCaskill D, Croteau RB (1995) Monoterpene and sesquiterpene biosynthesis in glandular trichomes of peppermint (Mentha x piperita) rely exclusively on plastid-derived isopentenyl diphosphate. Planta 197(1): 49–56

McCaskill D, Gershenzon J, Croteau R (1992) Morphology and monoterpene biosynthetic capabilities of secretory cell clusters isolated from glandular trichomes of peppermint (*Mentha piperita* L.). Planta 187: 445–454

McConkey ME, Gershenzon J, Croteau RB (2000) Developmental regulation of monoterpene biosynthesis in the glandular trichomes of peppermint. Plant Physiol 122(1): 215–224

Menges M, Pavesi G, Morandini P, Bogre L, Murray, JAH (2007) Genomic organization and evolutionary conservation of plant D-Type cyclins. Plant Physiol 145(4): 1558–1576

Munõz-Bertomeu J, Arrillaga I, Segura J (2007) Essential oil variation within and among natural populations of *Lavandula latifolia* and its relation to their ecological areas. Biochem Syst Ecol 35(8): 479–488

Nagegowda DA (2010) Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett 584(14): 2965–2973

Nagegowda DA, Gutensohn M, Wilkerson CG, Dudareva N (2008) Two nearly identical terpene synthases catalyze the formation of nerolidol and linalool in snapdragon flowers. Plant J 55(2): 224–239

Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J 53(4): 674–690

Qu J, Ye J, Fang R (2007) Artificial MicroRNA-mediated virus resistance in plants. J Virol 81(12): 6690–6699

Rodriguez-Concepcion M, Ahumada I, Diez-Juez E, Sauret-Gueto S, Lois LM, Gallego F, Carretero-Paulet L, Campos N, Boronat A (2001) 1-Deoxy-D-xylulose 5-phosphate reductoisomerase and plastid isoprenoid biosynthesis during tomato fruit ripening. Plant J 27: 213–222

Roeder S, Hartmann A, Effmert U, Piechulla B (2007) Regulation of simultaneous synthesis of floral scent terpenoids by the 1, 8-cineole synthase of *Nicotiana suaveolens*. Plant Mol Biol 65(1): 107–124

Santos FA, Rao VSN (2000) Antiinflammatory and antinociceptive effects of 1,8-cineole a terpenoid oxide present in many plant essential oils. Phytotherapy Res 14(4): 240–244

Schilmiller AL, Schauvinhold I, Larson M, Xu R, Charbonneau AL, Schmidt A, Wilkerson C, Last RL, Pichersky E (2009) Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. Proc Natl Acad Sci USA 106(26): 10865–10870

Scoville AG, Barnett LL, Bodbyl-Roels S, Kelly JK, Hileman LC (2011) Differential regulation of a MYB transcription factor is correlated with trans-generational epigenetic inheritance of trichome density in *Mimulus guttatus*. New Phytol doi: 10.1111/j.1469-8137.2011.03656.x

Sfara V, Zerba EN, Alzogaray RA (2009) Fumigant insecticidal activity and repellent effect of five essential oils and seven monoterpenes on first-instar nymphs of *Rhodnius prolixus*. J Med Ent 46(3): 511–515

Shimada T, Endo T, Fujii H, Hara M, Omura M (2005) Isolation and characterization of (E)-beta-ocimene and 1,8-cineole synthases in *Citrus unshiu* Marc. Plant Sci 168(4): 987–995

Southwell IA, Russell MF, Maddox CDA, Wheeler GS (2003) Differential metabolism of 1,8-cineole in insects. J Chem Ecol 29(1): 83–94

Tholl D (2006) Terpene synthases and the regulation, diversity and biological roles of terpene metabolism. Curr Opin Plant Biol 9: 1–8

Tholl D, Croteau RB, Gershenzon J (2001) Partial purification and characterization of the short-chain prenyltransferases, geranyl diphosphate synthase and farnesyl diphosphate synthase, from *Abies grandis* (grand fir). Arch Biochem Biophys 386(2): 233–242

Trapp SC, Croteau RB (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. Genetics 158: 811–832

Turner GJ, Gershenzon J, Nielson EE, Froehlich JE, Croteau RB (1999) Limonene synthase, the enzyme responsible for monoterpene biosynthesis in peppermint, is localized to leucoplasts of oil gland secretory cells. Plant Physiol 120: 879–886

Turner GW, Gershenzon J, Croteau RB (2000a) Development of peltate glandular trichomes of peppermint. Plant Physiol 124: 665–680

Turner GW, Gershenzon J, Croteau RB (2000b) Distribution of peltate glandular trichomes on developing leaves of peppermint. Plant Physiol 124: 655–664

Upson T (2002) The taxonomy of genus Lavandula *L.* In: Lis-Balchin M (ed) Lavender: the genus Lavandula. Taylor & Francis, London, pp 2–34

Upson T, Andrew S (2004) The genus Lavandula. Timber Press, Portland, Oregon. Pp 78–85

Von Heijne G, Steppuhn J, Herrmann RG (1989) Domain structure of mitochondrial and chloroplast targeting peptides. Eur J Biochem 180(3): 535–545

Wendt KU, Schulz GE (1998) Isoprenoid biosynthesis: manifold chemistry catalyzed by similar enzymes. Structure 6(2): 127–133

Williams DC, McGarvey DJ, Katahira EJ, Croteau R (1998) Truncation of limonene synthase pre-protein provides a fully active 'pseudomature' form of this monoterpene cyclase and reveals the function of the amino-terminal arginine pair. Biochemistry 37(35): 12213–12220

Wise ML, Croteau R (1999) Monoterpene biosynthesis. In: Cane DD (ed) Comprehensive natural products chemistry: isoprenoids including carotenoids and steroids, vol 2. Elsevier Science, Amsterdam, pp 97–153

Wise ML, Savage TJ, Katahira E, Croteau R (1998) Monoterpene synthases from common sage (*Salvia officinalis*). J Biol Chem 273: 14891–14899

Woronuck G, Demissie Z, Rheault M, Mahmoud S (2011) Biosynthesis and therapeutic properties of *Lavandula* essential oil constituents. Planta Med 77(1): 7–15

Zeng G (1998) Sticky-end PCR: new method for subcloning. Biotechniques 25: 206–208

Figure legends

Fig. 1 The MEP pathway of isoprenoid biosynthesis. 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, GPP: geranyl diphosphate, GPPS: geranyl diphosphate synthase, HDR: 4-hydroxy-3-methylbut-2-enyl diphosphate reductase, HDS: 4-hydroxy-3-methylbut-2-enyl diphosphate synthase, IPP: isopentenyl diphosphate, IPPI: isopentenyl diphosphate isomerase, LaLINS: *L. angustifolia* linalool synthase, LaLIMS: *L. angustifolia* limonene synthase, LaβPHLS: *L. angustifolia* β-phellandrene synthase, LiCINS: *L. x intermedia* 1,8-cineole synthase, MCT: 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase, MDS: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase, mTPs: monoterpene synthases, NPP: neryl diphosphate and NPPS: neryl diphosphate synthase.

Fig. 2 Alignment of *Lavandula* 1,8-cineole synthases with mTPSs of *L. angustifolia*, and 1,8-cineole synthases of *Salvia*. Bold letters indicate conserved motifs. The predicted transit peptide residues in LaCINS, LiCINS and LlCINS are underlined, and their polymorphic residues are underlined and bolded. 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*, LlCINS (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*.

Fig. 3 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.

Fig. 4 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.

Fig. 5 Schematic representation of LiCINS genomic DNA. Exons are denoted by rectangular boxes (Exon1 through Exon7), and introns are denoted by lines connecting adjacent exons (I_1 through I_6). 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”.

Fig. 6 The transcriptional activity of *LiCINS*, *LaβPHLS*, *LaLINS* and the reference gene (*β-actin*) in (a) floral tissue, and (b) leaf tissue of *L. angustifolia*, *L. x intermedia* and *L. latifolia* measured by standard PCR. (c) The 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).

Fig. 7 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, LlCINS and LaCINS were closely rooted with LaβPHLS. Accession numbers of terpene synthases used to generate the phylogenetic tree are: LiCINS: JN701459; LlCINS: 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._citriodora: 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_P._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._glauca: 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

Supplementary Fig. 1: Transcriptional activity of MEP-pathway genes in leaves and floral tissues of *L. angustifolia*, *L. x intermedia* and *L. latifolia.* FL: 30% flower and LF: leaf.

Supplementary Fig. 2: 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.

Supplementary Fig. 3: 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.

Supplementary Fig. 4: 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.

Supplementary Fig. 5 GC chromatogram of the recombinant *L. angustifolia* 1,8-cineole synthase (LaCINS) catalyzed products from GPP. Peaks correspond to: 1) sabinene, 2) α-phellandrene, 3) limonene, 4) 1,8-cineole, 5) linalool and 6) α-terpineol.

Table
Click here to download Table: Tables.doc

Table 1: Oligonucleotides used in this study.

Primer types	Target gene	Primers
Full length	*LiCinS* set I (without transit peptide)	F1: 5' - TATGATCCAAACGGGCCGACGAT-3'
		R1: 5' - CGATTCGTAGCGCTCGAACAAC-3'
	LiCinS set II (without transit peptide)	F2: 5' - TGATCCAAACGGGCCGACGAT-3'
		R2: 5' -AATTCGATTCGTAGCGCTCGAACAAC-3'
	LaLinS	F: 5' -CCGCATATGTCGATCAATATCAACAT-3'
		R: 5'-ATAGAATTCTGCGTACGGCTCGAACA-3'
	ß-Actin	F: 5' - GCACGGAATTGTGAGCAATTGGGA -3'
		R: 5' - TTATGTCCCTCACGATTTCCCGCT -3'
qRT-PCR	*LiCinS*	F1: 5' -CCAAGCCTCAGCCATGATAGA-3'
		R1: 5' -TTGCACATCGATGCTTATCGTA-3'
	LaLinS	F: 5' -ACACGCACGACAATTTGCCA-3'
		R: 5'-AGCCCTCCAATGAAGTGGGAT-3'
	ß-Actin	F: 5' -TGTGGATTGCCAAGGCAGAGT-3'
		R: 5'-AATGAGCAGGCAGCAACAGCA-3'

Table 2: Summary of the microarray analysis experiments for relative expression of EO biosynthetic genes in *Lavandula*.

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

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 "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.

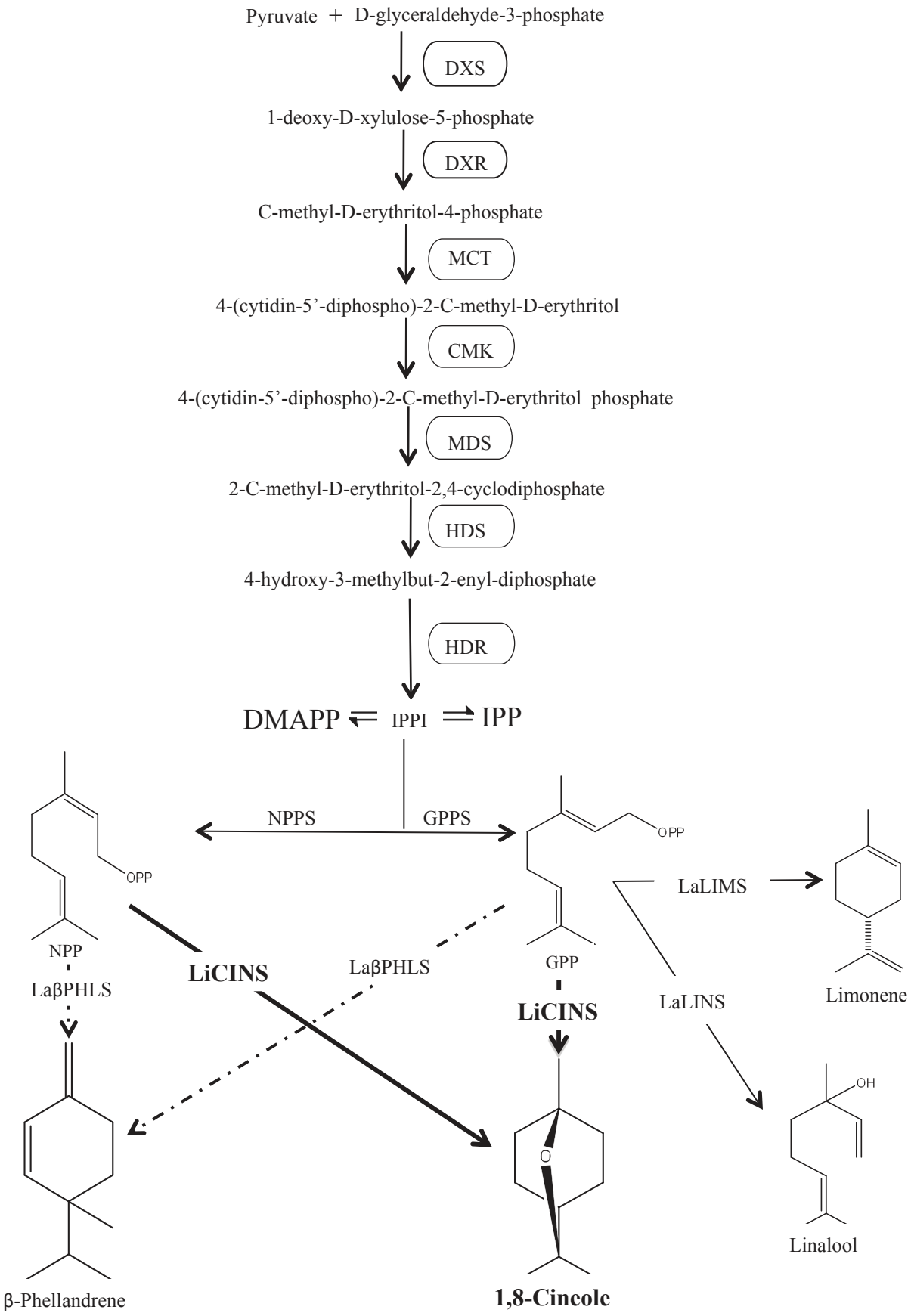
Table 3: Average 1,8-cineole and linalool content (mg per gm of fresh tissue) of oils distilled from three developmental stages of *L. x intermedia* cv Grosso.

Flower tissue	**Average mg 1,8-cineole per gm of fresh tissue**	**Average mg linalool per gm of fresh tissue**
Bud-I	2.9 ± 0.35	0.7 ± 0.2
Anthesis	2.8 ± 1.01	5.9 ± 0.14
30% flower	2.3 ± 0.35	11.3 ± 0.71

Screen/black and white figure

Click here to download Screen/black and white figure: Fig 1.pdf

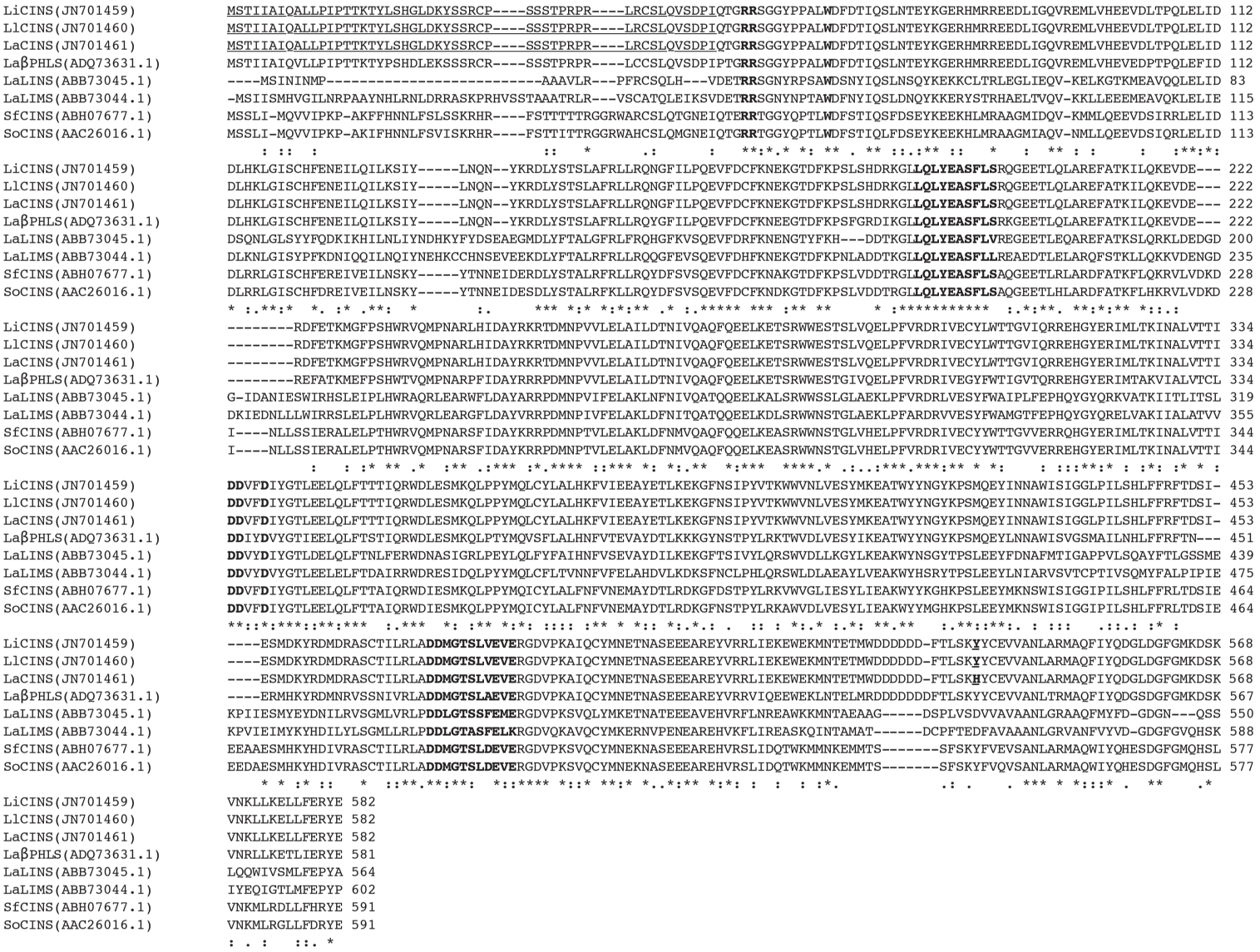
Fig 1



Screen/black and white figure

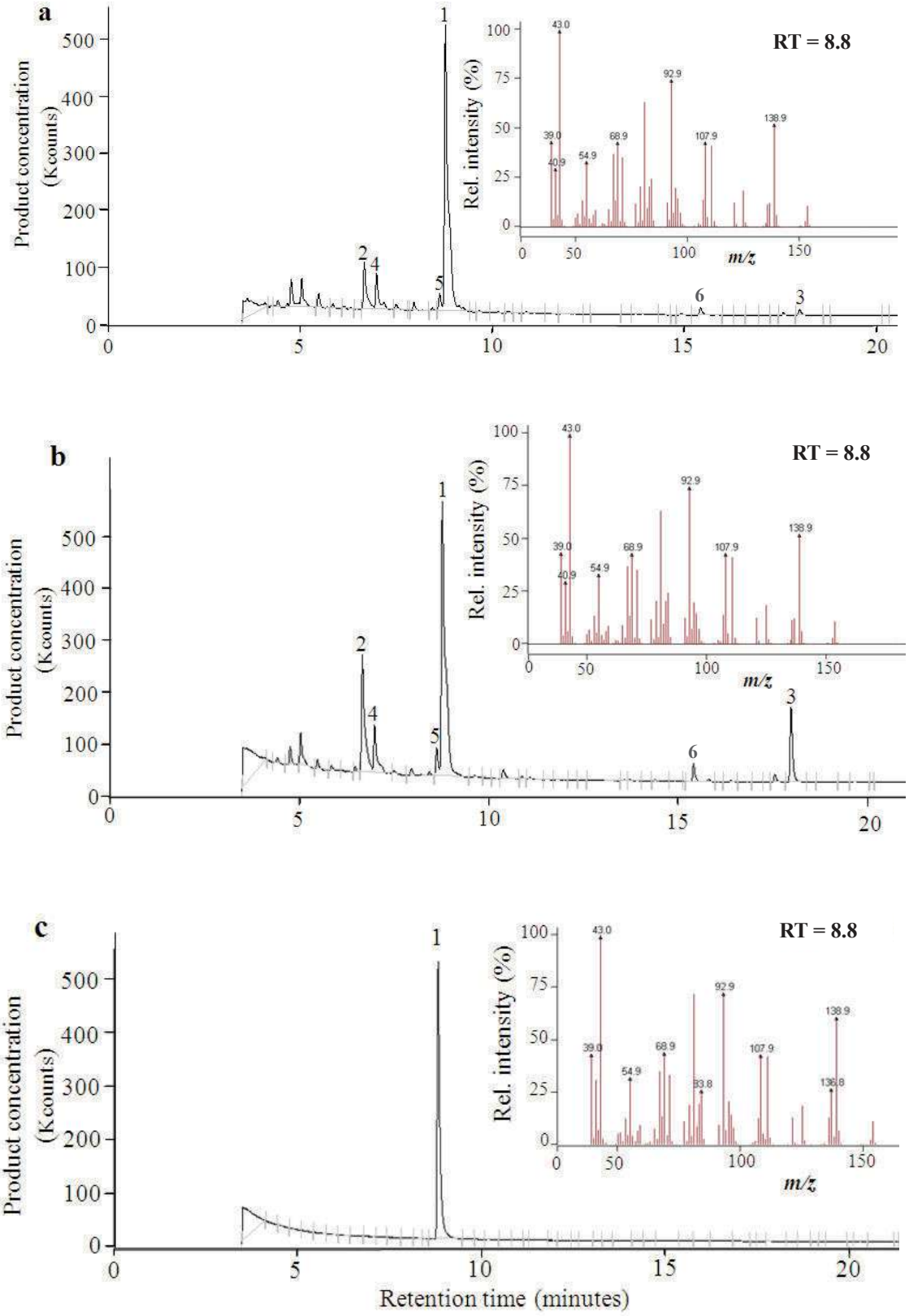
Click here to download Screen/black and white figure: Fig 2.pdf

Fig 2



Screen/black and white figure
Click here to download Screen/black and white figure: Fig 3.pdf

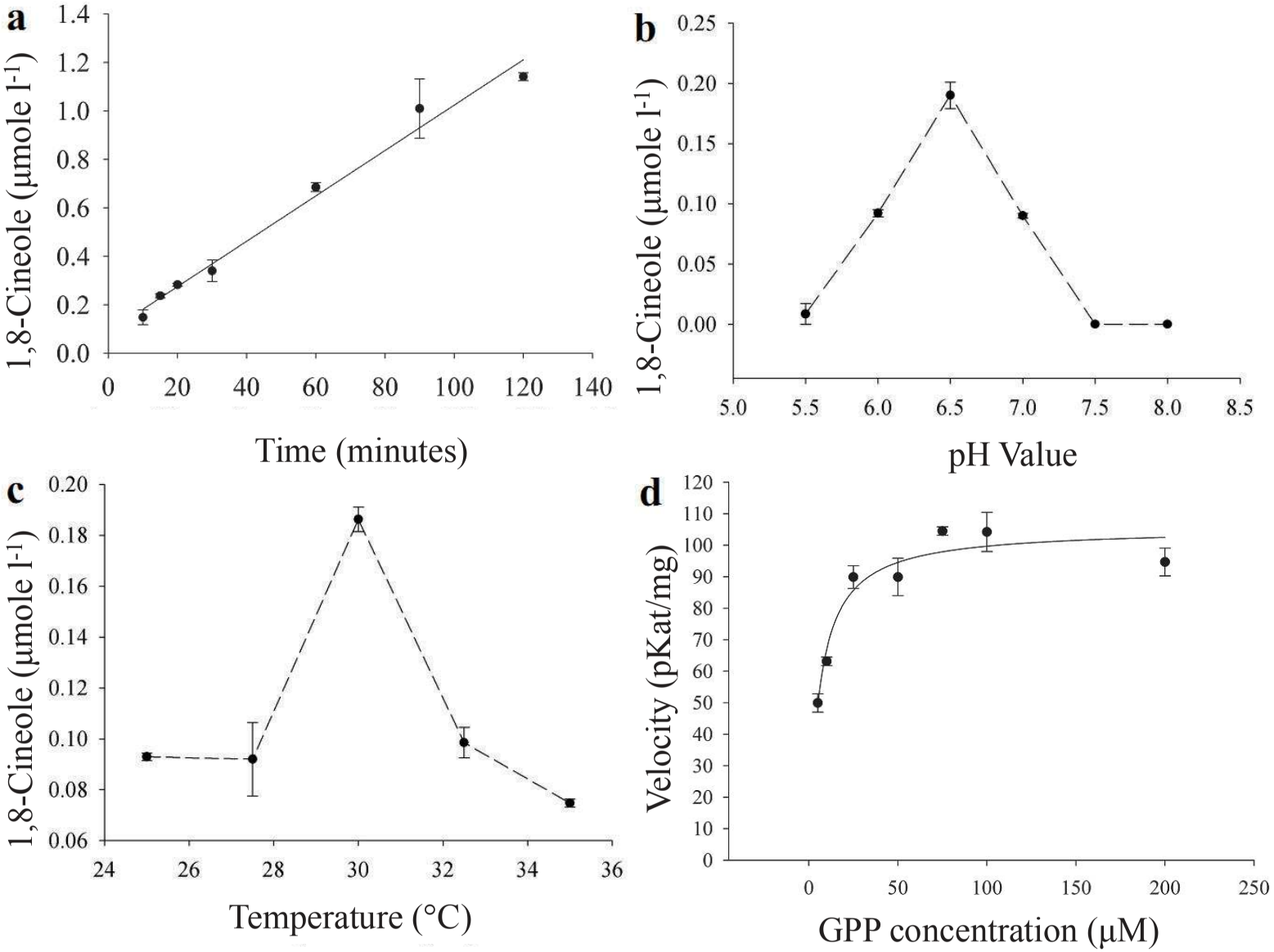
Fig 3



Screen/black and white figure

Click here to download Screen/black and white figure: Fig 4.pdf

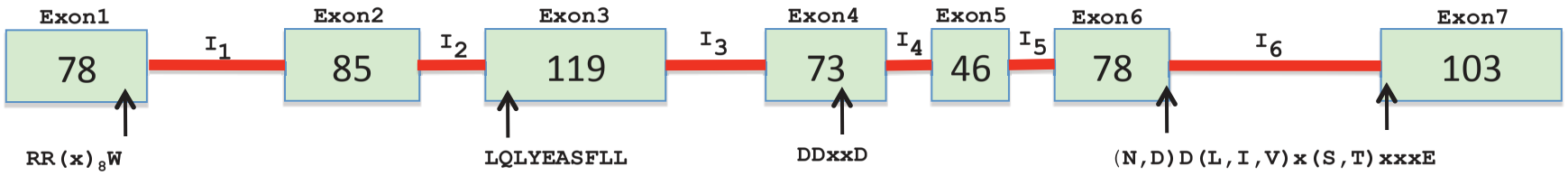
Fig 4



Screen/black and white figure

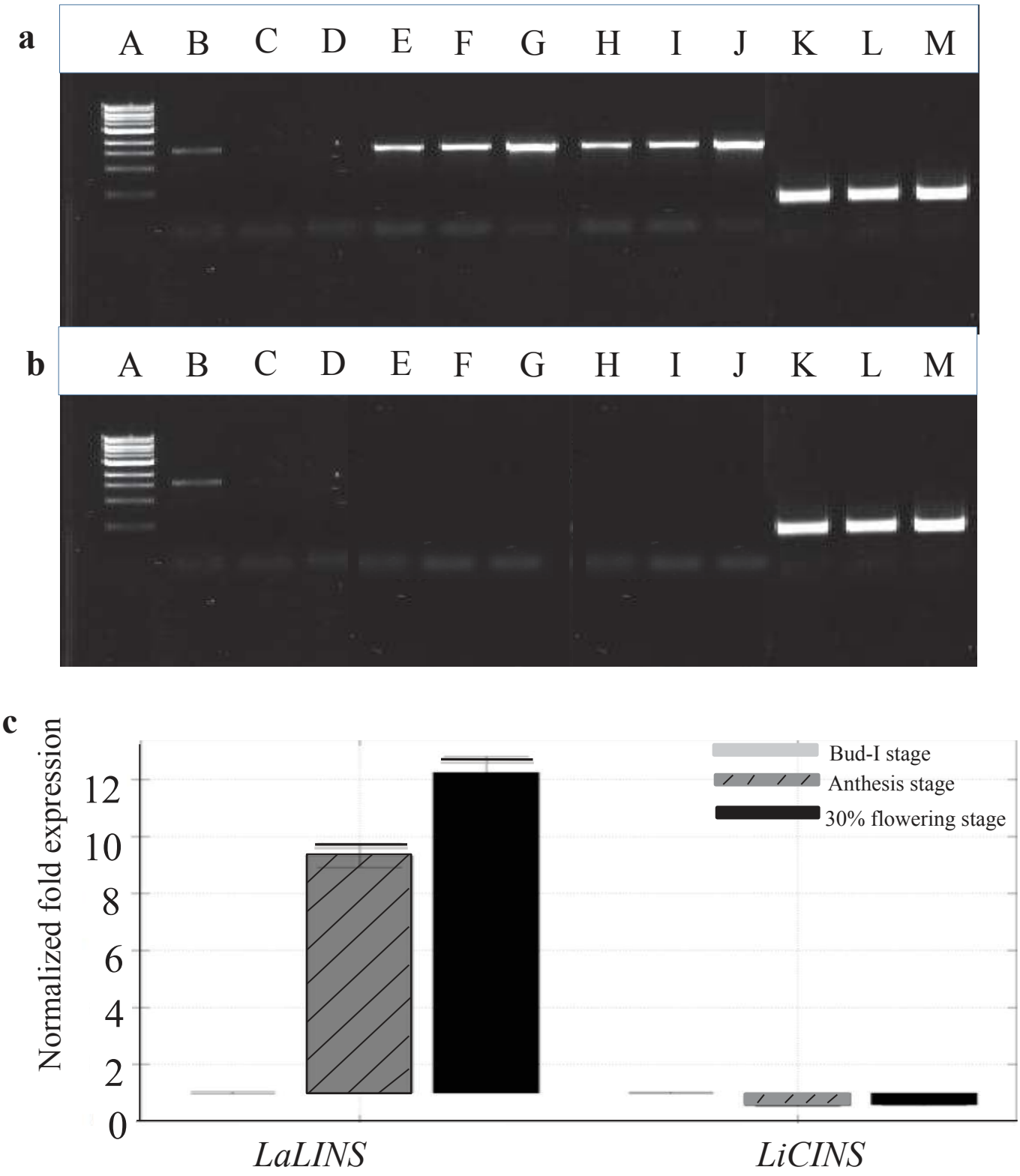
Click here to download Screen/black and white figure: Fig 5.pdf

Fig 5



Screen/black and white figure
Click here to download Screen/black and white figure: Fig 6.pdf

Fig 6



Screen/black and white figure

Click here to download Screen/black and white figure: Fig 7.pdf

Fig 7