ATATÜRK ÜNİVERSİTESİ YAYINLARI
ATATURK UNIVERSITY PUBLICATIONS

Research Article *Araştırma Makalesi* DOI: 10.5152/AUAF.2023.23150

Identification and Molecular Characterization of Genes Related to Terpene Biosynthesis From *Lavandula* × *intermedia*

Lavandula × *intermedia*'da Terpen Biyosentezi ile İlgili Genlerin Tanımlanması ve Moleküler Karakterizasyonu

Damla ÖNDER[1]
Sercan ÖNDER[2]
Yaşar KARAKURT[2]

[1]Department of Biology, Süleyman Demirel University, Faculty of Arts and Sciences, Isparta, Turkey
[2]Department of Agricultural Biotechnology, Isparta University of Applied Science, Isparta, Turkey

ABSTRACT

Lavandin (*Lavandula* × *intermedia* Emeric ex Loisel.) is an essential oil-bearing plant with high economic value for the cosmetics and pharmaceutical industries due to its fragrance components. These essential oils primarily contain monoterpenes, including geraniol, linalool, and linalyl acetate. The essential oil content of *Lavandula* × *intermedia* is five times higher than that of *Lavandula angustifolia* oil. However, due to its relatively high content of linalool and linalyl acetate, the highest quality and sought-after lavender oil is produced from *Lavandula angustifolia*. In this study, the genes encoding geranyl diphosphate synthase, linalool synthase, and linalyl acetyl transferase were isolated from glandular trichomes of lavandin flowers. The results showed that *Lavandula* × *intermedia* geranyl diphosphate synthase, *Lavandula* × *intermedia* linalool synthase, and *Lavandula* × *intermedia* linalyl acetyl transferase demonstrated expressions from the budding to the full blooming stage of flower development, and the expression levels of *Lavandula* × *intermedia* linalool synthase were lower in comparison with the expressions of *Lavandula* × *intermedia* geranyl diphosphate synthase and *Lavandula* × *intermedia* linalyl acetyl transferase. Phylogenetic analyses showed that *Lavandula* × *intermedia* linalool synthase and *Lavandula angustifolia* linalool synthase were at similar levels, while *Lavandula* × *intermedia* geranyl diphosphate synthase, *Lavandula* × *intermedia* linalyl acetyl transferase, and other lavender species geranyl diphosphate and linalyl acetyl transferase were different. The identification of these terpene synthases from lavandin forms the basis for improving the quality of essential oil composition in lavandin through metabolic engineering.

Keywords: Essential oil, gene cloning and expression, geranyl diphosphate synthase, lavandin, linalool synthase, linalyl acetyl transferase

ÖZ

Lavandin (*Lavandula* × *intermedia*), içerdiği koku bileşenleri nedeniyle kozmetik ve ilaç endüstrileri için ekonomik değeri yüksek olan bir uçucu yağ bitkisidir. Bu uçucu yağlar esas olarak geraniol, linalool ve linalil asetat dahil olmak üzere monoterpenler tarafından düzenlenir. Lavandinin uçucu yağ miktarı *L. angustifolia*'dan beş kat daha fazladır. Buna rağmen en kaliteli ve en çok talep edilen lavanta yağı nispeten yüksek linalool ve linalil asetat içeriği nedeniyle *L. angustifolia*'dan elde edilmektedir. Bu çalışmada, lavandin çiçeklerinin glandüler trikomlarından geranil difosfat sentaz (GPPS), linalool sentaz (LIS), linalil asetil transferaz (LAT) genlerini kodlayan genler izole edilmiştir. Sonuçlar *LiGPP (L.* × *intermedia* geranil difosfat sentaz), *LiLIS (L.* × *intermedia* linalool sentaz) ve *LiLAT (L.* × *intermedia* linalil asetil transferaz) genlerinin tomurcuklanmadan çiçek gelişiminin tam açma aşamasına kadar ifade edildiğini ve *LiLIS* geninin ifadesinin *LiGPP* ve *LiLAT* ifadelerine kıyasla daha düşük olduğunu göstermiştir. Filogenetik analizler, *LiLIS* geninin *L. angustifolia*'daki LIS geni ile benzer seviyelerde olduğunu, *LiGPP* ve *LiLAT* genlerinin ise diğer lavander türlerindeki GPP ve LAT genlerinden farklı seviyede bulunduğunu göstermiştir. Lavandinde terpen sentaz genlerinin tanımlanması, metabolik mühendislik yoluyla lavandindeki uçucu yağ bileşiminin kalitesinin iyileştirilmesi için temel oluşturma potansiyeline sahiptir.

Anahtar Kelimeler: Uçucu yağ, gen klonlama ve ifadesi, geranil difosfat sentaz, lavandin, linalool sentaz, linalil asetil transferaz

Received/Geliş Tarihi: 12.06.2023
Accepted/Kabul Tarihi: 24.07.2023
Publication Date/Yayın Tarihi: 20.08.2023

Corresponding Author/Sorumlu Yazar: Damla ÖNDER
E-mail: damlaguvercin@sdu.edu.tr

Cite this article as: Önder, D., Önder, S., & Karakurt, Y. (2023). Identification and molecular characterization of genes related to terpene biosynthesis from *Lavandula* × *intermedia*. *Research in Agricultural Sciences*, 54(3), 103-112.

Introduction

Lavender (*Lavandula* sp.) is a valuable perennial essential oil-bearing plant from the *Lamiaceae* family (Guenther, 1952). There are 39 lavender species (*Lavandula* sp.) species, mostly originating from the Mediterranean region. Lavender species contain a complex mixture of essential oil in glandular trichomes, making them highly aromatic plants (Blažeković et al., 2012). Among them, lavender, which is called "English lavender" (*Lavandula angustifolia* Mill. = *L. officinalis* L. = *L. vera* DC), spike lavender (*Lavandula spica* = *Lavandula latifolia* Medikus), and their hybrid lavandin, which is called "hybrid lavender" (*Lavandula* × *intermedia* Emeric ex Loisel. = *L. hybrida* L.), are the most economically important in the production of essential oils used for perfumes, cosmetics, aromatherapy, and pharmaceuticals (Tucker, 1985; Upson et al., 2004). The essential oils of the lavender species contain more than 100 compounds, but these components are present in different proportions. Therefore, the essential oil yield and composition of lavender and lavandin differ from each other, and the best quality lavender oil is obtained from lavender. Lavandin (7.1%–9.9%) has a higher rate of essential oil compared to lavender (2.8%–5.0%) (Renaud et al., 2001) but has lower essential oil quality (Baydar, 2009; Beetham & Entwistle, 1982). The essential oil produced from lavender is the most suitable for use in perfumes, due to its high linalool and linalyl acetate content (Lafhal et al., 2016; Weiss, 1997). The major difference between the lavender and lavandin oils lies in their relative contents of linalyl acetate, linalool, and camphor (Lafhal et al. 2016). Linalool, linalyl acetate, and camphor levels of essential oil are standard criteria for the determination of oil quality according to the International Organization for Standardization (ISO) 3515:2002 standard (Baydar & Kineci, 2009; Kıvrak, 2018). According to the ISO 8902:2009 standard, while the amount of linalool is 25%–38%, linalyl acetate is 25%–45%, and the camphor content must be between 0.5% and 1% in a quality lavender oil, lavandin essential oil contains 24%–35% linalool, 28%–38% linalyl acetate, and 6%–8% camphor (Kıvrak, 2018). It is important to increase the linalool and linalyl acetate content in lavandin to have as high-quality essential oil as lavender.

In lavandins, the biosynthesis of linalool, like other essential oil components, occurs in glandular trichomes or oil glands, via a series of biochemical reactions. It is well established that all terpenes are generated from five carbon units of isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP). Isopentenyl diphosphate and DMAPP are linear precursors of most isoprenoids, such as C_{10} geranyl diphosphate (GPP) and C_{15} farnesyl pyrophosphate. Geranyl diphosphate is synthesized in the presence of GPP synthase (GPPS) through biochemical condensation of IPP and DMAPP. Geranyl diphosphateGPP is the linear precursor to all regular monoterpenes, which is subsequently catalyzed to linalool and linalyl acetate through activities of linalool synthase (LIS) and linalyl acetyl transferase (LAT), respectively (Adal et al., 2017; Burke et al., 1999; Schilmiller et al., 2009; Tholl & Lee, 2011). Like other monoterpenes, linalool and linalyl acetate are synthesized from GPP and their synthesis is catalyzed by the enzymes LIS and LAT, respectively (Zaks et al., 2008) (Figure 1).

Increasing the expression levels of genes responsible for linalool and linalyl acetate synthesis may be an effective way to improve essential oil quality (Muñoz-Bertomeu et al., 2006). The genes synthesizing these components were cloned from cinnamon (*Cinnamomum osmophloeum)* (Lin et al., 2014), wintersweet (*Chimonanthus praecox* L.) (Kamran et al., 2020), lemon myrtle (*Backhousia citriodora*) (Sugiura et al., 2011), bergamot mint (*Mentha citrata* Ehrh.) (Crowell et al., 2002), tomato (*Lycopersicon esculentum*) (van Schie et al., 2007), *Rosa chinensis* (Magnard et al., 2018), and *Clarkia breweri* (Dudareva et al., 1996). Since lavandin, which is economically important, has not been investigated yet, it prevents the improvement of essential oil production by molecular techniques. Therefore, the aim of the present study was to investigate the cloning and characterization of GPPS (synthesize GPP), LIS (synthesize linalool), and LAT (synthesize linalyl acetate) that make substantial contributions to the flavor of lavandin essential oil.

Methods

Plant Material

Samples of lavandin (*Lavandula* × *intermedia* cv. Super A) flowers were collected from a commercial lavandin garden, located in Kuyucak village (latitude 37° 50' 45.2" (N) longitude 30° 15' 39.0" (E), 986 m altitude) in Isparta, Türkiye, between 8 and 10 a.m. on the 18th of June 2019. The fresh, fully blooming flowers were used for gene isolation. For comparative gene expression analysis, flowers at five developmental stages were collected according to distinct morphological characters: spike size (length of the spike in cm) and number of flowers post anthesis per flower spike. As described in Figure 2, there were five distinct stages (Figure 2). All fresh samples were immediately frozen in liquid nitrogen and preserved at –80°C for further use.

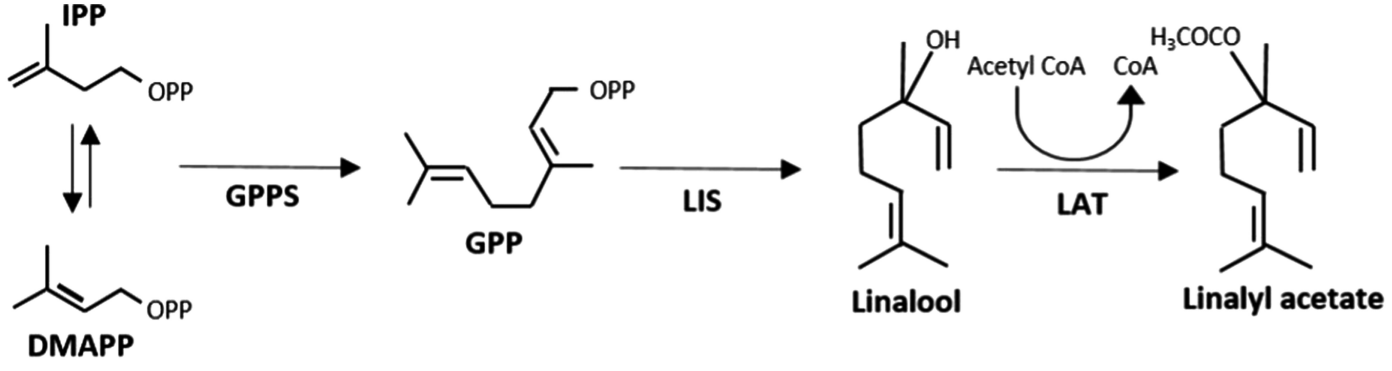


Figure 1.
Schematic Representation of the Biosynthesis of Geranyl Diphosphate, Linalool, and Linalyl Acetate Catalyzed by Linalool Synthase (LIS) and Linalyl Acetyl Transferase (LAT).

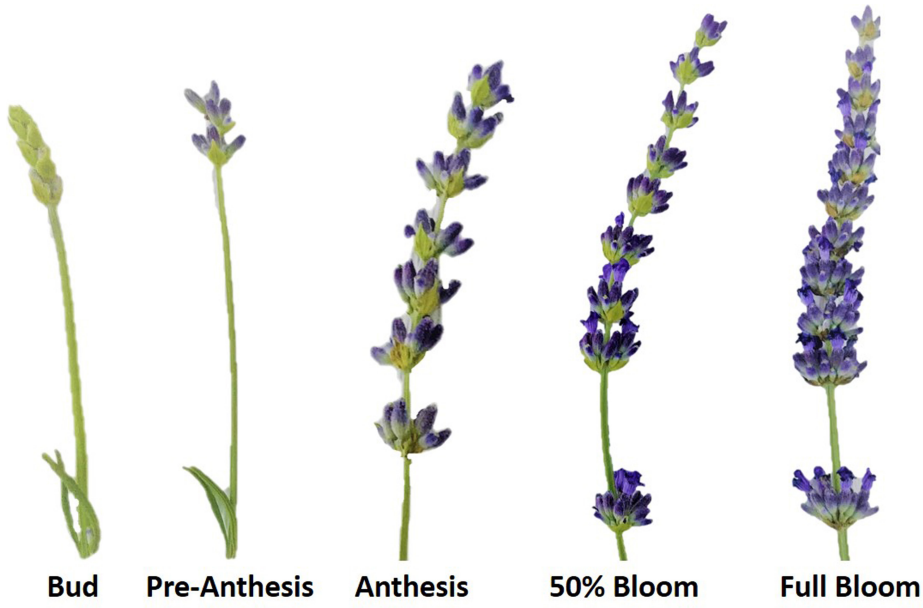


Figure 2.
Flower Developmental Stages of L. intermedia cv. Super A. Bud: 1–2 cm, Green; Pre-Anthesis: 3–4 cm, Violet; Anthesis: 5–8 cm, Violet; 50% Bloom: 50% of Flowers on Spike Opened; Full Bloom: All Flowers on Spike Opened.

Glandular Trichome Extraction

Glandular trichomes were isolated from mature (100% in bloom) lavandin flowers using a method modified from Gang et al. (2001), Gershenzon et al. (1992), and Wu et al. (2012). In brief, 10–20 g of fresh flowers were placed into a beaker (kept on ice) and filled with isolation buffer (25 mM MOPS (3-(N-Morpholino)propanes ulfonic acid), 5 mM thiourea, 2 mM dithiothreitol, 200 mM sorbitol, 10 mM KCl, 5 mM $MgCl_2$, 1% [w/v] polyvinylpyrrolidone, 10 mM sucrose, 0.5 mM sodium phosphate, and 0.6% [w/v] methylcellulose). The trichomes were gently removed from the flowers on ice with three pulses of 1 minute at 30 V, with 1 minute rest between pulses. The cellular extract was separated from the flower material by consecutively filtering through a 300 µm and a 105 µm mesh membrane (Small Parts Inc., Miami Lake, FL, USA). The solution of mixed cells was collected by centrifugation (15 minutes, 4°C, 8000 rcf) and was mixed in 25 mL wash buffer (isolation buffer without Polyvinyl pyrrolidone (PVP) and methylcellulose). The resulting solution of mixed cells was then concentrated by centrifugation (10 minutes, 4°C, 6000 rcf) and 5 mL of the concentrated mixture was mixed with 30 mL 40% sucrose solution and centrifuged at 4°C for 10 minutes at 800 rcf. The isolated glands were collected after three washes by adding ice-cold ribonuclease (RNase) free water and were rapidly frozen in liquid nitrogen and stored at –80°C until used.

Cloning of Partial Geranyl Diphosphate, Linalool Synthase, and Linalyl Acetyl Transferase Sequences

Total RNA was extracted from lavandin trichomes using a modified isothiocyanate method (Strommer et al., 1993) and RNA was quantified spectrophotometrically. DNase treatment of total RNA was performed using a DNA-Free kit (Qiagen, Valencia, California, United States) to remove any residual genomic DNA. The purified RNAs, whose quality and quantity were checked, were reverse transcribed using SuperScriptTM II RT reverse transcriptase (Invitrogen) with a commercial oligo-dT primer following the manufacturer's recommendations. The resulting complementary DNA (cDNA) product was directly used as a template in subsequent PCRs with Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific). The degenerate primers were designed from a conserved sequence of different terpene synthase genes previously identified from different plants. For the isolation of partial fragments of GPP, LIS, and LAT genes, PCR amplification was carried out with degenerate primers whose sequences were given in Table 1.

A basic 50 µL PCR mixture was prepared containing 0.2 mM of each of Deoxynucleotide triphosphates (dNTPs) (10 mM), 1.5 mM $MgCl_2$, 1 µM of each primer, 1 µg of cDNA template, 1.25 U Platinum *Taq* Polymerase (Invitrogen), 50 mM KCl, and RNase free

Table 1.
Primers Used for GPP, LIS, and LAT Genes Isolation in Lavandin

Gene	Primer	Oligonucleotide Sequence (5'–3')	
Geranyl diphosphate synthase (GPP)	*LiGPP*-F1	GCTGATGAARACAGCTCTGRATGTRCG	Forward
	LiGPP-R1	CCCTTTCCRAGAGAGGCWGATGTGCC	Reverse
Linalool synthase (LIS)	*LiLIS*-F1	GATGTKTTTAACMACTTYARRGA	Forward
	LiLIS-R1	GRTYCTCATCCTTKGCACTTCC	Reverse
Linalyl acetyl transferase (LAT)	*LiLAT*-F1	GTDTTYTACTAYCCDTTYGCHGG	Forward
	LiLAT-R1	ATVACCATWARATCYGCCAVHGA	Reverse

water. The thermal cycle program [BIO-RAD 580BR My Cycler Thermal Cycler (USA)] was performed with the following cycling conditions: 94°C for 5 minutes for initial denaturation, followed by 30 cycles of denaturation at 95°C for 1 minute, annealing at 60°C for 30 seconds, extension at 72°C for 30 seconds, and a final elongation step at 72°C for 5 minutes. Invitrogen PureLink® Quick gel extraction kit was used to purify PCR products that revealed the approximate size of partial sequences obtained by agarose gel electrophoresis. A second PCR using the same primers directly or in a nested position produced a larger amount of DNA that could be subcloned into the pGEM-T Easy Vector System according to the manufacturer's instructions (Promega, Madison, Wisconsin,USA) for sequencing. DNA sequencing of the positive clones was performed by the dideoxy method utilizing an ABI3730 automated sequencer (Iontek, Istanbul, Turkey). The NCBI database was used to verify the similarity of cloned sequences to known sequences using the BLASTn program, and contigs were constructed using ClustalX for each gene to identify the closest homologs.

Cloning of Full-Length Geranyl Diphosphate, Linalool Synthase, and Linalyl Acetyl Transferase Sequences

For full-length cDNA synthesis, partial cDNA sequences of GPP, LIS, and LAT were utilized to obtain gene-specific primers and used in the Rapid Amplification of cDNA Ends (RACE) analyses. RACE analysis was carried out with SMART RACE-cDNA Amplification Kit (Clontech). cDNA was synthesized from mRNA using Superscript III Reverse Transcriptase (Thermo Fisher Scientific) and gene-specific primers based on the known sequence fragments in order to clone the 5' ends of the transcripts (5'-RACE). A 50 µL PCR mixture was prepared and PCR was performed with a touchdown protocol starting with 12 cycles of 1 minute at 95°C and followed by 10 cycles with decreasing annealing temperature (30 seconds at 94°C, 40 seconds at 68°C to 1°C/cycle, and 1 minute at 72°C), 25 cycles (30 seconds at 94°C, 40 seconds at 66°C, and 120 seconds at 72°C). The amplified PCR products were purified by gel electrophoresis, then cloned into pGEM-T vector, and the resulting recombinant plasmids were sequenced. The 3' ends of the putative GPP, LIS, and LAT genes were amplified by 3'-RACE PCR. A touchdown protocol was used, similar to that used for 5'-RACE PCR. All three target genes showing the expected length of the 5'-terminal end were extracted from agarose gel, and the resulting amplicons were cloned into pGEM-T and sequenced.

Multiple Sequence Alignment and Phylogenetic Tree Analysis

The full-length cDNA sequences were used for BLAST searches of the GenBank databases. To convert nucleotide sequences into protein sequences, a translation tool (http://www.fr33.net/translator.php) was used. BioEdit (Hall, 2004) and ClustalW (version 2.1) (Jeanmougin et al., 1998; Thompson et al., 1994) methods were used to align nucleotide and protein sequences. For phylogenetic analysis, neighbor-joining (NJ) trees with 1000 bootstrap replicates were performed using MEGA6 software (Tamura et al., 2013). BLASTn was used to analyze homology at the nucleotide level, and BLASTp was used to search for homology at the amino acid (aa) level (http://www.ncbi.nlm.nih.gov/blast/). The coding sequences for *LiGPP* (MN641908.1), *LiLIS* (MN635786.1), and *LiLAT* (MN641909.1) have been submitted to NCBI.

Gene Expression Analysis

To analyze the expression profile of *LiGPP, LiLIS,* and *LiLAT*, total RNA was extracted from five different developmental stages of lavandin flower tissues and treated with DNase I enzyme to remove genomic DNA (Qiagen). Complementary DNA for relative transcript analysis was synthesized using SuperScript IV Reverse Transcriptase cDNA synthesis kit (Invitrogen) according to the manufacturer's instructions. The relative abundance of the GPP, LIS, and LAT genes was analyzed in flowers of all five developmental stages by quantitative real-time PCR (RT-qPCR), using the Applied Biosystems 7300 Real-Time PCR System (Applied Biosystems, Foster City, California, USA). Maxima SYBR Green/ROX qPCR Master Mix (Applied Biosystems) along with 25 ng of cDNA template and 500 nM of forward and reverse primers were used in a 20 µL total volume. Gene-specific primers (Table 2) used in qPCR experiments were designed manually as nucleotide sequence similarity. All experiments were carried out in triplicate under the same thermal cycling conditions (denaturation step at 94°C for 5 minutes, followed by 45 cycles of 94°C for 30 seconds, 60°C for 40 seconds, and a final elongation at 72°C for 5 minutes). Melting curve and gel electrophoretic analysis were used to confirm primer specificity. Because of its expression stability, β-actin was chosen as the reference gene. Relative abundance of transcripts was quantified using the comparative CT method (Schmittgen & Livak, 2008).

Table 2.
Gene-Specific Primer Sequences Used for the Detection of Expression Levels by Real-Time Quantitative Fluorescence PCR

Gene	Forward Primer (5'–3')	Reverse Primer (5'–3')
LiGPP	GGTTGCAGAGTTATTTATTCA	CAGCGCTGCAGCATGGAATAC
LiLIS	TTGATGACTCGCAGAATCCG	TCTTGGAAGGGCGGCGCAGT
LiLAT	GCCAACAGGTTCAACTCCAGA	GAGAGGATCCACGTTGCGAGC
Actin	TGTGGATTGCCAAGGCAGAGT	AATGAGCAGGCAGCAACAGCA

Results

Isolation and Bioinformatics Analysis of *LiGPP*, *LiLIS*, and *LiLAT* Genes Involved in the Biosynthesis of Terpenes

Degenerate primers were used to clone the conserved regions of GPP, LIS, and LAT from various plants, and RACE cDNA Amplification Kit procedures were used to extend the 5′ and 3′ DNA sequences of each gene. Full-length cDNAs of each gene from lavandin were identified as *LiGPP, LiLIS,* and *LiLAT* (GenBank Accession No: MN641908, MN635786, and MN641909, respectively). The entire *LiGPP* open reading frame (ORF) is 1284 bp long and encodes a protein of 427 aa, with a predicted molecular weight of 46.7 kDa. Compared with GPP in other plants, the sequences were highly similar, sharing 71%–97% similarity at the aa level. The most similar GPP belongs to that of *Salvia officinalis* where the aa has 97% similarity compared to *LiGPP*. The ORF for *LiLIS* was 1716 bp long, encoding a 571 aa protein with a predicted molecular weight of 65.8 kDa. Alignment of the aa sequences revealed that *LiLIS* shares 57%–90% similarity with LIS in other plants. Linalool synthase from *L. angustifolia* is the most similar where the aa has 90% similarity compared to *LiLIS*. The ORF of *LiLAT* consists of 972 bp, coding for proteins of 324 aa with predicted molecular weights of 33.2 kDa. Compared with LAT genes in other plants, the sequences were highly similar, sharing 77%–97% similarity at the aa level. The theoretical isoelectric point of the deduced LiGPP, LiLIS, and LiLAT proteins, which were predicted by ExPASy, is 5.61, 4.81, and 8.69, respectively (Table 3).

Multiple protein sequence alignments of target genes were performed with CLUSTALX and MEGA X. The results indicated that

Table 3.
Physicochemical Properties of Hypothetical Proteins in Lavandin

Proteins	Molecular Weight (kD)	Theoretical Isoelectric Point	Number of Acidic Amino Acids	Number of Basic Amino Acids	Instability Index	Aliphatic Index	Total Average Hydropathicity
LiGPP	46,661.22	5.61	50	42	44.25	101.22	0.011
LiLIS	65,840.98	4.81	90	56	48.73	84.54	–0.391
LiLAT	33,181.35	8.69	28	32	30.38	95.28	0.210

LiLAT, Lavandula x intermedia linalyl acetyl transferase; LiLIS, Lavandula x intermedia linalool synthase; LiLAT, Lavandula x intermedia linalool acetyl transferase

the aa sequence of *LiGPP* contains a highly conserved aspartate-rich region (DDXXD); however, it is free from the CxxxC (where "x" is any hydrophobic residue) motif, similar to other homomeric GPPSs characterized, which have two characteristic motives of terpene synthases. The *LiGPP* contained two DDXXD motifs, needed for Mg^{2+} or Mn^{2+} ion binding, which act as a cofactor in the substrate interaction (Lin et al., 2014; Whittington et al., 2002). The results showed that the aa sequence *LiGPP* contains the two conserved aspartate-rich motifs, containing FARM $(DD(X)_{2-4}D)$ (the first aspartate-rich motif, DDxxD) and SARM (DDXXD) (the second aspartate-rich motif, DDxxD) (Figure 3A). The *LiGPP* has the same ASLLHDDVLD "FARM signature" as mitochondrial diphosphate synthase genes. However, in *LiGPP,* valine (V) of the FARM signature was replaced with methionine (M). The aa alignment of GPPS is shown in Figure 3A. Since the number of acidic aa (Asp + Glu) in the sequence is 50 and the number of basic aa (Arg + Lys) is 42, the *LiGPP* protein shows acidic properties. The *LiGPP* is a stable hydrophobic protein. Conserved domain analysis against the GenBank CDD database showed that *LiGPP* proteins contain the conserved PLN02890 GPPS domain (PLN02890) from 20 to 427 aa (Figure 4A).

The aa alignment of LIS is shown in Figure 3B. The *LiLIS* contained the aspartate-rich DDxxD and $DDLxTx_3E$ motifs, which bind the substrate-divalent cations, and thus are essential for the binding and ionization of the substrate. Further, *LiLIS* contained the N-terminal arginine residues (R27 and R28) in the RRx_8W motif that are essential for the initial isomerization step in the conversion of GPP to the linalyl diphosphate intermediate. Another characteristic sequence LQLYEASFL, which is often partially conserved, is also present. Since the number of acidic aa (Asp + Glu) in the sequence is 90 and the number of basic aa (Arg + Lys) is 56, the *LiLIS* is a stable hydrophilic protein with acidic properties. The *LiLIS* has a plant terpene cyclases domain from aa 27 to 562 and has significant terpene synthase activity from 246 to 510 aa (Figure 4B).

A conserved CxxxC motif is present in *LiLAT*, but it did not contain the DDXXD terpene synthase signature motif, which was present in other terpene synthases. All of the conserved motifs present in typical plant alcohol acetyltransferases were found in *LiLAT*, including the HxxxxD and DFGWG motifs in the center of the protein believed to be involved in catalysis reactions (D'Auria, 2006). However, in *LiLAT, the* tryptophan (W) of the DFGWG motif was replaced with methionine (M) (Figure 3C). Since the acidic aa in the sequence is 28 and the basic aa is 32, the protein has a basic character. The *LiLAT* is a stable hydrophobic protein and has a thiolase (like) domain from 16 to 323 aa and acetyl-CoA acyltransferase activity (PLN02644) from aa 13 to 324 (Figure 4C). Residues of two cystines on the N and C terminal regions are essential for AACT's catalytic activity and are highly conserved in thiolases from various sources.

Phylogenetic Analysis

The NJ method was used to construct two phylogenetic trees using MEGAX. The phylogenetic analysis of the monoterpene synthase genes is shown in Figure 5. The various lineages are divided into sub-groups based on their relationship. The phylogenetic tree revealed that *LiGPP* exhibited a significant similarity to GPPs from sage (*Salvia officinalis*) (97%), pink trumpet tree (*Handroanthus impetiginosus*) (83%), patchouli (*Pogostemon cablin*) (81%), *Centranthera grandiflora* (80%), red witchweed (*Striga asiatica*) (78%), tobacco (*Nicotiana tabacum*) (77%), periwinkle (*Catharanthus roseus*) (77%), English ivy (*Hedera helix*) (76%), and tomato (*Solanum lycopersicum*) (73%) in the phylogenetic tree (Figure 5A). The BLASTp analysis indicated that *LiLIS* was closely related to LIS from lavender (*Lavandula angustifolia)* with a 95% aa similarity level in the phylogenetic tree. The *LiLIS* also exhibited sequence similarity with LIS of *Centranthera grandiflora* (86%), fragrant olive (*Osmanthus fragrans*) (83%), tea plant (*Camellia sinensis*) (81%), and white mulberry (*Morus alba*) (77%) (Figure 5B). The LiLAT exhibited a significant similarity to LATs from hara (*Isodon rubescens*) (98%), self-heal (*Prunella vulgaris*) (90%), peach sage (*Salvia dorisiana*) (60%), Indian borage (*Plectranthus amboinicus*) (60%), and spike lavender (*Lavandula latifolia*) (57%) (Figure 5C). These results verify that *LiGPP*, *LiLIS,* and *LiLAT* share a common evolutionary base with other plant GPP, linalool, and linalyl acetate synthase proteins primarily based on their conserved structure and sequence characteristics.

Analysis of *LiGPP*, *LiLIS* and *LiLAT* Gene Expression by Quantitative Real-Time Polymerase Chain Reaction

The standard PCR technique was employed to investigate expression patterns of *LiGPP, LiLIS, and LiLAT* transcripts at different developmental stages of lavandin flower tissues, including buds, anthesis, 50% bloom, and mature (100% in bloom) flowers. The qPCR method was used to determine the transcriptional activity of *LiGPP, LiLIS*, and *LiLAT* in floral tissues of lavandin. The results showed that the transcripts for these genes were abundant in flower tissues. The results of the qPCR assay also indicated that all genes had the lowest expression levels in the bud stage (Figure 6).

All monoterpene synthase gene expression levels increased from the bud stage and peaked in the 50% bloom stage, while their expression levels were significantly decreased in full bloom (Figure 6). The GPP, LIS, and LAT level of expressions were significantly increased in stage 4, which were approximately 264, 462, and 600 times higher than that of stage 1, respectively. All transcript levels decreased in full bloom stage (Figure 6).

Among the three genes, *LiGPP* showed the highest transcript, followed by *LiLAT* and *LiLIS*. In detail, *LiGPP* and *LiLAT* genes were highly expressed in the 50% bloom stage but had low expression levels in the bud stage. At all flowering stages, *LiLIS* gene

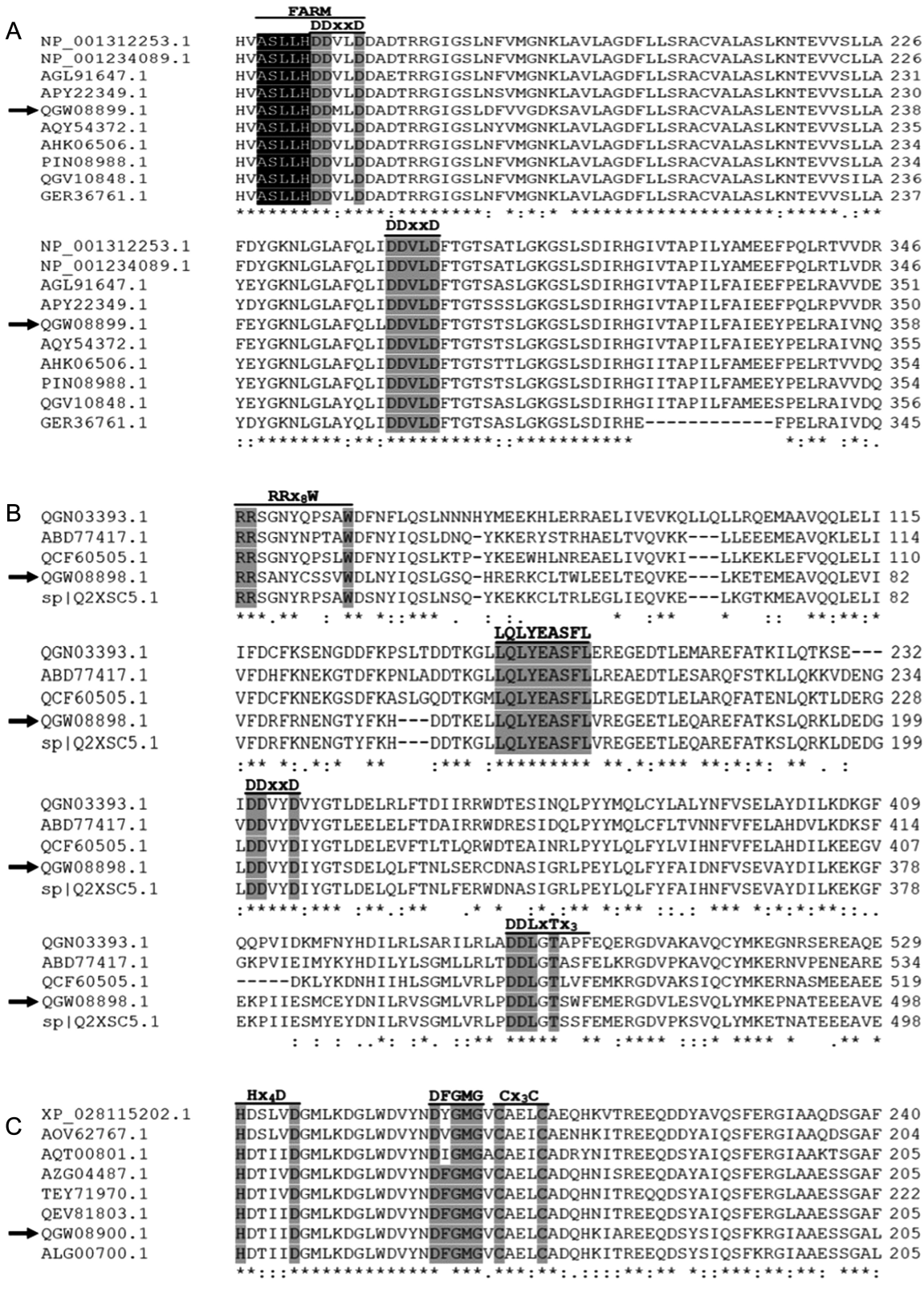


Figure 3.
Amino Acid Sequence Alignments of LiGPP (A), LiLIS (B), and LiLAT (C) With Other Corresponding Homologous Proteins. Enzymatic Motifs Are Marked With a Solid Line. CLUSTALX and MEGAX Provided a Comparative Diagram and Cluster Analysis of Target Gene Sequences With Similar Sequences From Other Species. The GenBank Accession Numbers Are Shown for Each Protein. NP_001312253.1, Nicotiana tabacum; NP_001234089.1, Solanum lycopersicum; AGL91647.1, Catharanthus roseus; APY22349.1, Hedera helix; AQY54372.1, Salvia officinalis; AHK06506.1, Pogostemon cablin; PIN08988.1, Handroanthus impetiginosus; QGV10848.1, Centranthera grandiflora; GER36761.1, Striga asiatica; QGN03393.1, Plectranthus amboinicus; ABD77417.1, Lavandula latifolia; QCF60505.1, Salvia dorisiana; Q2XSC5.1, Lavandula angustifolia; XP_028115202.1, Camellia sinensis; AOV62767.1, Morus alba; AQT00801.1, Osmanthus fragrans; AZG04487.1, Centranthera grandiflora; TEY71970.1, Salvia splendens; QEV81803.1, Prunella vulgaris; ALG00700.1, Isodon rubescens.

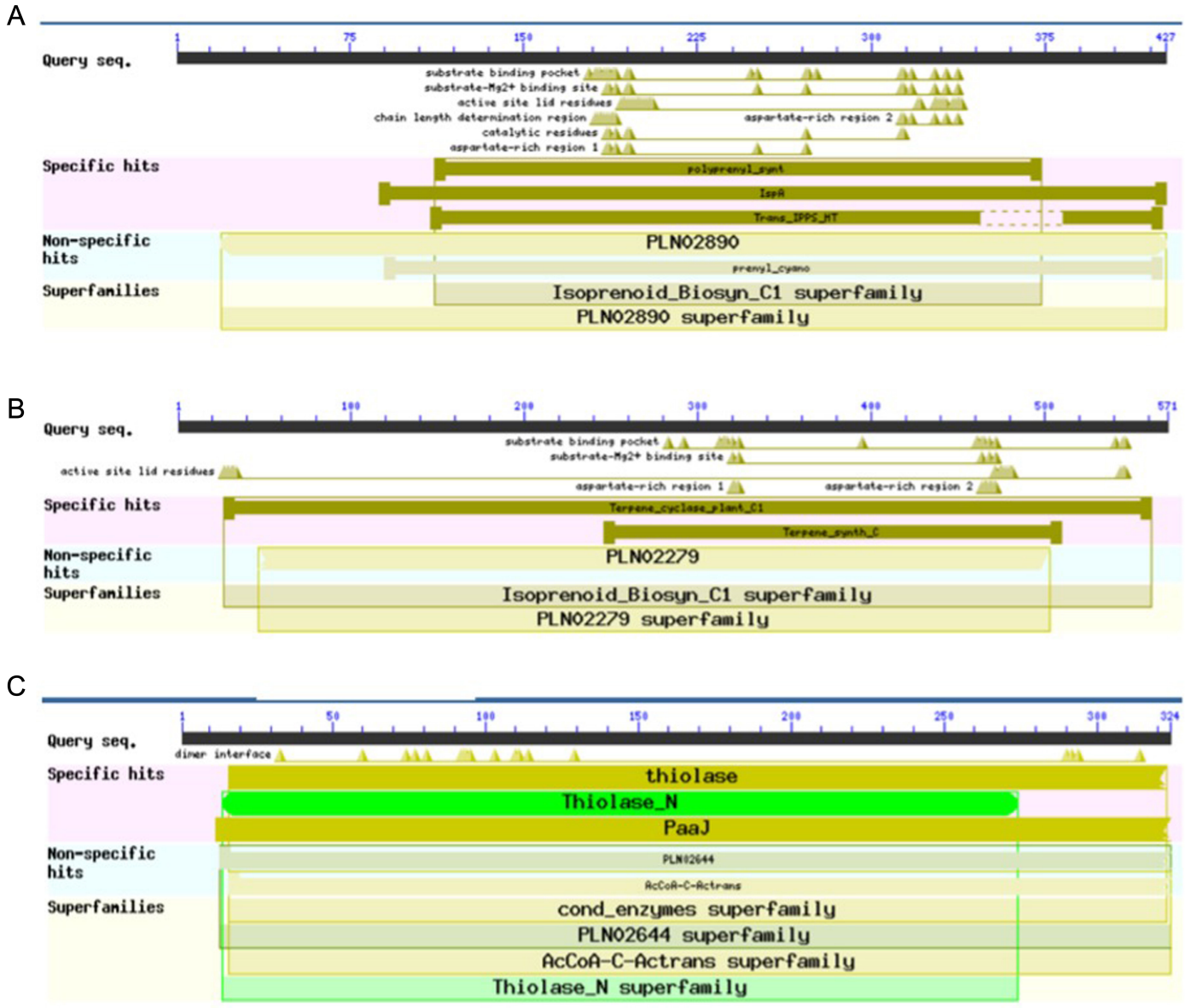


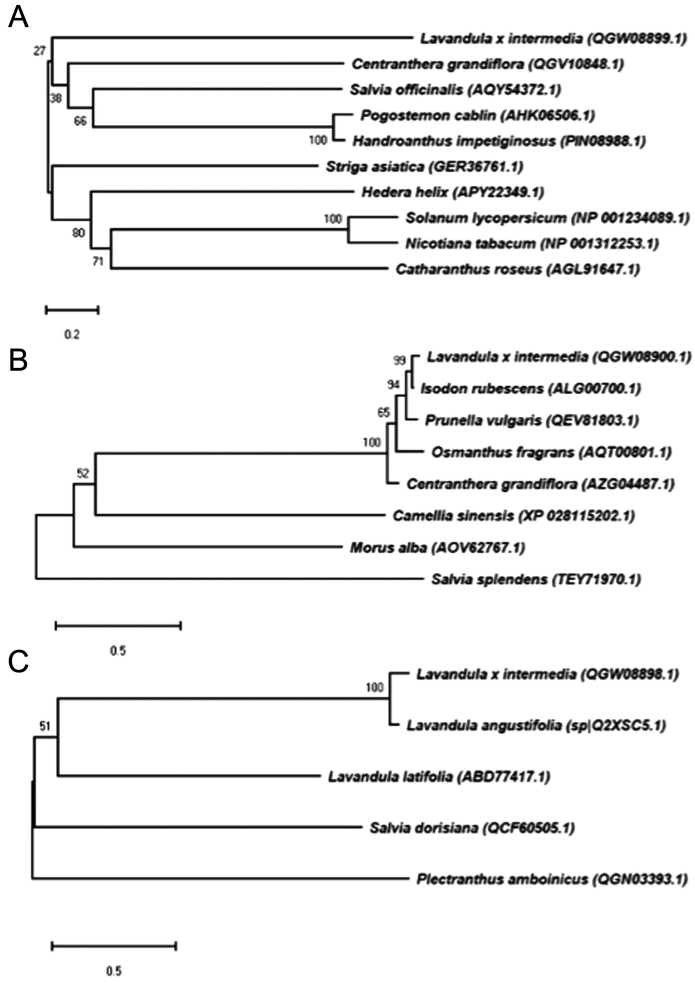
Figure 4.
Conserved Domains of LiGPP (A), LiLIS (B), and LiLAT (C) in L. intermedia Which Were Predicted by the Conserved Domains Database in NCBI Using Target Gene Sequences.

expression was lower than *LiGPP* and *LiLAT* expression. The *LiGPP* had the highest expression quantity at 50% bloom, and it was more than 244 times higher than the mean expression rate of *LiLIS*.

Discussion

Lavender species have a high market value due to their high-quality essential oils with high linalool and linalyl acetate content (Sarker et al., 2012). Lavandin is planted to produce high-quality oil in large quantities. It is important to increase linalool and linalyl acetate content in lavandin to have the same high-quality essential oil as lavender. The essential oils of lavandins, like those of many other plants, are regulated by irregular monoterpenes produced by the head-to-tail condensation of isoprene units. The GPPS, LIS, and LAT are enzymes of the isoprenoid pathway to produce linalool and linalyl acetate, respectively. Despite the importance of monoterpenes, the genes involved in their synthesis have not been identified in lavandins. The *LiGPP*, *LiLIS,* and *LiLAT* encoding linalool and linalyl acetate were cloned and characterized, and their expression levels were measured for the first time in the present research. Cloning these genes would allow metabolic experiments to improve the production of linalool and linalyl acetate in commercially valuable lavender species like *L.* × *intermedia* (Sarker, 2013). Overexpression of these three genes can increase the level of components that determine the quality of essential oil in many lavender varieties.

Bioinformatics analysis showed that *LiGPP*, *LiLIS,* and *LiLAT* have high similarity to other related genes characterized in other plants. The deduced aa sequences of *LiGPP*, *LiLIS,* and *LiLAT* have typical motifs and domains of GPP, LIS, and LAT proteins. Multiple alignments and phylogenetic analyses showed that *LiLIS* and *L. angustifolia* LIS were found to be in the same clade, while *LiGPP*, *LiLAT,* and other lavender species GPP and LAT were found to be in a different clade, suggesting a genetic association between these



D

LiGPP		
Species	**Accession number**	**Identity (%)**
Salvia officinalis	KY399788.1	97
Mentha x piperita	AJ249453.1	96
Pogostemon cablin	MN013340.1	81
Centranthera grandiflora	MK290850.1	80
Citrus unshiu	AY166797.2	77
LiLIS		
Species	**Accession number**	**Identity (%)**
Lavandula angustifolia	DQ263741.1	95
Protium alvarezianum	KC881144.1	89
Perilla citriodora	AY917193.1	77
Osmanthus fragrans	FJ645727.1	76
LiLAT		
Species	**Accession number**	**Identity (%)**
Isodon rubescens	KR154220.1	98
Prunella vulgaris	MK272796.1	90
Sesamum indicum	XM_011097603.2	86
Osmanthus fragrans	KY451760.1	83
Olea europaea	XM_023012505.1	83

Figure 5.
Phylogenetic Tree of Target L. intermedia Proteins (A: LiGPP; B: LiLIS, C: LiLAT) With Other Proteins.

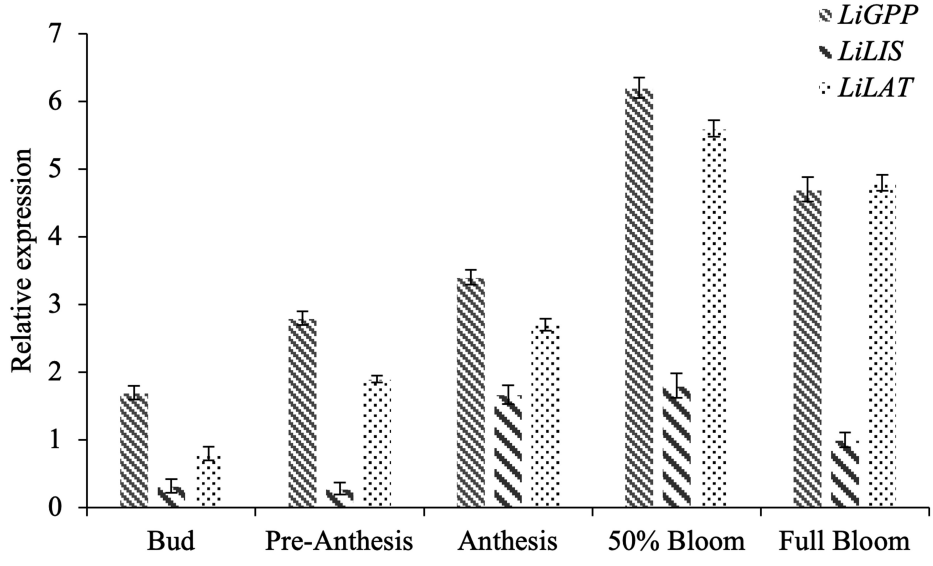


Figure 6.
Transcriptional Activity of LiGPP, LiLIS, and LiLAT in Different Developmental Stages of L. intermedia Flower Tissues. Bud: 1–2 cm, Green; Pre-Anthesis: 3–4 cm, Violet; Anthesis: 5–8 cm, Violet; 50% Bloom: 50% of Flowers on Spike Opened; Full Bloom: All Flowers on Spike Opened of L. × intermedia. To Normalize Gene Expression, β-actin Was Used As a Housekeeping Gene.

protein pairs. Based on their conserved structure and sequence characteristics, these results suggest that LiGPP and LiLAT are less conserved across plants. Subsequent sequence analysis indicated that *LiGPP* has two conserved aspartate-rich motifs, FARM ($DD(X)_{2-4}D$) and SARM (DDXXD), which are essential for catalysis and substrate linkage (Kellogg & Poulter, 1997). However, *LiGPP* lacks the small subunit binding CxxxC, which is necessary for the physical interaction between the two heterodimeric GPPS subunits (Hsiao et al. 2008; Wang & Dixon 2009), suggesting that the subunits bind together via hydrophobic interactions (Jones and Thornton, 1996). As previously reported, the CxxxC motif is required but not sufficient for physical interaction between the subunits (Wang & Dixon, 2009).

The *LiLIS* has a similar structure to other plant LISs and contains the characteristic conserved motifs found in these proteins such as the structural DDxxD and $DDLxTx_3E$ motifs, which are important for substrate linkage and ionization (Whittington et al., 2002). The *LiLIS* had strong sequence similarity to *L. angustifolia* LIS and had standard monoterpene synthase structural motifs, such as a transit peptide and the RRx_8W, DDxxD, and $DDLxTx_3E$ motifs.

The *LiLAT* was found to be associated with other characterized acetyltransferases from various plants. While both conserved motifs found in acyltransferases are present in these proteins, the tryptophan (W) of the DFGWG motif was replaced by methionine (M) in *LiLAT*. It has been hypothesized that in all acetyltransferase proteins, HxxxD and DFGWG motifs are highly conserved, and alterations of these motifs result in decreased activity of the enzymes (D'Auria, 2006; Unno et al., 2007). The *LiGPP*, *LiLIS,* and *LiLAT* genes were expressed in all developmental stages in flower tissues with varying levels (Figure 6). All genes were strongly expressed in flowers when the spikes opened, suggesting that they play important roles in essential oil composition.

In the present study, the highest expression of both target genes was observed at the 50% bloom stage, followed by a decrease at the full blooming stage, according to temporal expression analysis (Figure 6). These expression profiles were similar to the result of Boeckelmann (2008), who reported that the number of lavender LIS transcripts increased as the flower matured though linalool levels remained high, indicating that LIS transcription preceded the production of linalool, and linalool was retained after peak linalool synthesis. Moreover, Boeckelmann (2008) and Lane et al. (2010) reported that the expression of the English Lavender LIS gene increased with age till flower maturity. These results suggest that the roles of GPP, LIS, and LAT in scent biosynthesis in the lavandin flower should be investigated further. Also, when compared to bud-1, LIS transcript levels followed a previously reported trend (Lane et al., 2010), and were 9 and 12 times higher in the anthesis and 30% flowering stages, respectively (Demissie, 2014).

Conclusions and Recommendation

Since essential oils from lavandin are characterized by the presence of terpenes, the *LiGPP, LiLIS,* and *LiLAT* genes may be candidate genes for the regulation of essential oil components for lavandin. The artificial increase or stabilization of the abundance of linalyl acetate and linalool may be based on the isolation of putative GPP, LIS, and LAT genes from *L.* × *intermedia* cv. Super A. Taking these findings together, this research will help with improving essential oil quality and yield in lavandins through biotechnological modification and targeted breeding programs.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – D.Ö.; Design – D.Ö.; Supervision – Y.K.; Resources – Y.K.; Materials – D.Ö.; Data Collection and/or Processing – D.Ö., S.Ö.; Analysis and/or Interpretation – D.Ö., S.Ö.; Literature Search – D.Ö., S.Ö.; Writing Manuscript – D.Ö., S.Ö.; Critical Review – Y.K.

Declaration of Interests: The authors declare that they have no competing interest.

Funding: This work was supported by TÜBİTAK (The Scientific and Technological Research Council of Turkey), grant number 217O387.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Konsept – D.Ö.; Tasarım – D.Ö.; Denetim – Y.K.; Kaynaklar – Y.K.; Malzemeler – D.Ö.; Veri Toplama ve/veya İşleme – D.Ö., S.Ö.; Analiz ve/veya Yorum – D.Ö., S.Ö.; Literatür Taraması – D.Ö., S.Ö.; Yazma – D.Ö., S.Ö.; Eleştirel İnceleme – Y.K.

Çıkar Çatışması: Yazarlar çıkar çatışması bildirmemişlerdir.

Finansal Destek: Bu çalışma TÜBİTAK (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu) tarafından 217O387 proje numarası ile desteklenmiştir.

References

Adal, A. M., Sarker, L. S., Lemke, A. D., & Mahmoud, S. S. (2017). Isolation and functional characterization of a methyl jasmonate-responsive 3-carene synthase from Lavandula x intermedia. *Plant Molecular Biology*, *93*(6), 641–657. [CrossRef]

Baydar, H. (2009). *Tıbbi ve Aromatik Bitkiler Bilimi ve Teknolojisi* (Genişletilmiş 3. Baskı, Vol. *51*, pp. 194–212). SDÜ Yayınları.

Baydar, H., & Kineci, S. (2009). Scent composition of essential oil, concrete, absolute and hydrosol from lavandin (*Lavandula x intermedia* Emeric ex Loisel.). *Journal of Essential Oil Bearing Plants*, *12*(2), 131–136. [CrossRef]

Beetham, J., & Entwistle, T. (1982). *The cultivated lavenders* (1st ed, Entwistle). Royal Botanic Gardens.

Blažeković, B., Stabentheiner, E., Brantner, A., & Vladimir-Knezevic, S. (2012). A Comparative Study on Glandular Trichomes of *Lavandula x intermedia*, Budrovka' and *L. angustifolia*. In *Phyton-Annales Rei Botanicae 52*(2), 227–244.

Boeckelmann, A. (2008). *Monoterpene production and regulation in lavenders (Lavandula angustifolia and Lavandula x intermedia)* (Doctoral Dissertation). University of British Columbia.https://doi.org/10.14288/1.0066802

Burke, C. C., Wildung, M. R., & Croteau, R. (1999). Geranyl diphosphate synthase: Cloning, expression, and characterization of this prenyltransferase as a heterodimer. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(23), 13062–13067. [CrossRef]

Crowell, A. L., Williams, D. C., Davis, E. M., Wildung, M. R., & Croteau, R. (2002). Molecular cloning and characterization of a new linalool synthase. *Archives of Biochemistry and Biophysics*, *405*(1), 112–121. [CrossRef]

D'Auria, J. C. (2006). Acyltransferases in plants: A good time to be BAHD. *Current Opinion in Plant Biology*, *9*(3), 331–340. [CrossRef]

Demissie, Z. A. (2014). *Cloning and functional characterization of genes involved in the biosynthesis and secretion of essential oil constituents of Lavandula* (Doctoral Dissertation). University of British Columbia. https://doi.org./10.14288/1.0074341

Dudareva, N., Cseke, L., Blanc, V. M., & Pichersky, E. (1996). Evolution of floral scent in Clarkia: Novel patterns of S-linalool synthase gene expression in the C. breweri flower. *Plant Cell*, *8*(7), 1137–1148. [CrossRef]

Gang, D. R., Wang, J., Dudareva, N., Nam, K. H., Simon, J. E., Lewinsohn, E., & Pichersky, E. (2001). An investigation of the storage and biosynthesis of phenylpropenes in sweet basil. *Plant Physiology*, *125*(2), 539–555. [CrossRef]

Gershenzon, J., McCaskill, D., Rajaonarivony, J. I., Mihaliak, C., Karp, F., & Croteau, R. (1992). Isolation of secretory cells from plant glandular

trichomes and their use in biosynthetic studies of monoterpenes and other gland products. *Analytical Biochemistry*, *200*(1), 130–138. [CrossRef]

Guenther, E. (1952). *The essential oils* (Vol. *5*). Robert E. (pp. 3-38) Krieger publishing company.

Hall, T. (2004). BioEdit version 7.0. 0. Distributed by the Author. [CrossRef] Accessed June 13, 2020.

Hsiao, Y. Y., Jeng, M. F., Tsai, W. C., Chuang, Y. C., Li, C. Y., Wu, T. S., Kuoh, C. S., Chen, W. H., & Chen, H. H. (2008). A novel homodimeric geranyl diphosphate synthase from the orchid Phalaenopsis bellina lacking a DD (X) 2–4D motif. *Plant Journal*, *55*(5), 719–733.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., & Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Sciences*, *23*(10), 403–405. [CrossRef]

Jones, S., & Thornton, J. M. (1996). Principles of protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(1), 13–20. [CrossRef]

Kamran, H. M., Hussain, S. B., Junzhong, S., Xiang, L., & Chen, L. Q. (2020). Identification and molecular characterization of geranyl diphosphate synthase (GPPS) genes in wintersweet flower. *Plants*, *9*(5), 666. [CrossRef]

Kellogg, B. A., & Poulter, C. D. (1997). Chain elongation in the isoprenoid biosynthetic pathway. *Current Opinion in Chemical Biology*, *1*(4), 570–578. [CrossRef]

Kıvrak, Ş. (2018). Essential oil composition and antioxidant activities of eight cultivars of Lavender and Lavandin from western Anatolia. *Industrial Crops and Products*, *117*, 88–96. [CrossRef]

Lafhal, S., Vanloot, P., Bombarda, I., Kister, J., & Dupuy, N. (2016). Identification of metabolomic markers of lavender and lavandin essential oils using mid-infrared spectroscopy. *Vibrational Spectroscopy*, *85*, 79–90. [CrossRef]

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

Lin, Y. L., Lee, Y. R., Huang, W. K., Chang, S. T., & Chu, F. H. (2014). Characterization of S-(+)-linalool synthase from several provenances of *Cinnamomum osmophloeum*. *Tree Genetics and Genomes*, *10*(1), 75–86. [CrossRef]

Magnard, J. L., Bony, A. R., Bettini, F., Campanaro, A., Blerot, B., Baudino, S., & Jullien, F. (2018). Linalool and linalool nerolidol synthases in roses, several genes for little scent. *Plant Physiology and Biochemistry*, *127*, 74–87. [CrossRef]

Muñoz-Bertomeu, J., Arrillaga, I., Ros, R., & Segura, J. (2006). Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. *Plant Physiology*, *142*(3), 890–900. [CrossRef]

Renaud, E. N. C., Charles, D. J., & Simon, J. E. (2001). Essential oil quantity and composition from 10 cultivars of organically grown lavender and lavandin. *Journal of Essential Oil Research*, *13*(4), 269–273. [CrossRef]

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols*, *3*(6), 1101–1108. [CrossRef]

Sarker, L. S., Galata, M., Demissie, Z. A., & Mahmoud, S. S. (2012). Molecular cloning and functional characterization of borneol dehydrogenase from the glandular trichomes of Lavandula x intermedia. *Archives of Biochemistry and Biophysics*, *528*(2), 163–170. [CrossRef]

Sarker, L. S. (2013). *Cloning of Lavandula essential oil biosynthetic genes* (Doctoral Dissertation). University of British Columbia. https://dx.doi.org/10.14288/1.0073585

Schilmiller, A. L., Schauvinhold, I., Larson, M., Xu, R., Charbonneau, A. L., Schmidt, A., Wilkerson, C., Last, R. L., & Pichersky, E. (2009). Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10865–10870. [CrossRef]

Strommer, J., Gregerson, R., & Vayda, M. (1993). Isolation and characterization of plant mRNA. *Methods in Plant Molecular Biology and Biotechnology*, 49–65, Boca Raton, FL, USA: CRC

Sugiura, M., Ito, S., Saito, Y., Niwa, Y., Koltunow, A. M., Sugimoto, O., & Sakai, H. (2011). Molecular cloning and characterization of a linalool synthase from lemon myrtle. *Bioscience, Biotechnology, and Biochemistry*, *75*(7), 1245–1248. [CrossRef]

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, *30*(12), 2725–2729. [CrossRef]

Tholl, D., & Lee, S. (2011). Terpene specialized metabolism in Arabidopsis thaliana. *Arabidopsis Book*, *9*, e0143. [CrossRef]

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. [CrossRef]

Tucker, A. O. (1985). Lavender, spike, and lavandin. *Herbarist*, *51*, 44–50.

Unno, H., Ichimaida, F., Suzuki, H., Takahashi, S., Tanaka, Y., Saito, A., Nishino, T., Kusunoki, M., & Nakayama, T. (2007). Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. *Journal of Biological Chemistry*, *282*(21), 15812–15822. [CrossRef]

Upson, T., Andrews, S., Harriott, G., King, C., & Langhorne, J. (2004). *The genus Lavandula* (pp. 442). Timber Press.

van Schie, C. C., Haring, M. A., & Schuurink, R. C. (2007). Tomato linalool synthase is induced in trichomes by jasmonic acid. *Plant Molecular Biology*, *64*(3), 251–263. [CrossRef]

Wang, G., & Dixon, R. A. (2009). Heterodimeric geranyl(geranyl)diphosphate synthase from hop (Humulus lupulus) and the evolution of monoterpene biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(24), 9914–9919. [CrossRef]

Weiss, E. A. (1997). *Essential oil crops* (pp. 600). CAB International.

Whittington, D. A., Wise, M. L., Urbansky, M., Coates, R. M., Croteau, R. B., & Christianson, D. W. (2002). Bornyl diphosphate synthase: Structure and strategy for carbocation manipulation by a terpenoid cyclase. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(24), 15375–15380. [CrossRef]

Wu, T., Wang, Y., & Guo, D. (2012). Investigation of glandular trichome proteins in Artemisia annua L. using comparative proteomics. *PLoS One*, *7*(8), e41822. [CrossRef]

Zaks, A., Davidovich-Rikanati, R., Bar, E., Inbar, M., & Lewinsohn, E. (2008). Biosynthesis of linalyl acetate and other terpenes in lemon mint (Mentha aquatica var. citrata, Lamiaceae) glandular trichomes. *Israel Journal of Plant Sciences*, *56*(3), 233–244. [CrossRef]