IDENTIFICATION, CLONING, AND FUNCTIONAL CHARACTERIZATION OF *LAVANDULA* FLORAL DEVELOPMENT GENES

by

Rebecca Wells

BSc. (Honours) The University of British Columbia, 2017

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

MASTER OF SCIENCE

in

THE COLLEGE OF GRADUATE STUDIES

(Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Okanagan)

August 2020

The following individuals certify that they have read, and recommend to the College of Graduate Studies for acceptance, a thesis/dissertation entitled:

Identification, cloning, and functional characterization of *Lavandula* floral development genes

submitted by Rebecca Wells in partial fulfillment of the requirements of

the degree of Master of Science in Biology

Dr. Soheil Mahmoud, Biology, Irving K. Barber School of Arts and Sciences

Supervisor

Dr. Michael Deyholos, Biology, Irving K. Barber School of Arts and Sciences

Supervisory Committee Member

Dr. Louise Nelson, Biology, Irving K. Barber School of Arts and Sciences

Supervisory Committee Member

Dr. Hadi Mohammadi, Faculty of Applied Sciences, School of Engineering

University Examiner

Abstract

The molecular aspects of flower initiation and development have not been adequately investigated in lavender (*Lavandula*). In order to identify genes that control these processes, I employed RNA-Seq to obtain sequence information for transcripts originating from the vegetative shoot apical meristem (SAM) and developing inflorescence tissues of *Lavandula angustifolia* and *Lavandula* × *intermedia* plants, and assembled a comprehensive transcriptome of 105,294 contigs. Homology-based annotation provided gene ontology terms for the majority of transcripts, including over 100 genes homologous to those that control flower initiation and organ identity in *Arabidopsis thaliana*. Expression analysis revealed that most of these genes are differentially expressed during flower development. For example, *LaSVP*, a homolog of the floral repressor *SHORT VEGETATIVE PHASE (SVP*), was strongly expressed in vegetative SAM compared to developing flowers, implicating its potential involvement in flowering repression in lavender. In total, fourteen full-length lavender transcription factor homologs were identified. To investigate these genes further, I constitutively expressed each gene in transformed *A. thaliana* plants, evaluating the effects on flower initiation and morphology. Expression of the *Lavandula* genes in *A. thaliana* was gene- dependent. Phenotypic effects were observed for *LaCC*, *LaFT*, and *LaSVP* transformants. For example, constitutive expression of the *LaSVP* gene in *A. thaliana* delayed flowering and affected flower organ identity in a dosage- dependent manner. Four of the highest expressing lines produced sepals instead of petals and were sterile as they failed to develop proper seed pods. This study provides the foundation for future investigations aimed at elucidating flower initiation and development in lavender.

Lay Summary

This thesis provides insight into lavender floral developemnt. Floral development has been described in other flowering plants but has yet to be studied in lavenders. Lavenders are important both economically and medicinally due to their aromatic essential oils, which are widely used in a variety of settings. As these oils are mainly produced in the flowers of these plants, it is beneficial to improve understanding of floral development. This thesis identified genes involved in flowering initiation and pattern development in lavenders and examined the functions of these genes. This was achieved through the development of a transcriptome, identification of over 100 floral development genes, and through functional characterization of a subset of these genes in *A. thaliana*. This research begins a new area of research and the results of this thesis provide a better understanding of floral development in lavenders, which will facilitate future studies and improvements in oil production.

Preface

Design of this project was completed by Dr. Soheil Mahmoud (Professor, UBC Okanagan). All experimentation and analyses of research data in this thesis were completed by Rebecca Wells, with the exception of some data generated for *LaSVP*. An honours student, Lina Bauer, was given several *Lavandula* genes identified in this thesis to characterize for her honours thesis. Results achieved by Lina's *LaSVP* transformants have been included in this thesis in addition to the *LaSVP* results of the present study. Results of *LaSVP* transformant analysis are published in *Planta* (Wells et al. 2020). The heatmap and RT-PCR figures were prepared with assistance from Dr. Ayelign Adal. This thesis was written with the guidance of Dr. Soheil Mahmoud and was reviewed and approved by my supervisory committee: Dr. Soheil Mahmoud, Dr. Louise Nelson, and Dr. Michael Deyholos.

Chapters 1 through 4 contain text and figures which have been published in the journal *Planta*. Parts of these chapters were reprinted/adapted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia*, Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020, license 4900371164986.

The publication entitled "*Cloning and Functional Characterization of a Floral Repressor Gene from Lavandula angustifolia*" was primarily written by Rebecca Wells and Lina Bauer, with the aid of Dr. Soheil Mahmoud and Dr. Ayelign Adal and with contributions from Elaheh Najafianashrafi in collecting data for the phylogenetic analysis. Rebecca Wells was the primary author with the duty of preparing the manuscript for submission with the aid of Dr. Soheil Mahmoud and Dr. Ayelign Adal.

While completing my thesis, I also was first author on the following publication:

Wells R, Truong F, Adal AM, et al (2018) Lavandula essential oils: a current review of applications in medicinal, food, and cosmetic industries of lavender. Nat Prod Commun 13:1403–1417. doi: 10.1177/1934578x1801301038

In the preparation of "*Lavandula Essential Oils: A Current Review of Applications in Medicinal, Food, and Cosmetic Industries of Lavender*", Rebecca Wells and Felisha Truong were the primary authors, with contributions by Dr. Ayelign Adal and Dr. Lukman Sarker. Rebecca Wells was the primary author with the duty of preparing the manuscript for submission with the aid of Dr. Soheil Mahmoud and Dr. Ayelign Adal.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BD	Floral bud
cDNA	Complementary DNA
EO	Essential oil
FQM	Floral Quartet Model
F_1	Filial 1
FL	Flower
gDNA	Genomic DNA
LaAG	*Lavandula angustifolia Agamous*
LaAP1	*Lavandula angustifolia Apetala 1*
LaAP3	*Lavandula angustifolia Apetala 3*
LaARF5	*Lavandula angustifolia Auxin Response Factor 5*
LaCC	*Lavandula angustifolia Crab's Claw*
LaCAL	*Lavandula angustifolia Cauliflower*
LaEMF2	*Lavandula angustifolia Embryonic Flower 2*
LaFLC	*Lavandula angustifolia Flowering Locus C*
LaFT	*Lavandula angustifolia Flowering Locus T*
LaSEP3	*Lavandula angustifolia Sepallata 3*
LaSOC1	*Lavandula angustifolia Suppressor of CONSTANS1*
LaSVP	*Lavandula angustifolia Short Vegetative Phase*
LiAP2	*Lavandula x intermedia var. Grosso Apetala 2*
LD	Long day
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
mRNA	Messenger RNA
RT-PCR	Real-time Polymerase Chain Reaction
SAM	Shoot apical meristem
SD	Short day
TF	Transcription factor
TPS	Terpene Synthase

SVP	*Short vegetative phase*
VM	Vegetative meristem

Acknowledgements

I would like to thank all the people who have played a role during the completion of my graduate studies. I would like to thank Dr. Soheil Mahmoud for the opportunity to work in his lab for both my honours thesis and my graduate studies and for his guidance and support throughout my undergraduate and graduate studies. I would like to thank my supervisory committee for their support, feedback, and guidance throughout my graduate studies. I would also like to thank Dr. Ayelign Adal of Dr. Mahmoud's lab for his assistance and mentorship during my graduate studies. I would like to thank Melissa Larrabee and Elaheh Najafianashrafi for their friendship and support.

I would like to express my gratitude to the faculty members at UBC Okanagan that I had the privilege of being taught by during my undergraduate and graduate studies. I would like to thank my professors for nurturing my passion for learning and academia. I would also like to thank the professors I have had the opportunity to work with during my undergraduate and graduate studies. I would like to especially thank Dr. Brendan D'Souza, Dr. Richard Plunkett, Dr. Paul Shipley, Dr. Michael Russello, Dr. Janet Kluftinger, Dr. Soheil Mahmoud, Dr. Louise Nelson, Dr. Bruce Mathieson, Dr. Michael Deyholos, and Dr. Milka Beck. I would also like to extend thanks to Barb Lucente (Biology Unit Assistant) for her continued administrative support throughout my undergraduate and graduate studies at UBCO. I would like to thank UBC Okanagan for the generous financial assistance through scholarships, awards, and bursaries throughout my undergraduate and graduate studies at UBC Okanagan, including the University Graduate Fellowship scholarships and other awards I have been granted.

I would like to thank my family and friends who have supported me throughout my graduate studies. I would like to extend my sincere appreciation to Bill and Careyn Head for their endless guidance, love and support throughout my graduate studies.

I would like to especially thank Kevin Head, for his continuous encouragement, generosity, patience, guidance, love and support throughout my graduate studies and who has been instrumental in the completion of my thesis. I would like to sincerely thank him for always supporting me and believing in me. I would like to express my heartfelt appreciation to him for his unwavering love and dedication and for being my biggest support throughout my graduate studies. I sincerely thank him and I dedicate this thesis to him.

Dedication

…To Kevin…

Chapter 1: Introduction

1.1 Lavenders

1.1.1 The Genus *Lavandula*

Lavenders belong to the genus *Lavandula*, family *Lamiaceae*, which includes many flowering perennial species that produce aromatic essential oils (EO) (Lis-Balchin 2002; Vakili et al. 2014). Members of the genus exhibit distribution across Madeira, the Canary Islands, Cape Verde, North Africa, the Mediterranean, South West Asia, North East Africa, the Arabian Peninsula, and India, although native to the Mediterranean region (Lis-Balchin 2002; Ziaee et al. 2015). Lavenders, as with other members of the *Lamiaceae* family, exhibit two-lipped flowers comprised of five fused sepals, square-shaped stems, a bilobed stigma, opposite oriented leaves, four stamens, and produce nutlet fruits (Upson and Andrews 2004). The genus comprises 39 species and three subgenera; subgenus *Lavandula*, which is comprised of all woody shrubs, subgenus *Fabricia* comprised of woody shrubs, suffruticosa shrubs, or herbaceous perennials, and subgenus *Sabaudia*, which is comprised of woody shrubs (Lis-Balchin 2002; Upson and Andrews 2004). The three subgenera can be further classified based on leaf and floral morphology, growth pattern as well as chromosome number (Upson and Andrews 2004). For example, members of the genus *Lavandula* are polyploid and thus contain a variable number of chromosomes depending on the species, with *L. aristibracteata* containing 18 chromosomes, *L. x intermedia* containing 51 chromosomes, and *L. latifolia* containing 36, 54, or 75 chromosomes (Upson and Andrews 2004). Recent genomic analyses have provided evidence to suggest the lavender genome is 770 Mbp – 880 Mbp in size (Urwin et al. 2007; Malli et al. 2019). In addition, members of the genus *Lavandula* vary with regards to their leaf and floral morphology. For example, members of the subgenus *Lavandula* exhibit leaves that are linear to elliptic, which are occasionally shallowly lobed (also found in the subgenus *Sabaudia*), whereas all species of the *Lavandula* genus exhibit variation in leaf number, size, extent of dissection, with some being sessile, petiolate, or even lacking leaves entirely (Upson and Andrews 2004).

Lavenders are presently grown globally for their valuable ornamental flowers and essential oils (EO) in Spain, the United States, France, the UK, Portugal, Bulgaria, Australia and China (Ziaee et al. 2015). Cultivation of lavender and use of the plant dates back to ancient eras when the Greeks and Romans cultivated *Lavandula stoechas* for medicinal purposes and Victorian times when the EO was used in perfumes and cosmetics (Lis-Balchin 2002; Upson and

Andrews 2004). Lavenders can be cultivated in a variety of climates and can grow in rocky, arid, nutritionally poor soils, although there is variation in cold and frost hardiness across species (Upson and Andrews 2004). Currently, the most commonly cultivated varieties of lavenders include *Lavandula angustifolia* Mill (*L. ang*), *Lavandula latifolia* Medik., *Lavandula stoechas*, and *Lavandula x intermedia* Emeric ex Loisel (*L. x int)*, which is a sterile hybrid of *L. angustifolia* and *L. latifolia* (Koulivand et al. 2013; Wells et al. 2018). Lavenders are commonly cultivated for their flowers as well as for extraction of their valuable aromatic EO, which are produced in in structures called glandular trichomes found mostly in the flowers.

1.1.2 Lavender Floral Structure

The typical angiosperm flower exhibits a distinct organization of floral organs from the outermost whorl to the innermost whorl: sepals, petals, stamens, and carpels (Litt and Kramer 2010). The estimated 350,000 species of angiosperms share some floral features such as stamens and carpels; however, variation exists in organ number or appearance as well as addition of organs and variation in the presence or absence of certain organs (Litt and Kramer 2010). Like many flowering plants, lavenders exhibit the typical floral organs, although they exhibit a unique floral organization and growth pattern. An individual lavender flower is comprised of five fused sepals (the calyx), five fused petals (corolla), four stamens, and a pistil which contains a four-locular ovary arising from two fused carpels (Upson and Andrews 2004; Wells et al. 2020). Although most commonly violet, *Lavandula* flowers can be white, blue-violet, pink, and in some cases, black or yellow in colour (Upson and Andrews 2004). Lavenders grow in dense shrub-like bushes with flowers growing atop long, woody stems. Lavender flowers are densely packed on stem tops in structures called “spikes” and stems of lavender can exhibit varying degrees of branching (Upson and Andrews 2004). Each floral spike is comprised of numerous tiny flowers densely packed around the stem, arising from structures called “cymes”, which can each give rise to a single flower or to many flowers (Upson and Andrews 2004). These cymes give rise to flowers on each side of the stem (Upson and Andrews 2004). Below the main floral spike, some species of lavender exhibit an additional layer of flowers, referred to as the remote verticillaster (Upson and Andrews 2004). Further, ach lavender flower is considered hermaphroditic as it contains both male and female reproductive organs; however, lavenders rely on pollination via bees and other insects to maintain genetic diversity in the species (Upson and Andrews 2004).

Although a small amount of EO is produced in the leaves, the majority of lavender EO is produced in the flowers, making lavender flowers a valuable component of the plant.

1.1.3 Essential Oil Composition and Applications

The distinct aroma of lavenders is due to their production of fragrant EO in structures within the floral tissue. The fragrant EO of lavenders are produced mainly in the flowers and to a small extent in the leaves in structures called glandular trichomes (Sarker et al. 2013). *Lavandula* EO are complex, biologically active EO, comprised of varying amounts of over 50 monoterpenes and sesquiterpenes, such as linalyl acetate, cineole, 1,8-cineole, linalool, borneol, cis/trans-ocimene, lavandulyl acetate, α-pinene, β-pinene camphor, *β*-ocimene, and terpinen-4-ol, among others (Sarker et al. 2012; Koulivand et al. 2013; Lesage-Meessen et al. 2015; Wells et al. 2018; Adal et al. 2019). Although similar constituents are found in the commonly cultivated species' EO, the relative amount of these constituents in the EO can vary depending on the species which can also affect the aroma, chemical activity, and use of the EO of certain *Lavandula* species (Sarker et al. 2012). Variance in the percent composition of the constituents in the oils across species can affect the aroma, quality, chemical activity, as well as the suitability of the EO for certain commercial uses (Sarker et al. 2012). For example, *L. x intermedia* EO contain higher amounts of camphor compared to *L. angustifolia* EO (Kara and Baydar 2013). In contrast, *L. angustifolia* EO contain higher amounts of linalool and linalyl acetate (Lesage-Meessen et al. 2015). The differences in EO profiles account for the differences in aroma of the oils and implicate different species' oils in different products. Additionally, EO quality, composition and yield can be affected by growth conditions, climate, cultivation technique and extraction method and, as such, different wild populations of the same species have shown distinct oil profiles due to differences in these factors (Muñoz-Bertomeu et al. 2007). EO quality is affected by the relative ratios of certain mono- and sesquiterpenes in the oils, with the most desirable constituents being linalool and linalyl acetate and these compounds produce a pleasing aroma and are abundant in *L. angustifolia* EO (Sarker et al. 2012). In lower quality oil, such as that of *L. x intermedia*, there is relatively low (compared to oils of *L. angustifolia*) levels of linalool and linalyl acetate, and *L. x intermedia* EO typically contain high amounts of camphor which provides a less pleasant and stronger odor (Landmann et al. 2007; Sarker et al. 2012). This character of *L. x intermedia* EO makes it well-suited for use in cleaning products, whereas *L.*

angustifolia, being low in camphor, exhibits a more delicate aroma and is thus more commonly used in perfumes and personal care products (Lesage-Meessen et al. 2015). Further, the biologically active compounds in the EO have numerous beneficial properties including insecticidal, antifungal, antibacterial, sedative, analgesic, anti-inflammatory, and antioxidant properties. *Lavandula* essential oils are thus widely used in a variety of commercial products, from perfumes and soaps to food, flavorings, and medicines, making lavenders commercially and medicinally important plants (Wells et al. 2018).

Due to the high commercial value and diversity of applications, the world production of *L. angustifolia* EO is 200 tons per annum, with top *L. angustifolia* EO and *L. x intermedia* EO producing countries being Bulgaria and the UK (Lesage-Meessen et al. 2015). The world per annum production of *L. x intermedia* EO is approximately 1000 tons, with about 90% produced in France (Lesage-Meessen et al. 2015). In addition, lavender varieties differ in their per hectare production of EO as *L. x intermedia* produces approximately 100 kg/ha, while *L. angustifolia* produces approximately 15 kg/ha (Lesage-Meessen et al. 2015).

The biosynthetic pathways of lavender EO constituents have been well-characterized, although the development of the organs in which these EO are produced, the flowers, have yet to be examined (Woronuk et al. 2010; Demissie et al. 2011). There exists variation in the amount of EO produced as well as in the quality of EO produced by the various species of commonly cultivated lavenders.

1.2 *Arabidopsis thaliana*

1.2.1 About *Arabidopsis thaliana*

Arabidopsis thaliana, a member of the mustard family (Brassicaceae/Cruciferae), is a small angiosperm and a model plant used in a wide range of studies to effectively model eukaryotic systems (Meinke et al. 1998). The *A. thaliana* genome is comprised of only 22,000 genes encompassed on 5 chromosomes (Meinke et al. 1998). In addition, many *A. thaliana* genes have been studied and sequences are readily available for a wide range of genes. Arabidopsis plants are small, leafy plants which originate from Asia, Europe, and North America and produce tiny white flowers, with five petals and that produce many small seeds (Meinke et al. 1998). As such, *A. thaliana* is a well- studied model plant that lends well to use in a variety of studies and transformation procedures are well-documented (Meinke et al. 1998; Arabidopsis Genome

Initiative 2000). The most commonly utilised ecotypes include Landsberg erecta and Columbia, which lend well to studies (Meinke et al. 1998).

Arabidopsis plants can grow effectively in a variety of climates and under various growth conditions, making propagation and regeneration of these plants simple and efficient in both culture and soil. For example, *A. thaliana* can grow under a variety of light conditions such as under long day or short day conditions and in varied nutrient conditions such as in store-bought soil mixes or growth medium in greenhouses, growth chambers, or freely in the environment (ABRC Arabidopsis Biological Resource Center 2009). Although Arabidopsis can grow in a variety of settings, growth and progression to flowering in *A. thaliana* is highly dependent on light conditions due to the effects of light on the plants' internal rhythms (Searle and Coupland 2004a). Arabidopsis is photoperiod sensitive and flowers earlier under long days (LD) conditions than under short day (SD) conditions (Eriksson et al. 2006). As such, the length of the vegetative growth stage and timing of transition to the flowering stage can be altered through manipulation of the growth conditions. For example, *A. thaliana* are light-sensitive plants and highly responsive to alterations in light intensity as well as photoperiod. As such, the length of the vegetative growth stage and the timing of the initiation of flowering in *A. thaliana* can be manipulated by changing photoperiod. For example, when *A. thaliana* are grown in light conditions above 12 hours, flowering is initiated and early flowering can be induced in young plants when exposed to light periods above 16-hours (Searle and Coupland 2004a). Conversely, flowering initiation can be delayed when plants are exposed to photoperiods with short periods of light (less than 12 hours) and long periods of darkness (above 12 hours) (Searle and Coupland 2004a). The high dependency of *A. thaliana* flowering on photoperiod lends well to experimentation as flowering can be induced early or delayed by manipulation of light conditions to long-day or short-day conditions, respectively. As such, *A. thaliana* served as an excellent plant for the present study.

1.2.2 Life Cycle and Development

Arabidopsis plants have a short life cycle, which makes them an ideal model plant and facilitates ease of use for genetic studies across generations. The entirety of the life cycle of *A. thaliana* can be completed in approximately six weeks, from seed germination to the maturation and seed set of the plants (Meinke et al. 1998). The life cycle of *A. thaliana* begins with a small

brown-orange seed about 0.5 mm in diameter, which requires a period of cold exposure (vernalization) before germination may occur (Meinke et al. 1998). After vernalization, light exposure is required to induce germination, at which point, the seedling produces two small true cotyledons. As seedling maturation continues, the plant's vegetative mass increases and a rosette leaf pattern is produced (Meinke et al. 1998). After sufficient vegetative mass has been achieved, buds are produced which give rise to small white flowers at approximately 4-5 weeks of age, dependent on light conditions (Meinke et al. 1998).

At maturity, each Arabidopsis flower is comprised of four floral whorls (outermost to innermost) of sepals, petals, stamens, carpel and ovules (Honma and Goto 2001; Smaczniak et al. 2012; Theißen et al. 2016). Arabidopsis flowers thus exhibit both male and female floral organs and are capable of self-pollination (Meinke et al. 1998). Arabidopsis flowers are approximately 2 mm in length and give rise to long seed pod structures called siliques, which produce the seeds and can hold over 1000 seeds at maturity (Meinke et al. 1998). The siliques house the developing seeds and, at maturity, become dry and break open, releasing the seeds. The short life cycle and the production of numerous seeds makes *A. thaliana* an excellent model plant to examine processes across multiple generations. In addition, floral initiation and development have been extensively studied and Arabidopsis is the model plant for describing these processes in plants.

1.2.3 Floral Initiation and the Transition from Vegetative to Reproductive Growth

At some point in the angiosperm life cycle, the plant must determine the appropriate time to flower and initiate the transition from vegetative growth to reproductive growth. The appropriate timing of this stage is crucial as it is key to the plant's reproductive success and is dependent on the integration of a variety of intricate signalling pathways and environmental cues. Plants must, therefore, possess mechanisms to sense their external environment and internal status to ensure appropriate timing of reproduction initiation (Kim and Sung 2014). This process is well-described in *A. thaliana* and follows a predictable sequence of events to produce the intricately organized flower. The timing of initiation of flowering can be affected by external factors such as light quality and/or quantity, photoperiod, vernalization period, nutrient availability, temperature, and water availability (Levy and Dean 1998; Boss et al. 2004). In addition, flowering can be induced by environmental stress such as crowding, or drought conditions or by endogenous signals generated within the plant, such as hormone signalling, age,

or sufficient increase in biomass (Levy and Dean 1998; Boss et al. 2004). The various environmental and endogenous cues sensed by the plant impact the expression of a group of floral integrator genes which impact the transition from a vegetative meristem to a reproductive meristem (Boss et al. 2004). In addition, while some pathways serve to activate the expression of certain TFs or flowering promoting genes, others serve to regulate the expression of flowering repressor proteins (Boss et al. 2004). Many plants must reach a certain level of vegetative mass to allow the plant to establish itself prior to reproductive growth, and as such, juvenile plants may be unresponsive to signals encouraging reproductive growth before sufficient vegetative mass is achieved (Levy and Dean 1998; Boss et al. 2004).

During the stage of embryogenesis in flowering plants, the SAM is formed and is the meristem which will give rise to the entirety of the above-ground growth of the plant, including leaves, internodes, flowers, stems, and all lateral branches (Barton 1998; Lenhard and Laux 1999; Bowman and Eshed 2000). Initially, vegetative growth serving to increase vegetative biomass occurs at the rapidly dividing pluripotent cells of the shoot apical meristem (SAM), producing leaves in an intricate rosette pattern (Coen and Meyerowitz 1991; Yanofsky 1995; Liljegren et al. 1999). These undifferentiated cells within the SAM eventually differentiate and give rise to leaves and stems (Barton 1998). The SAM is not only a source of pluripotent stem cells which differentiate to distinct tissues, but is also the location at which the distinctive growth pattern of the stems and leaves is established (Barton 1998). Further, the SAM can be divided into 3 developmental zones: the peripheral zone containing cells which are differentiated and from which organs develop, the central zone, which contains pluripotent cells which give rise to new growth, and the rib meristem which is the site of production of the cells within the width of the stem (Barton 1998; Lenhard and Laux 1999). Cells within the SAM are pluripotent and thus have the unique capability of the potential to give rise to different tissue types as they have yet to be genetically assigned to a specific tissue (Barton 1998). As vegetative growth proceeds vertically, in the apical meristem, there exist specific bundles of cells on the sides of the meristem which are conserved to give rise to flower meristems or additional vegetative meristems (Coen and Meyerowitz 1991). This vegetative growth phase will continue until a signal is received, at which point, the SAM will convert to an inflorescence meristem which will produce an inflorescence meristem and, subsequently, a flower meristem, while suppressing leaf formation (Coen and Meyerowitz 1991; Yanofsky 1995; Koornneef et al. 1998a). Flowering is

bi-directional during this transition, as after the SAM has become a flowering meristem, all new primordia formed will produce flowers, but any leaf primordia formed prior to the transition will be affected by these signals and will become bracts rather than true rosette leaves (Hempel and Feldman 1994). The inflorescence meristem produces small leaves in the same rosette structure as before; however, the flower meristem produces flowers, with each meristem giving rise to a single flower (Coen and Meyerowitz 1991). Floral bud formation occurs at the edges of the floral meristem in a continuous manner, forming new buds in an indeterminate spiral pattern (Smyth et al. 1990). The inflorescence meristem will produce numerous floral meristems, whereas each floral meristem will produce only one flower with a distinct organizational pattern of the floral organs (Yanofsky 1995).

The molecular mechanisms which govern this transition have been well-described in *A. thaliana* and flowering time is governed by multiple factors. In Arabidopsis, a variety of factors, including environmental factors, such as photoperiod and ambient temperature, influence the stage at which the plant commences flowering (Hartmann et al. 2000; Andrés et al. 2015). Photoperiod is especially important as Arabidopsis plants commence flowering in response to long photoperiods as they are long-day (LD) plants and also require vernalization prior to germination (Levy and Dean 1998). In addition, overexpression of some floral meristem identity genes can lead to an early flowering phenotype in Arabidopsis (Levy and Dean 1998). Flowering initiation requires the suppression of floral repressors or genes that promote vegetative growth, expression of genes which promote transition to a flowering meristem, and later, genes which facilitate the development of floral organs (Levy and Dean 1998; Andrés et al. 2015). In addition, genes referred to as flowering time genes are either constitutively expressed or constitutively repressed or their expression can be influenced by a variety of environment conditions (Hartmann et al. 2000). The induction of expression of floral development genes is dependent on developmental cues and changing seasonal environmental cues such as temperature, moisture, stress, and photoperiod which aid plants in determining the appropriate time to begin flowering (Hayama and Coupland 2003; Searle and Coupland 2004a). Flower initiation and development in *A. thaliana* thus may occur in response to external cues such as temperature, photoperiod, day length, vernalization, and light quality (Simpson and Dean 2002). For example, the protein CONSTANS (CO) is sensitive to photoperiod and when activated in response to long day length, it activates expression of downstream genes, such *FLOWERING*

LOCUS T (FT) and *SUPRESSOR OF OVEREXPRESSION OF CO1 (SOC1)*, which promote flowering (Koornneef et al. 1998a; Samach et al. 2000a; Wells et al. 2020). In addition, genes such as *FLOWERING LOCUS C (FLC), SHORT VEGETATIVE PHASE (SVP)* and *EMBRYONIC FLOWER* 1 and 2 *(EMF 1* and *EMF2)* are repressors of flowering (Koornneef et al. 1998a; Hartmann et al. 2000; Samach et al. 2000a). Thus, while some genes stimulate flowering, others act as repressors

The initiation of flowering thus requires the activation and suppression of certain genes in pluripotent meristematic cells to convert the plant's vegetative meristem into a flowering meristem (Bowman and Eshed 2000; Andrés et al. 2015). For example, the genes *LEAFY* and *APETALA1* are expressed in the flower meristem and promote flower development via suppression of shoot identity genes and flowering repressors, as well as by facilitating expression of homeotic genes required for normal flower development (Maoileidigh et al. 2014) (Maoileidigh et al. 2014). Thus, flowering time genes can be classified as repressors or activators of flowering and can be autonomously expressed or their expression can be influenced by exogenous/endogenous signals (Levy and Dean 1998).

In *A. thaliana* , there are several types of genes involved in flowering within the floral meristem: floral meristem identity genes, floral integrators, and floral pattern development genes (Levy and Dean 1998; Theißen et al. 2016). The floral meristem identity genes include *LEAFY (LFY), APETALA1 (AP1), CAULIFLOWER (CAL),* and *UNUSUAL FLORAL ORGANS (UFO)* and serve to facilitate the transition of the meristem from a vegetative meristem to a flowering meristem by inducing certain cells in the inflorescence meristem to give rise to a floral meristem (Levy and Dean 1998; Kim et al. 2009). Floral integrator genes play a crucial role in the integration of endogenous and exogenous cues and affect the expression of floral meristem identity genes and include *FT*, *FD*, and *SUPPRESSOR OF CONSTANS1 (SOC1)* (Henderson and Dean 2004; Kim et al. 2009). The regulation of the expression of the floral integrator genes is dependent on external cues such as temperature, as well as internal cues, such as hormone levels (Kim et al. 2009). These genes integrate signals received to have an effect on the floral meristem identity genes based on information gathered from external and internal cues. It is of note, however, that some genes may have more than one role in flowering induction (Kim et al. 2009). The floral organ development genes require activation and function in floral pattern development by dictating which floral organ will develop in a given area of the meristem after

floral initiation (Levy and Dean 1998). Floral initiation is also governed by genes which are affected by signalling pathways that are acted upon by numerous endogenous and exogenous signals to either promote or inhibit flowering (Levy and Dean 1998). These pathways converge on a common set of integrator genes which activate downstream genes in response to the overlapping effects of the signalling pathways (Levy and Dean 1998; Boss et al. 2004). Thus, the initiation of flowering is commenced in response to the integration by the plant of many environmental and endogenous cues, enabling the plant to sense its environment and internal state to ensure flowering is initiated at the appropriate time (Hartmann et al. 2000; Boss et al. 2004).

1.3 Flowering Inductive Pathways

In *A. thaliana*, there exist four main pathways which govern the initiation of flowering (Levy and Dean 1998; Kim and Sung 2014). The photoperiodic pathway and the vernalization pathways initiate flowering in response to environmental factors, with the photoperiodic pathway inducing flowering in response to long photoperiod and the vernalization pathway activating flowering in response to a period of cold exposure (vernalization) (Levy and Dean 1998). These pathways are thus responsive to environmental conditions and activate or repress flowering based on external cues (Levy and Dean 1998). In addition, two other pathways have been described- an autonomous pathway which increasingly promotes flowering as the plant matures, and the gibberellin pathway (Levy and Dean 1998; Kim and Sung 2014).

1.3.1 Flowering in Response to Photoperiod

Seasonal changes in photoperiod play an important role in dictating when many plants like *A. thaliana* initiate flowering. Different plants flower under various light conditions. *A. thaliana*, for example, is a long day (LD) plant, which flowers earlier under LD conditions than under short day (SD) conditions (Searle and Coupland 2004a). Flowering time in *A. thaliana* is thus closely tied to photoperiod and Arabidopsis initiates flowering through the activation of certain flower development-inducing genes (Koornneef et al. 1998a; Hayama and Coupland 2003; Searle and Coupland 2004a). This photoperiodic response is governed by an internal clock which regulates the expression of flowering-related genes in response to environmental cues, allowing plants to respond to changes in photoperiod which varies on a seasonal basis (Hayama

and Coupland 2003; Searle and Coupland 2004a). Regulation of the rhythms in plants by this circadian clock occurs when the 24-hour circadian rhythm is generated and responds to cues, such as photoperiodic changes, by activating or repressing certain biochemical or developmental pathways, such as flowering-inducing pathways (Dunlap 1999). The transcriptional regulation and post-transcriptional regulation of certain genes is also impacted by light exposure (Hayama and Coupland 2003). The pathway senses photoperiod and relays this information to the internal signalling pathways, thus promoting flowering via the photoperiod pathway. This response to changes in light strength and length is translated into developmental changes through "flowering time genes" and their interaction with an internal circadian clock/rhythm system (Hayama and Coupland 2003). These "flowering-time genes" are governed by cues such as developmental stages, or environmental cues such as temperature, photoperiod or other seasonal changes and aid in the development of the floral meristem from which flowers arise (Becker and Theißen 2003; Searle and Coupland 2004a). It is this internal sensory system which integrates light information input and causes an effect on biochemical and signalling pathways in the plant, either promoting or inhibiting flowering (Dunlap 1999). The central system for responding to changes in photoperiod is comprised of the circadian rhythm which controls developmental pathways and the pathways which transmit the changing light signals to the internal system (Searle and Coupland 2004a). This system provides a response to environmental cues and controls the circadian rhythm and the responsive signalling pathways, allowing plants to respond to changes in photoperiod (Hayama and Coupland 2003). The internal circadian clock is responsible for producing the daily 24-hour internal rhythms and processes within the plant (Dunlap 1999).

In *A. thaliana* and other plants, the changes in photoperiod and light quality are sensed by photoreceptors which relay the information to an internal clock mechanism (Koornneef et al. 1998a; Hayama and Coupland 2003). These photoreceptors, called phytochromes and cryptochromes, detect and transmit the light signals to the internal circadian pathway to influence internal rhythms based on light quality and quantity changes (Hayama and Coupland 2003). This circadian pathway is an internal autonomous clock, acting to create an internal 24-hour rhythm (Hayama and Coupland 2003). The clock is comprised of temperature and light signalling pathways which are input pathways, an internal oscillator which integrates these signals, and output pathways whose activation leads to a range of biological effects (Hayama and Coupland

2003). Thus, flowering in response to photoperiodic changes is facilitated by the action of numerous floral development genes which are activated or repressed by the oscillator in response to changes in photoperiod (Searle and Coupland 2004a). In addition, mRNA levels of these floral initiation genes can vary, with accumulations and decreases in mRNA of certain genes occurring at different times of the day (Searle and Coupland 2004a). It has also been suggested that light activates some of these flowering time proteins post-translationally (Hayama and Coupland 2003).

The central pathway which translates light signals into developmental changes is comprised of the genes *LATE ELONGATED HYPOCOTYL (LHY)*, *TIMING OF CAB EXPRESSION1 (TOC1)*, *CIRCADIAN CLOCK ASSOCIATED1 (CCA1)* and mutations in any of these genes can affect the response to changing light conditions (Hayama and Coupland 2003). The phytochromes and cryptochromes sense the environmental light conditions and transmit this information to *TOC1* and *LHY*/*CCA1* (Hayama and Coupland 2003). These genes comprise the central oscillator which, in turn, regulates output pathways, such as flowering in response to photoperiod, via the activation of *CONSTANS* (*CO*) expression, which facilitates the activation of *Flowering Locus T (FT)* expression and floral initiation (Hayama and Coupland 2003). The gene *CO* facilitates initiation of flowering under long photoperiods, is highly expressed diurnally and can also become activated by exposure to light post-translationally (Robson et al. 2001; Hayama and Coupland 2003). The activation of *CO* in response to increased day length and favourable light conditions leads to flowering initiation through the downstream activation of additional genes such as *FLOWERING LOCUS T (FT)*, *AtP5CS2*, and *SOC1 (SUPPRESSOR OF CONSTANS1)* (Samach et al. 2000a; Hayama and Coupland 2003; Andrés et al. 2015). Further, the *Atp5CS2* gene leads to stem elongation, while the *FT* and *SOC1* genes promote flowering through the activation of the floral meristem identity genes (Samach et al. 2000a). The floral meristem identity genes in Arabidopsis include *LEAFY (LFY)*, *APETALA1 (AP1)*, *CAULIFLOWER (CAL)*, *APETALA2 (AP2)*, and *FRUITFUL* (*FUL*), whose expression is required for the transition from a vegetative meristem to a floral meristem (Henderson and Dean 2004; Andrés et al. 2015). In addition, the *FT* gene interacts with the transcription factor FD to activate the expression of the genes *AP1* and *LFY* in the meristem (Wigge et al. 2005; Andrés et al. 2015). As these genes both promote flowering and floral pattern development, the *FT* gene in this way mediates a response to a long daylength that subsequently initiates flowering (Andrés et

al. 2015). In addition, the *FT* gene works to activate other flowering genes such as *FRUITFUL (FUL), SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1), LEAFY (LFY), CAULIFLOWER (CAL), APETALA3 (AP3),* and *APETALA1 (AP1),* which are MADS-Box family transcription factors (Hayama and Coupland 2003; Mizoguchi et al. 2005; Andrés et al. 2015). Thus, the integration of external signals and the subsequent expression of certain genes together promotes the transition from a vegetative meristem to a floral meristem, induction of flowering, and floral development in *A. thaliana* (Wigge et al. 2005).

1.3.2 The Vernalization Pathway

Another important environmental aspect which plants sense to determine the appropriate time to flower is vernalization, which is the exposure of seeds to cold temperatures for some period of time (Kim et al. 2009). Plants, especially LD flowering plants, rely on cold exposure to sense the passing of winter and the commencement of spring/summer and impending LD photoperiod required for growth, although the temperature and duration of cold exposure may vary across species (Kim et al. 2009). Further, the requirement for cold exposure before floral initiation ensures that plants do not commence flowering in the fall prior to winter, which would thwart reproductive efforts due to inability to survive cold temperatures (Kim et al. 2009). In addition, many plants requiring a vernalization period also require the LD photoperiods of spring to flower (Kim et al. 2009).

Vernalization also impacts repressive pathways in *A. thaliana* which prevent flowering under unfavorable conditions. For example, vernalization reduces the level of transcripts of the repressor *Flowering Locus C* (*FLC*) and this repression is maintained throughout cold exposure (Sheldon et al. 2000). FLC and other related genes act as repressors of flowering through the repression of expression of the floral integrator genes, *SOC1*, *FT*, and *FD*. The FLC protein is a MADS-Box protein which represses flowering and is upregulated by some genes including (*FRIGIDA*) FRI and downregulated by other genes and vernalization (Michaels and Amasino 1999; Henderson and Dean 2004). As vernalization serves to promote flowering, the action of *FLC* as a repressor opposes vernalization. Once a plant has reverted to a vernalized state, this state is maintained and carried through to new growth, possibly due to epigenetic modifications (Henderson and Dean 2004). Further, overexpression of *FLC* can lead to delays in flowering (Michaels and Amasino 1999). Expression of the *FLC* gene is repressed by the vernalization

pathway via chromatin modifications (Kim et al. 2009). During long periods of cold exposure, transcripts of the *FLC* protein become near undetectable and expression of the *FLC* protein remains in an inactive state and is suggested to be in such a state due to stable epigenetic modifications (Michaels and Amasino 1999; Henderson and Dean 2004; Simpson 2004). Vernalization is a mitotically stable process, meaning that after cold exposure commences, the vernalization state is maintained in all cells until flowering is initiated (Boss et al. 2004).

1.3.3 The Autonomous Pathway

The autonomous pathway is comprised of several proteins which promote floral initiation through the downregulation of expression of the repressor protein, FLC (Henderson and Dean 2004; Simpson 2004). Although they function in different mechanisms, the genes of the autonomous pathway all function to ultimately repress *FLC* (Boss et al. 2004). In this way, the autonomous pathway functions additively with the vernalization pathway to promote flowering. The autonomous pathway promotes flowering and functions independently from other flower-inducing pathways, requiring no input signals (Henderson and Dean 2004; Simpson 2004). The autonomous pathway dictates the capacity which other flowering pathways can respond to promoting cues by controlling the level of *FLC* mRNA (Simpson 2004). In this way, the gene of the autonomous pathway dictate how responsive the plant is to the pathways that promote flowering (Simpson 2004). Thus, the levels of mRNA of *FLC* are reduced by vernalization, and repressed by autonomous pathway genes, mRNA processing, and through epigenetic modifications (Henderson and Dean 2004; Simpson 2004).

1.3.4 The Gibberellin Pathway

Gibberellic acid (GA) is a plant hormone which is required for the growth, development, and induction of flowering in many plants, including *A. thaliana*, which depends on GA signalling to induce flowering under SD conditions (Wilson et al. 1992; Gocal et al. 2001b; Moon et al. 2003a; Jung et al. 2012). The GA pathway thus promotes flowering initiation in *A. thaliana* and other plants (Henderson and Dean 2004). GA signalling promotes flowering through the activation of *SOC1* and through the expression of the meristem identity gene, *LFY*, to induce flowering under SD conditions (Moon et al. 2003a; Eriksson et al. 2006). The *LFY* gene is a floral meristem identity gene and GA signalling leads to increased transcription from

the LFY promoter, increasing concentration in the SAM and facilitating development of the flower (Blázquez et al. 1998). Further, some plants transition from vegetative growth to reproductive growth in the presence of exogenous GA and a form of GA_4 is seen to increase in concentration during SD at the SAM just prior to flowering (Wilson et al. 1992; Eriksson et al. 2006).

Thus, multiple pathways, including the vernalization, photoperiod, gibberellin (GA), ambient-temperature pathway, and the light quality pathway transmit information and regulate expression of the integrator genes *SOC1*, *FT*, and *LFY* (Levy and Dean 1998; Henderson and Dean 2004; Kim and Sung 2014). The integrator genes receive this information and initiate flowering through the activation of the floral meristem identity genes *AP1*, *AP2*, *FUL*, *CAL*, *LFY* which convert the vegetative meristem to a floral meristem (Henderson and Dean 2004). Transition to flowering is thus dependant on a variety of endogenous and exogenous factors (Koornneef et al. 1998a; Hartmann et al. 2000).

1.4 Transcription Factors

Transcription factors (TF) play an important role regulating gene expression in a variety of processes, including development, through activation or repression of transcription of mRNA from a gene's DNA (Latchman 1993). Transcription factors are protein elements of the cell which bind to distinct locations in the promoter region on the DNA to enhance or repress the transcription of a given gene via interactions with other transcription factors and RNA polymerase (Latchman 1993). As transcription is the first step of expression of a gene, transcription factors play a crucial role in the cell (Latchman 1993). Transcription factors can be agonistic, enhancing each others' activity, or antagonistic, interfering with each others' activity as a given transcription factor may repress a certain gene, while another activates the same gene. In this way, the cell can produce fewer TFs than if every gene required a unique set of transcription factors. Instead, a given TF may be a repressor of one gene, but an activator of another. In addition, some transcription factors require additional protein elements, such as coactivators to effectively activate gene expression. Additionally, the levels and types of transcription factors can vary during different developmental stages of an organism as well as across different tissue types (Eeckhoute et al. 2009). For example, certain genes may be activated only during certain stages of development, while others are activated at other stages.

Transcription factors bind to different sites in the regulatory region of a given gene on the DNA (Latchman 1993). These sites vary in nucleotide sequence as the DNA binding domain of a given transcription factor is distinct and thus capable of recognizing a specific recognition sequence. Further, many types of TFs have been identified and can be classified based on their DNA-binding domain structure and resulting dimerization pattern (Latchman 1993).

1.4.1 MADS-Box Transcription Factors

One such group of transcription factors is the MADS-Box transcription factors, which is a group of 107 TFs classified based on the presence of a highly conserved 180 base-pair DNA-binding domain called the MADS-box or MADS-domain at the amino terminus of the protein (Alvarez-Buylla et al. 2000; Pařenicová et al. 2003; Kaufmann et al. 2005). The MADS-box transcription factors are encoded by MADS-box genes and are so named based on the presence of the conserved DNA-binding domain which is named after the proteins MINICHROMOSOME MAINTENANCE 1, AGAMOUS, DEFICIENS, and SERUM RESPONSE FACTOR (Gramzow et al. 2010). The MADS-box proteins are found in a wide range of eukaryotic organisms including mammals, yeasts, insects, amphibians, and plants and play fundamental roles in important biological processes including metabolism, muscle development, regulation of oncogene expression and floral development (Shore and Sharrocks 1995; Theißen et al. 1996; Gramzow et al. 2010). In plants, the MADS-Box family of transcription factors is a developmentally important group with most of the members being genes of the floral meristem or floral identity genes which aid in specification of floral organ identity (Theißen et al. 1996; Becker and Theißen 2003). MADS-Box TFs can be classified as either Type I or Type II (Theißen et al. 1996). The type II MADS-box TFs are well-established as being implicated in regulating expression of genes involved in floral development (Pařenicová et al. 2003). MADS-Box TFs bind specifically to the CArG-box consensus sequence on DNA ($CC[A/T]_6GG$) of genes to be regulated (Kaufmann et al. 2005). Although there are several types of MADS-domain transcription factors, MIKC-type MADS-domain transcription factors are MADS-Box transcription factors exclusively found in plants and most are involved in organ identity or floral meristem identity (Theißen et al. 1996; Kaufmann et al. 2005). The MIKC-type MADS-domain transcription factors found in plants are so named due to the domains which they exhibit, containing a highly conserved MADS-Box domain at the amino terminus, an intervening (I)

sequence downstream of the MADS-Box, a weakly conserved K-domain, as well as a variable C-region or C-terminus domain (Ma et al. 1991; Theißen et al. 1996; Causier et al. 2010). The main DNA-binding domain is comprised of the I domain in combination with the MADS-box domain (Causier et al. 2010).

Consistent with other TFs, MADS-Box TFs affect gene expression via binding to specific sites on the DNA upstream of a given gene (Theißen et al. 1996). Most genes controlled via MADS-box genes are implicated in developmental processes in eukaryotic organisms and, more specifically, in floral development in plants (Theißen et al. 1996). The main DNA-binding domain, the MADS-Box, is a coiled coil comprised of two antiparallel, amphipathic α-helices which the DNA wraps around, allowing the helices to bind to the DNA minor groove to facilitate the TF binding (Pellegrini et al. 1995). In addition to facilitating appropriate DNA-binding, the MADS-Box of these TFs facilitates their localization to the nucleus as well as TF dimerization (Gramzow et al. 2010). In the developing flower, the Type II MADS-box TFs function not only through direct interactions with each other, forming complexes which directly regulate gene expression, but also through interactions with proteins such as chromatin remodelling proteins, among others (Smaczniak et al. 2012). These TFs form hetero- or homodimeric complexes of four TFs with other Type II MADS-box TFs through their K and I domains, whereas the C-domain is implicated in the formation of complexes with additional proteins (Causier et al. 2010). These TFs bind to CArG elements on the DNA of a given gene and promote bending of the DNA to facilitate transcription of the gene (Causier et al. 2010). The DNA binding domain binds to a sequence known as the CArG box with MADS-Box transcription factors binding in complexes of four, commonly known as “quartets” (Kaufmann et al. 2005). Although the roles of these transcription factors in floral pattern development have been described, there still remains more to discover, including the specific mechanism of action of these transcription factors and how and which genes are activated by these TFs (Smaczniak et al. 2012); However, it is known that the MADS-Box genes work through interactions with each other and other transcriptional coregulators to control the expression of these downstream genes of floral development and their expression is affected by upstream pathways (Smaczniak et al. 2012).

1.5 Floral Pattern Development in *Arabidopsis thaliana*

1.5.1 The ABC(DE) Model

Although flowers are highly variable in size, number, shape, aroma, and colour, most flowers exhibit a basic arrangement of common floral organs. The model describing floral pattern development is the ABC(DE) model of floral development (originally proposed as the ABC model, but later modified to include D and E-function genes) (Coen and Meyerowitz 1991; Weigel and Meyerowitz 1994; Causier et al. 2010; Theißen et al. 2016). The basic structure of a flower is comprised of four floral "whorls" with a distinct organ present in each whorl (Causier et al. 2010). Flowers most commonly exhibit sepals in the first outermost whorl, petals in the second whorl, stamens in the third whorl and carpels (containing ovules) in the innermost, fourth whorl (Litt and Kramer 2010; Theißen et al. 2016). The ABC(DE) model describes how differential expression of homeotic floral organ identity genes (Table 1) in areas of the meristem leads to the development of the appropriate floral organ in a given whorl (Weigel and Meyerowitz 1994; Causier et al. 2010). The model was originally postulated based on floral homeotic mutants, that is mutants which fostered a mutation which caused them to exhibit disorganization of the flower with fully developed floral organs in the incorrect whorl of the flower (Schwarz-Sommer et al. 1990; Bowman et al. 1991; Coen and Meyerowitz 1991; Weigel and Meyerowitz 1994; Theißen 2001). Further, the model describes the division of the angiosperm flower into four whorls and classifies the genes responsible for floral patterning in each respective whorl (Causier et al. 2010). The model further describes how in each whorl a unique combination of A, B, C, D, and E-type homeotic genes are expressed, giving rise to a unique floral structure (Coen and Meyerowitz 1991; Causier et al. 2010).

Table 1. *A. thaliana* genes comprising the ABC(DE) model in *Arabidopsis*. Genes are MIKC-type MADS-box transcription factors (with the exception of *AP2*) functioning in floral pattern development in Arabidopsis Bowman et al. 1991; Theißen 2001; Theißen et al. 2016).

Gene class	Gene(s)	Function	Whorl
A-class	*APETALA1* *APETALA2*	Sepal specification (in combination with E-class)	one
B-class	*APETALA3* *PISTILLATA*	Petal specification (in combination with A and E-class)	two
C-class	*AGAMOUS*	Stamen specification (in combination with E-class)	three
D-class	*SHATTERPROOF* *SEEDSTICK*	Carpel specification (in combination with D and E-class)	four
E-class	*SEPALLATA 1-3*	Specification of all floral organs	all

The ABC(DE) model explains that the expression of A-type genes gives rise to sepals in the first whorl, combined expression of A and B class genes produce petals in the second whorl, expression of B and C class genes together lead to stamen development in the third whorl, C-type genes facilitate carpel development in the fourth whorl, C and D- class genes give rise to ovules within the carpels, and E class genes are required in all floral whorls (Weigel and Meyerowitz 1994; Theißen et al. 2016). The A-class genes include *APETALA1 (AP1)* and *APETALA2 (AP2),* B-class genes include *APETALA3* (*AP3)* and *PISTILLATA (PI),* the C-class genes are represented by *AGAMOUS (AG),* the class D genes include *SHATTERPROOF (SHP)* and *SEEDSTICK (STK),* and the class E genes are represented by *SEPALLATA 1-3 (SEP1-3)* (Table 1 & Figure 1) (Theißen et al. 2016). It is of note that *AP1* and *AP2* also function as floral meristem identity genes, whose expression is required for the transition from a vegetative meristem to a flowering meristem (Henderson and Dean 2004). All of the genes described by the ABC(DE) model, with the exception of *AP2*, encode Type II MADS-Box transcription factors which promote the development of specific floral organs in the floral whorls through regulating the expression of other floral development genes (Causier et al. 2010; Theißen et al. 2016). Loss of expression of any of these genes leads to mutant phenotypes and deviations from the classic floral pattern (Pařenicová et al. 2003; Kram et al. 2009; Takeda et al. 2011; Wils and Kaufmann 2017). Although many plants exhibit the same floral organs, variation in structure and organization is observed across the over 350, 000 angiosperm species (Litt and Kramer 2010).

For example, some flowers may lack certain floral organs, may exhibit floral organ fusions, or may even exhibit novel floral organs (Litt and Kramer 2010). This high structural diversity observed across flowering plants resulting in deviations from the classical ABC(DE) model can be explained by alterations in the developmental programs such as gene duplications or deletions, or changes in gene expression patterns (Litt and Kramer 2010; Theißen et al. 2016). Thus, the appropriate expression and interaction of the A, B, C, D, and E-type genes is crucial to the proper organization and development of the angiosperm flower (Figure 1).

This thesis aimed to identify genes that control flower initiation and development in lavender, hypothesizing that these genes would be homologous to those that control these processes in *A. thaliana* and other flowering plants. A comprehensive transcriptome database corresponding to developing flowers of *L. angustifolia* and *L. x intermedia* plants was compiled and used to identify floral development-related genes. As a proof of concept, eight genes were cloned and functionally characterized through expression in *A. thaliana.* The results of this thesis indicate that the *Lavandula* genes perform a similar function to the *A. thaliana* homologs. Further, the database assembled in this thesis provides a valuable resource for cloning of genes that control flower development in lavenders. These genes may help improve flower timing, density (and hence EO yield) and visual appeal through plant biotechnology.

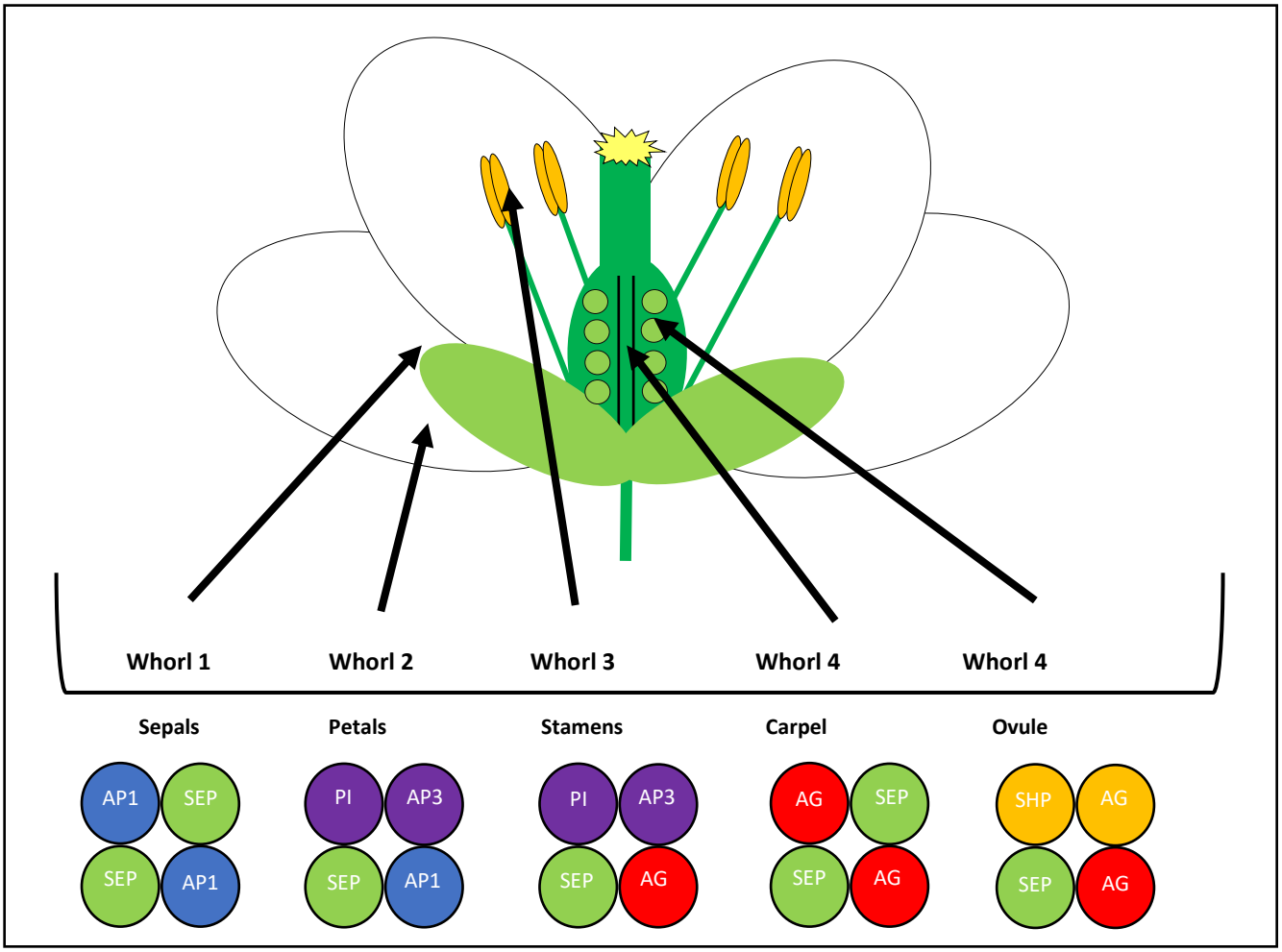

Figure 1. The ABC(DE) and the Floral Quartet models of floral development described in *A. thaliana* (Theißen et al. 2016). According to the ABC(DE) model, there exist A, B, C, D, and E-class genes which encode MADS-box transcription factors which govern floral pattern development in the different whorls of the flower. The A-class genes are *APETALA1* (AP1) and *APETALA2* (AP2), the B-class genes are *PISTILLATA* (PI) and *APETALA3* (AP3), the C-class gene is *AGAMOUS*, the D-class genes include *SHATTERPROOF* and *SEEDSTICK*, and the E-class genes include *SEPALLATA1-3* (Theißen and Saedler 2001; Theißen et al. 2016). The floral quartet model describes the mechanism in which these proteins dictate floral organ development in the floral meristem as tetrameric complexes. In whorl 1, two AP1 proteins and SEP proteins give rise to sepals. In the second whorl, the combination of one AP1 protein, one PI protein, one AP3 protein, and one SEP protein facilitates petal development. In the third whorl, one PI protein, one SEP protein, one AP3 protein, and one AG protein promote stamen identity. In the fourth whorl, carpels develop by the action of two AG proteins and two SEP proteins. Within the carpels, ovules develop via the combination of one SHP protein, one AG protein, one SEP protein, and one STK protein. These transcription factors form tetramers which bind to the promoter region of other genes to facilitate floral organ development.

1.6 Thesis Hypotheses and Objectives

The main objectives of this thesis were to gain an understanding of the molecular mechanisms which govern floral initiation and pattern development in lavenders.

1.6.1 Research Hypotheses

The hypotheses of this study were that:

1. Flower initiation and development in *Lavandula* are controlled by homologs of homeotic genes that control these processes in other plants including *Arabidopsis thaliana.*
2. Genes controlling flower initiation and development are expressed in *L. angustifolia*, but not in *L. x intermedia* plants grown under controlled light conditions in greenhouse (this hypothesis is based on the observation that under controlled light conditions, *L. angustifolia* plants produce flowers, but *L. x intermedia* plants do not). We further hypothesized that genes involved in flower initiation and development can be identified through differential gene expression analysis of greenhouse-grown *L. angustifolia* and *L. x intermedia* plants.
3. Genes involved in floral initiation and development in lavenders function similarly as in other flowering plants, and overexpression of these genes will visibly affect floral initiation and pattern development in transformed Arabidopsis plants.

1.6.2. Thesis Objectives

This thesis had the following specific objectives:

1. To assemble a comprehensive transcriptome representative of *L. angustifolia and L. x intermedia* tissues in CLC Genomics Workbench 11. This was achieved by growing *L. angustifolia* and *L. x intermedia* plants under controlled light conditions until L. angustifolia plants flowered. RNA was extracted at different stages of flower development from *L. angustifolia* plants, and from the shoot apical meristem of *L. x intermedia* plants to assemble a comprehensive transcriptome
2. To identify genes that potentially govern floral initiation and development in lavenders and to determine if they are homologous to those described in *A. thaliana* and other plants. The objective was met by i) using RNA-Seq to construct transcriptomes *for L.*

angustifolia and *L. x intermedia* plants ii) Searching the databases for homologues of *Arabidopsis thaliana* genes that are known to control flower initiation and development.

3. To examine differential expression of floral initiation and development genes across different lavender tissues through RNA-Seq and differential gene expression analysis in CLC Genomics Workbench 11. This objective was met through use of the RNA-Seq and Differential Expression functions of CLC Genomics Workbench 11 to examine gene expression patterns in the comprehensive lavender transcriptome assembled from the sequences from the various tissues.
4. To functionally characterize a subset of lavender floral initiation and development-related genes through cloning and constitutive ectopic expression in *Arabidopsis thaliana*. This objective was met through expression of *Lavandula* candidate genes in stably transformed *A. thaliana* plants under the control of the constitutive CaMV 35S promoter.

Chapter 2: Materials and Methods

2.1 Transcriptome Assembly

2.1.1. Plant Materials and Growth Conditions

L. angustifolia (dwarf lavender) and *L. x intermedia* plants were propagated from existing, mature plants growing in a greenhouse under controlled conditions of photoperiod of 16 hours light, 8 hours darkness at 25 ℃. Cuttings were obtained from the mature plants at axillary and terminal vegetative growth approximately three to four nodes below the meristem, with a total of 8 cuttings taken from *L. angustifolia* and 8 cuttings obtained from *L. x intermedia*. All leaves were trimmed from the cuttings except for the two uppermost leaves. The base of each cutting was dipped in rooting powder, planted in moistened ProMix soil, and covered with a plastic cover to maintain humidity. After approximately 1 week, covers were gradually removed. The young *L. angustifolia* and *L. x intermedia* plants were then grown in the greenhouse at 25℃ and a photoperiod of 16 hours light and 8 hours darkness. Plants were fertilized weekly with Miracle Gro fertilizer (Scott's Canada, Canada) prepared according to manufacturer's instructions. Young plants were transferred to larger pots, with a total of four *L. x intermedia* and four *L. angustifolia* plants transferred to growth to maturity (Wells et al. 2020).

Tissue was collected from plants once sufficient vegetative mass was established, after approximately 5 months. At the time of tissue collection, the *L. angustifolia* plants had produced flowers, whereas the *L. x intermedia* plants had yet to flower.

2.1.2 RNA Extraction and Sequencing

To obtain RNA from a variety of tissues, floral (FL), floral bud (BD), and shoot apical meristem (SAM) tissues were obtained from *L. angustifolia* and shoot apical meristem (SAM) tissues were obtained from *L. x intermedia* (Figure 2). *L. angustifolia* FL, BD and SAM tissues were obtained from one plant, and *L. x intermedia* SAM tissues were obtained from four plants. *L. x int* SAM tissues were obtained from four plants. Samples were cut and immediately flash-frozen in liquid nitrogen and stored at 80℃ (Wells et al. 2020).

Figure 2. *Lavandula* tissue from which RNA was extracted to assemble the lavender transcriptome. (a) *L. angustifolia* floral bud tissue (BD), (b) *L. angustifolia* shoot apical meristem tissue (SAM), and (c) *L. angustifolia* floral tissue (FL). SAM tissue was also obtained from *L. x intermedia* (not pictured) for transcriptome assembly, although BD and FL tissues were not available as flowering had not commenced in *L. x intermedia*. Arrow in (b) is indicative of location of SAM.

The total RNA was extracted in triplicates for each tissue using the QIAGEN RNEasy Plant Mini Prep Kit. Frozen tissues were ground to a fine powder using a pre-chilled mortar and pestle and extractions were performed according to the manufacturer's protocol. An amount of 5 µl of β-mercaptoethanol was added to the RLT buffer. An incubation with DNAse I was performed to remove any contaminating DNA from the sample. The Nanodrop-1000 Spectrophotometer was used to assess the integrity and concentration of the RNA. The RNA was run on a 1% agarose gel to visualize the RNA and determine RNA purity.

RNA was sequenced via Illumina Sequencing at the DNA Sequencing Core Facility of the University of British Columbia (Vancouver, B.C.). To prepare for sequencing, RNA samples were diluted to a concentration of 250 ng/µl and prepared in amounts of 20 µl. Prior to sequencing at the DNA Sequencing Core Facility, the RNA quality was ensured via the Agilent 2100 Bioanalyzer to ensure only samples of sufficient quality were used for sequencing. Subsequently, following the New England Biolabs standard protocol for NEBnext Ultra II stranded mRNA, the RNA samples were utilised for illumine library preparation. Using the Illumina NextSeq 500 via Paired End (42 bp x 42 bp reads), the RNA transcripts were sequenced. Finally, using Illumina's bcl2fastq, the transcript sequencing data was then demultiplexed and 40-43 base pair paired end reads were generated (Wells et al. 2020).

2.1.3 De Novo Transcriptome Assembly and Annotation

To assemble a transcriptome, the paired end reads generated by Illumina sequencing for *L. ang* FL, *L. ang* BD, *L. angustifolia* SAM, and *L. x intermedia* SAM were input as paired reads into CLC Genomics Workbench 11.0. Using the de novo sequence assembly tool, reads for each

tissue type were assembled into separate transcriptomes. Reads which had large stretches of unknown sequences, as indicated by "NNNNN", were omitted from the assemblies. The four tissue-specific assemblies were combined via *de novo* assembly to construct a single transcriptome, comprehensive of the *L. angustifolia* and *L. x intermedia* tissues.

The *De Novo* assembly parameters were set as word size of 22, bubble size of 50, and a minimum contig length of 200 bp (Wells et al. 2020). To remove redundant sequences from the assembled transcriptome, the CD-HIT-EST tool with 0.98 sequence identity cut-off was used (weizhonli-lab.org/jcvi/) (Li and Godzik 2006). The transcriptome was then annotated by BLASTx against an *A. thaliana* protein database. To verify completeness of the transcriptome, a BLASTx analysis was performed with the transcriptome and sequences of known *Lavandula* terpene synthase genes to identify these genes within the transcriptome. Transcripts within the *Lavandula* transcriptome showing high sequence similarity to *A. thaliana* floral development genes were searched for within the BLASTx results. *Lavandula* transcripts exhibiting high homology to *A. thaliana* proteins of known function, as indicated by low e-value ($\leq 10^{-10}$) were selected for further analysis. A low e-value was indicative of of a highly homologous match, wehereas higher e-values would indicate lower sequence homology between the given sequence and its match or matches (CLC bio a). Additionally, The BLAST2GO software in CLC Genomics Workbench 11 was then utilised for Gene Ontology mapping of the BLASTx results to provide functional analysis of the *Lavandula* transcripts within the transcriptome (Conesa et al. 2005). GO analyses were performed to classify transcripts into three main categories: molecular function, cellular component, or biological process and into one of many subcategories based on function (Figure 5)(Wells et al. 2020).

2.1.4 Differential Expression Analysis

Differential Expression Analysis for RNA-Seq data was performed using the CLC Genomics Workbench 11.0 software (Qiagen, USA) to compare expression patterns of floral development genes across different *Lavandula* tissues. Expression patterns were analyzed based on RPKM (reads per kilobase mapped) and fold change expression between the tissues. Sequences with high RPKM values and large negative or large positive fold change expression values were selected for further study.

2.1.5 Candidate Selection and ORF Identification

Candidate *Lavandula* floral development genes were selected based on high sequence homology to an *A. thaliana* floral development-related protein and by low e-value ($\leq 10^{-10}$). Candidates were further selected based on length and sequence completeness and differential expression pattern in the *Lavandula* tissues.

Each candidate transcript was then used in a BLASTx search against a previously-assembled *Lavandula* database to identify the most full-length sequence of the respective transcript. The open reading frame (ORF) of the *Lavandula* transcript was then determined by entering the transcript sequence into the NCBI Orf Finder at https://www.ncbi.nlm.nih.gov/orffinder/. The sequence and length of the ORF were compared with that of the respective *A. thaliana* protein and the ORF/protein sequence was determined. The corresponding cDNA sequence was then determined. If the *Lavandula* transcript did not generate the full-length ORF sequence, the *Lavandula* transcript was used in a BLASTx search against a previously- assembled *Lavandula* genomic DNA database to identify the full-length transcript within the genome. Once identified, the corresponding genomic sequence was entered into GeneScan (http://hollywood.mit.edu/GENSCAN.html) to identify all potential ORFs contained within the genomic sequence. The sequences of the generated ORFs were compared to the sequence of the *A. thaliana* protein of interest to identify the correct ORF of the homologous *Lavandula* protein. Once the full-length protein sequence was identified for each *Lavandula* candidate, the sequences were compared to the respective *A. thaliana* sequences to assess homology through Multiple Sequence Comparison by Log Expectation (MUSCLE) alignment (https://www.ebi.ac.uk/Tools/msa/muscle/). MUSCLE alignment was then performed to generate a percent identity value for each *Lavandula A. thaliana* homolog pair

2.1.6 Primer Design

Primers were designed based on the full-length sequence of the open reading frame (ORF) of each *Lavandula* candidate gene (Table 2). Primers were constructed to include the restriction enzyme cut sites for *Eco*RI and *Kpn*I at the 5’ and 3’ ends, respectively to facilitate cloning into the pGA482 (An 1986) vector for ectopic expression of the gene under the control of the constitutive CaMV promoter. The ORF sequence of each gene was entered at https://www.addgene.org/mol-bio-reference/restriction-enzymes/ to identify any restriction

enzyme cut sites present within the candidate gene sequences. Genes which contained either the *Eco*RI or *Kpn*I restriction enzyme recognition site within their ORF were identified and primers for these genes were designed for sticky end cloning (Yamabhai 2009). The length, %GC content, melting temperature, and annealing temperature of each primer were analyzed using OligoAnalyzer 3.1 and the Integrated DNA Technologies (IDT) website (https://www.idtdna.com/calc/analyzer).

Table 2. Gene-specific primers designed for *Lavandula* candidates. Primers were used in the PCR amplification of the respective candidate *Lavandula* floral development genes. Restriction enzyme cut site sequences for *Eco*RI and *Kpn*I in the forward and reverse primers, respectively, facilitated cloning into the pGA482 vector.

Gene Name	Forward	Reverse
Cloned		
CC	5'ATCTGGCGAATTCATGGATTTGGTTCAACAA-3'	5'-AACTTAGGTACCTCAGGCCTTAGTAATGCT-3'
CAL	5'AATATAGAATTCATGGGGAGGGGGAAGGTG-3'	5'TATATAGGTACCTCATGCGGCAAAGCATCC-3'
SVP	5'-TCCGCAGAATTCATGGCTAGAGAGAAGATT-3'	5'-ATTGTAGGTACCTCAGCCAGTGTAAGGAAG-3'
PI	5'-GCGCTCGAATTCATGAAAAAGGCTAAGGAG-3'	5'-TATATAGGTACCCTAGAACCGCTCCTGCAG-3'
AP1	5'-TACGCAGAATTCATGGGTAGAGGGAAAGTG-3'	5'-ATATAGGTACCTCATGCGGCAAAGCATCC-3'
FT	5'-TTCGCAGAATTCATGCCAAGAGATCGAGAC-3'	5'-TATATATATAGGTACCTCACCGTCTTCTCCCGCC-3'
Not Cloned		
SM	F1: 5'-AATTCATGCGTGTGTACAACAGTAGCACA-3' F2: 5'-CATGCGTGTGTACAACAGTAGCACAGA-3'	5'-ATATAGGTACCTTAGAGCAGCGCGGGCGA-3'
SEP	5'-ATTAGAATTCATGGGAAGAGGGAGAGTT-3'	R1: 5'-GTACCTTAGTTATCATGTAGCCACCCT-3' R2: 5'-CTTAGTTATCATGTAGCCACCCTGCTA -3'
AP2	5'-TATATAGAATTCATGGAATATTCCGGCAGC-3'	5'-TATTAGGTACCTCACCTCCAGTAGTAATG-3'
AP3	5'-TACTAGGAATTCATGGCTCGTGGGAAGATC-3'	5'-ACTAATGGTACCTCACTCAAGCAAGGCAAA-3'
SOC1	5'-ATCGTAGAATTCATGGTGAGGGGGAAGACT-3'	5'-TATTTAGGTACCCTATTTTTGGCACGGCCG-3'
AG	5'-CCCGTAGAATTCATGGAACAGCAAAGTGAT-3'	5'-GCTTCCGGTACCTCAGATTAACTGAAGAGA-3'
EMF 2	F1:5'-AATTCATGCCTGGCATTCCTCTCGTG -3' F2: 5'-CATGCCTGGCATTCCTCTCGTG-3'	5'-TATAAAGGTACCGCTCTGTGGGTCCG-3'
FLC	F1: 5'-AATTCATGGTGAGGGGGAAGGTGCAAATGAA-3' F2: 5'-CATGGTGAGGGGGAAGGTGCAAATGAAG-3'	5'-TATTTAGGTACC CGTCACCGCCTCCG-3'
LHY	5'-ACCGTAGAATTC ATGGATCCTTATTCTTCT-3'	5'-TATTTAGGTACC TCATGTAGAAGCCTCTC-3'
NPT-II	5'-AGAGGCTATTCGGCTATGACTG-3'	5' –TCAGAAGAACTCGTCAAGAAGGC-3'

2.2 Cloning and Functional Characterization

2.2.1 PCR Amplification of *Lavandula* Floral Genes

Total RNA was extracted from frozen *L. angustifolia* BD and FL tissue using the QIAGEN RNEasy Plant Mini Prep Kit and according to the manufacturer's protocol. *L. angustifolia* BD RNA was reverse transcribed to cDNA using the iScript cDNA Synthesis Kit (Biorad) according to the manufacturer's protocols and using the following cycle: 25 for 5 minutes, 42 ℃ for 1 hour, 85 ℃ for 5 minutes, and a 4 ℃ holding temperature. The *β-Actin* gene was PCR amplified from the cDNA to verify cDNA quality using the following cycle: 95°C for 5 minutes, then for 30 cycles, 95 ℃ for 5 minutes, 58°C for 30 seconds, 72°C for 45 seconds, followed by 72°C for 5 minutes and holding at 4°C. Floral bud cDNA was used as a template to amplify the *L. angustifolia* candidate genes, *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaGB*, *LaFT*, *LaAP1*, and *LaAP2* using the designed gene-specific primers (Table 2). Primers were centrifuged and resuspended in Tris-Hcl pH 7 to a 10 µM working stock. For sticky end primers, two primer working stocks were prepared (ie. forward primer 1 and reverse primer, forward primer 2 and reverse primer). For PCR amplification, a single Touchdown program was used to amplify all lavender candidate genes from the *L. angustifolia* BD cDNA. The PCR cycle was set as follows: 95 ℃ for 5 minutes for a single cycle, for 8 cycles: 95 ℃ for 30 seconds, 65 ℃ (-1) for 30 seconds, and 72 ℃ for 1 minute, followed by 33 cycles of 95 ℃ for 30 seconds, 57 ℃ for 30 seconds, 72 ℃ for 1 minute 30 seconds, followed by 1 cycle of 72 ℃ for 7 minutes with a final holding temperature of 4 ℃ after amplification was complete. All PCR reactions were subsequently verified on a 1% agarose gel to visualize the bands of RNA (Wells et al. 2020).

2.2.2 Cloning of *Lavandula angustifolia* Genes into the pGA482 Vector

Each amplified *L. angustifolia* candidate gene was cloned into a separate pGA482 vector to place the genes under control of the cauliflower mosaic virus (CaMV) 35S promoter of the vector and to facilitate transformation of separate *A. thaliana* plants. To clone the *L. angustifolia* genes into the multiple cloning site of the pGA482 vector, the vector and the cDNA of each gene were digested with the restriction enzymes, *Eco*RI and *Kpn*I. The digested vector and digested cDNA were then ligated together using DNA ligase to create the constructs, pGA482::*LaCC* , pGA482::*LaCAL*, pGA482::*LaSVP*, pGA482::*LaPI*, pGA482::*LaGB*, pGA482::*LaFT*, and pGA482::*LaAP1*. The constructs were then transformed into *Escherichia coli* JM109 cells.

2.2.3 Bacterial Growth and Plasmid Extraction

E. coli JM109 cells were grown in Liquid LB medium containing 30 mg/mL tetracycline. Bacteria was grown overnight at 37 °C in 50 mL LB medium with 30 mg/mL tetracycline at 200 rpm. Bacterial plasmids were extracted and purified from the JM109 *E. coli* cells using the Plasmid DNA Miniprep Kit (BIOBASIC INC) according to the manufacturer's protocol. Purity and quality of the extracted plasmids were verified via the Nanodrop Spectrometer 1000 (Table A.2) and via visualization on a 1% agarose gel.

Bacterial plasmids were extracted from the agarose gel using the E.Z.N.A.© Gel Extraction Kit (OMEGA). Plasmid DNA concentration and purity were verified using the Nanodrop Spectrometer 100 (Table A.3).

2.2.4 Restriction Digestion and Ligation

The extracted *E. coli* JM109 bacterial plasmids and the amplified *Lavandula* floral genes were digested with the restriction enzymes *Knp*I and *EcoR*I. A separate digestion reaction was prepared for each candidate gene. In a 1.5 mL tube, 5.0 µl of 10x NEB Cutsmart buffer, 3.0 µl plasmid DNA, 3 µl candidate cDNA, 2 µl high fidelity *Kpn*I, 2 µl high fidelity *EcoR*I and molecular grade water (to a total reaction volume of 50 µl) were combined. Reactions were incubated in a water bath at 37 °C for two hours and reactions were stored at -20 °C. An amount of 5 µl 10x loading dye was added to each 50 µl reaction and complete restriction digestion was verified through visualization of the reactions on a 1% agarose gel with a 1 kb ladder (FroggaBio). Digested plasmid DNA and *L. angustifolia* cDNA were extracted from the agarose gel using the E.Z.N.A.© Gel Extraction Kit (OMEGA).

Each gene's digested cDNA was ligated into a separate pGA482 plasmid. Each reaction was prepared in a PCR tube and contained a 1:5 ratio of vector to insert. Each reaction was comprised of 200 ng of candidate cDNA, 1 µl 10x T4 DNA Ligase buffer, 1.6 µl of pGA482 vector, 1 µl of T4 DNA Ligase and molecular grade water to bring the total reaction volume to 10 µl. A control reaction was prepared with 1 µl 10x T4 DNA Ligase buffer, 1.6 µl pGA482 vector, 1 µl T4 DNA Ligase and molecular grade water to a final reaction volume of 10 µl. Reactions were centrifuged briefly and mixed by pipette. Tubes were incubated at 16 °C overnight (approx. 15 hours) in a thermocycler. The tubes were then placed back in the

thermocycler with the following incubation program: 23 °C for 1 hour, followed by 65 °C for 20 minutes to complete ligation.

2.2.5 Heat-Shock Transformation of Competent *Escherichia coli* JM109

The ligation reactions were then used in heat-shock transformation of competent *Escherichia coli* (*E. coli*) JM109 cells via the method described by Froger and Hall (2008). Briefly, pipette tips and 2 mL microfuge tubes were placed in a -80 °C freezer to pre-chill the instruments. Competent cells were thawed on ice and 10 µl of the ligation mixture (10 µl mixture per gene) was chilled in the pre-chilled microfuge tube. For each gene, the chilled 10 µl ligation mixture was added to a separate tube of thawed competent cells. Without mixing, the mixture was placed on ice for 20 minutes. The tubes were then placed in a 42 °C water bath for 1 minute to heat-shock the competent cells. Tubes were then placed on ice for 3 minutes to recover, after which, 1 mL of room temperature S.O.C medium was added to each tube to resuspend. The tubes were then placed in the rotary shaker at 37 °C for 2 hours at 200 rpm.

2.2.6 Plating and Selection of Transformed *Escherichia coli* JM109

To select for transformed cells, Luria Bertani (LB) tetracycline antibiotic selection plates were prepared as follows: 25 g of LB medium were combined with 1 L deionized water and 11 g agar and autoclaved for 30 minutes. Once cool, 1 mL of 30 mg/mL tetracycline was added and plates were poured. For each gene, two plates were poured, for a total of 16 plates. Selection plates were warmed at 37 °C for 30 minutes prior to plating. For the first plate, 50 µl of the ligation/competent cells mixture was spread on the plates. For the second plate, the ligation/competent cells mixture tubes were centrifuged at 11000 rpm for 2 minutes. The supernatant was then removed with a pipette and the pellet was mixed and plated on the second plate to prepare a concentrated plate. Plates were incubated overnight at 37 °C.

For each candidate gene, 5 mL LB medium and 5 µl 30 mg/mL tetracycline were combined and allocated into three microfuge tubes. Three separate colonies were selected from the diluted plate of each gene and added to the microfuge tube. Tubes were shaken overnight at 37 °C at 200 rpm. Plasmids were then extracted using the Plasmid DNA Miniprep Kit (BIOBASIC INC) according to manufacturer's protocols. Concentration of plasmid DNA was assessed using the Nanodrop Spectrophotometer and two glycerol stocks were prepared for each

of the 7 genes using the two most highly concentrated plasmid samples. To confirm presence of the *L. angustifolia* gene, the plasmids were prepared for Illumina sequencing with their gene-specific primers. In addition, a restriction digestion was performed with all constructs using *EcoR*I and *Kpn*I and verified on a 1% agarose gel.

2.2.7 Preparation of Competent *Agrobacterium tumefaciens*

Competent *A. tumefaciens* cells were prepared to allow for bacterial uptake of the pGA482::lavender gene constructs. *A. tumefaciens* cells stored at -80 °C were combined with 7 mL LB medium and 25 mg/L rifampicin. Tubes were covered in aluminum foil (as *A. tumefaciens* and rifampicin are light-sensitive) and shaken at 28 °C overnight. The following day, 50 mL room temperature LB medium and 25 µl of 50 mg/mL rifampicin were combined in an autoclaved flask with 500 µl of the *A. tumefaciens* mixture and the flask was covered in aluminum foil and incubated overnight at 28 °C at 140 rpm. The following day, the optical density (OD) was measured using the spectrophotometer to determine the phase of bacterial growth. The flask was shaken at 200 rpm at 28 °C until the OD had reached 0.6. A 20 mM Calcium Chloride ($CaCl_2$) solution was prepared with 800 µl of 1 M $CaCl_2$, 40 mL deionized autoclaved water and was filter-sterilized into a sterile falcon tube and placed on ice with an additional sterile falcon tube. Thirty 1.5 mL microfuge tubes were placed in the -80 °C freezer. An amount of 45 mL of cell culture was added to the pre-chilled falcon tube and centrifuged at 4 °C for 10 minutes. The supernatant was discarded, and the pellet resuspended with 5.0 mL of ice-cold $CaCl_2$ solution. The tube was centrifuged for 5 minutes at 4000 rpm at 4 °C, the supernatant discarded, and the pellet resuspended in 3 mL ice cold $CaCl_2$. The solution was then divided amongst 30 1.5 mL microfuge tubes in amounts of 100 µl. Microfuge tubes were individually flash-frozen in liquid nitrogen for 30 seconds and stored at -80 °C.

2.2.8 *Agrobacterium tumefaciens* Transformation and Selection

Competent *A. tumefaciens* GV3101 cells were transformed with the pGA482 vectors containing the lavender genes (pGA482::lavender gene) via the freeze-thaw method (Jyothishwaran et al. 2007). Briefly, 1 tube of competent *A. tumefaciens* cells was thawed on ice for each gene and 2 ng of vector (pGA482::lavender gene) was added to each respective tube and tubes were incubated on ice for 5 minutes. Tubes were placed in liquid nitrogen for 5 minutes

and subsequently placed in a 37 °C water bath for 5 minutes. An amount of 1 mL 37 °C LB medium was added to each tube and tubes were placed on the rotary shaker at 200 rpm at 28 °C for 2.5 hours. LB antibiotic selection plates were prepared with 50 mg/mL kanamycin and 25 mg/mL rifampicin, with two plates prepared per gene. On the first plate, 100 µl of the dilute cell solution were plated and on the second plate, 100 µl of pelleted cell solution were plated. Plates were then incubated for two days at 28 °C.

For each candidate gene, 3 mL LB medium and 1.5 µl 25 mg/L rifampicin and 3 µl 50 mg/L kanamycin were combined in a test tube with 2 replicates per gene. Two separate colonies were selected from the diluted plate of each gene and added to each test tube. Tubes were shaken overnight at 28 °C at 200 rpm. Plasmids were then extracted using the Plasmid DNA Miniprep Kit (BIOBASIC INC) according to manufacturer's protocols. Glycerol stocks were prepared for each of the genes. A separate PCR reaction was then performed for each plasmid using the respective lavender gene primers to amplify each gene from its vector, including a control PCR reaction with an empty vector. Each PCR reaction included 2.5 µl 10x buffer, 0.5 µl 10 mM DNTPs, 0.5 µl Taq Polymerase, 1-1.5 µl plasmid DNA (for plasmids with concentration above 100, 1.0 µl was used and for plasmid below 100 µl, 1 µl was used), 1.5 µl 10 mM lavender gene primer, 13.5 µl molecular grade water to a total volume of 20 µl. The previous Touchdown PCR program was used to amplify the floral genes form the vector.

2.2.9 Plant Transformation

To facilitate constitutive ectopic expression of the lavender genes in *A. thaliana* via the CaMV 35S promoter of the pGA482 vector, *A. tumefaciens* GV3101 containing the pGA482::*Lavandula* gene constructs were used. In an autoclaved Erlenmeyer flask, 200 mL room-temperature LB medium, 100 µl of 25 mg/L rifampicin, 200 µl of 50 mg/L kanamycin and 1 mL of *A. tumefaciens* overnight stock were combined. Flasks were covered in aluminum foil and placed on the rotary shaker at 28 °C at 200 rpm overnight until OD was 1.5-2.0.

Wild-type *A. thaliana* var. *Landsberg erecta* were grown as previously mentioned until sufficient vegetative mass had been achieved and buds were apparent, but flowering had not commenced (Figure 3a). *A. tumefaciens* containing the pGA482::lavender gene constructs were used to transform wild-type *A. thaliana* via the floral dip method (Zhang et al. 2006a). Briefly, 1500 mL of infiltration medium were prepared which contained 600 µl of 200 mM

acetosyringone, 225 µl 0.015% Tween 80 (v/v) and 5% sucrose. For each gene, 2 falcon tubes of 50 mL overnight *A. tumefaciens* stock were added and centrifuged at 22 °C at 6100 rpm for 10 minutes. Supernatant was poured into a separate 200 mL beaker for each gene. An amount of 50 mL of the LB *A. tumefaciens* overnight stock was used to resuspend the pellet in each tube and the tubes were centrifuged again at 22 °C at 6100 rpm for 10 minutes. The pellets were resuspended with 50 mL infiltration medium which was combined with an additional 150 mL in an autoclaved beaker. Fourteen wild-type *A. thaliana* plants (2 per gene) were covered in aluminum foil, leaving all vegetative growth exposed, but completely covering the soil (Figure 3b). The plants were then individually inverted into the agrobacterium infiltration solution and both were placed in a desiccator for 10 minutes (figure 3c). After 10 minutes, the plants were placed on their side on a foil-lined sheet and covered to prevent light exposure. This process was repeated for all plants. Plants were left on their side in the dark for 48 hours, after which, they were uncovered and placed upright. Transformation was repeated 1 week later following the same process. For each gene, two separate transformations were performed, transforming two plants per gene and performing the transformation twice, with 4 days between transformations. A total of four plants were transformed for each gene in two transformation events, with a total of 28 plants transformed and each plant transformed twice. Transformant plants were grown simultaneously with wild-type controls until seed set and seeds were collected as available (Figure 3d). Seeds were dried and stored at 4 °C.

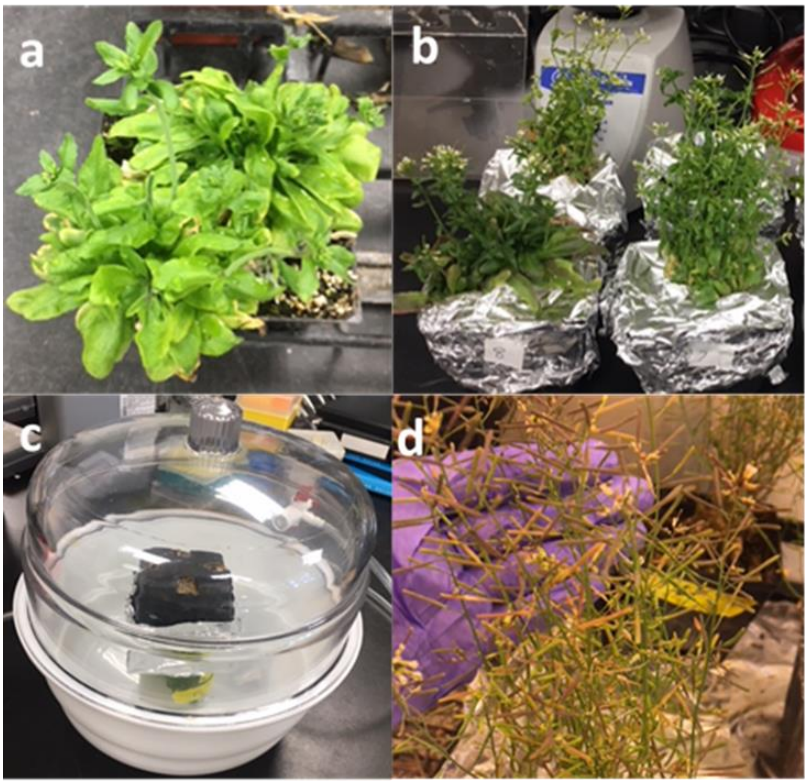

Figure 3. *Agrobacterium*-mediated transformation of wild-type *A. thaliana* via the floral dip. a) Wild-type plants were transformed at the stage of budding but prior to flowering. b) Plants were wrapped in foil to facilitate transformation. c) Vacuum infiltration in a desiccator facilitated the uptake of the *A. tumefaciens* containing the vector into the plant tissues. d) Seeds were collected from transformed plants at maturity when siliques had dried.

2.2.10 Selection of Transgenic Plants

To select for successful transformants, transformant and wild-type seeds were grown on kanamycin antibiotic selection medium. Seeds were plated for each of the 28 plants in conjunction with wild-type controls. Wild-type seeds were grown on medium containing no antibiotic to ensure medium viability. Seeds were sterilized prior to plating through a sterilization procedure adapted from Zhang et al. (2006b). Briefly, seeds were placed in in 96% ethanol for 1 minute followed by 20 minutes in 20% bleach aqueous solution containing 1% Triton X-100. In each solution, seeds were vortexed to ensure thorough sterilization of all seed surfaces. Seeds were subsequently washed once for 10 minutes in autoclaved deionized water (with vortexing) and twice for 5 minutes to ensure thorough washing and removal of detergent, bleach, and ethanol from the seeds. Deionized water was added to the tubes to facilitate plating. Selection medium was prepared with half-strength Murashige Skoog (MS) medium containing 0.5% gellan gum and 2% sucrose and at a pH of 5.7. Kanamycin was added to the medium at a concentration of either 50 mg/L or 75 mg/L Kanamycin. Wild-type seeds were plated on medium

containing 50 mg/L and 75 mg/L Kanamycin as well as medium lacking Kanamycin to serve as controls. All subsequent media were prepared at 75 mg/L kanamycin to facilitate most effective selection. Selection plates were sealed with parafilm and placed in a dark 5 ℃ fridge for four days for vernalization of the seeds to facilitate germination. Plates were then transferred to a growth chamber and grown at 25 ℃ with 16 hours light and 8 hours darkness. Seeds were grown until selection was visible on plates (approximately 3-4 weeks). Positive transformants were selected based on the ability to grow on plates containing 75 mg/L Kanamycin as well as the extent of growth and root biomass. Positive transformants were carefully removed from selection plates with forceps, removing any agar attached to the roots with water, and planted in moistened soil. Additionally, wild-type control seeds germinated on medium lacking Kanamycin were transferred for growth rate comparison and phenotypic analyses. The pots containing the seedlings were covered with plastic covers to maintain humidity for two-days and then gradually removed. Plants were then grown at 23 ℃ with 8 hours light, 16 hours darkness. Seedlings were grown to maturity and phenotype was monitored throughout growth.

2.2.11 Detection of Antibiotic Resistance and Floral Development Genes from cDNA

To confirm the presence of the Kanamycin resistance gene (*NPT-II)* and *LaSVP*, genomic DNA was extracted from leaf tissue using the Genomic DNA Mini Kit – Plant (Geneaid) according to the manufacturer's protocol. DNA concentration was then determined via the NanoDrop-1000 Spectrophotometer. The *NPT-II* gene was amplified from the genomic DNA using gene-specific primers. Primers used for PCR amplification were as follows: *NPT-II* forward 5'- AGAGGCTATTCGGCTATGACTG -3' and reverse 5'-TCAGAAGAACTCGTCAAGAAGGC-3'. For *NPT-II* gene amplification, the following program was used: DNA was denatured at 95°C for 5 minutes. Then for 40 cycles, 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute. Elongation occurred at 72°C for 7 minutes and the DNA was then kept at 4°C. The presence of the *LaSVP* gene in the transgenic plants was also confirmed through PCR amplification of the *LaSVP* gene from the corresponding genomic DNA. The *LaSVP* primer sequences were as follows: forward 5'-TCCGCAGAATTCATGGCTAGAGAGAAGATT-3' and reverse 5'-ATTGTAGGTACCTCAGCCAGTGTAAGGAAG - 3'(Wells et al. 2020).

To assess the relative expression of the transgene in transformed plants, RNA was extracted from flowers using the RNeasy Mini Plant Kit (Qiagen). The resulting RNA was quantified via the NanoDrop-1000 Spectrophotometer and synthesized into cDNA using the iScript cDNA Synthesis Kit (Biorad). The following thermocycler program was used for cDNA synthesis: 25°C for 5 minutes, 42°C for 60 minutes, 85°C for 5 minutes and then held at 4°C. To semi-quantitatively measure the expression level of *LaSVP* in the transformed plants, the cDNA was amplified using the previously indicated primer sequence for *LaSVP* and the *Actin* forward 5'-GCGACAATGGAACTGGAATGG-3' and reverse 5'-GGTGCCTCGGTAAGTAGAATAGG-3' primers. *Actin* was amplified using the same protocol as for *NPT-II* (indicated above), although the cycle number was reduced to 28. The Touchdown PCR program for *LaSVP* consisted of the following cycle: an initial denaturation at 95 °C for 5 minutes, for 8 cycles the cDNA was kept at 95°C for 30 seconds, 65 °C (-1) for 30 seconds, and 72 °C for 1 minute and 30 seconds. The next stage consisted of 20 cycles at 95 °C for 30 seconds, 57 °C for 30 seconds, and 72°C for 1 minute and 30 seconds and the sample was then held at 4 °C. Each cDNA sample was amplified in triplicates and the PCR products were subsequently run on a 1% agarose gel. Band intensities were measured by densitometry, using the ImageJ software (https://imagej.nih.gov/ij/download.html) (Wells et al. 2020).

2.2.12 Phenotypic Analyses

Transformed plants were watered and fertilized regularly with Miracle Gro fertilizer (Scott's Canada, Canada) prepared according to manufacturer's instructions for phenotypic comparisons between *A. thaliana* plants transformed with *LaSVP* and its counterpart wild- type. Plants were grown at 8 hours light/ 16 hours darkness, followed by 16 hours light/ 8 hours darkness at 23°C for vegetative growth and for flowering, respectively. Plants were assessed with regards to floral morphology, seed pod morphology, and flowering time.

Chapter 3: Results

3.1 Literature Review and Preliminary Homolog Identification

Through review of current literature describing the molecular mechanisms of flowering in *A. thaliana*, a subset of genes was identified which were screened for in lavenders. Preliminary BLASTx analyses were performed using a previously- assembled lavender transcriptome available in the Mahmoud lab against an annotated *A. thaliana* protein database. It was determined from these analyses that homologs of the previously- described *A. thaliana* floral initiation and developmental genes do exist in lavenders and provided preliminary evidence for the basis of this thesis.

3.2 RNA-Seq and Transcriptome Assembly and Analysis

To determine if homologs of the *A. thaliana* floral initiation and development genes exist in lavenders, a comprehensive lavender transcriptome was assembled. Pair-end sequence data were obtained for transcripts extracted from *L. angustifolia* flower (FL), floral bud (BD) and shoot apical meristem (SAM) tissues, and from *L. x intermedia* SAM tissue using Illumina Sequencing. A separate transcriptome for each tissue was assembled using CLC Genomics Workbench 11 (Qiagen) (Table 3). The individual *L. angustifolia* FL, BD, and SAM transcriptomes contained 47,630, 46,786, and 39,964 contigs, respectively, and that of the *L. x intermedia* SAM contained 26,821 contigs (Table 3). The N_{50} values of the transcriptomes ranged from 500 bp from the *L. x intermedia* SAM to 628 bp from the *L. angustifolia* FL transcriptome (Table 3). The N_{50} represents 50% of the largest contigs within the transcriptome (CLC bio b). Most of the contigs (~ 50%) were between 200-399 bp in length, with few contigs above 2500 bp (Table 3 & Figure 4). The minimum contig length in each transcriptome was 200 bp as this was set as the minimum during assembly, while the maximum contig length varied from 6,693 bp in the *L. x intermedia* SAM to 12,567 bp in the *L. angustifolia* FL (Table 3). Transcriptome results were published in Wells et al. 2020.

Table 3. Summary of the parameters of the *L. angustifolia* and *L. x intermedia* transcriptome assemblies. Note that minimum contig length was set to 200 bp during assembly of all transcriptomes. FL= flower, SAM = vegetative meristem, BD = floral bud tissue. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

Parameters	***L. angustifolia* FL**	***L. angustifolia* BD**	***L. angustifolia* SAM**	***L. x intermedia* SAM**
Total paired reads	36,630,196	35,690,350	28,380,848	17,543,766
Number of contigs	47, 630	46,786	39, 964	26,821
GC content (%)	45.1	45.3	45.3	45.8
Min length (bp)	200	200	200	200
Max length (bp)	12567	9937	10146	6693
Average contig length (bp)	497	493	476	410
N_{50} (bp)	628	618	617	500
Total assembled bases	23,650,730	23, 051, 014	19,008,841	11,005,731

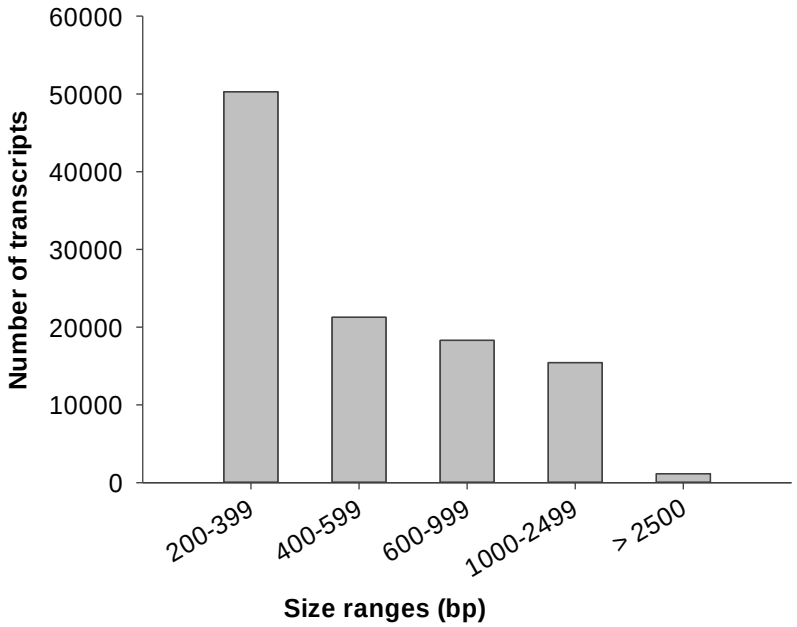

Figure 4. Distribution of transcript lengths in the comprehensive *Lavandula* transcriptome. Minimum and maximum contig lengths were 200 bp and 12,567 bp, respectively. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

The four tissue-specific assemblies were then combined into one comprehensive transcriptome and redundant contigs were removed using CD-HIT-EST, resulting in a transcriptome with a total of 105,294 contigs (Figure 4). Most of the transcripts generated had BLASTx hits against the *Arabidopsis* protein database, and the transcriptome included numerous full-length sequences, including those corresponding to ten previously reported *Lavandula* terpene synthases (Table 4) (Landmann et al. 2007; Demissie et al. 2011, 2012; Sarker et al.

2013; Jullien et al. 2014; Despinasse et al. 2017; Adal et al. 2019), which are enzymes vital to lavender EO synthesis (Aprotosoaie et al. 2017).

Table 4. Full-length, previously- reported *Lavandula* terpene synthase genes identified within the *Lavandula* transcriptome. A total of 10 full-length, previously- reported *Lavandula* terpene synthase genes were identified within the comprehensive transcriptome.

Accession Number	Gene/protein	Source Species
ADQ73631.1	beta-phellandrene synthase	*Lavandula angustifolia*
ABB73045.1	(R)-linalool synthase	*Lavandula angustifolia*
ABB73044.1	limonene synthase	*Lavandula angustifolia*
ABB73046.1	trans-alpha-bergamotene synthase	*Lavandula angustifolia*
AGL98420.1	germacrene-D synthase	*Lavandula angustifolia*
AGL98418.1	cadinol synthase	*Lavandula angustifolia*
AJW68082.1	bornyl diphosphate synthase	*Lavandula angustifolia*
AFL03421.1	1,8-cineole synthase	*Lavandula x intermedia*
AGU13712.1	9-epi-caryophyllene synthase	*Lavandula x intermedia*
AGL98419.1	B-caryophyllene synthase	*Lavandula angustifolia*

Gene ontology analyses were performed using the BLAST2GO function of CLC Genomics Workbench 11. Through this analysis, annotated transcripts were classified into biological process, molecular function, and cellular component categories of GO terms (Figure 5). Of the genes which mapped to cellular components, the top GO identities were cell (nearly 50,000 sequences), organelle (> 42,500 sequences), membrane (> 30,000 sequences), and other components (Figure 5). Of the genes that were mapped to biological processes, the top GO identities were cellular process (> 42,500 sequences), metabolic process (> 37,500 sequences), response to stimulus (> 32,500 sequences), biological regulation (> 27,500 sequences), and developmental process (> 25,000 sequences) (Figure 5). In the subset of genes which mapped to molecular function, the top GO identities were binding (> 37,500 sequences), catalytic activity (> 30,000 sequences), transporter activity (nearly 5,000 sequences), and transcription regulator activity (> 3,000 sequences) (Wells et al. 2020).

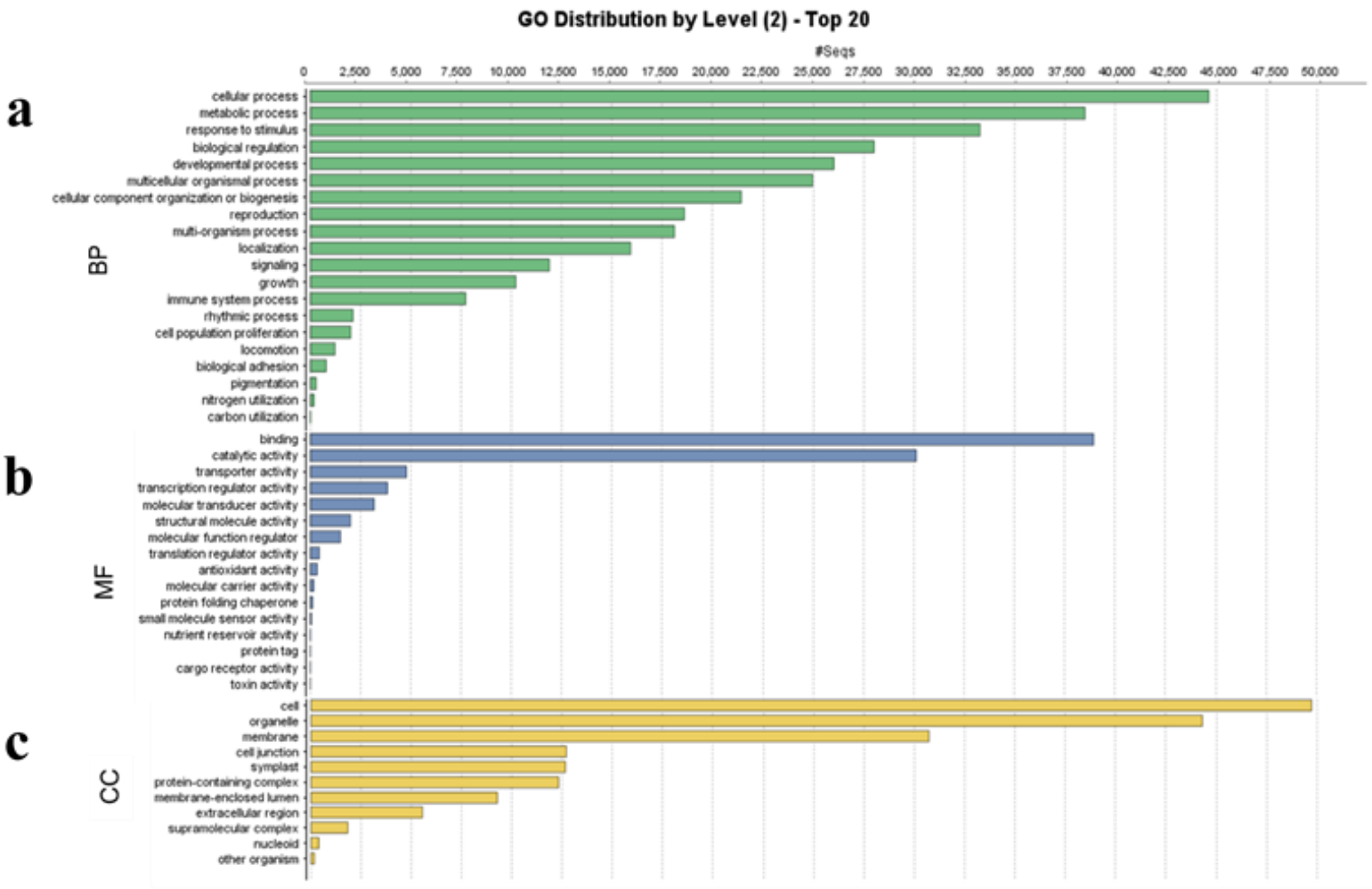

Figure 5. Classification of genes in the comprehensive *Lavandula* transcriptome. Contigs were classified based on (a) biological process, (b) molecular function, and (c), cellular component. Gene ontology studies were performed in CLC Genomics Workbench 11 to examine gene function and characterize transcripts. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

The transcriptome contained over 100 *Lavandula* contigs which exhibited high sequence homology to *A. thaliana* floral initiation and pattern development-related genes (Wils and Kaufmann 2017) (Table S1). These genes were previously- described proteins and included both non-repressor and repressor proteins. Particularly, the transcriptome contained representative sequences corresponding to all of the A, B, C, D, and E-type MADS-Box transcription factors described in *A. thaliana* (Causier et al. 2010; Theißen et al. 2016). Over 100 of these floral initiation and development-related genes previously- described in *A. thaliana* were identified within the assembled transcriptome (Table 5). In addition full-length sequences for some of the key *A. thaliana* floral development genes were identified in the *Lavandula* transcriptome

including *CAULIFLOWER, SHORT VEGETATIVE PHASE, PISTILLATA, FLOWERING LOCUS T, SHOOT MERISTEMLESS, APETALA1, APETALA2, APETALA3, SEPALLATA 3, LATE ELONGATED HYPOCOTYL, SUPPRESSOR OF OVEREXPRESSION OF CO1, FLOWERING LOCUS C, AGAMOUS,* and *EMBRYONIC FLOWER 2* (Table 6) (Wells et al. 2020).

Table 5. *Arabidopsis thaliana* floral development gene homologs identified within the comprehensive *Lavandula* transcriptome. A total of 104 *Lavandula* genes homologous to *A. thaliana* floral development genes were identified within the transcriptome, including full length and partial sequences for repressor and non-repressor proteins. LaBD = *L. angustifolia* bud, LaFL = *L. angustifolia* flower, LaSAM = *L. angustifolia* vegetative shoot apex, and LiSAM = *L. x intermedia* vegetative shoot apex.

***A.thaliana* Gene**	**e- value**	**LaSAM vs LiSAM**	**LaBD vs LiSAM**	**LaFL vs LaBD**	**LaSAM vs LaFL**	**Function**	**Reference**
AGL-like	1.34×10^{-58}	6.704	627.626	2.149	-174.384	Carpel development	(Theißen et al. 2016)
AGL1	4.16×10^{-40}	0.000	689.089	1.068	-963.661	Floral morphogenesis	(Ma et al. 1991)
AGL11	2.38×10^{-71}	0.000	197.384	6.858	-1770.699	Fertilization and seed development	(Mizzotti et al. 2012)
AGL12	2.88	0.000	8.896	6.404	-73.136	Floral transition and root proliferation	(Tapia-López et al. 2008)
AGL17	3.14×10^{-34}	1.445	1.923	-1.921	1.455	Flowering promotion in response to increased photoperiod	(Han et al. 2008)
AGL21	3.24×10^{-39}	1.342	2.203	1.992	-3.241	Lateral root development, seed germination	(Yu et al. 2014, 2017)
AGL24	3.19×10^{-11}	-2.245	-290.754	0.000	153.061	Promoter of flowering, component of vernalization response	(Michaels et al. 2003; Gregis 2006)
AGL6	2.16×10^{-74}	0.000	779.235	-1.090	-935.832	Promoter of flowering, regulator of FLC and FT genes	(Ma et al. 1991; Yoo et al. 2011)
AGL8	1.19×10^{-49}	163.277	537.480	-7.838	2.854	Expressed in inflorescence meristem and in developing carpels	(Mandel and Yanofsky 1995)
ANT	2.11×10^{-130}	1.192	-1.262	-59.055	88.909	Meristem establishment, organ development and growth	(Horstman et al. 2014)
AP1	8.72×10^{-38}	0.000	1492.208	-2.113	-924.701	Sepal development, petal development	(Theißen et al. 2016)

						when combined with B-class genes. Meristem identity gene	
AP2	9.73×10^{-117}	-2.028	1.171	-1.142	-2.064	Sepal development. Petal development when combined with B-class genes	(Theißen et al. 2016)
AP3	7.96×10^{-86}	1.089	51.591	1.034	-48.610	Stamen development when expressed with AG	(Theißen et al. 2016)
ARF2	0.000	-1.410	-1.440	-1.140	1.170	Cell division, seed and organ size	(Schruff et al. 2006)
ARF4	5.46×10^{-163}	2.220	1.650	-5.410	7.360	Male, female gametophyte development	(Liu et al. 2018c)
ARF5	3.55×10^{-138}	1.980	1.840	-3.380	3.680	Male, female gametophyte development	(Liu et al. 2018c)
ARF6	0.000	1.030	-1.180	-1.050	1.290	Regulation of floral organ development	(Tabata et al. 2010)
ARF8	0.000	4.550	9.850	-2.220	1.040	Regulation of floral organ development	(Tabata et al. 2010)
BEL1	1.93×10^{-106}	-4.990	-3.470	1.570	-2.240	Ovule development	(Brambilla et al. 2007)
BHLH10	3.87×10^{-28}	6.700	1037.380	-34.250	-3.930	Flowering promotion via photoperiodic pathway	(Ito et al. 2012)
BLH1	1.93×10^{-106}	-4.990	-3.470	1.570	-2.240	Early seedling development	(Kim et al. 2013)
BLH2	2.83×10^{-124}	1.660	1.760	-1.620	1.530	Leaf shape establishment	(Kumar et al. 2007)
BLH4	7.73×10^{-104}	-1.720	-5.600	1.340	2.450	Leaf shape establishment	(Kumar et al. 2007)
BPC1	2.40×10^{-1}	-18.810	-16.380	3.050	-3.510	Aid in tissue-specific gene expression during development	(Simonini et al. 2012)
BPC2	1.34×10^{-93}	-1.080	-1.380	-1.390	1.780	Aid in tissue-specific gene expression during development	(Simonini et al. 2012)
BPC3	1.67×10^{-34}	-1.330	-1.400	1.120	-1.050	Aid in tissue-specific gene expression during development	(Simonini et al. 2012)
BPC4	1.75×10^{-96}	-8.210	-8.480	1.350	-1.300	Aid in tissue-specific gene expression during development	(Simonini et al. 2012)

BPC6	5.62×10^{-144}	-1.280	-1.240	1.270	-1.300	Aid in tissue-specific gene expression during development	(Simonini et al. 2012)
BPE	3.01×10^{-76}	2.470	4.310	1.420	-2.470	Petal growth and development	(Brioudes et al. 2009)
BZIP	0.000	1.410	1.990	-2.000	1.430	Flower development, defense, light and stress signalling	(Dröge-Laser et al. 2018)
CAL	4.35×10^{-45}	10.540	187.240	-1.080	-16.190	Interaction with AP1, affects floral meristem identity	(Bowman et al. 1993)
COL9	2.70×10^{-103}	1.360	-1.220	1.000	1.680	Flowering regulation in response to photoperiod	(Cheng and Wang 2005)
CC	1.43×10^{-58}	12.500	500.600	6.240	-212.330	Carpel development	(Bowman and Smyth 1999)
CUC2	3.35×10^{-52}	12.503	123.628	-2.056	-4.090	Boundary formation, organ development	(Nikovics et al. 2006)
CUC3	4.6	3.420	2.820	-1.440	1.770	Boundary formation, organ development	(Nikovics et al. 2006)
ELF3	5.24×10^{-42}	-1.200	-1.290	1.140	-1.050	Flowering initiation in response to photoperiod	(Zagotta et al. 1996b)
ELF7	9.18×10^{-131}	1.170	-1.070	1.120	1.120	Regulation of *FLC* expression via chromatin modification	(He et al. 2004)
EMF2	0.000	-1.060	1.050	-1.380	1.260	Maintains vegetative development, flowering repressor	(Moon et al. 2003b)
FCA	1.09×10^{-81}	-2.270	-1.930	2.840	-3.320	Flowering time control, repressor	(Bäurle et al. 2007)
FLC	2.08×10^{-26}	-4.020	-1.550	-1.570	-1.650	Flowering time gene responsive to vernalization, repressor	(Hepworth et al. 2002) (Hepworth and Dean 2015)
FLD	1.46×10^{-81}	-1.740	1.870	1.290	-1.190	Flowering control in autonomous pathway	(He et al. 2003)
FPA	4.52×10^{-68}	-1.090	-1.210	1.010	1.110	Floral induction, regulation of floral initiation in autonomous pathway	(Schomburg et al. 2001)
FT	7.55×10^{-102}	401.040	2791.130	-2.710	-2.140	Flowering promotion	(Teper-Bamnolker and Samach 2005)

FY	0.000	-1.394	-1.430	-1.258	1.301	Flowering transition regulator	(Simpson et al. 2003)
GI	6.48×10^{-162}	-4.500	-7.580	1.440	1.180	Regulation of flowering in response to photoperiod, inducer of *FT* and *CO*	(Jung et al. 2007)
GNL	0.000	18.300	-1.040	1.330	-1.280	Flowering time, senescence, repressed during flowering	(Richter et al. 2013)
HDG1	2.22×10^{-103}	1.280	1.360	-1.850	1.750	Stamen development, petal and stamen identity	(Kamata et al. 2013)
HDG12	0.000	2.350	1.720	-3.850	5.320	Stamen development, petal and stamen identity	(Kamata et al. 2013)
HDG2	1.65×10^{-141}	3.140	6.520	-3.650	1.770	Stamen development, petal and stamen identity	(Kamata et al. 2013)
HDG5	0.000	1.240	1.400	-3.610	3.220	Stamen development, petal and stamen identity	(Kamata et al. 2013)
HEC1	2.06×10^{-16}	12.503	90.847	-18.977	3.032	Female organ development, gynoecium development	(Gremski et al. 2007)
HEC2	4.90×10^{-01}	-1.100	-1.170	1.800	-1.690	Female organ development, gynoecium development	(Gremski et al. 2007)
HEN	0.000	-1.340	-2.230	1.140	1.480	Floral organ pattern development	(Chen et al. 2002)
JAG	3.56×10^{-35}	18.300	82.650	-1.370	-2.770	Organ growth and patterning	(Schiessl et al. 2014)
KAN1	4.72×10^{-59}	2.800	3.760	-1.210	-2.320	Regulates lateral organ polarity	(Kerstetter et al. 2001)
KAN2	2.36×10^{-63}	-1.580	1.200	-1.130	-1.670	Regulates lateral organ polarity	(Kerstetter et al. 2001; Wu et al. 2008)
KAN4	7.07×10^{-49}	-1.890	-4.610	-1.200	2.960	Integument development	(McAbee et al. 2006)
KNAT2	4.30×10^{-71}	2.700	6.200	1.030	-2.350	Carpel development	(Pautot et al. 2001)
LDL1	0.000	-1.670	-1.120	1.090	-1.610	Repression of seed dormancy	(Zhao et al. 2015)
LFY	1.05×10^{-29}	0.000	62.160	-117.830	0.000	Flowering promotion, interaction with	(Parcy et al. 1999)

						homeotic floral organ genes	
LD	2.75×10^{-68}	1.120	1.020	1.040	1.060	Promotion of floral transition	(Aukerman et al. 1999)
MAF5	1.00×10^{-2}	-3.112	-2.421	1.261	-1.610	Flowering time protein	(Ratcliffe et al. 2003)
MYB21	1.13×10^{-25}	0.000	275.240	28.660	-10319.740	Stamen development	(Qi et al. 2015)
MYB24	1.58×10^{-50}	1.270	5.270	12.490	-51.510	Stamen development	(Qi et al. 2015)
MYC2	7.59×10^{-115}	2.090	-1.430	-1.370	2.090	Stamen development, seed production	(Qi et al. 2015)
MYC3	2.64×10^{-51}	4.080	2.940	-1.080	1.520	Stamen development, seed production	(Qi et al. 2015)
MYC4	1.03×10^{-103}	-1.240	-1.230	-1.070	1.060	Stamen development, seed production	(Qi et al. 2015)
NGA1	3.00×10^{-66}	2.690	5.230	-1.750	-1.100	Style development	(Trigueros et al. 2009)
PAN	5.44×10^{-87}	3.990	1.010	-9.990	39.890	Affects floral architecture and expression of homeotic genes, regulator of AG	(Maier et al. 2009)
PDF2	0.000	2.120	2.400	-2.260	2.010	Embryo development	(Ogawa et al. 2014)
PHYB	0.000	-1.190	-1.080	1.320	-1.450	Regulation of flowering time	(Hajdu et al. 2015)
PI	1.64×10^{-48}	1.572	223.785	-1.586	-89.107	Stamen development when expressed with AGAMOUS	(Theißen et al. 2016)
PIE1	0.000	-1.390	-1.430	-1.260	1.300	FLC activation, floral repression	(Noh and Amasino 2003)
PTL	2.63×10^{-53}	12.500	131.820	-1.400	-6.400	Petal development, boundary formation between sepal and petal	(Lampugnani et al. 2013)
RAV1	4.13×10^{-56}	1.011	-1.800	-1.154	2.119	Seed germination, seedling development	(Feng et al. 2014)
REV	0.000	-1.110	-1.070	2.190	-2.260	Meristem development, cell division	(Talbert et al. 1995)
SEP1	1.77×10^{-70}	-2.249	4.910	1.374	-15.164	Expressed in all whorls, needed for expression in all whorls	(Theißen 2001)
SEP2	3.65×10^{-67}	47.297	3606.543	1.788	-114.145	Expressed in all whorls, needed for	(Theißen 2001)

						expression in all whorls	
SEP3	1.09×10^{-88}	24.101	4323.614	1.014	-153.090	Expressed in all whorls, needed for expression in all whorls	(Theißen 2001)
SHN2	2.31×10^{-33}	1.940	1.510	-3.270	4.220	Expression in anthers and siliques	(Shi et al. 2011)
SOC1	4.07×10^{-77}	-1.120	-8.920	1.040	7.730	Floral transition gene, flowering promotion	(Liu et al. 2008)
SPL3	7.64×10^{-34}	2.680	2.960	-2.130	1.950	Transition from vegetative phase to flowering	(Wu and Poethig 2006)
SPL8	2.59×10^{-78}	3.230	6.260	-1.470	-1.300	Pollen sac development	(Unte et al. 2003)
SPT	2.89×10^{-42}	35.699	156.408	-1.017	-3.609	Carpel and fruit development	(Groszmann et al. 2008)
STM	6.80×10^{-133}	-1.720	15.930	1.380	-3.430	Leaf shape establishment	(Aida et al. 1999)
SUP	2.69	-1.810	-1.240	2.060	-2.990	Carpel/stamen boundary conservation	(Sakai et al. 2000)
SVP	1.17×10^{-80}	-1.390	-5.630	-4.580	18.710	Floral transition repressor, floral meristem identity specification	(Gregis et al. 2013)
TCP10	1.70×10^{-44}	-6.210	-5.050	1.560	-1.900	Lateral organ development	(Koyama et al. 2007)
TCP13	3.69×10^{-44}	1.750	1.040	1.180	1.440	Lateral organ development	(Koyama et al. 2007)
TCP15	1.32×10^{-49}	2.460	3.310	-2.930	2.190	Lateral organ development	(Koyama et al. 2007)
TCP3	2.05×10^{-12}	-1.170	-1.210	-3.350	3.500	Lateral organ development	(Koyama et al. 2007)
TCP4	8.67×10^{-58}	-1.490	-2.370	-3.250	5.220	Lateral organ development	(Koyama et al. 2007)
TCP5	2.89×10^{-42}	1.630	1.310	-1.350	1.700	Lateral organ development	(Koyama et al. 2007)
TGA10	7.95	-1.420	-1.030	-1.120	-1.220	Anther development	(Murmu et al. 2010)
TOE3	3.67×10^{-38}	1.260	1.697	2.775	-3.702	Repressor of *AG* and AP2, repression of TOE3 required for pattern development	(Jung et al. 2014)
UFO	6.21×10^{-61}	58.900	258.850	-28.650	7.740	Regulation of floral meristem and floral organ identity	(Samach et al. 1999)

ULT	0.26	1.340	3.760	1.080	-3.000	Shoot and floral meristem size and development	(Fletcher 2001)
VIN3	3.76×10^{-23}	-3.130	-1.660	-1.250	-1.510	Epigenetic modifications during vernalization response	(Kim and Sung 2013)
VRN1	7.20×10^{-19}	-1.090	-1.070	-1.250	1.230	Vernalization response, repressor of FLC	(Levy et al. 2002)
WOX1	1.31×10^{-24}	2.060	6.480	-2.670	-1.170	Petal, carpel, leaf development	(Costanzo et al. 2014)
WOX13	3.63×10^{-86}	-1.060	-1.240	1.880	-1.600	Replum development	(Romera-Branchat et al. 2013)
WOX3	6.27×10^{-7}	-3.030	1.160	-10.180	2.900	Sepal and petal development	(Costanzo et al. 2014)
WOX8	2.40×10^{-1}	-1.100	1.590	-2.380	1.360	Embryo patterning, maintenance of the SAM	(Costanzo et al. 2014)
WOX9	3.89×10^{-20}	0.000	66.260	-125.620	0.000	Embryo development, maintenance of vegetative SAM	(Costanzo et al. 2014)

Note: AGL-Like = *Agamous-like*, AGL(1,6,8,11,12,17,21,24) = *Agamous-Like MADS Box Protein*, ANT = *Aintegumenta*, AP(1-3) = *Apetala*, ARF(2,4,5,6,8) =*Auxin Response Factor*, BEL1 = *Bell 1*, BHLH010 = *Basic Helix Loop Helix Protein 10*, BLH(1,2,4) = *Bel1-Like Homeodomain*, BPC(1-4, 6) = *Basic Pentacysteine*, BPE = *Big Petal P*, BZIP = *BZIP2* , CAL = *Cauliflower*, COL9 = *Constans- Like 9*, CRC = *Crab's Claw*, CUC(2,3)= *Cup-Shaped Cotyledon 2*, ELF(3,7) = *Early Flowering*, EMF2 = *Embryonic Flower 2*, FCA = *Flowering Control Locus A* , FLC = *Flowering Locus C*, FLD = *Flowering Locus* D, FPA = *FPA*, FT = *Flowering Locus T*, FY = *FY*, GI = *Gigantea* , GNL = *GNC – Like*, HDG(1,2,5,12) = *Homeodomain Glabrous*, HEC(1,2) = *Hecate*, HEN = *HUA Enhancer 2*, JAG = *Jagged*, KAN(1,2,4) = *Kanada*, KNAT2 = *Knotted – Like 2*, LD = *Luminidependens*, LDL1 = *Lysine-Specific Histone Demethylase*, LFY = *Leafy*, MAF5 = *MADS Affecting Flowering 7*, MYB(21,24) = *MYB*, MYC = *MYC*, NGA1 = *Ngatha 1*, PAN = *Periantha*, PDF2 = *Protodermal Factor 2*, PHYB = *Phytochrome B*, PI =*Pistillata*, PIE1 = *Photoperiod- Independent Early Flowering*, PTL = *Petal Loss*, RAV1 = *Related to ABI3/VP1*, REV = *Revoluta*, SEP(1-3) = *Sepallata*, SHN2 = *Shine 2*, SOC1 = *Suppressor of CONSTANS1*, SPL(3,8) = *Squamosa Promoter Binding Protein-Like* , SPT = *Spatula*, STM = *Shoot Meristemless*, SUP = *Superman*, SVP = *Short Vegetative Phase*, TCP = *TCP*, TGA10 = *TGA10*, TCP(3,4,5,10,13,15) = named from *Teosinte Branched1* (TB1), *Cycloidea* (CYC) *Proliferating Cell Nuclear Antigen Factor* (PCF1) , TOE3 = *Target of Early Activation Tagged 3*, UFO = *Unusual Flower Organs*, ULT = *Ultrapetala*, VIN3 = *Vrenalization Insensitive* 3, VRN1 *Vernalization 1* , WOX(1,3,8,9,13) = *Wuschel-Related Homeobox*

Using the Differential Expression for RNA-Seq function in CLC Genomics Workbench 11, a Differential Expression analysis was performed to examine expression patterns of the over 100 identified genes. Genes of particular interest were those that showed differential expression patterns between *L. ang* SAM and *L. ang* BD as well as *L. ang* SAM and *L. x int* SAM. Differential expression patterns were recorded for all identified *A. thaliana* floral development gene homologs and differential expression analysis revealed that most of the floral initiation and identity genes were differentially expressed across tissues (Figure 6). For example, the *LaSVP* gene was more strongly expressed in both *L. angustifolia* and *L. x intermedia* SAM tissue compared to buds and later floral stages. Expression pattern was consistent with gene function.

Based on the differential expression analysis, literature review of floral development gene functions, and identification of sufficiently low e- values ($\leq 10^{-10}$) 14 full-length lavender candidate genes homologous to *A. thaliana* floral initiation and pattern development genes were successfully identified within the transcriptome and selected for functional characterization (Table 6). An e-value of $\leq 10^{-10}$ was selected as a cut off for sufficient sequence similarity. The full-length ORF sequence of the protein was determined for each of the candidates and protein length was similar between each *Lavandula* candidate and its respective homolog (Table 6). Candidates were further examined and a subset of 7 candidates were selected for functional characterization.

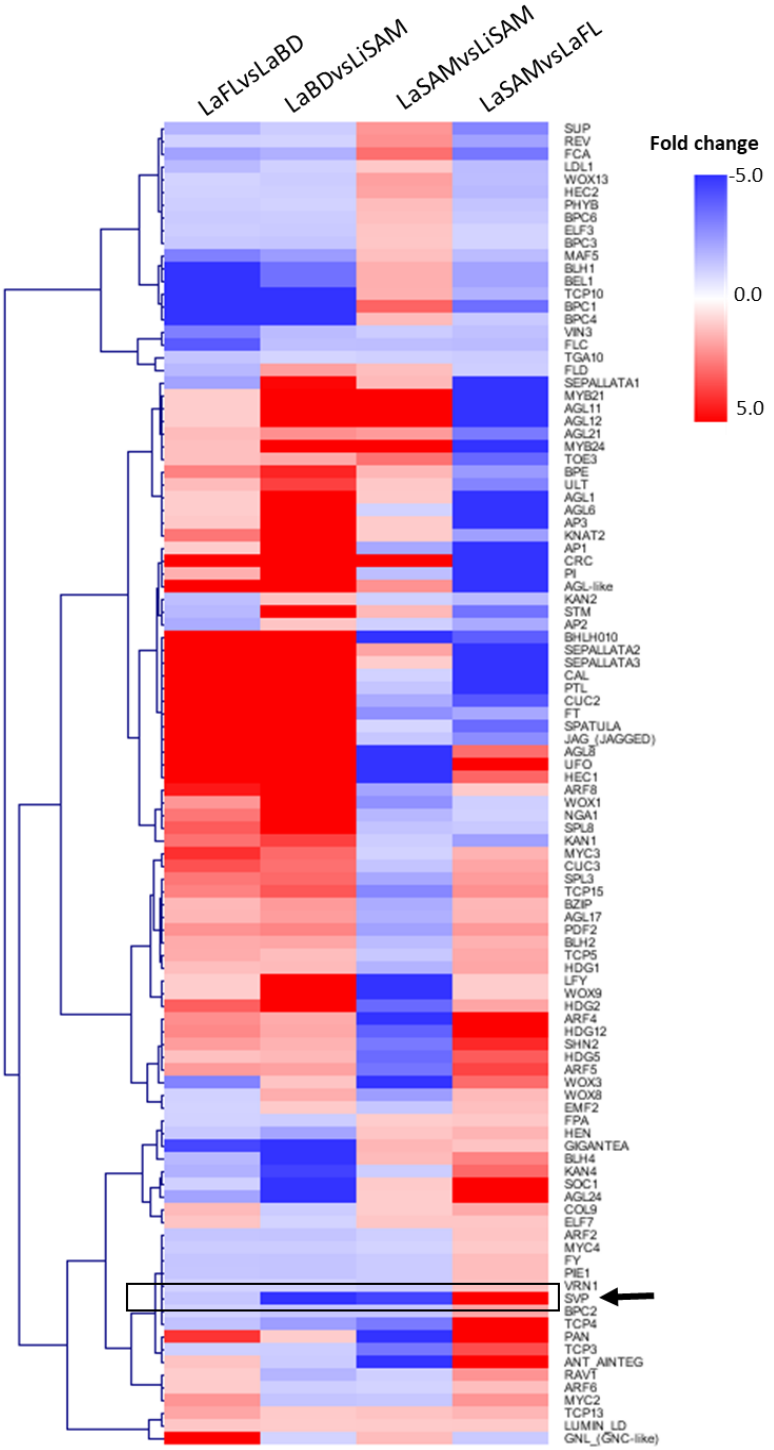

Figure 6. Heatmap showing the relative expression levels of selected floral development genes. Relative expression of the genes (rows) was assessed in various tissues of *L. angustifolia* and *L. x intermedia* plants. Expression values used for clustering are fold changes. LaBD = *L. angustifolia* bud, LaFL = *L. angustifolia* flower, LaSAM = *L. angustifolia* vegetative shoot apex, and LiSAM = *L. x intermedia* vegetative shoot apex. An arrow and a box show the expression patterns of the *L. angustifolia SVP* gene. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

Table 6. Full-length *Lavandula* floral initiation and development genes identified within the comprehensive *Lavandula* transcriptome. Genes were identified based on full-length sequence, high sequence homology to the respective *A. thaliana* protein homolog and gene function.

Gene	*Lavandula* protein length (aa)	*A. thaliana* protein length (aa)	e-value	*A. thaliana* Function	Reference
Shoot Meristemless (*LaSTM*)	391	382	6.80E-133	SAM cell proliferation	(Byrne et al. 2002)
Crab's Claw (*LaCC*)	163	181	1.43E-58	Carpel development, nectary formation, termination of floral meristem	(Gross et al. 2018)
Cauliflower (*LaCAL*)	253	255	4.35E-45	Interaction with *AP1* and *LFY*, floral meristem identity	(Bowman et al. 1993)
Short Vegetative Phase (*LaSVP)*	230	240	1.17E-80	Floral transition repressor, meristem identity gene	(Gregis 2006)
Pistillata (*LaPI*)	187	208	1.64E-48	Stamen development, petal development	(Theißen et al. 2016)
Flowering Locus T (*LaFT*)	174	175	7.55E-102	Floral induction in response to photoperiod	(Andrés et al. 2015)
Apetala1 (*LaAP1*)	253	256	8.72×10^{-38}	Sepal development Petal development when combined with B-class genes.	(Theißen et al. 2016)
Apetala2 (*LiAP2*)	433	432	9.73×10^{-117}	Sepal development. Petal development when combined with B-class genes	(Theißen et al. 2016)
Apetala3 (*LaAP3*)	232	234	7.96×10^{-86}	Stamen development when expressed with AGAMOUS	(Theißen et al. 2016)
Suppressor of CONSTANS1 (*LaSOC1*)	221	214	4.07×10^{-77}	Integration of gibberellin and vernalization pathway during floral transition	(Moon et al. 2003a)
Sepallata 3 (*LaSEP3*)	241	251	1.09×10^{-88}	Expressed in all whorls; required for development of organs in all whorls	(Theißen et al. 2016)
Agamous (*LaAG*)	247	252	1.34×10^{-58}	Carpel development	(Theißen et al. 2016)
Embryonic Flower 2 (*LaEMF2*)	627	631	0.00	Repressor of flowering	(Henderson and Dean 2004; Theißen et al. 2016)
Flowering Locus C (*LaFLC*)	190	196	2.08×10^{-26}	Repressor of flowering	(Henderson and Dean 2004)

3.3 Sequence Analysis and Phylogeny

To examine the percent similarity of the protein sequences of the *Lavandula* candidates and their respective *A. thaliana* homologs, MUSCLE sequence alignment with the homologous amino acid sequences was performed for each gene (Table 7, Table 8, & Figures 8, B1-B13). MUSCLE protein sequence alignment of the 14 *Lavandula* candidate proteins revealed that all candidates had >50% sequence similarity to their *A. thaliana* homolog , with LaFT being the most similar to its homolog with 78.16 % similarity, and LaFLC being the least similar to its homolog at 41.53% sequence similarity (Table 7). Further, MUSCLE alignment revealed the presence of conserved domains characteristic of the class of TFs each lavender gene belonged to. LaCC, the *Lavandula* homolog of the zinc finger domain TF AtCC, contained the conserved zinc finger domain and the helix loop helix domain (Figure B.1). LaSM, the homolog of AtSM, a TALE/KNOX homeobox TF, exhibited the ELK domain and the homeobox domain present in AtSM (Figure B.2). LaEMF2, the homolog of AtEMF2, exhibited the VEFS-Box and the Zinc finger domains found in AtEMF2 (Figure B.3). LiAP2, the homolog of AtAP2, exhibited the AP2/ERF2 domains found in AtAP2. In addition, LaAG, LaSVP, LaCAL, LaFLC, LaPI, LaAP1, LiAP3, LaSEP3, LaSOC1, homolog of the *A. thaliana* MADS-Box TFs, AAtAG, AtSVP, AtCAL, AtFLC, AtPI, AtAP1, AtAP3, AtSEP3, AtSOC1, respectively, exhibited the MADS-Box and K-Box conserved domains (Figure B.6-B.13) (Wells et al. 2020).

Table 7. The percent similarity of candidate *Lavandula* proteins and their *A. thaliana* homolog based on MUSCLE alignment of protein sequences. Matrix was created by Clustal 2.1.

***Lavandula* Protein**	**% Identity to *A. thlaliana* Homolog**
LaAG	68.16
LaAP1	70.56
LiAP2	50.12
LaAP3	56.90
LaCAL	64.26
LaCC	66.88
LaEMF2	62.97
LaFLC	41.53
LaFT	78.16
LaPI	55.75
LaSEP3	74.17
LaSM	72.35
LaSOC1	68.22
LaSVP	67.39

To examine the relationship between *LaSVP* and other plant *SVP*-like genes, a phylogenetic tree was constructed using the amino acid sequences of these proteins (Figure 7). Although the amino acid sequence of *LaSVP* exhibited higher sequence similarity to *EgSVP-like*, *BrSVP*, *BjSVP*, and *PkSVP* homologs compared to *AtSVP*, all of these genes contained the K-Box and MADS-Box conserved domains (Figure 7). The LaSVP protein was thus found to

cluster most closely with the SVP homologs PkSVP (*Paulowinia kawakamii*) (76.21% sequence identity), EgSVP-like (*Eucalyptus grandis*) (69.78% sequence identity), BrSVP (*Brassica rapa*) (68.26% sequence identity), AtSVP (*Arabidopsis thaliana*) (67.39% sequence identity), CmSVP (*Chrysanthemum x mortifolium*) (59.46% sequence identity) and BjSVP (*Brassica juncea*) (67.83 % sequence identity) (Figure 8 & Table 8). Thus, LaSVP clustered most closely with *PkSVP*, and most distantly with CmSVP (Wells et al. 2020).

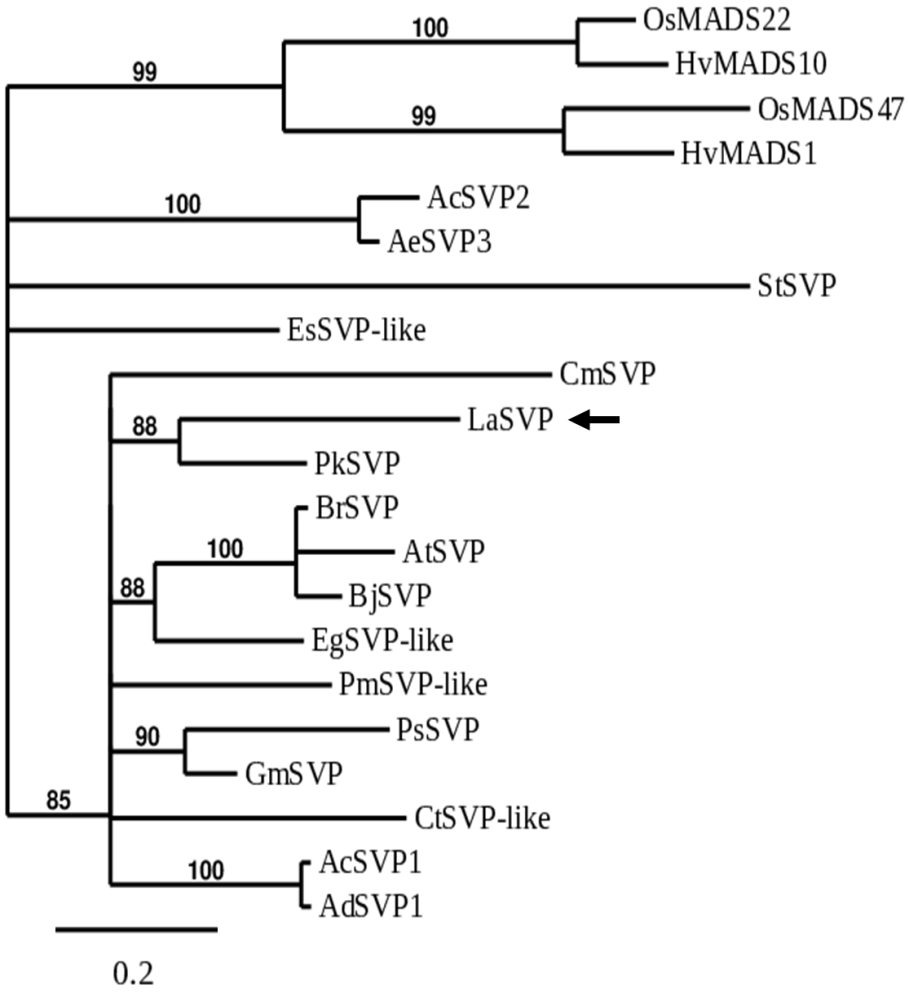

Figure 7. Phylogenetic analysis of the LaSVP protein and other plant SVP-like proteins. OsMADS22, *Oryza sativa*; HvMADS10, *Hordeum vulgare*; OsMADS47, *Oryza sativa*, HvMADS1, *Hordeum vulgare;* AcSVP2, *Actinidia chinensis*; AeSVP3, *Actinidia eriantha*; StSVP, *Solanum tuberosum*; EsSVP-like, *Epimedium sagittatum;* CmSVP, *Chrysanthemum × morifolium*; LaSVP, *Lavandula angustifolia,* PkSVP, *Paulownia kawakamii Actinidia chinensis,* BrSVP, *Brassica rapa;* AtSVP, *Arabidopsis thaliana*; BjSVP, *Brassica juncea;* EgSVP-like, *Eucalyptus grandis*; PmSVP-like, *Prunus mume*; PsSVP *Pisum sativum;* GmSVP, *Glycine max;* CtSVP-like, *Citrus trifoliata*; AcSVP1, *Actinidia chinensis;* AdSVP1, *Actinidia deliciosa*. The scale bar indicates a divergence of 0.2 amino acid substitutions per site. The arrow indicates *L. angustifolia SVP (LaSVP)* studied in this thesis. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

Table 8. Percent identity matrix of the % similarity of *SVP* protein sequences derived from different plants. Matrix was created by Clustal2.1, where *LaSVP = Lavandula angustifolia, AtSVP = Arabidopsis thaliana* SVP, *CmSVP = Chrysanthemum x morifolium* SVP, *PkSVP = Paulownia kawakamii* SVP, *BrSVP = Brassica rapa* SVP, *BjSVP = Brassica juncea* SVP, and *EgSVP-like = Eucalyptus grandis* SVP. *AtSVP* and *LaSVP* show 67.39% sequence similarity, with highest similarity between the *LaSVP* and *EgSVP-like* protein sequences at 69.78%.(Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

No.	Gene	1	2	3	4	5	6	7
1	**LaSVP**	100.00						
2	**AtSVP**	67.39	100.00					
3	**CmSVP**	59.46	62.61	100.00				
4	**PkSVP**	76.21	69.60	65.91	100.00			
5	**BrSVP**	68.26	91.67	64.86	69.60	100.00		
6	**BjSVP**	67.83	90.83	63.51	69.16	97.51	100.00	
7	**EgSVP-like**	69.78	70.35	65.44	74.44	72.57	73.01	100.00

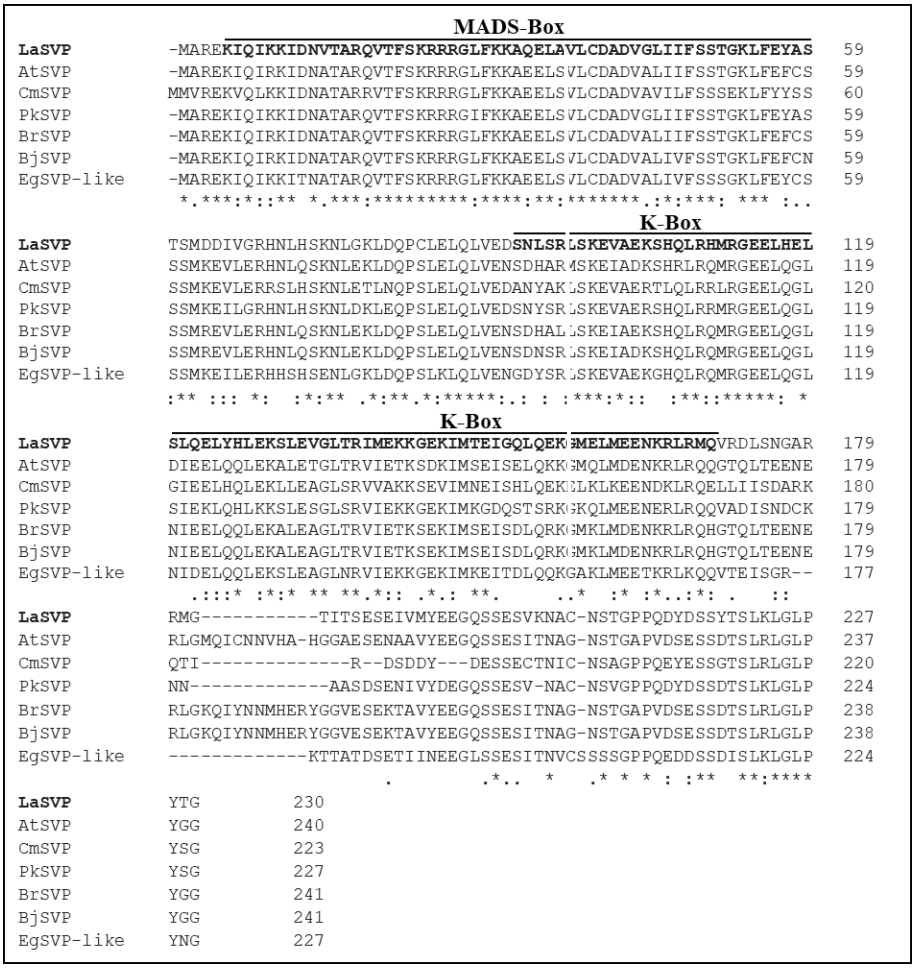

Figure 8. Multiple Sequence Alignment of protein sequences of *LaSVP* and other plant *SVP* homologs. *LaSVP, AtSVP, CmSVP, PkSVP, BrSVP, BjSVP*, and *EgSVP-like* protein sequences, where *LaSVP =Lavandula angustifolia* SVP, *AtSVP =Arabidopsis thaliana* SVP, *CmSVP = Chrysanthemum x morifolium* SVP, *PKSVP = Paulownia kawakamii* SVP, *BrSVP = Brassica rapa* SVP, *BjSVP = Brassica juncea* SVP, and *EgSVP-like = Eucalyptus grandis* SVP-like. The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

3.4 Cloning and Transformation

3.4.1. Cloning of Candidates in the pGA482 Vector

Based on differential expression pattern and high sequence homology to an *A. thaliana* floral development protein as indicated by an e-value $\leq 10^{-10}$, a subset of eight full-length *Lavandula* genes was selected for cloning and functional characterization in wild-type *A. thaliana*. Using gene-specific primers, *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, *LaAP1*, and *LiAP2* were successfully PCR amplified from *L. angustifolia* BD cDNA using gene-specific primers. The PCR reactions were visualized by 1 % agarose gel electrophoresis and bands were observed at the expected band size based on the amino acid length of each protein size at 492 bp, 762 bp, 693 bp, 564 bp, 525 bp, 762 bp, and 1302 bp for *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, *LaAP1*, and *LiAP2*, respectively (Figure 9).

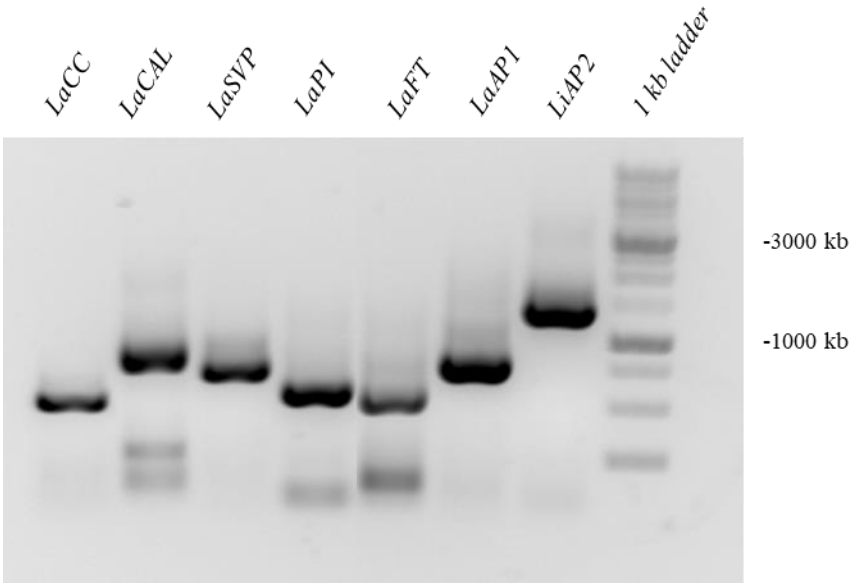

Figure 9. Detection of candidate genes with PCR using cDNA derived from *L. angustifolia* BD tissue. All 8 candidate genes, *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, *LaAP1*, and *LiAP2*, were amplified via PCR from *L. angustifolia* BD cDNA and were detected at the appropriate band when compared to a 1 kb ladder.

All 7 of the amplified candidate genes were successfully extracted from the agarose gel, although *LiAP2* extraction was not of sufficient purity and thus was excluded from the remainder of the study. The *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, and *LaAP1* DNA and the pGA482 plasmid were digested with the restriction enzymes *Eco*RI and *Kpn*I. All six candidate genes were successfully ligated into separate pGA482 vectors and the pGA::*LaCC* , pGA::*LaCAL* , pGA::*LaSVP* , pGA::*LaPI* , pGA::*LaFT*, and pGA::*LaAP1* constructs were successfully prepared. The constructs were successfully transformed into competent *E. coli* JM109 cells and

colonies grew well for each candidate. Positive colonies were successfully selected for via antibiotic selection on ½ MS medium containing tetracycline antibiotic (Figure 10). For each gene construct, two plates were prepared; a 50 µl plate and a concentrated plate (Figure 10). For each construct, both the 50 µl plate and the concentrated plate exhibited colony growth, although colony growth was significantly higher on the concentrated plates (Figure 10). Control plates containing an empty pGA482 vector exhibited no growth on the 50 µl plates, but exhibited sparse growth on the concentrated plates.

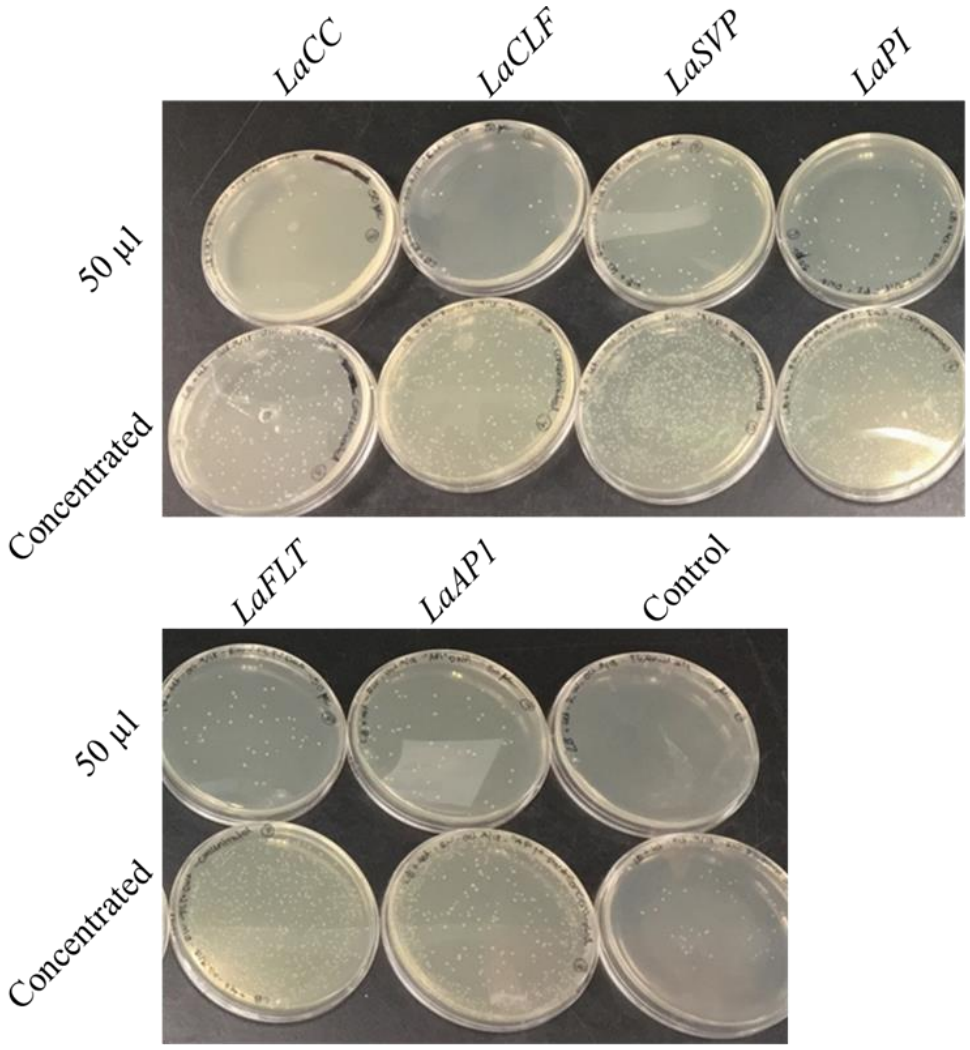

Figure 10. Antibiotic selection of transformed *E. coli* JM109. Selection was performed on Luria Bertani 30 mg/mL tetracycline selection plates. For each gene, two plates were prepared, including a 50 µl plate and concentrated plate. Two control plates were prepared which included untransformed *E.coli* JM109 containing the empty vector housing the tetracycline antibiotic resistance gene. All plates exhibited growth and 3 colonies were selected from each 50 µl plate.

Plasmids were successfully extracted from the *E. coli* colonies with good purity and Illumina sequencing with gene-specific primers confirmed the presence of the correct *Lavandula*

gene insert in each respective pGA482 vector. In addition, restriction digestion of the constructs further confirmed the presence of each of the lavender gene inserts in the pGA482 vectors extracted from the *E.coli* colonies as bands were detected at the appropriate size for each of the eight genes on an agarose gel (Figure 11). For each gene, two bands were observed on the agarose gel, indicative of a lavender gene insert of variable size (gene-specific) and the digested vector at >10,000 bp (Figure 11). The construct for each gene was successfully digested via restriction digest with *Kpn*I and *Eco*RI and the lavender genes *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, *LaAP1*, and *LiAP2* were detected via 1% agarose gel electrophoresis at 492 bp, 762 bp, 693 bp, 564 bp, 525 bp, and 762 bp, respectively.

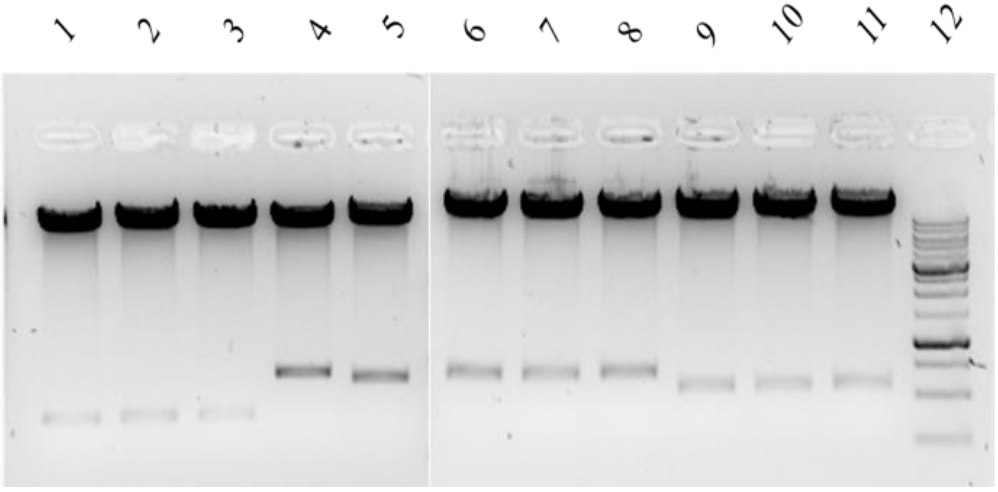

Figure 11. Agarose gel of restriction digestion of the pGA482:: *Lavandula* gene constructs extracted from *E. coli* JM109 cells. Lanes 1-3 indicate *LaCC* at 492 bp, lanes 4- 5 indicate *LaCAL* at 762 bp, lanes 6-8 indicate *LaSVP* at 693 bp, and lanes 9-11 show *LaPI* at 564 bp. All lanes show a dark band at the top of the gel above 10,000 bp, which is indicative of the digested pGA482 vector.

After validation of the presence of the respective *Lavandula* genes in the pGA482 vector, *Agrobacterium tumefaciens* GV3101 cells were transformed with the constructs via the freeze-thaw method (Jyothishwaran et al. 2007). All six constructs were successfully transformed into *A. tumefaciens* and positive colonies were selected for via antibiotic selection on kanamycin and rifampicin selection medium (Figure 12). For each gene construct, two plates were prepared, including a 50 µl plate and a concentrated plate (Figure 12). The 50 µl plate and the concentrated plate exhibited colony growth in all cases, although colony growth was significantly higher on the concentrated plates (Figure 12). Control plates containing an empty pGA482 vector exhibited no growth on the 50 µl plates, but exhibited sparse growth on the concentrated plates.

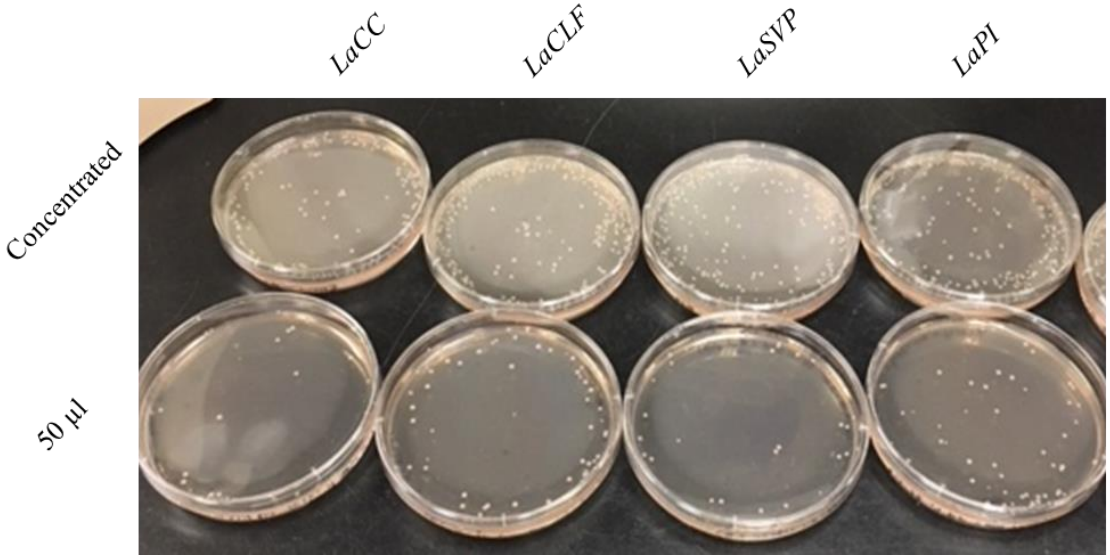

Figure 12. Representative plates of selection of transformed *A. tumefaciens* GV3101 with the pGA482::lavender gene constructs. Selection occurred on Luria Bertani selection plates containing 25 mg/mL rifampicin and 50 mg/mL kanamycin. For each gene, two plates were prepared, including a 50 µl plate and concentrated plate. Two control plates were prepared which included untransformed *A. tumefaciens* GV3101 containing the empty vector housing the antibiotic resistance genes, but without a gene insert. All plates exhibited growth and 3 colonies were isolated from each 50 µl plate for each gene.

Constructs from the transformed *A. tumefaciens* colonies were successfully extracted and *Lavandula* candidate genes were PCR amplified using the extracted constructs as a template and gene-specific primers. All candidates were successfully amplified from their respective construct (Figure 13). It is of note that during the extraction of the *LiAP2* plasmid from *A.* tumefaciens, there was insufficient sample and thus this gene was not used in the PCR reaction, nor was it further examined in the remainder of the study due to time constraints. The PCR reaction was performed to verify the successful transformation of *A. tumefaciens* via the freeze-thaw method. The genes *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, and *LaAP1* were successfully PCR amplified from the respective extracted *A. tumefaciens* plasmids and visualized on a 1% agarose gel. Bands indicative of *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, and *LaAP1* were observed at 492 bp, 762 bp, 693 bp, 564 bp, 525 bp, and 762 bp, respectively (Figure 13). A PCR reaction was performed with an *E. coli* plasmid for each gene to serve as a control as candidate gene presence had been confirmed previously (Figure 13). All *E. coli* plasmids again showed successful amplification of candidate genes.

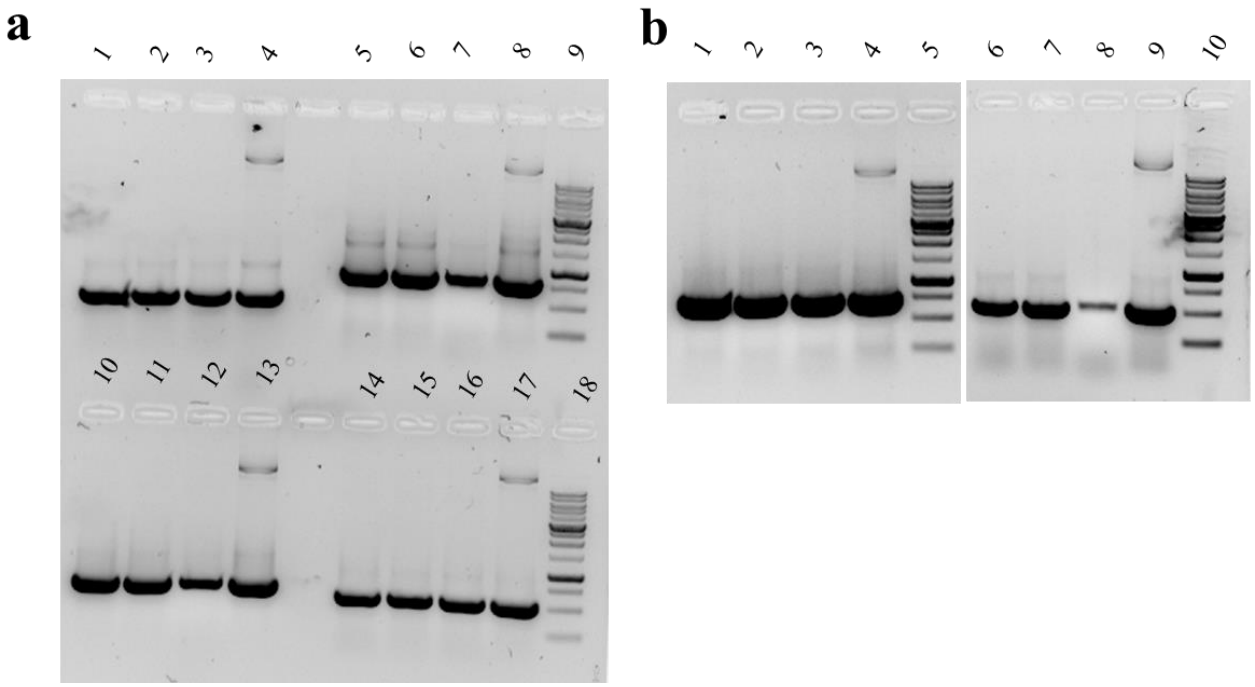

Figure 13. PCR amplification of lavender candidate genes from the pGA482::lavender gene constructs. Constructs were extracted from transformed *A. tumefaciens* and were PCR amplified using gene-specific primers. a) Lanes 1-3 indicate pGA482::*LaCC*, lanes 5-7 show pGA482::*LaCAL*, lanes 10-12 indicate pGA::*LaSVP*, lanes 14-16 indicate pGA482::*LaPI*, lanes 4, 8, 13, and 17 show the respective digested *E.coli* JM109 plasmids for verification and lanes 9 and 18 show the 1 kB ladder. b) Lanes 1-3 show amplification from pGA482::*LaAP1*, lanes 6-8 show pGA482::*LaFT*, Lanes 4 and 9 show the respective restriction digested *E. coli* JM109 plasmids and lanes 5 and 10 indicate the 1 kB ladder. Lanes showing amplification from the digested *E. coli* JM109 plasmids show a dark band at the top of the gel above 10,000 bp, which is indicative of the pGA482 vector.

3.4.2 Transformation and Selection of Transgenic *Arabidopsis thaliana*

To functionally characterize the lavender candidate genes, each gene was expressed in wild-type *A. thaliana* var. *Landsberg erecta* via the constitutive CaMV 35S promoter of the pGA482 vector. Transformation of *A. thaliana* was successfully performed via the floral dip method (Zhang et al. 2006a). For each pGA482::lavender gene construct, a total of four plants were transformed, with a total of 24 plants transformed. Plants were all successfully grown to maturity and seeds were collected from all mature plants at the point at which the siliques had dried and released seeds (Figure 14d). All plants grew well after transformation and produced normal siliques and viable seeds that germinated well. Viable seeds were successfully collected from each transformed plant.

Plant transformation was first assessed through germination of F_1 seeds of transformants on selection medium containing kanamycin alongside wild-type controls. Seeds were grown for approximately 3-4 weeks until selection was apparent (Figure 14). Initially, plates were prepared

with either 50 mg/mL kanamycin or 75 mg/mL kanamycin, but the 75 mg/mL plates proved to provide better selection and thus, 75 mg/mL kanamycin was used for all subsequent plates. While wild-type seedlings became purple or yellow and died soon after germination, some of the seeds originating from the transformed plants germinated and grew well on the selection medium (Figure 14). Seedlings from transformed plants which were able to grow on selection medium were green and exhibited increased root and leaf biomass in comparison to wild- type (Figure 14). Plates which exhibited growth of seedlings often exhibited 1-4 transformants per plate. Wild-type seeds grew well on the medium containing no antibiotic but failed to grow well on the medium containing kanamycin. Positive growth of F_1 seedlings was observed for *LaSVP*, *LaPI*, *LaAP1*, *LaCAL*, *LaCC* and *LaFT*, indicating successful transformation as seedlings thus harbored the kanamycin resistance gene. For each gene, at least 10 seedlings exhibiting positive growth were identified. Seedlings growing on kanamycin selection medium were transferred to soil and were analyzed in parallel with wild-type controls. Over 40 transformant seedlings, in addition to wild-type controls, were transferred from the selection plates for genetic analyses. Although many seedlings survived to transfer, some died soon after transfer or before sufficient tissue was available.

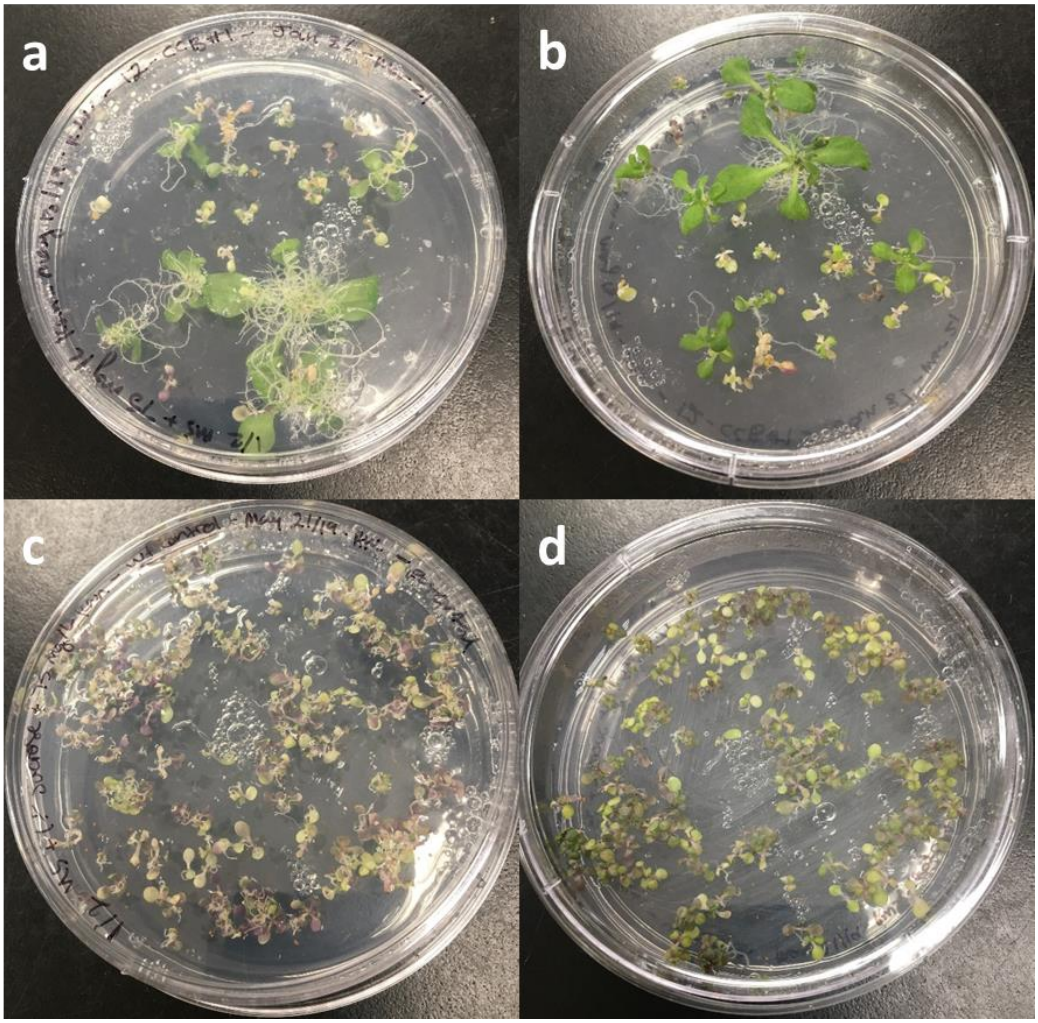

Figure 14. Representative plates of kanamycin antibiotic selection of F_1 transformant seedlings. Selection was performed on ½ MS 75 mg/mL kanamycin selection medium with wild-type *A. thaliana* controls of the same age grown under the same conditions. (a) bottom view of T2CCB1 seedlings germinated on 75 mg/mL kanamycin medium, (b) top view of T2CCB1 seedlings germinated on 75 mg/mL kanamycin selection medium. T1CCB2 transformants able to grow in the presence of kanamycin exhibited a large root biomass and broad, green leaves. (c), (d) Wild-type *A. thaliana* seeds. Wild-type seeds germinated but died soon after as they did not confer the antibiotic resistance gene, *NPT-II*.

3.5 Genetic Analyses of F_1 transformants

3.5.1 DNA Detection of the *NPT-II* Gene

Successful plant transformation was further verified by PCR amplification of the *Lavandula* and neomycin phosphotransferase (*NPT-II*) genes from leaf-derived genomic DNA (gDNA) of the F_1 transformant plants using gene-specific primers for each gene. The gDNA was extracted from F_1 plants with sufficient quality and purity and the gene-specific primers for *NPT-II* successfully amplified the *NPT-II* gene from the leaf gDNA of the transformants, T1SVPB1, T1SVPB4, T1SVPB5, SVP1, SVP2, SVP3, T2CALB1, T1CALA1, T1PIB1, TIPIB3, T1PIB4,

T1PIB5, T2PIB1, T2PIB2, T2PIA1, T2PIA2, T2AP1A3, T2AP1A4, T2AP1A8, T1FTA2. In some instances, there was insufficient leaf tissue, so cDNA only was obtained and used to amplify the *NPT-II* gene as in T1FTA2 and T1FTA8 (Table 9). In addition, amplification was verified through visualization of the PCR reaction for each transformant's cDNA on a 1% agarose gel and for each transformant, the *NPT-II* gene was detected at the appropriate length, 693 bp (Figure 15). This result was observed for all candidates in Table 9.

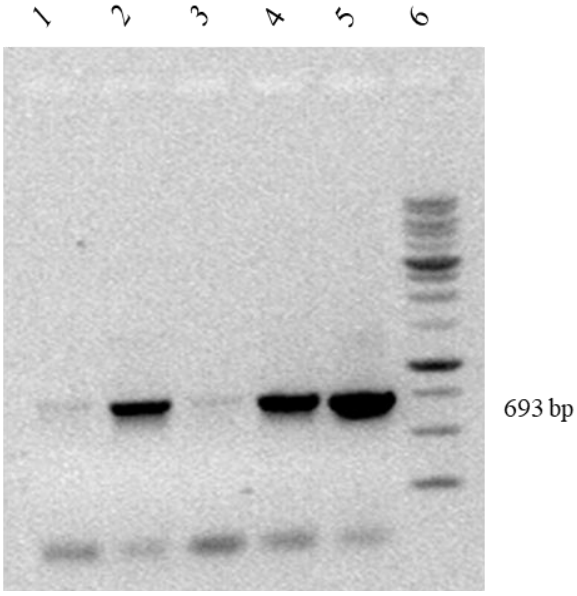

Figure 15. PCR amplification of the *NPT-II* gene from leaf-derived gDNA of F1 transformants using gene-specific primers. The *NPT-II* gene was successfully amplified from (1) T1AP1A3, (2) T1AP1A4, (3) T1SVPB4, (4) T1SVPB5, and (5) T1PIA1 and (6) 1 kB DNA ladder. Detection of the *NPT-II* gene in all candidates confirmed successful transformation with the pGA482:lavender gene constructs.

3.5.2 cDNA Detection of *Lavandula* Candidate Genes

To confirm expression of the respective lavender gene in each transformant, flower and/or leaf RNA were extracted from each transformant plant and RNA was reverse transcribed to cDNA. cDNA was used for expression analysis as detection in gDNA would not confirm if the gene was expressed. Using the cDNA as a template, transformation was verified via PCR amplification of the respective lavender genes using gene-specific primers for each gene. The respective lavender genes were successfully amplified from the cDNA of the transformants, T1SVPB1, T1SVPB4, T1SVPB5, SVP1-3, T1CALA1, T1CALA2, T1CALB3, T1PIB3, T1PIB4, T1PIB5, T2PIB1, T2PIB2, T1PIA1, T2PIA2, T2CCB1, T2AP1A1, T2AP1A3, T2AP1A4, T2AP1A8, T1FTA2 (Table 9 & Figure 16). Successful amplification of the genes from cDNA indicated the gene was actively being expressed in these transformant tissues. In

some cases, there was insufficient tissue to obtain floral RNA due to altered growth. In these instances, whole plant or mixed leaf/flower RNA was obtained and analyzed. For example, transformant T2CCB1 did not produce sufficient tissue and thus cDNA derived from floral and leaf tissue was used to detect the *LaCC* gene in this plant. In addition, PCR amplification of the *NPT-II* and respective lavender gene was unsuccessful from transformants T1CALB3, T2CCB2, and T2AP1A6, although they successfully grew on selection medium (Table 9).

Gene-specific primers for each *Lavandula* gene were also used to PCR amplify the *NPT-II* gene and the *Lavandula* genes from wild-type *A. thaliana* as a control. The *LaSVP* and *LaAP1* genes were detected in the wild-type leaf cDNA and the *LaSVP* and *LaFT* genes were detected in the wild-type flower cDNA at the appropriate band weight; however, the *NPT-II* gene was detected only in transgenic lines and was absent in all wild-type plants.

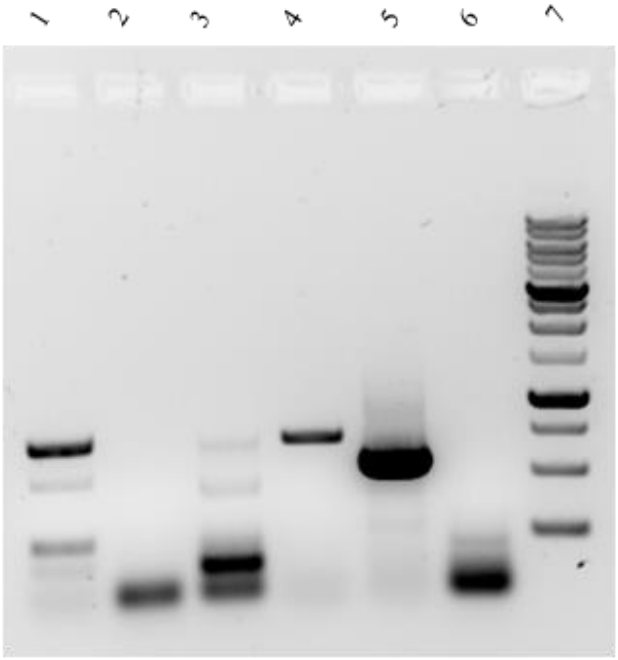

Figure 16. Detection of lavender candidate genes in leaf-derived cDNA from wild-type and F_1 *A. thaliana* and transformants via PCR with gene-specific primers. (1) *LaSVP* primers amplified the *A. thaliana* SVP gene due to high sequence homology and nonspecific amplification in wild-type *A. thaliana*, (2) *LaPI* was not amplified from wild-type *A.thaliana* cDNA ,(3) *LaCAL* was amplified from wild-type *A. thalaina* cDNA (4), *LaSVP* was amplified from T1SVPB1 cDNA (5) *LaPI* was amplified from T1PIB cDNA, (6) *LaCAL* was amplified from T1CALA1 cDNA.

Table 9. Genetic analysis of F_1 transformants. Results of gene detection and expression analyses of F_1 transformants and wild-type *A. thaliana*. Transformant plants listed based on the genes detected and confirmed expression of the *NPT-II* and *Lavandula* candidate floral gene in a given tissue. T1 = seed derived from a plant from transformation 1, T2= seed derived from a plant from transformation 2, A/B = plant A or B, number at the end is indicative of the replicate, SVP1 through 3 = *LaSVP* transformants prepared by an honors student. Detection of genes in gDNA confirmed gene presence, while detection in cDNA confirmed active gene expression. Wild-type *A. thaliana* exhibited amplification of *A. thaliana* homologs of *SVP*, *FT*, and *AP1* due to high sequence homology.

Transformant Name	**Sample**	**Genes Amplified**
T1SVPB1	Leaf gDNA Leaf RNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in flower and leaf cDNA
T1SVPB4	Leaf DNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in leaf gDNA *LaSVP* in flower cDNA
T1SVPB5	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in flower cDNA
SVP1	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in flower cDNA
SVP2	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in flower cDNA
SVP3	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaSVP* in flower cDNA
T2CALB1	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA
T2CALB3	Insufficient tissue	None
T1CALA1	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaCAL* from flower cDNA
T1CALA2	RNA whole plant	*LaCAL* from plant cDNA
T1CALB3	Leaf gDNA	None
T1PIB1	Leaf gDNA Leaf RNA	*NPT-II* in leaf gDNA *LaPI* in leaf cDNA
T1PIB3	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaPI* in floral cDNA
T1PIB4	Leaf gDNA Leaf RNA	*NPT-II* in gDNA *LaPI* in flower cDNA
T1PIB5	Leaf gDNA Leaf RNA	*NPT-II* in gDNA *LaPI* in flower cDNA
T2PIB1	Leaf gDNA Leaf RNA	*NPT-II* in leaf gDNA *LaPI* in leaf cDNA

T2PIB2	Leaf gDNA Leaf RNA	*NPT-II* in leaf gDNA *LaPI* in leaf cDNA
T2PIA1	Leaf gDNA Mixed leaf and flower RNA	*NPT-II* in leaf gDNA *LaPI* in leaf/flower cDNA
T2PIA2	Leaf gDNA Leaf RNA	*NPT-II* in leaf gDNA *LaPI* in leaf cDNA
T1CCA1	Insufficient tissue	None
T2CCB1	Mixed leaf and flower RNA	*LaCC* in cDNA
T2CCB2	Leaf RNA	None
T2AP1A1	Mixed leaf and flower RNA	*LaAP1* in cDNA
T2AP1A3	Leaf gDNA	*NPT-II in* leaf gDNA *LaAP1* in leaf gDNA
T2AP1A4	Leaf gDNA Leaf RNA	*NPT-II* in leaf gDNA *NPT-II* in leaf cDNA *LaAP1* in leaf cDNA
T2AP1A6	Leaf RNA	None
T2AP1A8	Leaf gDNA Flower RNA	*NPT-II* in leaf gDNA *LaAP1* in flower cDNA
T1FTA1	Insufficient tissue	None
T1FTA2	Whole plant RNA	*NPT-II* and *LaFT* from plant cDNA
T1FTA5	Insufficient tissue	None
T1FTA8	Whole plant RNA	*NPT-II* and *LaFT* from plant cDNA
Wild-type A. thaliana	Leaf RNA Flower RNA	*LaSVP* in leaf cDNA *LAPI* in leaf cDNA *LaSVP* in flower cDNA *LaFT* in flower cDNA

3.6 Phenotypic Analyses of Transformants

3.6.1. *LaSVP* Transformant Analysis

A total of seven kanamycin resistant lines expressing the *LaSVP* gene and originating from independent transformation events were identified (Table 9). Compared to wild-type, plants expressing the *LaSVP* gene exhibited a delay in flowering (Figure 17). The *LaSVP* primers weakly amplified the *AtSVP* gene from wild-type floral cDNA in one RT-PCR and failed to amplify in another reaction due to high sequence homology between *AtSVP* and *LaSVP* in the primer flanking regions. The *LaSVP* transformants started bolting 5-6 days later than their wild

type counterparts (Figure 17). In addition, *LaSVP* expressers produced abnormal carpels and seed pods and failed to develop petals (Figure 18a & c). The four transgenic lines which expressed the *LaSVP* the strongest were SVP-1, SVP-2, T1SVPB-4, and T1SVPB-5 (Figure 18 & 19). Further, the four strongest *LaSVP* expressers did not produce seeds and were sterile. Some of the flowers in other transformants, exemplified by the SVP-3 line, were similar to wild type flowers, while some flowers resembled those of the SVP-1, SVP-2, T1SVPB-4, and T1SVPB-5 lines (Figure 18b & 19). Additionally, these transgenic lines developed some normal seed pods and produced viable seeds in the phenotypically normal parts of the plant, while in other parts of the plants, the seed pods were abnormal and failed to produce viable seeds (Wells et al. 2020).

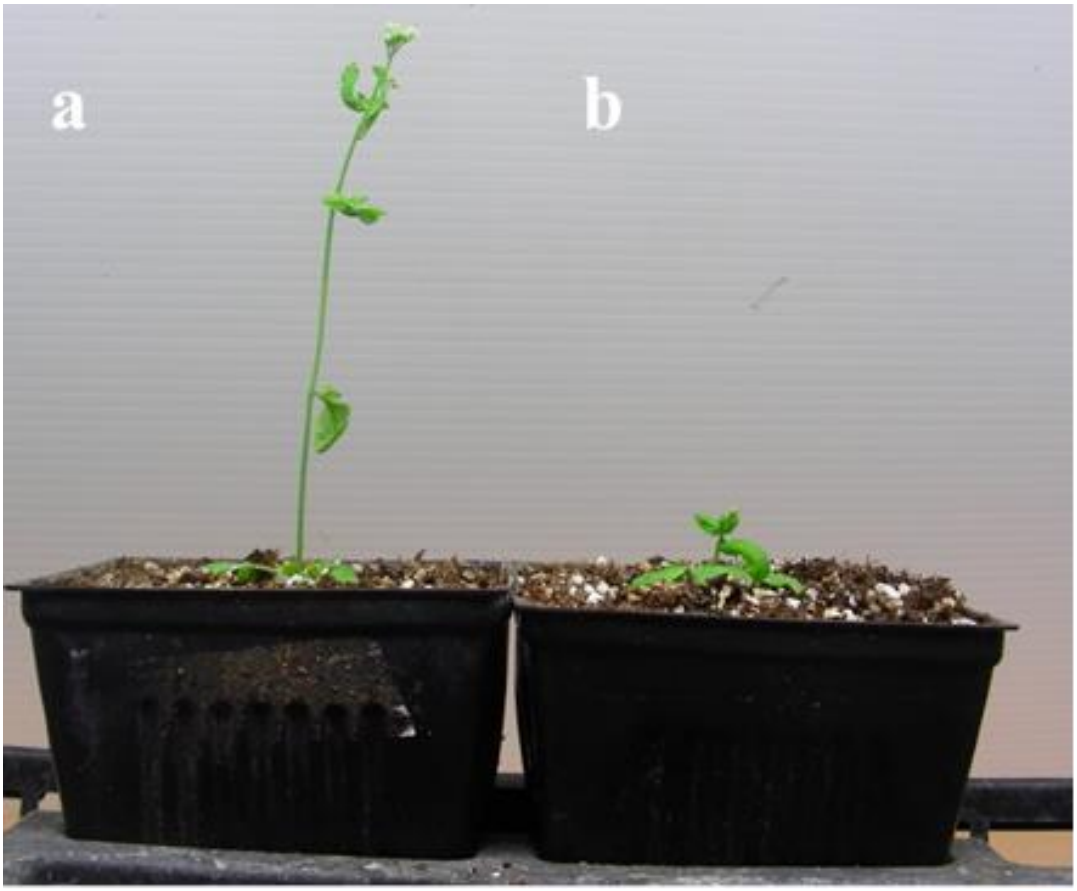

Figure 17. Effects of *LaSVP* overexpression on *A. thaliana* growth and flowering. The *LaSVP* transformant exhibited a delay in the transition to the flowering phase in comparison to wild-type *A. thaliana* when grown under long day conditions. (a) Wild-type (Wt) and (b) *LaSVP* transformant. The apparent difference in plant size (compared to WT plants) was not consistent among *LaSVP* transformants**.** (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

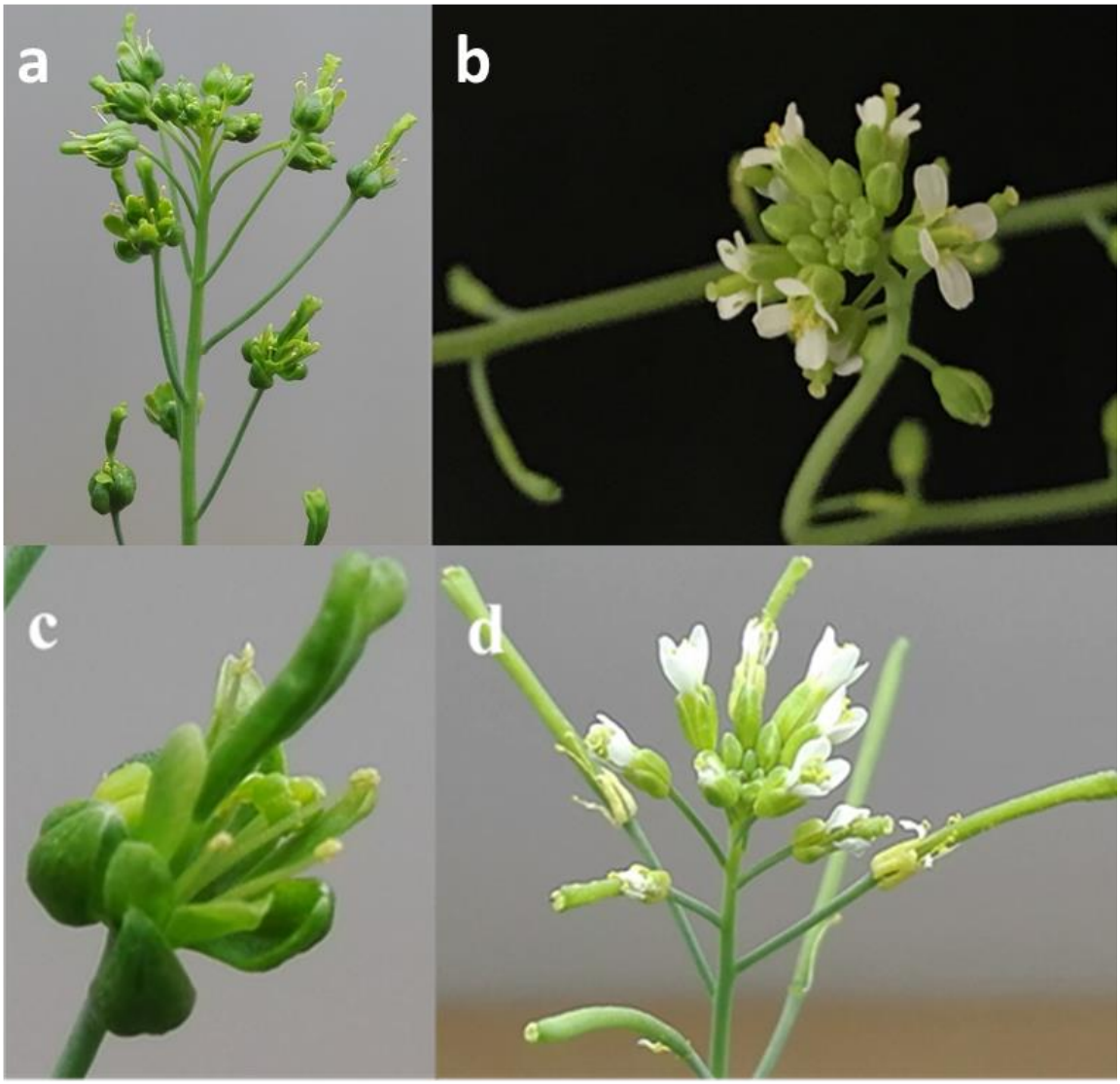

Figure 18. Differences in floral organ structures of *A. thaliana LaSVP* transformants versus wild-type *A. thaliana.* (a) Plants exhibiting high *LaSVP* expression, such as SVP-2, lacked petals, produced more sepals, developed abnormal seed pods, and were sterile. (b) and (c) Plants exhibiting lower expression of the *LaSVP* gene, (e.g., SVP-3) exhibited normal floral development in some portions of the plant, but abnormal floral development on other stems and produced seeds. (d) Wild-type *A. thaliana* var. *Landsberg erecta* flowers and seed pods. (Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

3.6.1.1 Expression of *LaSVP* and its Effects on Floral Transition and Morphology

Varying degrees of phenotypic abnormalities were observed across the independent transgenic lines of *LaSVP* transformants. To confirm *LaSVP* expression in transgenic plants, a fragment of the *LaSVP* coding sequence was amplified from floral cDNA of three representative transgenic plants by Reverse Transcriptase PCR (RT-PCR). The results demonstrated that the

LaSVP gene was expressed in all transgenic plants (Figure 19). RT-PCR results showed higher expression of *LaSVP* in SVP-1, SVP-2, and T1SVPB5 transformants than in other transgenic lines (Figure 19). The results also showed higher expression of *LaSVP* in T1SVPB4 (data not shown). Transformants T1SVPB1, and SVP-3 showed weak amplification of the *LaSVP* gene from floral cDNA (Figure 23). The *A. thaliana Actin* reference gene was amplified in all wild-type and transformant samples successfully. The *LaSVP* gene was weakly amplified from wild-type floral cDNA in one RT-PCR and failed to amplify in another reaction, although *Actin* was successfully amplified from wild type floral cDNA in both cases (Figure 19).

In addition to varying levels of expression observed across transformants, varying levels of mutant phenotype was observed in transformants. The strength of expression of *LaSVP* in the transformants was associated with the intensity of the altered phenotype. In comparison to wild-type *A. thaliana*, *LaSVP* transformants exhibited a delay in flowering (Figure 17). In addition, the *LaSVP* expressers began bolting 5-6 days later than their counterpart wild-type plants (Figure 17). In addition, *LaSVP* transformants produced abnormal carpels and seed pods and failed to develop petals (Figure 18 a & c). Further, four of the strongest *LaSVP* expressers (SVP-1, SVP-2, T1SVPB4, and T1SVPB5) did not produce seeds and were completely sterile. Some of the flowers in other transformants, including the SVP-3 line and T1SVPB1, were similar to wild type flowers, while other flowers within the same plant resembled those of the SVP1, SVP2 and T1SVPB-4 and T1SVPB-5 lines (Figure 18 a & c, Figure 20). Thus, these plants exhibited a partially altered phenotype, with some parts of the plant being sterile and exhibiting a transgenic phenotype, and other parts of the plant resembling wild- type. These transgenic lines thus also developed some normal seed pods and produced viable seeds in the phenotypically normal parts of the plant (Wells et al. 2020).

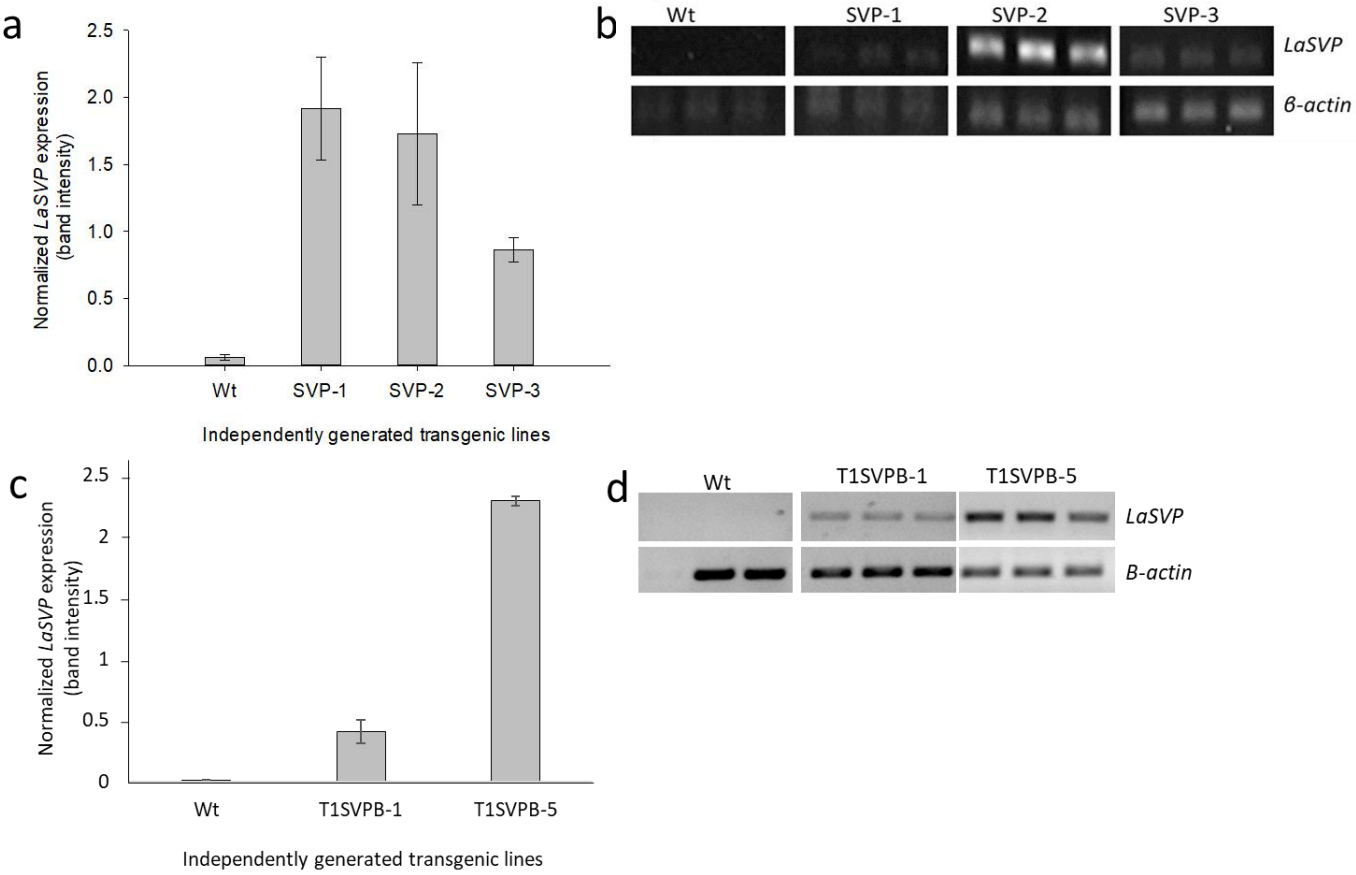

Figure 19. Normalization (to *β-actin*) of *LaSVP* expression and RT-PCR detection of the *β-actin* and *LaSVP* genes in wild-type *A. thaliana* and *LaSVP* transformant floral cDNA. a) Normalized (to β-actin) expression of *LaSVP* in wild-type (Wt) and SVP-1, SVP-2, and SVP-3, assessed by RT-PCR. The *LaSVP* and *β-actin* genes were amplified by RT-PCR, and PCR products were resolved on an agarose (1%) gel. Band intensities were measured by densitometry, using the ImageJ software. No *LaSVP* amplification was observed in Wt; the apparent low expression in Wt is due densitometry background reading. b) RT-PCR detection of *β-actin* and the *LaSVP* gene in wild-type *A. thaliana* and SVP-1, SVP-2, and SVP-3 leaf cDNA. *LaSVP* was detected (~693 bp) in the transformants SVP-1, SVP-2, and SVP-3, while *Actin* was detected in all samples (~250 bp). SVP-2 and SVP-3 exhibited strong expression of the *LaSVP* gene in comparison to SVP-1. c) Normalized (to β-actin) expression of the *LaSVP* gene in wild-type (Wt) and two representative transformed Arabidopsis plants, T1SVPB1 and T1SVPB2 assessed by RT-PCR. The LaSVP and β-actin genes were amplified by RT-PCR, and PCR products were resolved on an agarose (1%) gel. Band intensities were measured by densitometry, using the ImageJ software. No *LaSVP* amplification was observed in Wt; the apparent low expression is due densitometry background reading. d) RT-PCR detection of *β-actin* and *LaSVP* in wild-type *A. thaliana* and T1SVP-1 and T1SVPB-5 floral cDNA**.** *LaSVP* was detected (~693 bp) in the transformants T1SVPB-1 and T1SVPB-5, while *β -Actin* was detected in all samples (~250 bp). T1SVPB-5 exhibited strong expression of the *LaSVP* gene in comparison to T1SVPB-1. ((a) and (b) Reprinted by permission from Springer Nature: *Planta* 251(41): 1-11, Cloning and functional characterization of a floral repressor gene from *Lavandula angustifolia* by Rebecca S. Wells, Ayelign M. Adal, Lina Bauer, Elaheh Najafianashrafi and Soheil S. Mahmoud, © 2020).

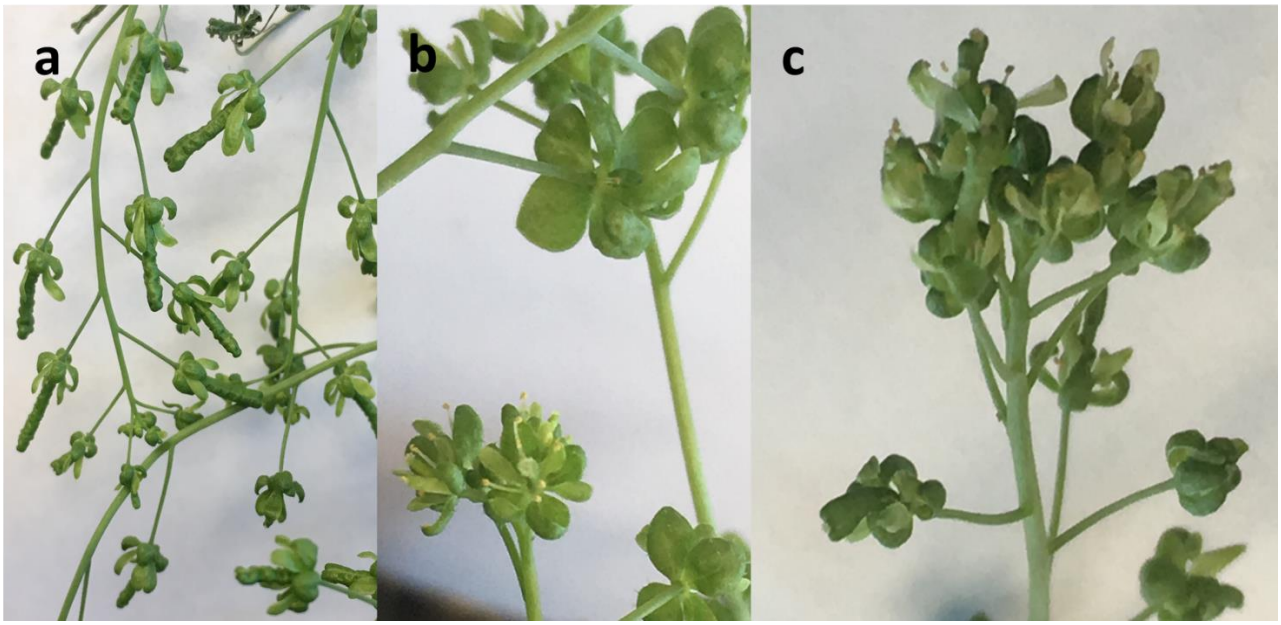

Figure 20. Floral organ structures of *A. thaliana LaSVP* transformants. (a), (b), Transformant T1SVPB4, and (c) transformant T1SVPB5 lacked petals and exhibited altered carpel shape. Both T1SVPB4 and T1SVPB5 were sterile as they failed to produce viable seeds.

3.6.2 *LaCC* Transformant Analysis

Three kanamycin resistant transformant lines which had arisen from independent transformations with pGA4882::*LaCC* were identified (Table 9). Several F_1 seedlings grew in the presence of kanamycin but died shortly after transfer to soil. Three seedlings successfully grew in soil, although the *LaCC* gene was only successful amplified from one plant, T2CCB1, which differed significantly from the wild-type phenotype (Figure 21a). All *LaCC* transformants exhibited significantly stunted growth and a decrease in vegetative growth compared with wild-type (Figure 21). In addition, *LaCC* transformants exhibited leaves which curled towards the abaxial surface and had an increase in carpel abundance, but a decrease in carpel length (Figure 21). In addition, the T2CCB1 transformant exhibited a reduction in the number of cotyledons.

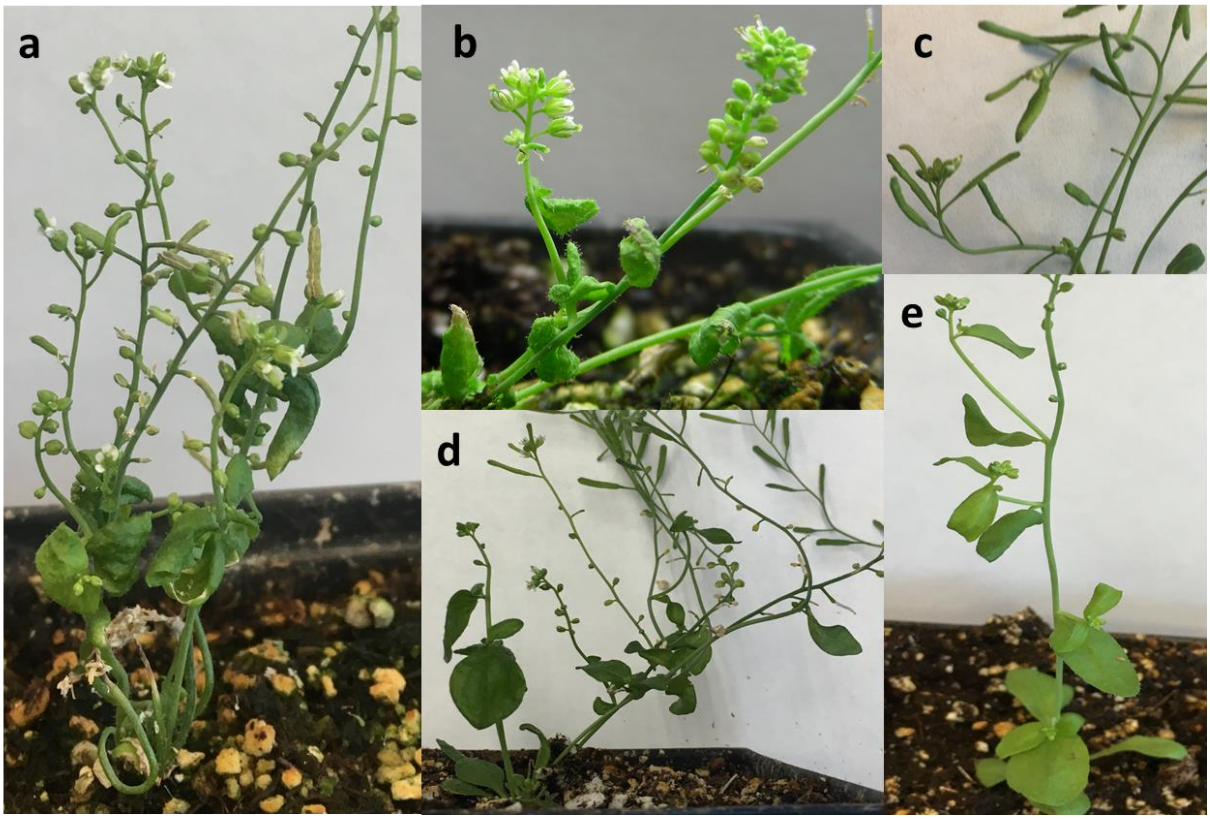

Figure 21. Differences in floral organ structures of *A. thaliana LaCC* transformants versus wild-type *A. thaliana.* (a) Transformant T2CCB1, (b) transformant T1CCA1, (c) wild-type *A. thaliana* var. *Landsberg erecta* siliques (d) transformant T2CCB2, (e) wild-type *A. thaliana* var. *Landsberg erecta*. Transformants exhibited curled leaves, increased carpel number, and a reduction in carpel length compared to wild- type.

3.6.3. *LaPI* Transformant Analysis

A total of eight kanamycin resistant lines expressing the *LaPI* gene and originating from independent transformation events were identified (Table 9). All transformants exhibited expression of both the *NPT-II* and *LaPI* genes. In most cases, *LaPI* transformants remained phenotypically the same as wild-type plants; however, some transformants presented slight phenotypic differences (Figure 22). Transformants T1PIB1, T2PIB2, T2PIA2, and T2PIB1 exhibited a slight increase in stem and leaf growth (Figure 22 e, f, g, h). In addition, transformant T2PIA2 exhibited abnormal leaves, buds, carpels and siliques in some areas of the plant, but remained phenotypically similar to wild type in other areas of the plant (Figure 22 a). Transformants T1PIB3, T1PIB4, T1PIB5, and T1PIB1 exhibited leaves which appeared to be slightly enlarged (Figure 22, b, c, d, e). Transformant T1PIB1 exhibited both a slight increase in vegetative growth and leaf size in many leaves, although results were insignificant. Some transformants also exhibited early flowering, although many of the early flowering plants did not survive.

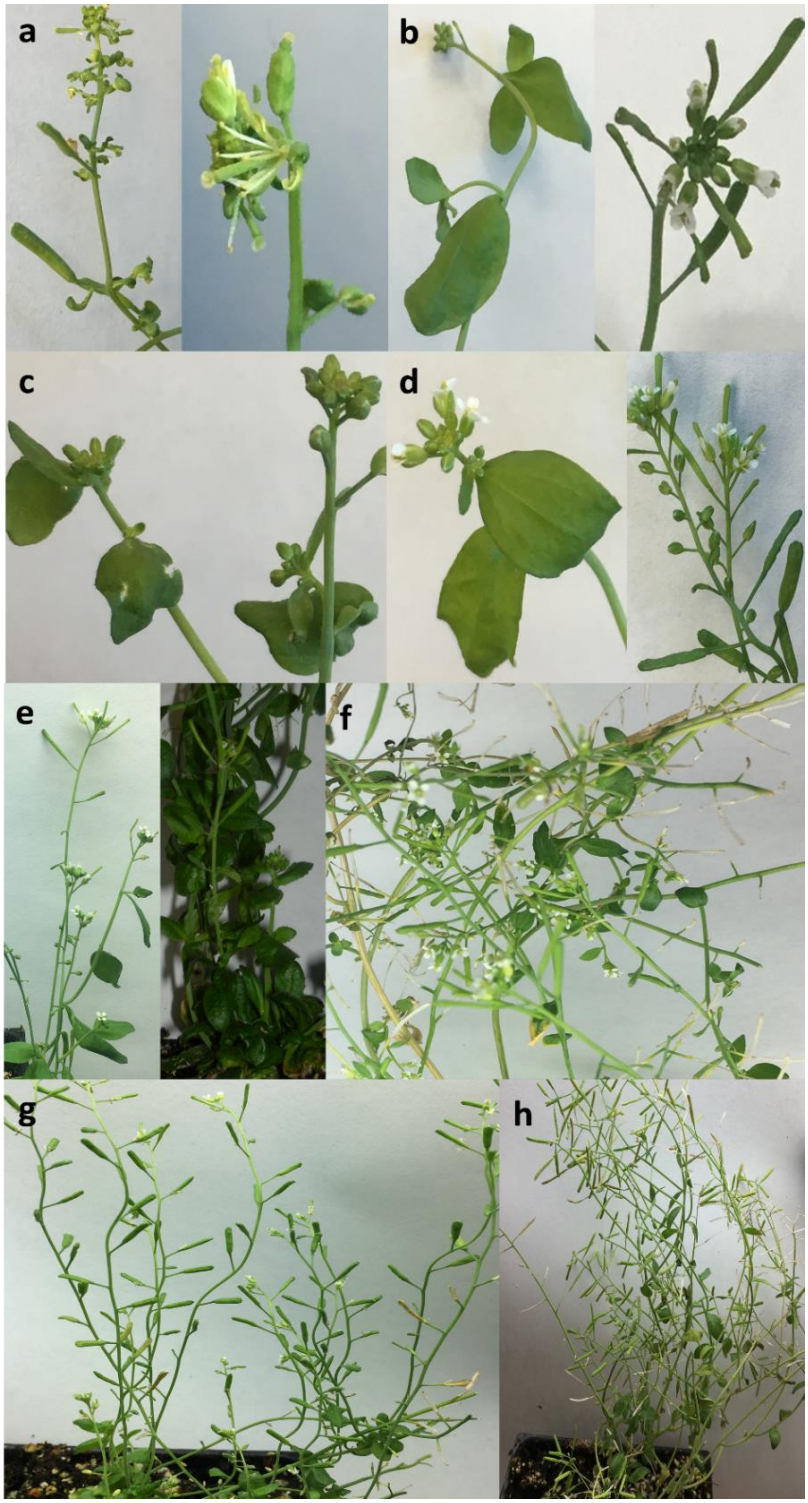

Figure 22. Phenotype and floral organ structures of *A. thaliana LaPI* transformant F_1 plants. (a) T2PIA2, (b) T1PIB3, (c) T1PIB4 (d), T1PIB5, (e) T1PIB1, (f) T2PIB2, (g) T2PIA2, (h) T2PIB1. The *LaPI* transformants exhibited phenotypes similar to wild type *A. thaliana*, with some exceptions. T2PIA2 exhibited abnormal leaves, buds and carpels in some areas of the plant, but resembled wild type in other areas. Transformants T2PIB3, T1PIB4, T1PIB5, and T1PIB1 exhibited an apparent overall increase in leaf size. Transformants T1PIB1, T1PIB2, T1PIA2, and T2PIB1 exhibited an apparent increase in stem number and number of leaves.

3.6.4 *LaAP1* Transformant Analysis

A total of 4 plants which expressed the *LaAP1* gene were identified, including T2AP1A1, T2AP1A3, T2AP1A4 and T2AP1A8. All transformants expressed both the *NPT-II* and the *LaAP1* genes, with the exception of T2AP1A1, from which only *LaAP1* was detected (Table 9). Most of the transformants exhibited a phenotype very closely resembling that of the wild- type. Although some transformants appeared to have a slight increase in bud number, this result was insignificant (Figure 23 & 24). In addition, it appeared that the *LaAP1* transformants flowered at a young age, although this result was inconsistent, and it is unclear if this result was due to growth conditions.

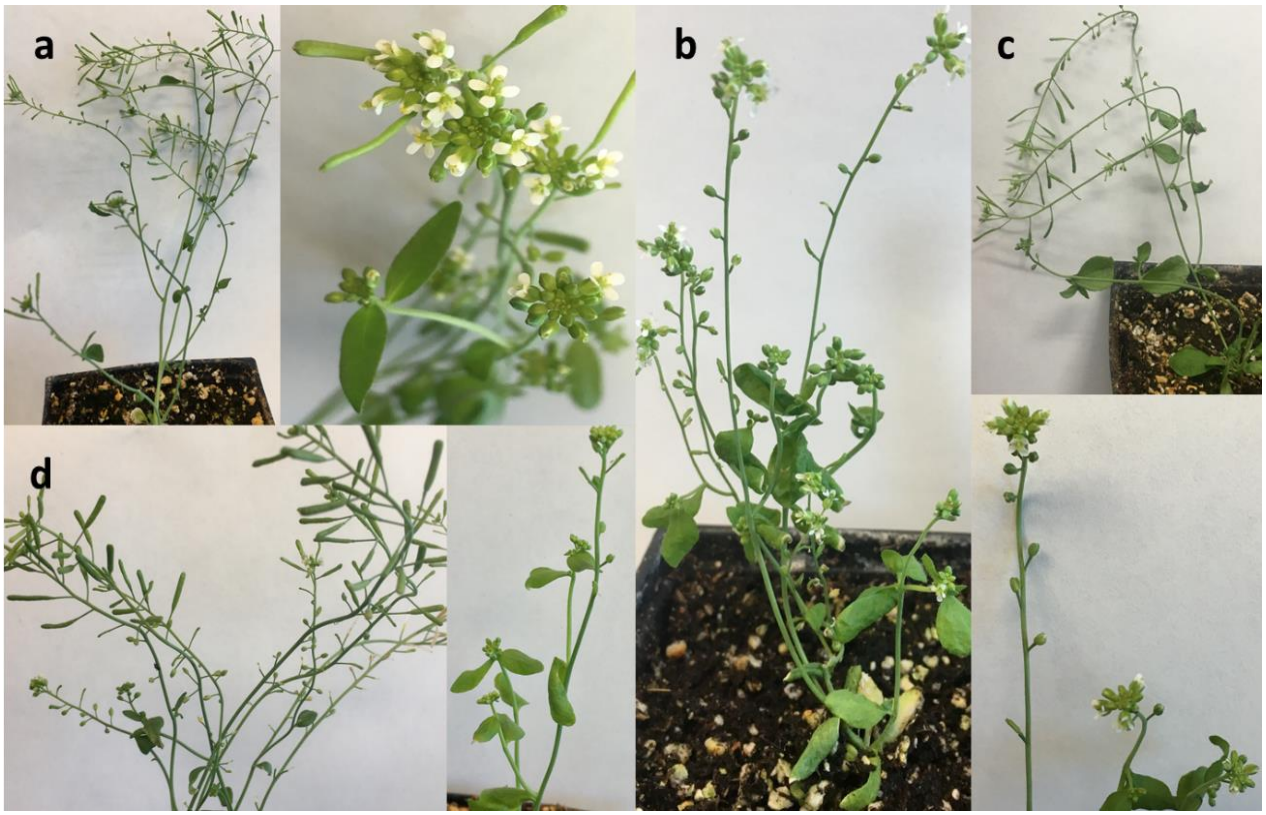

Figure 23. Phenotypes of the *LaAP1* F_1 transformants. (a) T2AP1A4, (b) T2AP1A1, (c) T2AP1A8, (d) T2AP1A3. *LaAP1* transformants remained phenotypically similar to wild- type *A. thaliana*, although T2AP1A1 exhibited altered carpel structure.

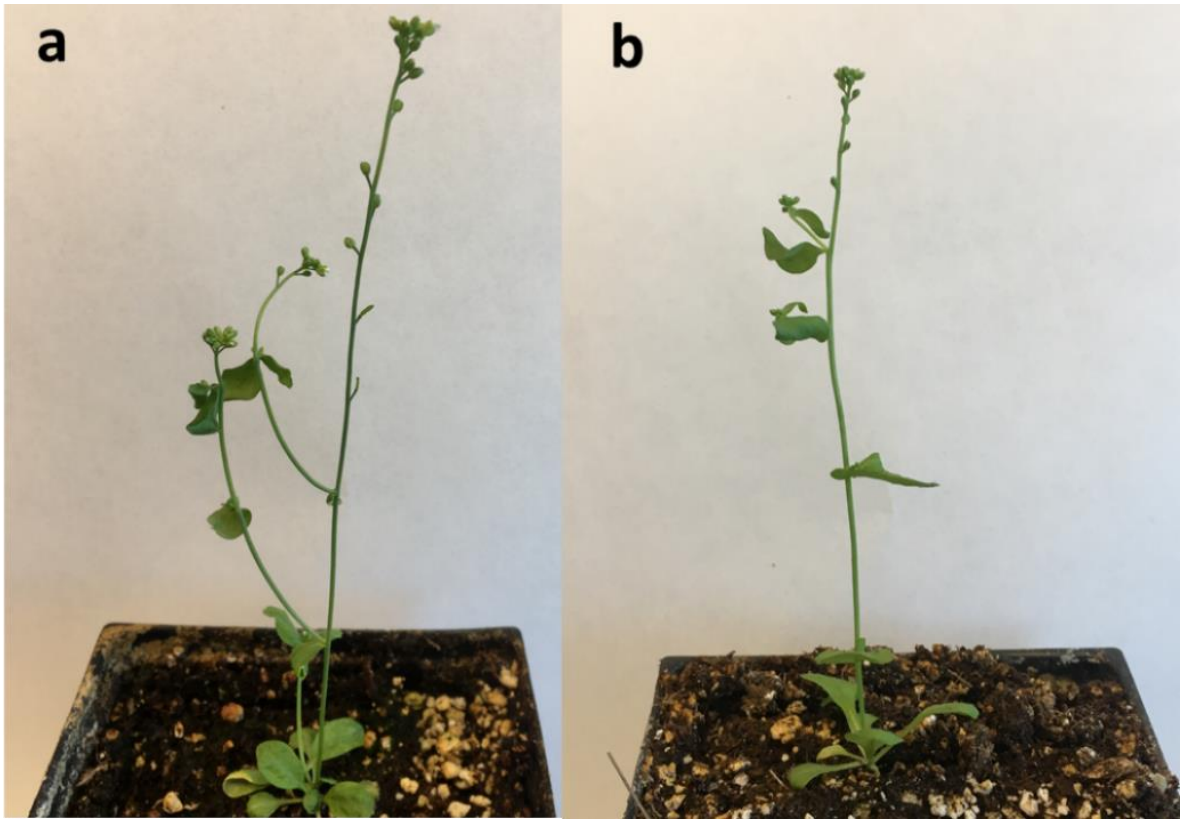

Figure 24. Phenotypic comparison of *LaAP1* transformant and wild-type *A. thaliana.* (a) Transformant T2AP1A8, (b) wild-type *A. thaliana* var. *Landsberg erecta*. No significant phenotypic difference was observed between *LaAP1* transformant and wild type. Plants are shown at 5 weeks, 4 days of age.

3.6.5 *LaCAL* Transformant Analysis

A total of 3 *LaCAL* transformants originating from 3 independent transformation events were identified through expression analysis. Over 10 seedings derived from pGA::*LaCC* transformants were selected from antibiotic selection plates and transferred to soil. Several transformants had insufficient tissue for testing prior to dying and were thus not tested. Of the tested transformants, T1CALA1, T2CALB1, and T1CALA1 were confirmed transgenic (Figure 25). Transformants T1CALA1, and T2CALB1, and T1CALA1 exhibited expression of the *NPT-II* and *LaCal* gene, the *NPT-II* gene, and *LaCAL*, respectively (Table 9). Of the transformants, none exhibited significant phenotypic differences compared to wild-type *A. thaliana*, although they all seemed to exhibit a slight reduction in vegetative growth and reduced height.

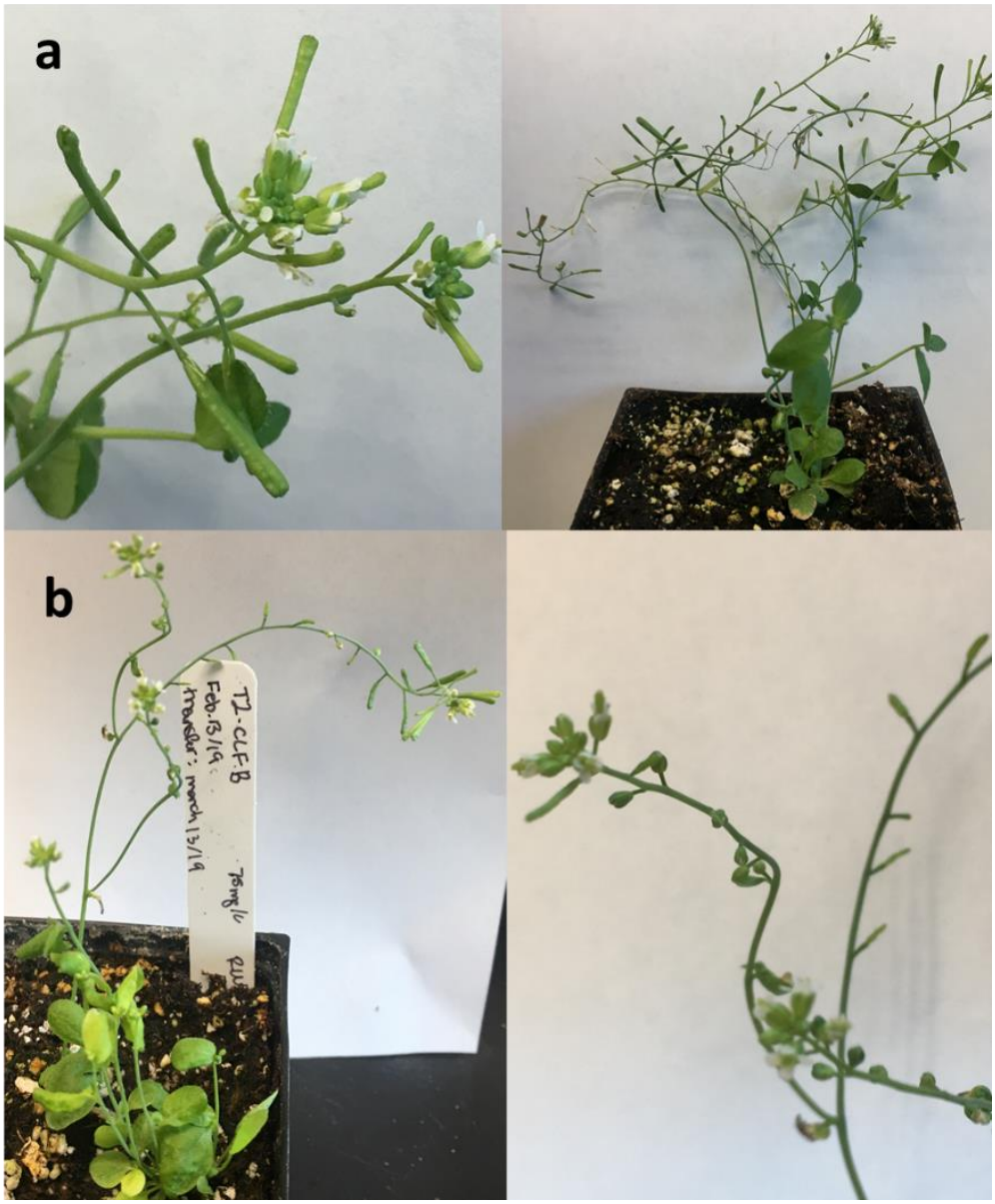

Figure 25. Phenotypes of the *LaCAL* transformants. (a) T1CALA1, (b) T2CALB1. Note that during the experiment, *LaCAL* transformants were referred to as *LaCALF*, as seen in (b). *LaCAL* transformants exhibited a slight reduction in height and vegetative growth compared to wild- type, although remained phenotypically very similar.

3.6.6. *LaFT* Transformant Analysis

A total of two transformants which successfully expressed the *LaFT* and *NPT-II* genes, T1FTA2 and T1FTA8, were identified. Although over 8 transformants were selected from the antibiotic selection plates, only several transformants survived and/or had sufficient tissue for gene expression testing. Both T1FTA2 and T1FTA8 exhibited expression of *NPT-II* and *LaFT* in cDNA derived from the entire plant (Table 9). Transformants T1FTA2 and T1FTA8 both exhibited decreased vertical growth, reduced flower production, and abnormal carpels in some

areas of the plant when compared to wild type (Figure 26 a, b, & d). These transformants also appeared to have an overall reduction in leaf growth and flowered at an early age. In addition, transformants T1FTA1 and T1FTA5 exhibited a similar altered phenotype, although there was insufficient tissue to obtain mRNA for testing, as occurred with other *LaFT* F_1 seedlings (Figure 26). Both T1FTA1 and T1FTA5 exhibited significantly decreased vegetative growth at 1 month of age (Figure 26 c). Both T1FTA1 and T1FTA5 exhibited a significant reduction in height and leaf biomass and had begun flowering at 1 month of age.

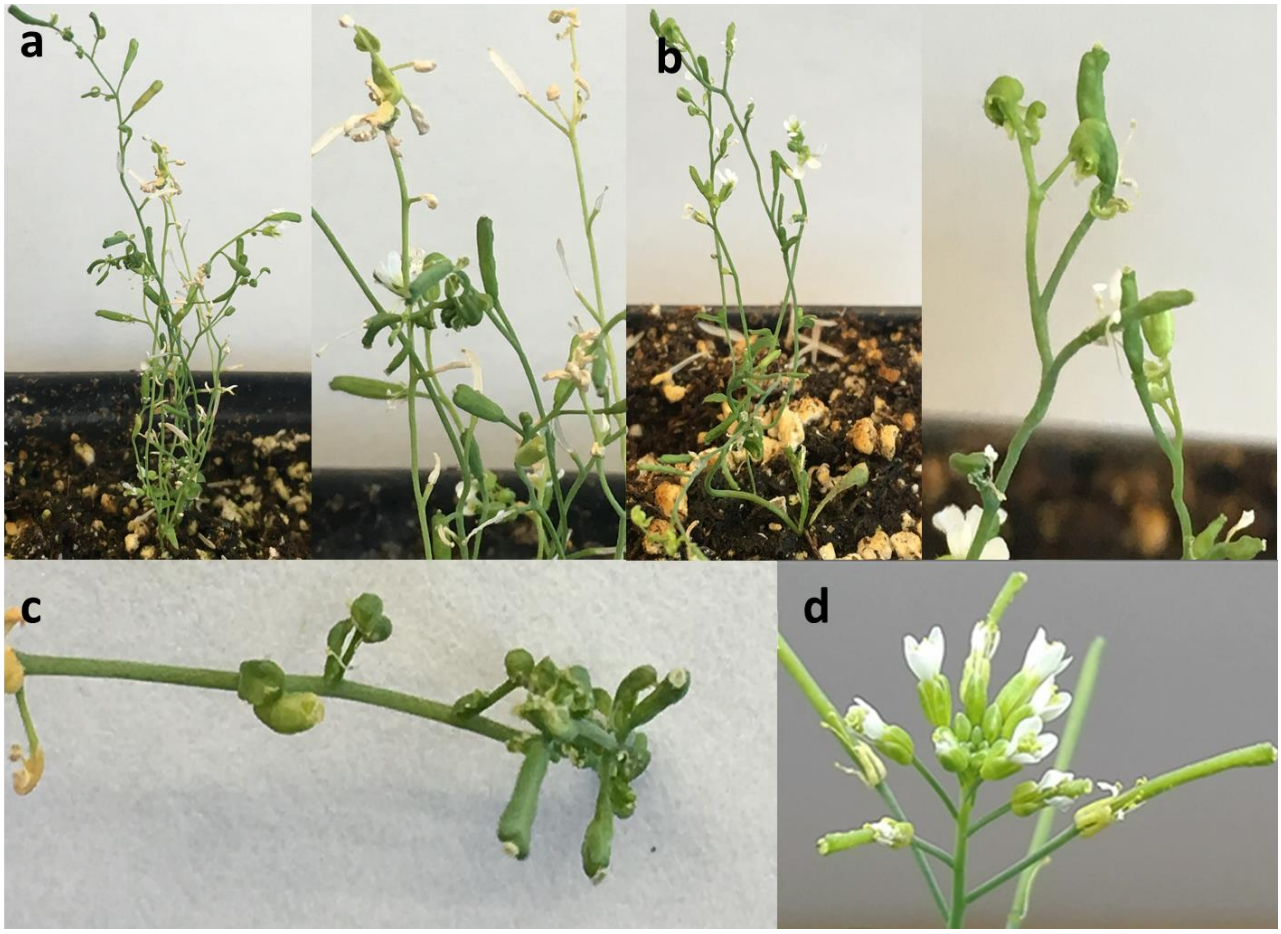

Figure 26. Phenotype of the *LaFT* transformants. (a) T1FTA2, (b) T1FTA8, (c) T1FTA5, (d) wild-type *A. thaliana*. *LaFT* transformants exhibited reduced height and decreased flower production and abnormal carpels in some areas of the plants when compared to wild-type *A. thaliana.* Both T1FTA2 and T1FTA8 exhibited expression of *NPT-II* and *LaFT* in cDNA derived from the entire plant. Transformant T1FTA5 exhibited a transformant phenotype but had insufficient tissue for genetic testing.

Chapter 4: Discussion

Lavenders are commercially and medicinally important plants that have been implicated for use in a variety of settings from medicinal use as an analgesic or anxiolytic to commercial use in perfumes and personal care products (Field et al. 2005; Smith et al. 2011; Soltani et al. 2013; Wells et al. 2018). Previous studies have examined the biosynthesis and biological activities of lavender EO (Landmann et al. 2007; Zuzarte et al. 2009; Woronuk et al. 2010; Demissie et al. 2011; Demissie et al. 2012; Sarker et al. 2012; Soltani et al. 2013; Lesage-Meessen et al. 2015). However, the molecular processes governing the initiation and development of flowers – the main source of essential oils in lavenders– have not been adequately investigated in this plant. Floral development has, however, been well-studied in *A. thaliana*, and many genes involved in floral initiation and development have been documented (Coen and Meyerowitz 1991; Theißen et al. 1996; Koornneef et al. 1998a; Levy and Dean 1998; Bowman and Eshed 2000; Hartmann et al. 2000; Samach et al. 2000a; Theißen 2001; Causier et al. 2010; Smaczniak et al. 2012; Andrés et al. 2015). The ABC(DE) model described in *A. thaliana* is the most well-defined model of floral development which describes the genes involved in flowering and their roles in floral pattern development (Weigel and Meyerowitz 1994; Causier et al. 2010). Additionally, many pathways governing floral initiation and repression of flowering in *A. thaliana* have been documented (Samach et al. 2000a; Hayama and Coupland 2003). Homologsof the *A. thaliana* floral development genes have been identified and studied in other plants but, to our knowledge, have yet to be formally identified in lavenders (Kang et al. 1998; Teeri et al. 2006; Bendahmane et al. 2013; Wu et al. 2017). Here, we proposed that the genes governing floral initiation and development in *A. thaliana* and other flowering plants are also present in the genus *Lavandula.* To bridge the gap in knowledge, this thesis employed a genomic approach to identify and functionally characterize genes that control flower initiation and pattern development in lavenders. We used RNA-Seq to obtain information on genes expressed during various stages of flower development in *L. angustifolia* and *L. x intermedia* plants. The information was used to develop a comprehensive *Lavandula* transcriptome within which *Lavandula* genes homologous to *A. thaliana* floral development genes were identified. A subset of 6 genes was successfully functionally characterized via expression in wild-type *A. thaliana*. The present study is the first of its kind in lavenders and has

thus made a significant scientific contribution and marks the start of new area of research in lavender development.

4.1 Transcriptome Assembly and Identification of *Lavandula* Homologs

4.1.1 Transcriptome Assembly, Annotation, and Gene Identification

To examine the genes expressed at various stages of flower development, total RNA was extracted from SAM, BD, and FL tissues of mature *L. x intermedia* and *L. angustifolia* plants and was sequenced by Illumina sequencing. The *L. angustifolia* FL transcriptome exhibited the most contigs at 47,630 contigs, while the *L. x intermedia* SAM gave rise to the least amount of contigs at 26,821 (Table 3). This is likely due to the vast number of genes which are active in the flower and the floral meristem, including not only developmental genes, but genes involved in terpenoid metabolism and biosynthesis in FL tissues which are inactive in SAM tissues (Klepikova et al. 2015; Adal et al. 2019; Malli et al. 2019). RNA-Seq was then used to compile this information and assemble it into a comprehensive transcriptome of 105, 294 transcripts from a total of 118,245,160 paired end reads, representative of genes expressed in the developing tissues. CD-HIT-EST was used to remove redundant sequences which had >98% sequence similarity, ensuring that the transcriptome was comprised of unique genes and minimal duplicate transcripts. All transcriptome parameters (including N_{50} and transcript length distribution) were within a reasonable range when compared to previously-reported transcriptomes assembled by similar means (Duan et al. 2012; Lulin et al. 2012; Wang et al. 2012; Xu et al. 2013; Lu et al. 2014; Adal et al. 2019). Thus, the present transcriptome can be considered a reliable genomic resource for genetic studies.

Using CLC Genomics Workbench 11, BLASTx analyses of the transcriptome against an annotated *A. thaliana* protein database assigned putative functions to the majority of transcripts by comparing the *Lavandula* sequences to the sequences of *A. thaliana* proteins with known functions. BLASTx analyses also revealed the presence of ten previously- described *Lavandula* terpene synthase genes within the transcriptome, which verified the comprehensiveness and quality of the transcriptome as these genes are vital for the synthesis of lavender EO constituents (Table 4) (Landmann et al. 2007; Demissie et al. 2011; Sarker et al. 2013; Jullien et al. 2014; Despinasse et al. 2017; Adal et al. 2019).

Annotation of the transcriptome via the BLAST2GO function of CLC Genomics Workbench 11 provided gene ontology terms for the majority of transcripts, classifying the transcripts based on cellular component, molecular function, or biological process, with few transcripts lacking annotation (Figure 5). This thorough classification of transcripts provided support for the quality and comprehensiveness of the transcriptome as the majority of transcripts could be assigned a function. The thorough gene classification was consistent with other similar RNA-Seq assembly classifications (Gao et al. 2013; Adal et al. 2019). Regarding biological process, the majority of the contigs were classified as being involved in cellular process, followed by metabolic process, response to stimulus, biological regulation and developmental process (Figure 5). These classification categories were consistent with the functions of the *A. thaliana* proteins to which the *Lavandula* transcripts matched via BLASTx analyses. As many of the identified genes within the transcriptome were development-related genes or transcription factors involved in floral initiation or pattern development, the classification of these genes functionally into response to stimulus (>32,500), biological regulation (>27,500) and developmental process (>25,000) is consistent with the corresponding *A. thaliana* genes' functions in the SAM and floral meristem. With regards to molecular function, the most common category was binding (>37,000), followed by catalytic activity (>30,000) and transporter activity (approximately 5000). The abundance of transcripts classified based on binding activity is consistent with the abundance of TFs identified within the transcriptome and with the function of many TFs in protein-protein interactions and in DNA-binding to regulate gene expression in the SAM during floral initiation (Yant et al. 2010; Smaczniak et al. 2012; Theißen et al. 2016; Dinh et al. 2017). The majority of contigs matching to cellular component were classified as part of the cell (>50,000), followed by organelle component (>42,500) and membrane component (>30,000) (Fig. 4). This result is consistent with the identification of genes within the transcriptome involved in floral initiation based on the sensing and integration of endogenous and exogenous cues via regulatory pathway at the SAM to regulate flowering timing (Koornneef et al. 1998a; Boss et al. 2004; Kim et al. 2009). Gene ontology studies thus provided insight into the functions of the transcripts within the transcriptome and showed the presence of a variety of genes involved in gene regulation during floral initiation. Future studies should examine gene ontology terms of specific tissues as well as at shorter time increments to gain a more comprehensive view of these genes over time.

The assembled transcriptome can thus be considered a comprehensive representation of the genes expressed in *Lavandula* floral tissues at various developmental stages. In lieu to this, the transcriptome assembled in this study is a valuable genomic resource that can be used in future studies for the identification of additional floral initiation and development-related genes as well as genes involved in other developmental processes in lavenders.

BLASTx analyses and Gene Ontology studies facilitated the identification of *Lavandula* candidate genes homologsto *A. thaliana* floral initiation and development genes within the transcriptome. The transcriptome "quality" was high as it contained full-length transcripts for numerous genes, including partial or full-length sequences for over 100 previously- described *A. thaliana* floral initiation and floral organ identity genes (representing all of the A, B, C, D, and E-type gene classes), and those for most of the previously reported lavender terpene synthases. Candidate genes identified within the transcriptome included homologs of previously- examined *A. thaliana* repressors and initiators of flowering (Table S1). Further, a total of 14 full-length lavender genes homologous to *A. thaliana* floral development and initiation genes were identified within the transcriptome (Table 6). From the 14 full length genes, a subset of 6 candidate genes, including *LaCC*, *LaCAL*, *LaSVP*, *LaPI*, *LaFT*, *LaAP1*, and *LiAP2*, were selected for functional characterization in *A. thaliana*. These genes were selected based on high sequence homology, full length sequence and given that these genes were homologous to genes which play vital roles in floral development in *A. thaliana*. For example, *LaCC* is the homolog of *AtCC*, which functions in the regulation of carpel development in *A. thaliana*, whereas *LaAP1*, the homolog of *AtAP1*, is vital to sepal and petal development (Bowman and Smyth 1999; Theißen et al. 2016). The presence of these genes in lavenders supported the hypothesis that there exist genes in lavenders which are homologous to genes which govern floral initiation and development in *A. thaliana.*

4.1.2 Sequence Analysis

MUSCLE sequence alignment of the full-length *Lavandula* protein sequences and their respective *A. thaliana* homolog's protein sequence revealed sequence similarities and differences between the pairs (Tables 7- 8, Figures 8 & B1-B13). MUSCLE alignment of the homologous pairs revealed the presence of conserved sequences respective to the TF class in each of the protein pairs. *Lavandula* protein sequences were similar to their respective *A. thaliana*

homologous both in amino acid sequence and protein length (Table 6). For example, the repressor protein, LaFLC, was determined to be 190 amino acids in length, while its homolog, AtFLC, is 196 amino acids in length (Table 6). This trend of similarity was observed between all homolog pairs and was an expected result as the homologous pairs contain conserved domains and are hypothesized to serve similar functions. Protein sequence similarity between the homologs ranged from 41.53-78.16%, which indicated the presence of conserved domains within the proteins, but differences in protein sequences due to being derived from different species (Table 7, Figures 8 & B1-B13). These results were consistent with homologs of the *A. thaliana* floral development genes described in other plants, such as homologs of *PI* and *AG* described in *Oryza sativa* (rice) and *Lilium longiflorum* (lily*)*, respectively, which also exhibited conserved domains, but exhibited different amino acids and sequence length (Kang et al. 1998; Kyozuka et al. 2000).

Many of the TFs identified in the present study, including LaAG, LaSVP, LaCAL, LaFLC, LaPI, LaAP1, LiAP3, LaSEP3, LaSOC1, are homologs of *A. thaliana* MADS-Box TFS, which play important roles in floral development. Alignment of these proteins with their respective homologs revealed the presence of the MADS-Box and K-Box domains characteristic to these TFs (Figures 8, B1-B13) (Shore and Sharrocks 1995; Gramzow et al. 2010).

For example, alignment of genes with their *A. thaliana* homologs and other plant MADS-Box transcription factors revealed the presence of the conserved MADS-Box and K-box domains of the MADS-Box TFs in several of the *Lavandula* proteins (Figure B.1-B.12). MADS-Box genes are derived from common ancestral genes which underwent gene duplication events that caused divergence, giving rise to the plant and animal MADS-Box TFs observed today (Alvarez-Buylla et al. 2000). These gene duplication events in ancient MADS-Box genes also led to the diversification of functions of the MADS-Box genes, including the ABCDE-type TFs (Kramer and Hall 2005). It is proposed that the differences observed between the MADS-Box proteins are due to gene duplications, divergence, and changes in regulatory sequences (Force et al. 1999; Kramer and Hall 2005). Thus, although high sequence similarity between homologous MADS-Box genes is observed, sequence differences are expected between species and the MADS-Box genes identified in this study likely exhibit sequence differences outside the conserved domains for these reasons. In addition, the presence of these conserved domains provides evidence for a conserved function of these lavender proteins as MADS-Box TFs regulating the expression of

other genes involved in floral initiation and pattern development (Theißen 2001; Pařenicová et al. 2003; Kaufmann et al. 2005). Further, results support the hypothesis that the lavender proteins function similarly as their *A. thaliana* homologs Future studies should examine the 3D protein structure of these TFs and identify specific domains and interactions with other proteins.

The identification of homologous genes supported the hypothesis that there exist lavender genes which are homologous to the floral initiation and pattern development genes present in *A. thaliana* and other plants, thus providing support for the function of the ABC(DE) model in *Lavandula* species. The ABC(DE) model of floral development is the most predominant and widely accepted model of floral pattern development in *A. thaliana* and other higher plants (Coen and Meyerowitz 1991; Kang et al. 1998; Levy and Dean 1998; Causier et al. 2010; Theißen et al. 2016). To date, numerous studies have identified homologs to these *A. thaliana* floral development genes in a variety of other flowering plants, including rice, wheat, and soybeans, amongst others, providing support for this model across many plants(Coen and Meyerowitz 1991; Kaufmann et al. 2005; Shitsukawa et al. 2006; Huang et al. 2014; Irish 2017; Wu et al. 2017). The identification of lavender homologs of the previously-described *A. thaliana* floral initiation and development genes, including the A, B, C, D, and E-class genes, provides support for the hypothesis that floral development in lavenders is similar to that in *A. thaliana* and other flowering plants. Further, as the ABC(DE) model is well-described in *A. thaliana*, homologs identified within the *Lavandula* transcriptome can be assigned putative functions based on the function of the respective *A. thaliana* homologs, thus providing insight into the genes involved in floral development in lavenders and an abundance of vital developmental genes to be examined.

4.1.3 Expression Analysis of *Lavandula* Transcripts

To examine expression patterns of floral initiation and development genes during various stages of flower development in *L. angustifolia* and *L. x intermedia* FL, BD, and SAM tissues, the Differential Expression for RNA-Seq function of CLC Genomics was employed. Differential expression analysis revealed that many of the identified genes within the transcriptome exhibited varying levels of expression across different developmental stages (Figure 6). Some genes exhibited higher expression in the SAM than the FL or BD tissues, whereas other genes exhibited higher expression in the SAM in comparison to FL or BD tissues. Differential

expression across developmental stages was due to different genes being active at a given time based on their role. For example, genes which are involved in floral initiation, are more likely to be expressed in the bud or shoot apical meristem, whereas floral patterning and flower maturity genes are more likely to be active at later stages (Figure 6). The results were thus consistent with the diversity of the developmental roles of the identified genes and were consistent with other studies examining expression patterns of floral development related genes over time (Kram et al. 2009; Yant et al. 2010; Klepikova et al. 2015). For example, the gene for the repressor of flowering, *FLC*, was expressed in low amounts in the LaFL in comparison to the LaBD, which was consistent with studies that found levels of expression of this gene to decrease throughout floral initiation (Klepikova et al. 2015). Similarly, results showed that *LaSVP*, the *Lavandula* homolog of the *A. thaliana* floral repressor, *AtSVP*, exhibited higher expression in the *La*SAM in comparison to *La*FL tissues, lower expression in *La*SAM versus *Li*SAM, lower expression in *La*BD versus *Li*SAM, and similar expression in *La*FL and *La*BD tissues. This result is consistent with the role of *AtSVP* as a floral repressor previously-described to be active in vegetative tissues, but absent in floral tissues (Hartmann et al. 2000). Differential expression studies thus provided support for the differing roles of the identified *Lavandula* floral development genes and insight into their expression patterns at various stages of development.

To verify expression of the selected candidates in *L. angustifolia* and to facilitate cloning in the pGA482 vector, all seven candidates were PCR amplified from *L. angustifolia* BD cDNA using gene-specific primers. All genes were successfully amplified from *L. angustifolia* BD cDNA, indicating that all selected candidate genes were actively expressed in *L. angustifolia* BD tissue (Figure 9).

4.2 Cloning and Selection of Candidates

4.2.1 Cloning in the pGA482 Vector and Bacterial Selection

Due to the high protein sequence similarity between the *Lavandula* and *A. thaliana* homolog, it was proposed that the lavender genes would likely function similarly to their *A. thaliana* counterpart and that floral development in lavenders is governed by similar processes as in *A. thaliana.* To explore this hypothesis and as a proof of concept, 7 *Lavandula* candidate genes were functionally characterized via cloning and overexpression in wild-type *A. thaliana* driven by the CaMV 35S promoter of the pGA482 vector. Functional characterization of the

Lavandula candidate genes, *LaCC, LaCAL, LaSVP, LaPI, LaFT,* and *LaAP1*, was facilitated through cloning and expression via the pGA482 vector. Each gene was successfully ligated into a pGA482 vector via sticky end cloning, giving the constructs pGA::*LaCC* , pGA::*LaCAL* , pGA::*LaSVP* , pGA::*LaPI* , pGA::*LaFT*, and pGA::*LaAP1*, which were transformed into *E.coli* via heat-shock. For each construct, both the 50 µl plate and the concentrated plate exhibited colony growth, although in all cases, the concentrated plate exhibited more growth. Positive colony growth in the presence of tetracycline indicated that the cells giving rise to the colonies harbored the pGA482 vector containing the tetracycline resistance gene and the respective *Lavandula* gene insert. Control plates containing *E. coli* transformed with an empty pGA482 exhibited growth on the concentrated plates, confirming that the tetracycline resistance gene within the pGA482 vector was functional (Figure 10). All pGA482::lavender gene constructs were successfully transformed into *A. tumefaciens*.

4.2.2 Plant Transformation and Selection

To functionally characterize the lavender candidate genes, wild-type *A. thaliana* var. *Landsberg erecta* were transformed with the pGA482::lavender gene constructs via the floral dip method (Zhang et al. 2006b). Agrobacterium-mediated plant transformation via the floral dip is a well-established protocol which has proven to be an effective means for transformation of *A. thaliana* in a variety of studies, as well as in other plants, such as Finger millet and Sorghum (Clough and Bent 1998; Zhang et al. 2006b; Sharma et al. 2011; De Pater et al. 2013; Huang et al. 2014; Che et al. 2018). Expression of these genes in *A. thaliana* enabled the study of the functions of these genes in an effective model plant.

Four wild-type *A. thaliana* plants were transformed for each construct, for a total of 24 transformed plants. All transformed plants produced viable seeds which were grown on antibiotic selection medium to identify positive transformants. This was an effective means of selection as only seeds which had successfully taken up the pGA482::lavender gene construct harboring the kanamycin resistance gene were able to survive in the presence of antibiotic. Successful selection of transformant seedlings was observed on plates for all six genes, indicating that transformation had been successful. Initially, selection took place on plates containing either 50 mg/mL kanamycin and 75 mg/mL kanamycin, but as selection was more efficient on 75 mg/mL kanamycin, all subsequent plates had 75 mg/mL kanamycin.

Transformant seedlings able to grow in the presence of kanamycin had significant phenotypic differences from those which could not. Surviving seedlings exhibited green leaves and an expansive, branching root system within the medium in comparison to wild- type, which appeared yellow and failed to develop roots (Figure 14). Some plates exhibited only seedlings which died soon after germination, indicating no positive transformants. Conversely, wild-type seeds grew well on medium containing no antibiotic, thus serving as an effective control. Most seeds were able to germinate but died soon after (Figure 14). This indicated that seeds could germinate in the presence of antibiotic and use the reserves within the seed, but once the seed began drawing nutrients from the medium, the antibiotic killed the seeds lacking the *NPT-II* gene. Successful germination of the seeds indicated that transformation had been successful and that F_1 seeds were viable.

Seedlings which grew successfully in the presence of kanamycin and which were presumed to be kanamycin resistant were transferred to soil for growth and genetic analysis. For each gene, there were 5-15 seedlings transferred from plates to soil, with over 40 seedlings selected from the antibiotic plates. This result indicated that transformation success varied widely across the plants. This result is consistent with reports that transformation efficiency can be highly variable and largely dependent on the number of unopened floral buds at the time of transformation of the parental plants (Ghedira et al. 2013). Further, although numerous seedlings were transferred for each gene, some failed to survive. For example, over 14 *LaAP1* seedlings were transferred from plates to soil, but few survived and/ or had sufficient tissue for testing (Table 9). Although transformation protocols were closely followed, the number of positive transformants was lower than the expected 1/100 seeds predicted by Zheng et al. (2006). This result may also be due to the use of the *A. thaliana* ecotype used in the present study as it has been found that *A. thaliana* var. *Landsberg erecta* exhibits lower transformation efficiency compared to other ecotypes (Zhang et al. 2006b). Future studies should examine the transformation efficiencies of other ecotypes and the effects of transformation of lavender genes into other ecotypes.

4.2.3 Expression Analysis of F_1 Transformants

Tissue was collected from presumed kanamycin-resistant seedlings (Table 9). The *NPT-II* gene was PCR amplified from gDNA extracted from transformant tissue. PCR amplification of

the *NPT-II* gene from gDNA and/or cDNA of transformants indicated that the kanamycin resistance gene was present within the transformant's DNA and transformation had been successful. As there was a possibility that an empty vector could have been transferred to a plant, the amplification of each respective lavender candidate gene was vital.

Amplification of the *Lavandula* genes from floral cDNA confirmed that the gene was not only present, but also was actively being expressed as mRNA. Thus, amplification of a given gene from gDNA confirmed the presence of the gene but would did not confirm expression of the gene. In some cases, such as with T2AP1A1, there was insufficient tissue for full expression analysis, but amplification of the *LaAP1* gene from the cDNA provided support that the lavender gene was present and being expressed in the transformant. From this result, it could be inferred that the *NPT-II* gene was also present as it was transferred with the lavender gene in the vector. There were, however, some seedlings which were able to grow in the presence of kanamycin, but amplification of *NPT-II* and the respective lavender gene was unsuccessful. This result may have been observed due to uneven mixing of medium in some cases causing pockets lacking antibiotic within the medium, insufficient RNA quantity, or low expression of these genes in these plants below a level which was detectable. Wild-type *A. thaliana* plants also exhibited a low level of expression of *LaSVP* and *LaAP1* as well as *LaSVP* and *LaFT* in the leaf cDNA and flower cDNA, respectively (Table 9). This result was likely due to the high sequence similarity between the lavender and *A. thaliana* gene homolog pairs in the 5' and 3' regions to which the primers were designed (Figures 8, B4 & B8).

Similarly, *LiAP2*, although originating from *L. x intermedia*, was successfully amplified from *L. angustifolia* FL BD tissue due to high sequence homology in the 5' and 3' regions of the *L. x intermedia* and *L. angustifolia AP2* genes.

Although positive transformants were identified for all genes, identifying additional replicates in future studies would be beneficial. In the present study there were six genes that were characterized, and 24 plants transformed and, as such, many plates were prepared to screen seeds for transformants. Futures studies should focus on 1-3 genes at a time to allow for the more rapid identification of transformant seeds and increased replicates per gene.

The successful PCR amplification of the lavender genes from cDNA and gDNA of F_1 seedlings indicated that these genes were successfully transformed into these plants. Further, it showed that these genes could be successfully expressed in *A. thaliana* via the 35S CaMV

promoter. This method can thus be used in future studies to characterize additional lavender developmental genes in *A. thaliana*.

Future studies should also examine the effects of the manipulation of these genes in lavenders via transformation. For example, it would be valuable to examine the effects of overexpression of certain genes responsible for petal development or to examine the effects of silencing the expression of floral repressors. Successful *Agrobacterium*-mediated transformation of lavenders using lavender explants hasbeen reported; however, the long generation time of lavenders poses limitations on the ability to examine the effects of transformation in short amounts of time (Nebauer et al. 2000; Mishiba et al. 2015). Thus, improving upon current lavender transformation procedures would be beneficial in many areas of study, including the study of floral development. Currently in the Mahmoud lab, such transformation procedures are being examined and optimized.

4.3 Functional Characterization of *Lavandula* Candidates

A total of 24 F_1 transformant plants were identified as indicated by growth on antibiotic selection medium and PCR amplification of *NPT-II* and the respective lavender gene. Although all six candidate genes were successfully expressed in wild-type *A. thaliana* plants, the phenotypic effects of expression of each gene varied. For example, the expression of the MADS-Box proteins LaAP1, LaCAL and LaPI did not elicit significant phenotypic effects (Figures 21 & 22). *LaAP1* and *LaPI* are homologs of *AtAP1* and *AtPI*, which function as MADS-Box TFs during floral development in *A. thaliana* (Theißen et al. 2016). In *A. thaliana*, these proteins function as quartets, thus binding to three other MADS-Box proteins and binding to the promoter region of a given gene (Figure 1) (Theißen et al. 2016). Although *LaPI* and *LaAP1* were expressed in several F_1 transformants, expression of these genes did not significantly impact the phenotype of these transformants. This result could be due to the nature by which these proteins function in quartets as both *LaPI* and *LaAP1* elicit their effects on gene expression through interactions with other proteins (Theißen et al. 2016). These proteins thus rely on the ability to bind to and interact with other proteins to form a quartet, which regulates gene expression through promoter binding (Theißen et al. 2016). Each TF must thus be able to productively bind to other proteins in the quartet. For example, AtPI is involved in the regulation of petal and stamen development in *A. thaliana* (Theißen et al. 2016). When one AtPI protein combines with

one AtAP3 protein, one AtSEP protein, and one AtAP1 protein, a quartet is formed that leads to petal development in whorl 2 of the *A. thaliana* flower (Figure 1) (Theißen et al. 2016). In contrast, when one AtPI protein combines with one AtAP3 protein, one AtSEP protein, and one AtAG protein, the quartet binds to the promoter of genes that promote stamen development in whorl 3 of the *A. thaliana* (Figure 1) (Theißen et al. 2016). Thus, the ability for these proteins to effectively bind one another and a specific gene's promoter is vital in order to elicit the desired effect on gene expression in the appropriate whorl.

In the current study, expression of the *Lavandula* genes via the constitutive CaMV 35S promoter led to production of many copies of each protein in each *A. thaliana* transformant. It was hypothesized that an increase in concentration of the proteins via expression of the *Lavandula* orthologs in *A. thaliana* would lead to an increase in the processes these TFs govern.

In the present study, F_1 transformants of three of the *Lavandula* proteins, LaCC, LaSVP, and LaFT exhibited phenotypic differences from wild-type plants (Figures 18-20 & 24). In contrast, transformants expressing the MADS-Box proteins LaPI, LaAP1, and LaCAL exhibited no significant phenotypic differences in comparison to wild-type (Bowman et al. 1993; Bowman and Smyth 1999; Ferrándiz et al. 2000; Kaufmann et al. 2005) (Figures 21-23). It was hypothesized that an increase in expression of a given protein would lead to an increase in the activity that protein performs; however, this was not observed with LaPI, LaAP1, and LaCAL transformants. Thus, for these proteins an increase in protein concentration may not be associated with increase in activity. This result also may have been due to limited availability of gene promoters for the additional proteins to bind. In addition, TFs function through protein-protein interactions, so an increased concentration of one protein may not increase a given effect if the protein must interact with other proteins to elicit this effect. Additionally, the activity of the *Lavandula* proteins may have been limited by an incompatibility and inability to bind to available *A. thaliana* proteins and promoters of *A. thaliana* genes due to differences in the a.a. sequence between the homologou proteins. These a.a. differences may have led to incompatibilities in binding domains such that they were not sufficiently similar that they could interact with each other and form the quartets. In addition, although the concentration of LaPI, and LaAP1 proteins increased, the proteins' effects on downstream genes did not increase as the other proteins of the quartet did not increase in concentration. In this case, there could have been a saturation effect in which the number of PI proteins increased as both AtPI and LaPI were

present, but the number of the other required proteins remained constant. Thus, although there were many copies of the PI protein, their effect was limited due to a limiting number of the other members of the TF quartet and a finite number of proteins to form the quartet (Deitchman et al. 2018). Further, it has been shown that *A. thaliana* plants which overexpress multiple MADS-Box gene exhibit a phenotypic effect which is not observed in plants in which only one MADS-Box gene is overexpressed (Krizek and Meyerowitz 1996; Honma and Goto 2001). The other proteins required for the quartet and the binding sites of the *Lavandula* proteins may be different than those of the *A. thaliana* proteins due to amino acid differences. Such differences may have led to incompatibilities in binding domains of the proteins, and thus saturation effect (Deitchman et al. 2018). Thus, an increase in concentration of some of the *A. thaliana* proteins via expression of the *Lavandula* homolog, did not have an impact in all cases. Future studies should examine the effects of overexpression of multiple *Lavandula* MADS-Box homologs to examine the effects of the increase in certain MADS-Box complexes, as examined in *A. thaliana* (Honma and Goto 2001).

4.3.1 *LaSVP*

4.3.1.1. Phylogenetic analysis of *LaSVP* and other plant homologs

To examine the relationship of the *LaSVP* gene with other plant *SVP*-like genes, a phylogenetic tree was constructed using the protein sequences of plant SVP genes. Although the amino acid sequence of *LaSVP* showed higher sequence similarity to *EgSVP-like*, *BrSVP*, *BiSVP*, and *PkSVP* homologs *AcSVP2*, *AcSVP3*, and *AdSVP3* homologs compared to *AtSVP*, all of these genes contained the K-Box and MADS-Box conserved domains characteristic of MADS-Box transcription factors, suggesting that they share a similar DNA-binding function as MADS-Box transcription factors (Figure 20). The MADS domain of these TFs is highly conserved, facilitating DNA binding and the K-box is a known conserved region in plant MADS-Box proteins found downstream of the MADS domain, and is an important region facilitating protein-protein interactions in these TFs (Alvarez-Buylla et al. 2000). The presence of these conserved domains in the LaSVP protein provides support for *LaSVP* functioning as a MADS-box TF in lavenders, homologous to other plant *SVP* genes. Interestingly, *LaSVP* was found to cluster closely with the *SVP* homologs, *EgSVP-like* and *BrSVP*, but was further away in relatedness to *AtSVP* (Fig. S2 & Table 2). This result suggests that the processes of floral

initiation and development in lavenders may differ from those described in *A. thaliana*, although this postulate requires additional investigation (Wells et al. 2020).

4.3.1.2. Ectopic Expression of *LaSVP* Alters Floral Morphology and Flowering Time

The role of *SVP* as a repressor in the suppression of the transition from vegetative growth to flowering, promoting vegetative growth and altering floral morphology has been widely studied in *A. thaliana* (Hartmann et al. 2000; Gregis et al. 2009). Similar *AtSVP* homologous genes were isolated from different plant species including *Poncirus trifolate*, *Eucalyptus grandis*, *Actinidia deliciosa*, *Chrysanthemum morifolium*, and *Brassica campestris*. The ectopic expression of these genes in *A. thaliana* was observed to delay flowering time and negatively affected petal development (Brill and Watson 2004; Lee et al. 2007; Li et al. 2010; Wu et al. 2012; Gao et al. 2017).

To examine the role of the *LaSVP* gene in floral initiation and development in lavenders, the gene was expressed in *A. thaliana* under the control of the constitutive ectopic CaMV 35S promoter (Wells et al. 2020). Although flower development was drastically altered in all transgenic plants, there was an apparent correlation between *LaSVP* expression level and flower morphology (Wells et al. 2020). To confirm *LaSVP* expression in transgenic plants, a fragment of the *LaSVP* coding sequence was amplified from flowers of three representative transgenic plants by Reverse Transcriptase PCR (RT-PCR). The results demonstrated that the *LaSVP* gene was expressed in all transgenic plants (Figure 20). RT-PCR results showed higher expression of *LaSVP* in SVP-2 than in other transgenic lines). The *Actin* reference gene was amplified in all samples. Further, SVP2, T1SVPB4, and T1SVPB5, which expressed the transgene stronger than the other transformants, exhibited a phenotype most drastically different than the wild-type plants, while SVP-3, SVP-4, and T1SVPB1 exhibited a milder altered phenotype in some portions of the plant (Figure 18 & 19). For example, some portions of the T1SVPB1 plant exhibited normal flowers which gave rise to normal seeds pods and produced viable seeds. In contrast, other portions of the plant exhibited the transformant phenotype, lacking petals, exhibiting additional sepals, and having wrinkled carpels failing to produce viable seeds. Thus, when affected, the flowers of the transgenic plants lacked petals, and exhibited the formation of additional sepals. When the *LaSVP* expression increased, the flowers lacked petals, exhibited additional sepals, and the carpels were vastly altered appearing wrinkled, wide, and failing to produce viable seeds. These results imply that the effects

of *LaSVP* on floral development may be dosage- dependent as increased expression of *LaSVP* coincided with a more drastic change in phenotype (Wells et al. 2020). This is consistent with studies examining the overexpression of *AtSVP*, which have shown that *AtSVP* acts as a repressor of flowering in *A. thaliana* in a dosage-dependent manner (Hartmann et al. 2000; Jaudal et al. 2014). Other studies in *A. thaliana* also indicate that altered expression of class B and C genes in the floral meristem results in homeotic conversion in floral whorls (Weigel and Meyerowitz 1994; Gregis 2006; Gregis et al. 2009; Theißen et al. 2016). However, studies on *SVP* mutants show that the absence of the gene does not lead to obvious morphological differences nor changes to fertility (Hartmann et al. 2000). Given this observation, the absence of expression of the *SVP* gene may not impact on fertility in lavenders, but the overexpression of the *LaSVP* gene may lead to sterility in these plants due to the absence of viable seeds. Additionally, the effects of *LaSVP* on floral organ structure are similar to the findings of the functions of *AtSVP* in *A. thaliana* in the control of the floral meristem and repression of B, C and E-class floral homeotic genes whose overexpression leads to morphological changes in the flower (Liu et al. 2007; Li et al. 2008; Gregis et al. 2009). In the present study, the floral morphology differed from wild-typ *A. thaliana* and the B, C, and E-type genes appeared to be affected; however, there was no observable effects on C-type genes as stames developed in the transformant flowers. The B, C, and E class floral homeotic genes play important roles in the organizational development of flowers, and are required for the normal development of all five floral organs (Gregis et al. 2009; Wells et al. 2020).

In addition to changes in floral phenotype, flowering time was also affected by the expression of *LaSVP* in *A. thaliana* plants (Figure 17). The transgenic plants exhibited a late-flowering phenotype, compared to the wild-type counterpart. This outcome is not surprising as it has been shown that *AtSVP* acts as a repressor of floral initiation in *A. thaliana*, preventing the transition from a vegetative meristem to a flowering meristem (Gregis et al. 2013; Andrés et al. 2014). A similar result was obtained when the *Chrysanthemum morfolium SVP (CmSVP)* gene was expressed in *A. thaliana*, causing a significant delay in the onset of flowering (Gao et al. 2017). Thus, the expression of *LaSVP* in *A. thaliana* causing a delay in flowering suggests that this gene functions similarly to other plant *SVP* homologs.

4.3.2 *LaCC*

The *AtCC* gene is an important gene involved in carpel development in *A. thaliana* and is involved in interactions with numerous proteins to regulate aspects of this process (Gross et al. 2018). The *AtCC* gene encodes a YABBY domain TF which acts as both an activator and a repressor (Gross et al. 2018). The LaCC TF acts as a repressor of genes which promote the floral meristem to be terminated, but acts as an activator of genes which promote development of the carpel (Gross et al. 2018). To examine the role of *LaCC* in lavenders, the gene was expressed in *A. thaliana* under the control of the constitutive ectopic CaMV 35S promoter. Although several *LaCC* F_1 seedlings were identified on the antibiotic selection medium, only one plant that successfully expressed the *LaCC* gene was identified (Table 9). In several cases, the plants had insufficient tissue to extract sufficient mRNA or DNA and, in some cases, plants failed to survive. T2CCB1 (Figure 21) was the only plant which was confirmed to express *LaCC.* This plant exhibited curled leaves, and severely stunted growth. In addition, the morphology of the flower pattern differed from wild type in that the flowers were more dispersed along the stem rather than concentrated at the tip of the stem. This phenotype was also observed in T2CCB2 and T1CCA1, although the *LaCC* gene failed to be amplified from these plants.

The *LaCC* gene is homologous to the YABBY TF *AtCC* (Gross et al. 2018). Previous studies have identified similar YABBY homologs in other plants, such as in maize and soybeans, which are involved in flowering processes (Strable et al. 2017; Yang et al. 2019). For example, the soybean YABBY TF, GmFILa, was altered leaf polarity and growth and increased the flowering period length when overexpressed in *A. thaliana* (Yang et al. 2019). Future studies should further examine the effects of overexpression of *LaCC* in *A. thaliana* as well as the effects of overexpression of *AtCC* and *LaCC* in lavenders.

4.3.3 *LaPI*

The *LaPI* gene is homologous to the *A. thaliana PI* gene, a B-class MADS-Box TF that is vital for petal and stamen development in these plants (Figure 1) (Theißen et al. 2016). According to the ABC(DE) model, the *A. thaliana* PI protein forms a quartet with other MADS-Box proteins to elicit its function by combining with one AP3 protein, one SEP protein, and one AP1 protein to promote petal development in whorl 2 of the *A. thaliana* flower (Theißen et al. 2016). In contrast, when the PI protein combines with one AP3, one SEP, and one AG protein,

stamen development occurs in the third floral whorl (Figure 1) (Theißen et al. 2016). Mutations in this homeotic gene have been observed to cause sepals and carpels to develop in place of petals and stamens, respectively (Goto and Meyerowitz 1994). Homologs of the *PI* gene have been described in other flowering plants such as lily (*Lilium longoflorum*), pea (*Pisum sativum*), apple, *Magnolia wufengensis*, snapdragon (*Antirrhinum majus*), rice, and in the Ranunculaceae family (Kang et al. 1998; Kramer et al. 2003; Berbel et al. 2005; Tanaka et al. 2007; Wu et al. 2010; Liu et al. 2018a).

To examine the role of *LaPI* in lavenders, the pGA482::*LaPI* construct was transformed into wild-type *A. thaliana*. A total of 8 F_1 seedlings were identified which expressed both *NPT-II* and *LaPI* (Table 9). Most transformants exhibited no phenotypic differences from wild-type *A. thaliana,* although some transformants appeared to exhibit a slight increase in stem and leaf mass (Figure 21). One transformant, T2PIA2 exhibited abnormal carpels and siliques in some areas of the plant but remained the same as wild type in other parts of the plant (Figure 22 a). This phenotypic difference could be due to growth conditions or mutations within the plant unrelated to the expression of *LaPI*, as this was not observed in other transformants.

It is likely that no significant phenotypic effects were observed due to the way in which the LaPI protein likely functions in a quartet as the AtPI protein does (Theißen et al. 2016). Studies have indicated that overexpression of one MADS-Box TF is insufficient to elicit a phenotypic change (Krizek and Meyerowitz 1996; Honma and Goto 2001). For example, *A. thaliana* overexpressing both AP3 and PI exhibited increased petal and stamen growth, whereas overexpression of one gene did not have a significant effect (Krizek and Meyerowitz 1996).In addition, within the quartet in *A. thaliana*, the AtPI protein forms a heterodimer with the AtAP3 protein via the proteins' K-domains (Yang et al. 2003). In lieu of this, it is also possible that due to amino acid sequence differences in the K-domain of the LaPI protein and the AtPI protein, the LaPI proteins were unable to bind to the AtAP3 proteins to elicit an effect (Figure B.7). Future studies should examine the effects of knocking out the *LaPI* gene as well as the effects of concurrent overexpression of *LaPI* and the other MADS-Box TFs it interacts with in *A. thaliana* and in lavenders to fully elucidate the role this gene plays in lavenders.

4.3.4 *LaAP1*

In *A. thaliana*, the *AP1* gene encodes a MADS-Box TF involved in the development of sepals in the outermost whorl of the *A. thaliana* flower (Theißen et al. 2016). Two AP1 proteins combine with two SEP proteins to form a quartet which binds to the promoter region of genes which promote sepal development in the first floral whorl (Theißen et al. 2016) (Figure 1). In addition, the AP1 protein combines with AP3, PI, and SEP in the second floral whorl to give rise to petals (Theißen et al. 2016). Genes similar to the *A. thaliana AP1* gene have been described in other flowering plants such as aspen trees, grass (*Lolium temulentum*), poppy (*Papaver somniferum* and *Eschscholzia californica*), and bread wheat (*Triticum aestivum*) (Gocal et al. 2001a; Murai et al. 2003; Pabón-Mora et al. 2012; Azeez et al. 2014). Homologs of the *A. thaliana AP1* gene have shown differing functions from roles in floral organ identity to effects on flowering time and growth (Gocal et al. 2001a; Murai et al. 2003; Pabón-Mora et al. 2012; Azeez et al. 2014). It was hypothesized that the lavender homolog, *LaAP1*, would function similarly to previously described genes and that its overexpression would impact these processes. Further, it was hypothesized that an increase in expression of *PI* in wild-type *A. thaliana* facilitated by expression of *LaAP1* would enhance the effects of this gene via additional copies of the protein being present and functioning. To test this hypothesis, wild-type *A. thaliana* were transformed with the pGA::*LaAP1* construct to constitutively and ectopically express the *LaAP1* gene. The pGA::*LaAP1* construct was transformed into wild-type *A. thaliana*. A total of 4 F_1 seedlings were identified which expressed the *LaAP1* gene (Table 9). It was hypothesized that an increase in expression of the *AP1* gene woud result in phenotypic changes; however, this was not observed (Figure 23 & 24). Interestingly, it has been found that overexpression of *A. thaliana* AP1 leads to a decrease in time to flowering in tomato and citrus. Further, it was reported that overexpression of the *MdMADS5* gene found in apple, a gene similar to AP1, led to phenotypic changes and early flowering (Kotoda et al. 2002). This was not observed in the present study as there was no significant phenotypic difference between *LaAP1* transformants and wild-type *A. thaliana* (Figure 23 & 24). Although there appeared to be a slight increase in floral bud mass in some transformants, such as in T2AP1A4, this increase was not significant or consistent between transformants. In addition, although many *LaAP1* transformants flowered at a young age, this was not consistent between all transformants it is unclear if this effect was due to light conditions in the growth chamber. It is possible that there were no significant phenotypic effects due to

incompatibility of the binding domains of the LaAP1 protein and the proteins which must interact with or due to saturation effects described previously. Further, the *LaAP1* gene could function differently than that of *A. thaliana AP1*. Future studies should examine in-depth the effects of overexpression of *LaAP1* in a large-scale transformation with only *LaAP1* to determine if *LaAP1* significantly impacts flowering time in these plants.

4.3.5 *LaCAL*

In *A.* thaliana, the *CAL* gene is thought to interact similarly to *AP1* and is involved in the enhancing *LFY* gene expression, promoting the initiation of flowering and the floral meristem identity (Bowman et al. 1993; Ferrándiz et al. 2000). The *CAL* gene encodes a MADS-Box TF and functions similarly to *AP1* and plays a significant role in the floral transition (Bowman et al. 1993; Ferrándiz et al. 2000; Alvarez-Buylla et al. 2006; Kempin et al. 2017). Further it has been shown that the sequences of *AP1* and *CAL* are similar and the genes are likely products of a gene duplication event as the genes also serve mostly redundant functions during floral initiation (Yanofsky 1995; Alvarez-Buylla et al. 2006). Interestingly, it has been shown that *CAL* cannot function in place of *AP1* in *A. thaliana*, but, *AP1* can function in place of *CAL* (Alvarez-Buylla et al. 2006). In addition, mutations in *AP1* and *CAL* in combination with other MADS-Box genes have been shown to prevent flowering in *A. thaliana Landsberg erecta* (Ferrándiz et al. 2000). It was hypothesized that *LaCAL* would function similarly to *AtCAL*. Further, it was hypothesized that in increase in expression of *CAL* in wild-type *A. thaliana* facilitated by expression of *LaCAL* would enhance the effects of this gene via additional copies of the protein being present and functioning. To test this hypothesis, wild-type *A. thaliana* were transformed with the pGA::*LaCAL* construct to constitutively and ectopically express the *LaCAL* gene.

Three F1 seedlings were identified which expressed genes via the CaMV 35S promoter of the pGA vector (Table 9). Two plants expressed the *NPT-II* gene, and two plants expressed *LaCAL*, with one plant expressing both (Table 9). Several other plants were identified, but there was insufficient tissue or plant did not amplify either the *NPT-II* or *LaCAL* genes. The *LaCAL* transformants did not exhibit any significant phenotypic changes (Figure 25). A few seedlings, including T1CALB3, appeared to flower at a young age, but this plant was not confirmed to express either the *NPT-II* or the *LaCAL* gene and as such, this was not a significant result. The lack of impact of overexpression of *LaCAL* may have been due to the redundancy of the *CAL*

and *AP1* gene functions or due to an inability of the *LaCAL* homologto interact with the proteins which *A. thaliana CAL* interacts (Bowman et al. 1993; Ferrándiz et al. 2000; Alvarez-Buylla et al. 2006). Future studies should examine the effects of knockout studies of *LaCAL* in lavenders to examine the function of *LaCAL* and *LaAP1* in lavenders to determine if the redundant relationship between these genes in *A. thaliana* is observed in *Lavandula*.

4.3.6 *LaFT*

The *A. thaliana FT* gene is a floral pathway integrator and plays a vital role in the induction of flowering in response to photoperiod (Kobayashi et al. 1999; Simpson and Dean 2002; Mizoguchi et al. 2005; Ho and Weigel 2014; Andrés et al. 2015). In *A. thaliana*, *FT* promotes flowering in response to long days and is activated by the protein CO in the floral meristem (Koornneef et al. 1998b; Kardailsky et al. 1999; Samach et al. 2000b; Yoo et al. 2016). Further, AtFT promotes flowering through the activation of several downstream floral meristem identity genes and mutations in *FT* and *CO* cause a delay in the transition to flowering (Samach et al. 2000b; Mizoguchi et al. 2005). Homolog of the *AtFT* gene have been identified and described to function in flowering and development in other flowering plants such as Medicago (*Medicago trunculata*), Garden pea (*Pisum sativum*), Norway Spruce (*Picea abies* L. Karst.), and soybean (*Glycine max* L. Merr.) (Gyllenstrand et al. 2007; Hecht et al. 2011; Laurie et al. 2011; Sun et al. 2011). It was hypothesized that the *Lavandula* homolog of *AtFT*, *LaFT*, would function similarly in lavenders. Further, it was hypothesized that an increase in the expression of *FT* in wild-type *A. thaliana* facilitated by expression of *LaFT* would enhance the effects of this gene via additional copies of the protein being present and functioning. To test this hypothesis, wild-type *A. thaliana* were transformed with the pGA::*LaFT* construct to constitutively and ectopically express the *LaFT* gene.

Over 8 *LaFT* transformants were identified from antibiotic selection plates and grown in soil. Of the plants transferred, there were only two (T1FTA2 and T1FTA8) from the genomic DNA of which the *NPT-II* and *LaFT* genes could be amplified (Table 9). Several transformants, including T1FTA1 and T1FTA5 had insufficient tissue to extract cDNA and mRNA. The *NPT-II* and *LaFT* genes was also amplified from the cDNA obtained from T1FTA2 and T1FTA8 transformants, indicating that these genes were indeed overexpressed in the transformed plants. (Figure 26). The transformed plants exhibited abnormal sepals, abnormal siliques which varied

in size and length, and abnormal flowers (Figure 26). In addition, all transformants exhibited a significant reduction in vegetative growth and leaf mass. Due to the phenotypic differences observed in T1FTA1 and T1FTA5, it is possible that these plants were transformed, but that *NPT-II* and *LaFT* were not amplified due to low quality or concentration of mRNA available due to lack of tissue. All *LaFT* transformants exhibited a phenotype significantly different than that of wild type, with decreased vegetative growth and an apparent reduced time to flowering. A similar reduction in height and decrease in time to flowering was found when *AtFT* was overexpressed in cassava (*Manihot esculenta,* Crantz), (Adeyemo et al. 2017). A similar decrease in time to flowering was found when *AtFT* was overexpressed in *Eucalyptus* trees and in apples, with apples also showing perpetual flowering (*Malus X domestica* Borkh.) (Tanaka et al. 2014; Klocko et al. 2016). Thus, FT overexpression leads to early flowering in a variety of plants. As *LaFT* overexpression also had significant effects on the phenotype of transformants in the present study, further analysis of the *LaFT* gene would be beneficial for elucidating this gene's function in lavenders. In addition, future studies should examine the effects of overexpression of *AtFT* in lavenders in addition to the overexpression of *LaFT* in lavenders. Further, as many *LaFT* transformants did not survive or exhibited significantly decreased overall growth, it would be valuable to examine the effects of overexpression of *LaFT* in *A. thaliana* using a method in which mRNA and cDNA can be efficiently extracted from minimal tissue. Future studies should examine the effects of overexpression of *LaFT* in lavenders to determine if similar effects are observed. An early flowering phenotype in *Lavandula* would be beneficial as it may increase floral production early on, which could allow for increased production of lavender EO.

Chapter 5: Conclusion

Lavenders are a commercially and medicinally important plant, with many species being cultivated globally for their valuable EO. Lavender EOs have been implicated in a variety of settings from use medicinally as an anxiolytic agent to use in perfumes and personal care products (Rastogi et al. 2001; Herrera 2004; Kasper et al. 2010; Grunebaum et al. 2011). The biological activities and diverse applications of lavender EO constituents and the identification of genes involved in the biosynthetic pathways of important oil constituents have been examined in previous studies, with numerous important genes being identified (Woronuk et al. 2010; Demissie et al. 2011; Sarker et al. 2012; Mendoza-Poudereux et al. 2014; Adal et al. 2019). Thus, significant advances have been made in elucidating the details of important pathways and enzymes involved in terpenoid biosynthesis in lavenders. Although flowers are the main site of EO in lavenders, the molecular mechanisms governing floral initiation and pattern development in these plants have not been adequately investigated. These processes have, however, been examined in other flowering plants and are best described in the model plant *A. thaliana* (Kang et al. 1998; Searle and Coupland 2004a; Lee et al. 2007; Litt and Kramer 2010; Hecht et al. 2011; Wu et al. 2012, 2017; Melzer and Ru 2016; Gao et al. 2017). To bridge the gap in knowledge, a transcriptomic approach was employed in this thesis to identify and functionally characterize genes that control flower initiation and pattern development in lavenders. Specifically, the objectives of this thesis were to assemble a comprehensive transcriptome database representative of various developmental stages in lavenders, to identify lavender homologs of *A. thaliana* floral developmental genes, and to functionally characterize a subset of these floral development related genes.

The objectives of this thesis were achieved through the assembly of a comprehensive lavender transcriptome from FL, BD, and vegetative SAM tissues of *L. x intermedia* and *L. angustifolia* tissues, identification of *Lavandula* homologs of previously- described *A. thaliana* floral development genes within the transcriptome via CLC Genomics Workbench 11, and through the expression of *Lavandula* homologs in wild-type *A. thaliana* driven by the CaMV 35S promoter of the pGA482 vector. As a proof of concept, several *Lavandula* homologs were functionally characterized through their overexpression in wild-type *A. thaliana*. The method employed in this thesis demonstrates a comprehensive procedure for the identification, cloning, and characterization of lavender floral development genes which can be extended to future

studies to examining developmental or regulatory genes. This study is the first of its kind in lavenders and the results provide a rich genomic resource for a novel area of research in lavenders. Additionally, the results of this thesis provide valuable insight into a previously unstudied process in lavenders with the identification and characterization of key lavender floral development genes and provides the potential for the manipulation of these genes to improve EO production or yield.

5.1 Development of a Comprehensive *Lavandula* Transcriptome

This thesis first aimed to identify genes involved in floral development by assembling a comprehensive *Lavandula* transcriptome representative of *L. angustifolia* and *L. x intermedia* tissues at different developmental stages to determine if gene homologs to the *A. thaliana* floral development genes existed in lavenders. A comprehensive literature review was conducted to identify floral initiation pathway genes and floral pattern development genes previously described in *A. thaliana.* Based on the presence of homologous floral initiation and pattern development genes in *A. thaliana* and other flowering plants, it was hypothesized that homologsof these genes would exist and function similarly in lavenders (Kang et al. 1998; Theißen 2001; Causier et al. 2010; Ito et al. 2012; Huang et al. 2014; Jaudal et al. 2014; Theißen et al. 2016). Preliminary BLASTx analysis using a previously assembled lavender transcriptome against an annotated *A. thaliana* protein database revealed the presence of lavender genes with high sequence similarity to many *A. thaliana* floral initiation and floral development genes. This result provided support for the hypothesis that these genes exist in lavenders.

To discover the floral development-related genes in *Lavandula*, a comprehensive transcriptome database derived from flowers, buds and shoot apical meristem tissues was developed. Assembly of a comprehensive *Lavandula* transcriptome was then achieved by extracting total RNA from *L. x intermedia* vegetative SAM and *L. angustifolia* vegetative SAM, FL, and BD tissues and transcripts were sequenced via Illumina sequencing. Using the RNA-Seq function of CLC Genomics Workbench 11, a comprehensive lavender transcriptome database of 105,294 contigs representative of developing lavender floral tissues was assembled. The parameters of the resulting transcriptome were comparable to other plant transcriptomes and the transcriptome was validated through the identification of 10 important lavender terpene synthase genes (Duan et al. 2012; Lulin et al. 2012; Xu et al. 2013; Lu et al. 2014; Liu et al. 2018b; Adal

et al. 2019). This transcriptome thus provides a valuable genomic resource for future studies. Additionally, gene ontology studies classified transcripts based on biological process, molecular function, and cellular component (Figure 5).

Although the transcriptome assembled in the present study was comprehensive and reliable, these parameters could be enhanced in future studies by obtaining RNA from additional lavender replicates and from developing floral tissues at smaller increments of time. These additions would allow for an increased comprehensiveness of the transcriptome and thorough representation all genes involved incrementally in floral development, including those which may be expressed transiently, or in low levels in some developmental stages. In addition, species-specific transcriptomes could be assembled to examine the interspecific differences in these genes across lavender species.

The comprehensive *Lavandula* transcriptome constructed in this thesis provides insight into the genes involved in flowering and floral pattern development in *Lavandula* and provides a valuable genomic resource for future studies. Gene ontology studies provide putative functions for the majority of transcripts within the transcriptome, which may be useful in identifying genes of interest to analyze in future studies. Thus, the transcriptome assembled in this study provides a comprehensive database for the identification and functional characterization of additional important floral initiation and development related genes in lavenders. The assembly of this transcriptome has made a significant contribution to what is known about floral development in lavenders. The identification and characterization of additional genes from this transcriptome will provide further insight into floral development in lavenders as well as the potential for manipulation of the expression of these genes to alter floral development.

5.2 Candidate Identification and Differential Expression Analysis

The second objective of this thesis was to identify *Lavandula* floral initiation and development related genes within the transcriptome and to identify candidates for functional characterization in *A. thaliana*. The *Lavandula* transcriptome was annotated via BLASTX analysis in CLC Gemonics Workbench 11 against an annotated *A. thaliana* protein database. These analyses resulted in the identification of *Lavandula* transcripts within the transcriptome showing high sequence similarity to *A. thaliana* floral development related genes. Specifically, over 100 lavender genes homologous to *A. thaliana* floral initiation and pattern development

genes, including repressors and non-repressor proteins, as well as the MADS-Box ABC(DE)-type genes described in *A. thaliana* were identified within the transcriptome (Table 5). This result supported the hypothesis that there exist floral initiation and pattern development genes in lavenders which are homologous to those described in *A. thaliana* and other flowering plants. The identification of over 100 *Lavandula* genes with high sequence homology to *A. thaliana* floral initiation and development genes provided excellent support for the hypothesis that lavenders develop in a similar way to other flowering plants.

Differential expression analyses were performed to examine the expression patterns of the identified genes in *L. angustifolia* and *L. x intermedia* SAM, FL, and BD tissues. Results of these analysese provided insight into the expression patterns of the over 100 floral development related genes identified within the transcriptome and revealed that the expression pattern of many genes varied significantly across developmental stages (Figure 6). For example, *LaSVP* exhibited high expression in the LaSAM in comparison to the LaFL, but low expression in the LaBD in comparison to the LiSAM. This result was consistent with the role of *LaSVP* as a floral repressor (Hartmann et al. 2000; Liu et al. 2007; Wells et al. 2020). Thus, the results of the differential expression analyses in this study provided insight into the expression patterns of *Lavandula* floral development genes and was reflective of the variety of roles of these genes. In addition, the results demonstrate that the expression patterns of these lavender genes similarly varied across developmental stages and certain genes may be active or inactive at a given stage, depending on their role in floral development. This result is consistent with what is known about gene expression and was also consistent with other studies showing differential gene expression during the various stages floral development (Kram et al. 2009; Yant et al. 2010; Klepikova et al. 2015).

The analysis of expression patterns of the identified genes provides a starting point for subsequent studies examining important floral development genes. Further analysis of these expression patterns may facilitate the identification of additional key genes, thus providing a valuable genomic resource for future studies. In addition, it would be valuable to examine expression patterns of the *Lavandula* floral developmental genes parallel to their homologs through short increments of time to examine expression patterns through time as well as the expression patterns at the moment of floral transition.

Based on homology to an *A. thaliana* protein sequence and differential expression pattern, 14 full-length *Lavandula* candidate genes were identified within the transcriptome (Table 6). The identification of lavender genes homologous to those described in *A. thaliana* indicated that the ABC(DE) floral pattern development model and the initiation processes described in *A. thaliana* may also be used to describe flowering in lavenders (Koornneef et al. 1998a; Searle and Coupland 2004a, b; Cheng and Wang 2005; Melzer and Ru 2016; Theißen et al. 2016; Irish 2017).

To examine sequence similarity between the homologous proteins, MUSCLE alignment of the protein sequences was performed. Results of the analysis indicated high sequence similarity of the *Lavandula* proteins and each respective *A. thaliana* homolog and the presence of conserved domains found in other MADS-Box proteins (Alvarez-Buylla et al. 2000; Becker and Theißen 2003). In addition, there were differences in the amino acid sequences between the homologous pairs, as expected in different species. It is hypothesized that the sequence differences may lead to functional differences in these proteins. Together, these results supported the hypothesis that these genes exist and function similarly in lavenders.

The identification of *Lavandula* floral initiation and development-related genes in this study is the first formal description of these genes in lavender and provides the first glimpse into floral initiation and pattern development in lavenders. These genes have not been formally examined in lavenders and the results of this thesis serve to bridge this gap in knowledge and begin a new area of research in lavenders. With the identification of over 100 floral development-related genes, including numerous full-length genes, this transcriptome serves as a rich genomic resource for future studies examining floral development and other developmental processes in lavenders, thus providing a starting point for many future studies.

5.3 Cloning and Functional Characterization of *Lavandula* Homologs

Given the high sequence similarity between lavender floral development genes and those of other flowering plants, it was hypothesized that these genes would function in lavenders as they do in *A. thaliana.* As a proof of concept, 6 of the full-length genes were functionally characterized through overexpression in wild-type *A. thaliana* var. *Landsberg erecta* driven by the CaMV 35S promoter. It was hypothesized that given the high sequence similarity to *A. thaliana* floral development related genes, the lavender homologs would function similarly to the

A. thaliana genes. Transformation was successful for each of the lavender genes, *LaPI*, *LaAP1*, *LaCC*, *LaFT*, *LaCAL*, and *LaSVP* and F_1 transformants expressing the transformed gene were identified for each gene construct. Functional characterization of these genes provided insight into the function of these genes in lavenders and provided evidence that these genes likely function similarly in lavenders as they do in *A. thaliana* and other plants. Ectopic expression of these genes in *A. thaliana* had varying phenotypic effects, with the most drastic phenotypic effects observed in *LaSVP* and *LaFT* mutants. The *LaSVP* gene, was shown to be a repressor of flowering in lavenders through functional annotation in Arabidopsis. As found in previous overexpression studies, this gene repressed the development of flowers and delayed time to flowering. Ectopic expression of the *LaSVP* gene in *A. thaliana* caused a marked phenotypic change, including suppression of petal formation, increased sepal development, sterility, as well as a delay in flowering in comparison to wild- type plants. The results of this study indicate that *LaSVP* acts as a floral repressor in a dosage-dependent manner, functioning similarly to other homologs of *SVP*, including *AtSVP* (Lis-Balchin and Hart 1999; Jaudal et al. 2014). Further, the effects of *LaSVP* on floral organ structure are consistent with the function of *AtSVP* in control in the floral meristem and its role in repression of B, C, D, and E-type floral homeotic genes, having morphological implications (Liu et al. 2007; Li et al. 2008; Gregis et al. 2009).

LaFT transformants exhibited an early onset of flowering and a significant reduction in vegetative and floral growth. Transformants *LaPI*, *LaCAL* and *LaAP1* did not exhibit significant phenotypic differences compared to wild type. It was hypothesized that these genes also function similarly to their *A. thaliana* homologs. It is possible that the lack of phenotypic effects could be due to the manner in which these genes function as MADS-Box TFs in protein complexes (Figure 1) (Theißen et al. 2016). For example, *AtPI* is a B-class gene which binds to the DNA in a quartet with *AtSEP*, *AtAP3*, and *AtAP1* to promote petal development and with *AtSEP*, *AtAP3*, and *AtAG* to promote stamen development (Causier et al. 2010). It is also possible that amino acid differences in the binding domains of the *Lavandula* and *A. thaliana* impacted the ability of the proteins to interact productively. It is also possible that the increase in concentration of a given TF did not increase functions elicited by that TF as the concentrations of the other MADS-box TFs and other interacting proteins did not change. Thus, it is possible that there was no phenotypic change due to a saturation effect with there being a finite number of the other TFs available to bind with the overexpressed lavender TFs (Deitchman et al. 2018). In addition, there

may be differences in post-translational modifications of the proteins in *A. thaliana* and *Lavandula.*

One limitation of this study thus may be that some of the MADS-Box TFS must be overexpressed in combination with other TFs in order to fully determine the effects of overexpression on floral morphology, as the quartets are limited by the TF with the lowest concentration. Several studies examining double mutants of flowering development genes to have significant effects (Pautot et al. 2001; Mizoguchi et al. 2005; Gregis 2006; Hecht et al. 2011; Zhao et al. 2015). As such, future studies should examine the effects of overexpression of multiple *Lavandula* floral development genes in *A. thaliana.* Future studies should further functionally characterize the genes examined in the present study as well as functionally characterize additional repressors and initiators of floral development identified within the *Lavandula* transcriptome. It would be of interest to simultaneously overexpress all member of a quartet in a given whorl to examine the effects on flowering (Figure 1). The genes identified within this thesis should also be further functionally characterized through knockout studies in *A. thaliana* to examine the developmental effects of the absence of expression of these genes and further establish their role in floral development in lavenders. These genes should then be examined through overexpression and knockout studies in *Lavandula*. Although transformation methods have been described in lavenders, the consistency and efficacy of transformation is reported to be inconsistent and the long generation time makes short-term gene expression studies difficult (Mishiba et al. 2000; Nebauer et al. 2000). Recent progress has been made in the Mahmoud lab towards the optimization of effective *Lavandula* transformation protocols.

5.4 Study Limitations

Although these analyses provided valuable insight into the expression patterns and provided a starting point for future studies, some limitations existed. Limitations of the analyses included lack of biological replicates used for transcriptome assembly, lack of replication of the differential expression analysis, and absence of qPCR of all identified genes. Future studies should include an increased number of biological replicates from which RNA is extracted for transcriptome assembly to increase accuracy of the representation of the totality of the genes expressed in these tissues, thus increasing the strength of the transcriptome and the expression analysis results. In addition, future studies may include additional *Lavandula* species in the

transcriptome assembly to examine expression patterns across tissues in other species and examine differences in the sequences in the floral development genes across various lavender species. As there were few biological replicates used, there could exist bias in the expression patterns of the genes due to lack of replicate source tissue. An increase in the number of biological replicates would thus strengthen the transcriptome. In addition, it would be beneficial to assemble multiple transcriptomes and perform Differential Expression analyses on the various transcriptomes to provide replicates of the differential expression results, thus strengthening the results. The observation of the same or similar expression patterns of the identified genes across numerous tissues in numerous transcriptomes would provide additional support for the expression patterns elucidated in this thesis. Thus, increased biological replicates and increased replication of expression analyses would provide further support for the observed expression patterns. In addition, a large-scale qPCR analysis of all identified candidate genes could be performed to validate the results of the differential expression analyses. This large-scale qPCR was beyond the scope of this study but may be beneficial to provide further support for the expression patterns elucidated in this study. Lastly, it would be beneficial in future studies to examine additional developmental stages and identify additional genes which were not identified in this study. In addition, it would be interesting to examine the expression patterns of these genes at smaller time intervals.

Future studies should further characterize the initiation pathways which govern flower initiation in lavenders. As the generation time between *A. thaliana* and lavenders differs greatly, it would be valuable to examine the differences in the initiation pathways in lavenders and *A. thaliana* and the photoperiodic response in lavenders. A further understanding of the initiation pathways and the genes involved in the photoperiodic response in lavenders would potentially lead to the manipulation of flowering time in lavender through the manipulation of expression of these genes either through overexpression or through silencing. In addition, pattern development should also be further examined in lavenders. The identification of the genes which lead to the duplications of flowers along the lavender spike and the division of the SAM are an important area of research. A further understanding of the division of the lavender SAM as well as the processes which govern the development of the lavender floral pattern may lead to the ability to manipulate the expression of genes within the SAM which govern floral pattern in lavenders.

5.5 Summary

The objective of this thesis was to examine the previously unstudied area of floral development in lavenders. The objective was successfully achieved by the completion of three main goals. First, a comprehensive transcriptome database was assembled from RNA of lavender tissues at different developmental stages. The parameters of this transcriptome were comparable to other plant transcriptomes assembled in a similar manner. Second, BLASTx analysis of the transcriptome against an annotated *A. thaliana* protein database facilitated the identification of over 100 *Lavandula* homologs of previously- described *A. thaliana* floral development genes and assigned putative functions to these genes. Lastly, a subset of these *Lavandula* genes were functionally characterized through ectopic expression in wild-type *A. thaliana.* Of the characterized genes, the floral repressor, *LaSVP*, was identified and ectopic expression of this gene in *A. thaliana* caused marked phenotypic changes including suppression of petal formation, increased sepal development, sterility, and delayed flowering. In addition, *LaFT* transformants exhibited a significant decrease in vegetative growth and decreased time to flowering, which was consistent with other studies (Hecht et al. 2011; Azeez et al. 2014; Tanaka et al. 2014; Klocko et al. 2016). Similarities in the function of *LaSVP* and *AtSVP* and *LaFT* and *AtFT*, provides evidence that floral development in lavenders may closely resemble that of *A. thaliana* and that the genes described in *A. thaliana* likely function similarly in lavenders.

The present study has made a significant contribution to research in floral development in lavenders. This study is the first of its kind and a valuable lavender genomic resource was generated. The work in this thesis and the genes identified in the present study have given rise to additional studies. Specifically, this thesis provided the basis and the genes for three undergraduate theses and the start of several additional studies in the Mahmoud lab, providing a significant contribution to research in lavenders. Several genes of the over 100 genes discovered in this study were given to undergraduate students to classify in three separate honors theses. These genes were characterized, providing further significant contributions to research in lavender floral development. Thus, the results of this study have significantly contributed to the Mahmoud lab and developmental research in lavenders.

The results of this thesis mark the first formal identification and characterization of floral initiation and pattern development genes in lavenders and provide a significant contribution research in these plants. Further, the identification and characterization of the floral repressor,

LaSVP, has provided valuable insight into a previously- unstudied process in lavenders with the publication of these results in *Planta* (Wells et al. 2020). The results of this study provide evidence that flowering in lavenders is governed by similar processes as are observed in *A. thaliana* and other plants as the floral development genes in *Lavandula* function similarly to those reported in *A. thaliana*. In addition, the identification of over 100 floral development-related genes within the transcriptome provides a resource for future studies to functionally characterize other floral development-related lavender genes. The successful expression of these *Lavandula* genes in *A. thaliana* provides an efficient means of functional characterization of other lavender floral development related genes. The techniques utilised in this thesis may thus be used in future studies to functionally characterize other floral development genes. The results of this thesis provide compelling evidence that the floral development genes in *Lavandula* likely function similarly to those reported in *A. thaliana* and other plants. Furthermore, this study provides a rich genomic resource for future identification of other developmental genes and insight into the genes involved in flowering and floral pattern development in the genus *Lavandula.* Further, understanding the mechanisms underlying floral initiation and development in lavenders may also lead to future improvements in *Lavandula* EO quality and yield via the manipulation of the expression of these developmental genes in lavenders to induce early flowering, increase petal size or number, or increase oil production. Improvements in quality and yield of *Lavandula* EO would have widespread commercial, therapeutic, and medicinal benefits.

Bibliography

ABRC Arabidopsis Biological Resource Center (2009) Seed Handling. https://abrc.osu.edu/seed-handling. Accessed 30 Jan 2018

Adal AM, Sarker LS, Malli RPN, et al (2019) RNA-seq in the discovery of a sparsely expressed scent-determining monoterpene synthase in lavender (Lavandula). Planta 249:271–290. doi: 10.1007/s00425-018-2935-5

Adeyemo OS, Chavarriaga P, Tohme J, et al (2017) Overexpression of Arabidopsis FLOWERING LOCUS T (FT) gene improves floral development in cassava (Manihot esculenta, Crantz). PLoS One 12:1–15. doi: 10.1371/journal.pone.0181460

Aida M, Ishida T, Tasaka M (1999) Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: Interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. Development 126:1563–1570

Alvarez-Buylla ER, García-Ponce B, Garay-Arroyo A (2006) Unique and redundant functional domains of APETALA1 and CAULIFLOWER, two recently duplicated Arabidopsis thaliana floral MADS-box genes. J Exp Bot 57:3099–3107. doi: 10.1093/jxb/erl081

Alvarez-Buylla ER, Pelaz S, Liljegren SJ, et al (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proc Natl Acad Sci 97:5328–5333. doi: 10.1073/pnas.97.10.5328

An G (1986) Development of plant promoter expression vectors and their use for analysis of differential activity of nopaline synthase promoter in transformed tobacco cells. Plant Physiol 81:86–91. doi: 10.1104/pp.81.1.86

Andrés F, Porri A, Torti S, et al (2014) SHORT VEGETATIVE PHASE reduces gibberellin biosynthesis at the Arabidopsis shoot apex to regulate the floral transition. Proc Natl Acad Sci 111:E2760–E2769. doi: 10.1073/pnas.1409567111

Andrés F, Romera-Branchat M, Martínez-Gallegos R, et al (2015) Floral induction in arabidopsis by FLOWERING LOCUS T requires direct repression of BLADE-ON-PETIOLE genes by the homeodomain protein PENNYWISE. Plant Physiol 169:2187–2199. doi: 10.1104/pp.15.00960

Aprotosoaie AC, Gille E, Trifan A, et al (2017) Essential oils of Lavandula genus : a systematic review of their chemistry. Phytochem Rev 16:761–799. doi: 10.1007/s11101-017-9517-1

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the Flowering plant Arabidopsis thaliana. Nature 408:796–815. doi: 10.1038/35048692

Aukerman MJ, Lee I, Weigel D, Amasino RM (1999) The arabidopsis flowering-time gene LUMINIDEPENDENS is expressed primarily in regions of cell proliferation and encodes a nuclear protein that regulates LEAFY expression. Plant J 18:195–203. doi: 10.1046/j.1365-313X.1999.00442.x

Azeez A, Miskolczi P, Tylewicz S, Bhalerao RP (2014) A tree ortholog of APETALA1 mediates photoperiodic control of seasonal growth. Curr Biol 24:717–724. doi: 10.1016/j.cub.2014.02.037

Barton MK (1998) Cell type specification and self renewal in the vegetative shoot apical meristem. Curr Opin Plant Biol 1:37–42

Bäurle I, Smith L, Baulcombe DC, Dean C (2007) Widespread Role for the Flowering-Time Regulators FCA and FPA in RNA-Mediated Chromatin Silencing. Image (Rochester, NY) 318:109–112. doi: 10.1126/science.1146565

Becker A, Theißen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol 29:464–489. doi: 10.1016/S1055-7903(03)00207-0

Bendahmane M, Dubois A, Raymond O, Le Bris M (2013) Genetics and genomics of flower initiation and development in roses. J Exp Bot 64:847–857. doi: 10.1093/jxb/ers387

Berbel A, Navarro C, Ferrándiz C, et al (2005) Functional conservation of PISTILLATA activity in a pea homolog lacking the PI motif. Plant Physiol 139:174–185. doi: 10.1104/pp.104.057687

Blázquez MA, Green R, Nilsson O, et al (1998) Gibberellins promote flowering of Arabidopsis by activating the LEAFY promoter. Plant Cell 10:791–800. doi: 10.1105/tpc.10.5.791

Boss PK, Bastow RM, Mylne JS, Dean C (2004) Multiple pathways in the decision to flower: Enabling, promoting, and resetting. Plant Cell 16:18–31. doi: 10.1105/tpc.015958

Bowman JL, Alvarez J, Weigel D, et al (1993) Control of flower development in Arabidopsis thaliana by APETALA 1 and interacting genes. Development 119:721–743

Bowman JL, Eshed Y (2000) Formation and maintenance of the shoot apical meristem. Trends Plant Sci 5:110–115. doi: 10.1016/S1360-1385(00)01569-7

Bowman JL, Smyth DR (1999) CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix-loop-helix domains. Development 126:2387–2396

Bowman JL, Smyth DR, Meyerowitz EM (1991) Genetic interactions among floral homeotic genes of Arabidopsis. Development 112:1–20

Brambilla V, Battaglia R, Colombo M, et al (2007) Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in arabidopsis. Plant Cell 19:2544–2556. doi: 10.1105/tpc.107.051797

Brill EM, Watson JM (2004) Ectopic expression of a eucalyptus grandis SVP orthologue alters the flowering time of arabidopsis thaliana. Funct Plant Biol 31:217–224. doi: 10.1071/fp03180

Brioudes F, Joly C, Szécsi J, et al (2009) Jasmonate controls late development stages of petal growth in Arabidopsis thaliana. Plant J 60:1070–1080. doi: 10.1111/j.1365-313X.2009.04023.x

Byrne ME, Simorowski J, Martienssen RA (2002) ASYMMETRIC LEAVES1 reveals know gene redundancy in Arabidopsis. Development 129:1957–1965

Causier B, Schwarz-Sommer Z, Davies B (2010) Floral organ identity: 20 years of ABCs. Semin

Cell Dev Biol 21:73–79. doi: 10.1016/j.semcdb.2009.10.005

Che P, Anand A, Wu E, et al (2018) Developing a flexible, high-efficiency Agrobacterium - mediated sorghum transformation system with broad application. Plant Biotechnol J 16:1388–1395. doi: 10.1111/pbi.12879

Chen X, Liu J, Cheng Y, Jia D (2002) HEN1 Functions Pleiotropically in Arabidopsis Development and Acts in C Funciton in the Flower. Development 129:1085–1094. doi: 10.1016/j.physbeh.2017.03.040

Cheng XF, Wang ZY (2005) Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in Arabidopsis thaliana. Plant J 43:758–768. doi: 10.1111/j.1365-313X.2005.02491.x

Clark SE, Running MP, Meyerowitz EM (1993) CLAVATA1 , a regulator of meristem and flower development in Arabidopsis. Development 119:397–418

CLC bio a (no date) BLAST table. CLC bio, a QIAGEN Company. http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/755/index.php?manual=BLAST_table.html. Accessed 31 August, 2020

CLC bio b (no date) De novo assembly report. CLC bio, a QIAGEN Company. http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/755/index.php?manual=De_novo_assembly_report.htmlAccessed 31 August, 2020

Clough SJ, Bent AF (1998) Floral dip : a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16:735–743. doi: 10.1046/j.1365-313x.1998.00343.x

Coen ES, Meyerowitz EM (1991) The war of the whorls: genetic interactions controlling flower development. 353:31–37. doi: 10.1038/353031a0

Conesa A, Götz S, García-Gómez JM, et al (2005) Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676. doi: 10.1093/bioinformatics/bti610

Costanzo E, Trehin C, Vandenbussche M (2014) The role of WOX genes in flower development. Ann Bot 114:1545–1553. doi: 10.1093/aob/mcu123

De Pater S, Pinas JE, Hooykaas PJJ, Van Der Zaal BJ (2013) ZFN-mediated gene targeting of the Arabidopsis protoporphyrinogen oxidase gene through Agrobacterium -mediated floral dip transformation. Plant Biotechnol J 11:510–515. doi: 10.1111/pbi.12040

Deitchman AN, Singh RSP, Derendorf H (2018) Nonlinear Protein Binding: Not What You Think. J Pharm Sci 107:1754–1760. doi: 10.1016/j.xphs.2018.03.023

Demissie ZA, Cella MA, Sarker LS, et al (2012) Cloning, functional characterization and genomic organization of 1, 8-cineole synthases from Lavandula. Plant Mol Biol 79:393–411. doi: 10.1007/s11103-012-9920-3

Demissie ZA, Sarker LS, Mahmoud SS (2011) Cloning and functional characterization of β-phellandrene synthase from Lavandula angustifolia. Planta 233:685–96. doi:

10.1007/s00425-010-1332-5

Despinasse Y, Fiorucci S, Antonczak S, et al (2017) Bornyl-diphosphate synthase from Lavandula angustifolia: A major monoterpene synthase involved in essential oil quality. Phytochemistry 137:24–33. doi: 10.1016/j.phytochem.2017.01.015

Dinh JL, Farcot E, Hodgman C (2017) The logic of the floral transition: Reverse-engineering the switch controlling the identity of lateral organs. PLoS Comput Biol 13:1–25. doi: 10.1371/journal.pcbi.1005744

Dröge-Laser W, Snoek BL, Snel B, Weiste C (2018) The Arabidopsis bZIP transcription factor family — an update. Curr Opin Plant Biol 45:36–49. doi: 10.1016/j.pbi.2018.05.001

Duan J, Xia C, Zhao G, et al (2012) Optimizing de novo common wheat transcriptome assembly using short-read RNA-Seq data. BMC Genomics 13:1–12. doi: 10.1186/1471-2164-13-392

Dunlap JC (1999) Molecular Bases for Circadian Clocks. Cell 96:271–290. doi: 10.1016/S0092-8674(00)80566-8

Eeckhoute J, Métivier R, Salbert G (2009) Defining specificity of transcription factor regulatory activities. J Cell Sci 122:4027–4034. doi: 10.1242/jcs.054916

Eriksson S, Böhlenius H, Moritz T, Nilsson O (2006) GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. Plant Cell 18:2172–2181. doi: 10.1105/tpc.106.042317

Feng CZ, Chen Y, Wang C, et al (2014) Arabidopsis RAV1 transcription factor, phosphorylated by SnRK2 kinases, regulates the expression of ABI3, ABI4, and ABI5 during seed germination and early seedling development. Plant J 80:654–668. doi: 10.1111/tpj.12670

Ferrándiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development 127:725–734

Field T, Diego M, Hernandez-Reif M, et al (2005) Lavender fragrance cleansing gel effects on relaxation. Int J Neurosci 115:207–22. doi: 10.1080/00207450590519175

Fletcher JC (2001) The ULTRAPETALA gene controls shoot and floral meristem size in Arabidopsis. Development 128:1323–1333

Force A, Lynch M, Pickett FB, et al (1999) Preservation of Duplicate Genes by Complementary , Degenerative Mutations. Genetics 151:1531–1545

Froger A, Hall JE (2008) Transformation of plasmid DNA into E. coli using the heat shock method. J Vis Exp 2007. doi: 10.3791/253

Gao Y, Gao Y, Fan M, et al (2017) Overexpression of Chrysanthemum morifolium SVP gene delays blossoming and regulates inflorescence architecture in transgenic arabidopsis. Can J Plant Sci 97:1130–1139. doi: 10.1139/cjps-2017-0007

Gao Y, Xu H, Shen Y, Wang J (2013) Transcriptomic analysis of rice (Oryza sativa) endosperm using the RNA-Seq technique. Plant Mol Biol 81:363–378. doi: 10.1007/s11103-013-0009-4

Ghedira R, De Buck S, Nolf J, Depicker A (2013) The efficiency of Arabidopsis thaliana floral dip transformation is determined not only by the Agrobacterium strain used but also by the physiology and the ecotype of the dipped plant. Mol Plant-Microbe Interact 26:823–832. doi: 10.1094/MPMI-11-12-0267-R

Gocal GFW, King RW, Blundell CA, et al (2001a) Evolution of floral meristem identity genes. Analysis of Lolium temulentum genes related to APETALA1 and LEAFY of arabidopsis. Plant Physiol 125:1788–1801. doi: 10.1104/pp.125.4.1788

Gocal GFW, Sheldon CC, Gubler F, et al (2001b) GAMYB-like Genes , FlowerinGocal, G. F. W., Sheldon, C. C., Gubler, F., Moritz, T., Bagnall, D. J., Macmillan, C. P., … King, R. W. (2001). GAMYB-like Genes , Flowering , and Gibberellin Signaling in Arabidopsis 1. doi:10.1104/pp.010442.1682g , and Gibber. Plant Physiol 127:1682–1693. doi: 10.1104/pp.010442.1682

Goto K, Meyerowitz EM (1994) Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. Genes Dev 8:1548–1560. doi: 10.1101/gad.8.13.1548

Gramzow L, Ritz MS, Theißen G (2010) On the origin of MADS-domain transcription factors. Trends Genet 26:149–153. doi: 10.1016/j.tig.2010.01.004

Gregis V (2006) AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in arabidopsis. Plant Cell Online 18:1373–1382. doi: 10.1105/tpc.106.041798

Gregis V, Andrés F, Sessa A, et al (2013) Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in Arabidopsis. Genome Biol 14:1–26. doi: 10.1186/gb-2013-14-6-r56

Gregis V, Sessa A, Dorca-Fornell C, Kater MM (2009) The Arabidopsis floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. Plant J 60:626–637. doi: 10.1111/j.1365-313X.2009.03985.x

Gremski K, Ditta G, Yanofsky MF (2007) The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana. Development 134:3593–3601. doi: 10.1242/dev.011510

Gross T, Broholm S, Becker A (2018) CRABS CLAW acts as a bifunctional transcription factor in flower development. Front Plant Sci 9:1–13. doi: 10.3389/fpls.2018.00835

Groszmann M, Paicu T, Smyth DR (2008) Functional domains of SPATULA, a bHLH transcription factor involved in carpel and fruit development in Arabidopsis. Plant J 55:40–52. doi: 10.1111/j.1365-313X.2008.03469.x

Grunebaum LD, Murdock J, Castanedo-Tardan MP, Baumann LS (2011) Effects of lavender olfactory input on cosmetic procedures. J Cosmet Dermatol 10:89–93. doi: 10.1111/j.1473-2165.2011.00554.x

Gyllenstrand N, Clapham D, Källman T, Lagercrantz U (2007) A Norway spruce FLOWERING LOCUS T homolog is implicated in control of growth rhythm in conifers. Plant Physiol 144:248–257. doi: 10.1104/pp.107.095802

Hajdu A, Ádám É, Sheerin DJ, et al (2015) High-level expression and phosphorylation of phytochrome B modulates flowering time in Arabidopsis. Plant J 83:794–805. doi: 10.1111/tpj.12926

Han P, García-Ponce B, Fonseca-Salazar G, et al (2008) AGAMOUS-LIKE 17, a novel flowering promoter, acts in a FT-independent photoperiod pathway. Plant J 55:253–265. doi: 10.1111/j.1365-313X.2008.03499.x

Hartmann U, Höhmann S, Nettesheim K, et al (2000) Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. Plant J 21:351–360. doi: 10.1046/j.1365-313x.2000.00682.x

Hayama R, Coupland G (2003) Shedding light on the circadian clock and the photoperiodic control of flowering. 6:13–19. doi: 10.1016/S1369-5266(02)00011-0

He Y, Doyle MR, Amasino RM (2004) PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. Genes Dev 18:2774–2784. doi: 10.1101/gad.1244504

He Y, Michaels SD, Amasino RM (2003) Regulation of Flowering Time by Histone Acetylation in Arabidopsis. Science (80-) 302:1751–1754. doi: 10.1126/science.1091109

Hecht V, Laurie RE, Vander Schoor JK, et al (2011) The pea GIGAS gene is a FLOWERING LOCUS T homolog necessary for graft-transmissible specification of flowering but not for responsiveness to photoperiod. Plant Cell 23:147–161. doi: 10.1105/tpc.110.081042

Hempel FD, Feldman LJ (1994) Bi-directional inflorescence development in Arabidopsis thaliana: Acropetal initiation of flowers and basipetal initiation of paraclades. Planta 192:276–286. doi: 10.1007/BF01089045

Henderson IR, Dean C (2004) Control of Arabidopsis flowering: The chill before the bloom. Development 131:3829–3838. doi: 10.1242/dev.01294

Hepworth J, Dean C (2015) Flowering locus C's lessons: Conserved chromatin switches underpinning developmental timing and adaptation. Plant Physiol 168:1237–1245. doi: 10.1104/pp.15.00496

Hepworth SR, Valverde F, Ravenscroft D, et al (2002) Antagonistic regulation of flowering-time gene SOC1 by CONSTANS and FLC via separate promoter motifs. EMBO J 21:4327–4337. doi: 10.1093/emboj/cdf432

Herrera AG (2004) Microbiological analysis of cosmetics. Methods Mol Biol 268:293–5. doi: 10.1385/1-59259-766-1:293

Ho WWH, Weigel D (2014) Structural features determining flower-promoting activity of Arabidopsis FLOWERING LOCUS T. Plant Cell 26:552–564. doi: 10.1105/tpc.113.115220

Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 409:525–529. doi: 10.1038/35054083

Horstman A, Willemsen V, Boutilier K, Heidstra R (2014) AINTEGUMENTA-LIKE proteins: Hubs in a plethora of networks. Trends Plant Sci 19:146–157. doi:

10.1016/j.tplants.2013.10.010

Huang F, Xu G, Chi Y, et al (2014) A soybean MADS-box protein modulates floral organ numbers, petal identity and sterility. BMC Plant Biol 14:1–14. doi: 10.1186/1471-2229-14-89

Irish V (2017) The ABC model of floral development. Curr Biol 27:R887–R890. doi: 10.1016/j.cub.2017.03.045

Ito S, Song YH, Josephson-Day AR, et al (2012) FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. Proc Natl Acad Sci U S A 109:3582–3587. doi: 10.1073/pnas.1118876109

Jaudal M, Monash J, Zhang L, et al (2014) Overexpression of medicago SVP genes causes floral defects and delayed flowering in Arabidopsis but only affects floral development in Medicago. J Exp Bot 65:429–442. doi: 10.1093/jxb/ert384

Jullien F, Moja S, Bony A, et al (2014) Isolation and functional characterization of a τ-cadinol synthase, a new sesquiterpene synthase from Lavandula angustifolia. Plant Mol Biol 84:227–241. doi: 10.1007/s11103-013-0131-3

Jung JH, Ju Y, Seo PJ, et al (2012) The SOC1-SPL module integrates photoperiod and gibberellic acid signals to control flowering time in Arabidopsis. Plant J 69:577–588. doi: 10.1111/j.1365-313X.2011.04813.x

Jung JH, Lee S, Yun J, et al (2014) The miR172 target TOE3 represses AGAMOUS expression during Arabidopsis floral patterning. Plant Sci 215–216:29–38. doi: 10.1016/j.plantsci.2013.10.010

Jung JH, Seo YH, Pil JS, et al (2007) The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. Plant Cell 19:2736–2748. doi: 10.1105/tpc.107.054528

Jyothishwaran G, Kotresha D, Selvaraj T, et al (2007) A modified freeze-thaw method for efficient transformation of Agrobacterium tumefaciens. Curr Sci 93:770–772

Kamata N, Okada H, Komeda Y, Takahashi T (2013) Mutations in epidermis-specific HD-ZIP IV genes affect floral organ identity in Arabidopsis thaliana. Plant J 75:430–440. doi: 10.1111/tpj.12211

Kang HG, Jeon JS, Lee S, An G (1998) Identification of class B and class C floral organ identity genes from rice plants. Plant Mol Biol 38:1021–1029. doi: 10.1023/A:1006051911291

Kara N, Baydar H (2013) Determination of lavender and lavandin cultivars (Lavandula sp.) containing high quality essential oil in Isparta, Turkey. Turkish J F Crop 18:58–65. doi: 10.17557/tjfc.86804

Kardailsky I, Shukla VK, Ahn JH, et al (1999) Activation tagging of the floral inducer FT. Science (80-) 286:1962–1965. doi: 10.1126/science.286.5446.1962

Kasper S, Gastpar M, Müller WE, et al (2010) Silexan, an orally administered Lavandula oil preparation, is effective in the treatment of “subsyndromal” anxiety disorder: a randomized,

double-blind, placebo controlled trial. Int Clin Psychopharmacol 25:277–87. doi: 10.1097/YIC.0b013e32833b3242

Kaufmann K, Melzer R, Theißen G (2005) MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. Gene 347:183–198. doi: 10.1016/j.gene.2004.12.014

Kempin SA, Savidge B, Yanofsky MF (2017) Molecular Basis of the cauliflower Phenotype in Arabidopsis. Science (80-) 267:522–525

Kerstetter RA, Bollman K, Taylor RA, et al (2001) KANADI regulates organ polarity in Arabidopsis. Nature 411:706–709. doi: 10.1038/35079629

Kim D-H, Doyle MR, Sung S, Amasino RM (2009) Vernalization: Winter and the Timing of Flowering in Plants. Annu Rev Cell Dev Biol 25:277–299. doi: 10.1146/annurev.cellbio.042308.113411

Kim D-H, Sung S (2014) Genetic and Epigenetic Mechanisms Underlying Vernalization. Arab B 12:e0171. doi: 10.1199/tab.0171

Kim D, Cho YH, Ryu H, et al (2013) BLH1 and KNAT3 modulate ABA responses during germination and early seedling development in Arabidopsis. Plant J 75:755–766. doi: 10.1111/tpj.12236

Kim DH, Sung S (2013) Coordination of the vernalization response through a VIN3 and FLC gene family regulatory network in Arabidopsis. Plant Cell 25:454–469. doi: 10.1105/tpc.112.104760

Klepikova A V., Logacheva MD, Dmitriev SE, Penin AA (2015) RNA-seq analysis of an apical meristem time series reveals a critical point in Arabidopsis thaliana flower initiation. BMC Genomics 16:. doi: 10.1186/s12864-015-1688-9

Klocko AL, Ma C, Robertson S, et al (2016) FT overexpression induces precocious flowering and normal reproductive development in Eucalyptus. Plant Biotechnol J 14:808–819. doi: 10.1111/pbi.12431

Kobayashi Y, Kaya H, Goto K, et al (1999) A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. Science (80-) 286:1960–1962

Koornneef M, Alonso-Blanco C, Peeters AJM, Soppe W (1998a) Genetic control of flowering time in Arabidopsis. Annu Rev Plant Biol 49:345–370. doi: 10.1146/annurev.arplant.49.1.345

Koornneef M, Alonso-Blanco C, Peeters AJM, Soppe W (1998b) Genetic Control of Flowering Time in Arabidopsis. Annu Rev Plant Biol Plant Mol Biol 49:345–370

Kotoda N, Wada M, Kusaba S, et al (2002) Overexpression of MdMADS5, an APETALA1-like gene of apple, causes early flowering in transgenic Arabidopsis. Plant Sci 162:679–687. doi: 10.1016/S0168-9452(02)00024-9

Koulivand PH, Khaleghi Ghadiri M, Gorji A (2013) Lavender and the nervous system. Evidence-Based Complement Altern Med 2013:681304–681310. doi: 10.1155/2013/681304

Koyama T, Furutani M, Tasaka M, Ohme-Takagi M (2007) TCP transcription factors control the morphology of shoot lateral organs via negative regulation of the expression of boundary-specific genes in Arabidopsis. Plant Cell 19:473–484. doi: 10.1105/tpc.106.044792

Kram BW, Xu WW, Carter CJ (2009) Uncovering the Arabidopsis thaliana nectary transcriptome : investigation of differential gene expression in floral nectariferous tissues. BMC Plant Biol 9:1–16. doi: 10.1186/1471-2229-9-92

Kramer EM, Di Stilio VS, Schlüter PM (2003) Complex patterns of gene duplication in the APETALA3 and PISTILLATA lineages of the ranunculaceae. Int J Plant Sci 164:1–11. doi: 10.1086/344694

Kramer EM, Hall JC (2005) Evolutionary dynamics of genes controlling floral development. Curr Opin Plant Biol 8:13–18. doi: 10.1016/j.pbi.2004.09.019

Krizek BA, Meyerowitz EM (1996) The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. 22:11–22

Kumar R, Kushalappa K, Godt D, et al (2007) The Arabidopsis BEL1-like homeodomain proteins SAW1 and SAW2 act redundantly to regulate KNOX expression spatially in leaf margins. Plant Cell 19:2719–2735. doi: 10.1105/tpc.106.048769

Kyozuka J, Kobayashi T, Morita M, Shimamoto K (2000) Spatially and Temporally Regulated Expression of Rice MADS Box Genes with Similarity to Arabidopsis Class A , B and C Genes. Plant Cell Physiol 41:710–718. doi: 10.1093/pcp/41.6.710

Lampugnani ER, Kilinc A, Smyth DR (2013) Auxin controls petal initiation in Arabidopsis. Development 140:185–194. doi: 10.1242/dev.084582

Landmann C, Fink B, Festner MD, et al (2007) Cloning and functional characterization of three terpene synthases from lavender (Lavandula angustifolia). Arch Biochem Biophys 465:417–429. doi: 10.1016/j.abb.2007.06.011

Latchman DS (1993) Transcription factors: an overview Function of transcription factors. Int J Exp Pathol 74:417–422

Laurie RE, Diwadkar P, Jaudal M, et al (2011) The medicago Flowering locus T homolog, MtFTa1, is a Key Regulator of Flowering Time. Plant Physiol 156:2207–2224. doi: 10.1104/pp.111.180182

Lee JH, Park SH, Lee JS, Ahn JH (2007) A conserved role of SHORT VEGETATIVE PHASE (SVP) in controlling flowering time of brassica plants. Biochim Biophys Acta - Gene Struct Expr 1769:455–461. doi: 10.1016/j.bbaexp.2007.05.001

Lenhard M, Laux T (1999) Shoot meristem formation and maintenance. Curr Opin Plant Biol 2:44–50. doi: 10.1016/S1369-5266(99)80009-0

Lesage-Meessen L, Bou M, Sigoillot J-C, et al (2015) Essential oils and distilled straws of lavender and lavandin: a review of current use and potential application in white biotechnology. Appl Microbiol Biotechnol 99:3375–3385. doi: 10.1007/s00253-015-6511-7

Levy YY, Dean C (1998) The Transition to Flowering. Plant Cell 10:1973–1989. doi:

10.1105/tpc.10.12.1973

Levy YY, Mesnage S, Mylne JS, et al (2002) Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control. Science (80-) 297:243–246. doi: 10.1126/science.1072147

Li D, Liu C, Shen L, et al (2008) A repressor complex governs the integration of flowering signals in Arabidopsis. Dev Cell 15:110–120. doi: 10.1016/j.devcel.2008.05.002

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659. doi: 10.1093/bioinformatics/btl158

Li ZM, Zhang JZ, Mei L, et al (2010) PtSVP, an SVP homolog from trifoliate orange (Poncirus trifoliata L. Raf.), shows seasonal periodicity of meristem determination and affects flower development in transgenic Arabidopsis and tobacco plants. Plant Mol Biol 74:129–142. doi: 10.1007/s11103-010-9660-1

Liljegren SJ, Gustafson-Brown C, Pinyopich A, et al (1999) Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. Plant Cell 11:1007–1018. doi: 10.1105/tpc.11.6.1007

Lis-Balchin M (ed) (2002) Lavender: the genus Lavandula, 1st edn. Taylor & Francis Inc., New York

Lis-Balchin M, Hart S (1999) Studies on the mode of action of the essential oil of lavender (Lavandula angustifolia P. Miller). Phytother Res 13:540–2

Litt A, Kramer EM (2010) The ABC model and the diversification of floral organ identity. Semin Cell Dev Biol 21:129–137. doi: 10.1016/j.semcdb.2009.11.019

Liu C, Chen H, Er HL, et al (2008) Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. Development 135:1481–1491. doi: 10.1242/dev.020255

Liu C, Zhou J, Bracha-Drori K, et al (2007) Specification of Arabidopsis floral meristem identity by repression of flowering time genes. Development 134:1901–1910. doi: 10.1242/dev.003103

Liu W, Shen X, Liang H, et al (2018a) Isolation and functional analysis of PISTILLATA homolog from magnolia wufengensis. Front Plant Sci 871:1–9. doi: 10.3389/fpls.2018.01743

Liu X, Sun Z, Dong W, et al (2018b) Expansion and functional divergence of the SHORT VEGETATIVE PHASE (SVP) genes in eudicots. Genome Biol Evol 10:3026–3037. doi: 10.1093/gbe/evy235

Liu Z, Miao L, Huo R, et al (2018c) ARF2-ARF4 and ARF5 are Essential for Female and Male Gametophyte Development in Arabidopsis. Plant Cell Physiol 59:179–189. doi: 10.1093/pcp/pcx174

Lu X, Kim H, Zhong S, et al (2014) De novo transcriptome assembly for rudimentary leaves in Litchi chinesis Sonn. and identification of differentially expressed genes in response to reactive oxygen species. BMC Genomics 15:1–14. doi: 10.1186/1471-2164-15-805

Lulin H, Xiao Y, Pei S, et al (2012) The first illumina-based de novo transcriptome sequencing and analysis of safflower flowers. PLoS One 7:1–11. doi: 10.1371/journal.pone.0038653

Ma H, Yanofsky MF, Meyerowitz EM (1991) AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. Genes Dev 5:484–495. doi: 10.1101/gad.5.3.484

Maier AT, Stehling-Sun S, Wollmann H, et al (2009) Dual roles of the bZIP transcription factor PERIANTHIA in the control of floral architecture and homeotic gene expression. Development 136:1613–1620. doi: 10.1242/dev.033647

Malli RPN, Adal AM, Sarker LS, et al (2019) De novo sequencing of the Lavandula angustifolia genome reveals highly duplicated and optimized features for essential oil production. Planta 249:251–256. doi: 10.1007/s00425-018-3012-9

Mandel MA, Yanofsky MF (1995) The Arabidopsis AGL8 MADS box gene is expressed in inflorescence meristems and is negatively regulated by APETALA1. Plant Cell 7:1763–1771. doi: 10.1105/tpc.7.11.1763

Maoileidigh DSO, Graciet E, Wellmer F (2014) Chapter Six: Genetic Control of Arabidopsis Flower Development. Elsevier

McAbee JM, Hill TA, Skinner DJ, et al (2006) Aberrant testa shape encodes a KANADI family member, linking polarity determination to separation and growth of Arabidopsis ovule integuments. Plant J 46:522–531. doi: 10.1111/j.1365-313X.2006.02717.x

Meinke DW, Cherry JM, Dean C, et al (1998) Arabidopsis thaliana : A Model Plant for Genome Analysis. Science (80-) 282:662–682. doi: 10.1126/science.282.5389.662

Melzer R, Ru F (2016) MADS-domain transcription factors and the floral quartet model of flower development : linking plant development and evolution. 3259–3271. doi: 10.1242/dev.134080

Mendoza-Poudereux I, Muñoz-Bertomeu J, Navarro A, et al (2014) Enhanced levels of S-linalool by metabolic engineering of the terpenoid pathway in spike lavender leaves. Metab Eng 23:136–44. doi: 10.1016/j.ymben.2014.03.003

Michaels SD, Amasino RM (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949–956. doi: 10.1105/tpc.11.5.949

Michaels SD, Ditta G, Gustafson-Brown C, et al (2003) AGL24 acts as a promoter of flowering in Arabidopsis and is positively regulated by vernalization. Plant J 33:867–874. doi: 10.1046/j.1365-313X.2003.01671.x

Mishiba K-I, Ishikawa K, Tsujii O, Mii M (2015) Efficient transformation of lavender (Lavandula latifolia Medicus) mediated by Agrobacterium. 0316:. doi: 10.1080/14620316.2000.11511239

Mishiba KI, Ishikawa K, Tsujii O, Mii M (2000) Efficient transformation of lavender (Lavandula latifolia Medicus) mediated by Agrobacterium. J Hortic Sci Biotechnol 75:287–292. doi: 10.1080/14620316.2000.11511239

Mizoguchi T, Wright L, Fujiwara S, et al (2005) Distinct Roles of GIGANTEA in Promoting Flowering and Regulating Circadian Rhythms in Arabidopsis. Plant Cell 17:2255–2270. doi: 10.1105/tpc.105.033464

Mizzotti C, Mendes MA, Caporali E, et al (2012) The MADS box genes SEEDSTICK and ARABIDOPSIS Bsister play a maternal role in fertilization and seed development. Plant J 70:409–420. doi: 10.1111/j.1365-313X.2011.04878.x

Moon J, Suh SS, Lee H, et al (2003a) The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J 35:613–623. doi: 10.1046/j.1365-313X.2003.01833.x

Moon YH, Chen L, Pan RL, et al (2003b) EMF Genes Maintain Vegetative Development by Repressing the Flower Program in Arabidopsis. Plant Cell 15:681–693. doi: 10.1105/tpc.cor831

Muñoz-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 35:479–488. doi: 10.1016/j.bse.2007.03.006

Murai K, Miyamae M, Kato H, et al (2003) WAP1, a Wheat APETALA1 Homolog, Plays a Central Role in the Phase Transition from Vegetative to Reproductive Growth. Plant Cell Physiol 44:1255–1265. doi: 10.1093/pcp/pcg171

Murmu J, Bush MJ, de Long C, et al (2010) Arabidopsis basic leucine-zipper transcription factors TGA9 and TGA10 interact with floral glutaredoxins ROXY1 and ROXY2 and are redundantly required for anther development. Plant Physiol 154:1492–1504. doi: 10.1104/pp.110.159111

Nebauer SG, Arrillaga I, Castillo-agudo L, Segura J (2000) Agrobacterium tumefaciens-mediated transformation of the aromatic shrub Lavandula latifolia. Mol Breed 6:539–552. doi: 10.1023/A:1011332904024

Nikovics K, Blein T, Peaucelle A, et al (2006) The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. Plant Cell 18:2929–2945. doi: 10.1105/tpc.106.045617

Noh Y-S, Amasino RM (2003) PIE1, an ISWI Family Gene, Is Required for FLC Activation and Floral Repression in Arabidopsis. Plant Cell 15:1671–1682. doi: 10.1105/tpc.012161

Ogawa E, Yamada Y, Sezaki N, et al (2014) ATML1 and PDF2 play a redundant and essential role in arabidopsis embryo development. Plant Cell Physiol 56:1183–1192. doi: 10.1093/pcp/pcv045

Pabón-Mora N, Ambrose BA, Litt A (2012) Poppy APETALA1/FRUITFULL orthologs control flowering time, branching, perianth identity, and fruit development. Plant Physiol 158:1685–1704. doi: 10.1104/pp.111.192104

Page T, Macknight R, Yang C, Dean C (1999) Genetic interactions of the Arabidopsis flowering time gene FCA , with genes regulating floral initiation. Plant J 17:231–239. doi: 10.1046/j.1365-313X.1999.00364.x

Parcy F, Nils O, Busch MA, et al (1999) A genetic framework for floral patterning. Chemtracts 12:956–959

Pařenicová L, De Folter S, Kieffer M, et al (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: new openings to the MADS world. Plant Cell 15:1538–1551. doi: 10.1105/tpc.011544.these

Pautot V, Dockx J, Hamant O, et al (2001) KNAT2: Evidence for a link between Knotted-like genes and carpel development. Plant Cell 13:1719–1734. doi: 10.1105/tpc.13.8.1719

Pellegrini L, Tan S, Richmond TJ (1995) Structure of serum response factor core bound to DNA. Nature 376:490–498. doi: 10.1038/376490a0

Qi T, Huang H, Song S, Xie D (2015) Regulation of jasmonate-mediated stamen development and seed production by a bHLH-MYB complex in arabidopsis. Plant Cell 27:1620–1633. doi: 10.1105/tpc.15.00116

Rastogi SC, Heydorn S, Johansen JD, Basketter DA (2001) Fragrance chemicals in domestic and occupational products. Contact Dermatitis 45:221–5

Ratcliffe OJ, Kumimoto RW, Wong BJ, Reichmann JL (2003) Analysis of the Arabidopsis MADS AFFECTING FLOWERING Gene Family: MAF2 Prevents Vernalization by Short Periods of Cold. Plant Cell 15:1159–1169. doi: 10.1105/tpc.009506

Richter R, Behringer C, Zourelidou M, Schwechheimer C (2013) Convergence of auxin and gibberellin signaling on the regulation of the GATA transcription factors GNC and GNL in Arabidopsis thaliana. Proc Natl Acad Sci U S A 110:13192–13197. doi: 10.1073/pnas.1304250110

Robson F, Costa MMR, Hepworth SR, et al (2001) Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. Plant J 28:619–631. doi: 10.1046/j.1365-313x.2001.01163.x

Romera-Branchat M, Ripoll JJ, Yanofsky MF, Pelaz S (2013) The WOX13 homeobox gene promotes replum formation in the Arabidopsis thaliana fruit. Plant J 73:37–49. doi: 10.1111/tpj.12010

Sakai H, Krizek BA, Jacobsen SE, Meyerowitz EM (2000) Regulation of SUP expression identifies multiple regulators involved in Arabidopsis floral meristem development. Plant Cell 12:1607–1618. doi: 10.1105/tpc.12.9.1607

Samach A, Klenz JE, Kohalmi SE, et al (1999) The UNUSUAL FLORAL ORGANS gene of Arabidopsis thaliana is an F-box protein required for normal patterning and growth in the floral meristem. Plant J 20:433–445. doi: 10.1046/j.1365-313X.1999.00617.x

Samach A, Onouchi H, Gold SE, et al (2000a) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science (80-) 288:1613–1616. doi: 10.1126/science.288.5471.1613

Samach A, Onouchi H, Gold SE, et al (2000b) Distinct Roles of CONSTANS Target Genes in Reproductive Development of Arabidopsis. 288:1613–1617

Sarker LS, Demissie ZA, Mahmoud SS (2013) Cloning of a sesquiterpene synthase from Lavandula x intermedia glandular trichomes. Planta 238:983–989. doi: 10.1007/s00425-013-1937-6

Sarker LS, Galata M, Demissie ZA, Mahmoud SS (2012) Molecular cloning and functional characterization of borneol dehydrogenase from the glandular trichomes of Lavandula x intermedia. Arch Biochem Biophys 528:163–170. doi: 10.1016/j.abb.2012.09.013

Schiessl K, Muiño JM, Sablowski R (2014) Arabidopsis JAGGED links floral organ patterning to tissue growth by repressing Kip-related cell cycle inhibitors. Proc Natl Acad Sci U S A 111:2830–2835. doi: 10.1073/pnas.1320457111

Schomburg FM, Patton DA, Meinke DW, Amasino RM (2001) FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. Plant Cell 13:1427–1436. doi: 10.1105/tpc.13.6.1427

Schruff MC, Spielman M, Tiwari S, et al (2006) The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. Development 133:251–261. doi: 10.1242/dev.02194

Schwarz-Sommer Z, Huijser P, Nacken W, et al (1990) Genetic control of flower development by homeotic genes in Antirrhinum majus. Science (80-) 250:931–936. doi: 10.1126/science.250.4983.931

Searle I, Coupland G (2004a) Induction of flowering by seasonal changes in photoperiod. EMBO J 23:1217–1222. doi: 10.1038/sj.emboj.7600117

Searle I, Coupland G (2004b) New EMBO Member ' s Review Induction of flowering by seasonal changes in photoperiod. 1217–1222. doi: 10.1038/sj.emboj.7600117

Sharma M, Kothari-Chajer A, Jagga-Chugh S, Kothari S (2011) Factors influencing Agrobacterium tumefaciens -mediated genetic transformation of Eleusine coracana (L .) Gaertn. Plant Cell Tissue Organ Cult 105:93–104. doi: 10.1007/s11240-010-9846-x

Sheldon CC, Rouse DT, Finnegan EJ, et al (2000) The molecular basis of vernalization: The central role of FLOWERING LOCUS C (FLC). Proc Natl Acad Sci U S A 97:3753–3758. doi: 10.1073/pnas.97.7.3753

Shi JX, Malitsky S, de Oliveira S, et al (2011) SHINE transcription factors act redundantly to pattern the archetypal surface of arabidopsis flower organs. PLoS Genet 7:. doi: 10.1371/journal.pgen.1001388

Shitsukawa N, Takagishi A, Ikari C, et al (2006) WFL, a wheat FLORICAULA/LEAFY ortholog, is associated with spikelet formation as lateral branch of the inflorescence meristem. Genes Genet Syst 81:13–20. doi: 10.1266/ggs.81.13

Shore P, Sharrocks AD (1995) The MADS-Box Family of Transcription Factors. Eur J Biochem 229:1–13. doi: 10.1111/j.1432-1033.1995.tb20430.x

Simonini S, Roig-Villanova I, Gregis V, et al (2012) BASIC PENTACYSTEINE proteins mediate MADS domain complex binding to the DNA for Tissue-Specific expression of target genes in Arabidopsis. Plant Cell 24:4163–4172. doi: 10.1105/tpc.112.103952

Simpson GG (2004) The autonomous pathway: Epigenetic and post-transcriptional gene regulation in the control of Arabidopsis flowering time. Curr Opin Plant Biol 7:570–574. doi: 10.1016/j.pbi.2004.07.002

Simpson GG, Dean C (2002) Arabidopsis, the rosetta stone of flowering time. Science (80-) 296:285–289. doi: 10.1126/science.296.5566.285

Simpson GG, Dijkwel PP, Quesada V, et al (2003) FY is an RNA 3′ end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113:777–787. doi: 10.1016/S0092-8674(03)00425-2

Smaczniak C, Immink RGH, Muiño JM, et al (2012) Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. Proc Natl Acad Sci U S A 109:1560–1565. doi: 10.1073/pnas.1112871109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1112871109

Smith CA, Collins CT, Crowther CA (2011) Aromatherapy for pain management in labour. Cochrane database Syst Rev CD009215. doi: 10.1002/14651858.CD009215

Smyth DR, Bowman JL, Meyerowitz EM (1990) Early Flower Development in. Plant Cell 2:755–767. doi: 10.1105/tpc.2.8.755

Soltani R, Soheilipour S, Hajhashemi V, et al (2013) Evaluation of the effect of aromatherapy with lavender essential oil on post-tonsillectomy pain in pediatric patients: a randomized controlled trial. Int J Pediatr Otorhinolaryngol 77:1579–81. doi: 10.1016/j.ijporl.2013.07.014

Strable J, Wallace JG, Unger-Wallace E, et al (2017) Maize YABBY genes drooping leaf1 and drooping leaf2 regulate plant architecture. Plant Cell 29:1622–1641. doi: 10.1105/tpc.16.00477

Sun H, Jia Z, Cao D, et al (2011) GmFT2a, a Soybean homolog of Flowering Locus T, is involved in Flowering Transition and Maintenance. PLoS One 6:18–20. doi: 10.1371/journal.pone.0029238

Tabata R, Ikezaki M, Fujibe T, et al (2010) Arabidopsis AUXIN RESPONSE FACTOR6 and 8 regulate jasmonic acid biosynthesis and floral organ development via repression of class 1 KNOX genes. Plant Cell Physiol 51:164–175. doi: 10.1093/pcp/pcp176

Takeda S, Ogata Y, Kojima H, et al (2011) Arabidopsis thaliana: a novel biocatalyst for asymmetric reductions. Plant Biotechnol 28:77–82. doi: 10.5511/plantbiotechnology.10.0927a

Talbert PB, Adler HT, Parks DW, Comai L (1995) The REVOLUTA gene is necessary for apical meristem development and for limiting cell divisions in the leaves and stems of Arabidopsis thaliana. Development 121:2723–2735

Tanaka N, Ureshino A, Shigeta N, et al (2014) Overexpression of Arabidopsis FT gene in apple leads to perpetual flowering. Plant Biotechnol 31:11–20. doi: 10.5511/plantbiotechnology.13.0912a

Tanaka N, Wada M, Komori S, et al (2007) Functional analysis of MdPI, the PISTILLATA gene

homologue of apple, in Arabidopsis. J Japanese Soc Hortic Sci 76:125–132. doi: 10.2503/jjshs.76.125

Tapia-López R, García-Ponce B, Dubrovsky JG, et al (2008) An AGAMOUS-related MADS-box gene, XAL1 (AGL12), regulates root meristem cell proliferation and flowering transition in Arabidopsis. Plant Physiol 146:1182–1192. doi: 10.1104/pp.107.108647

Teeri TH, Uimari A, Kotilainen M, et al (2006) Reproductive meristem fates in gerbera. J Exp Bot 57:3445–3455. doi: 10.1093/jxb/erl181

Teper-Bamnolker P, Samach A (2005) The flowering integrator FT regulates SEPALLATA3 and FRUITFULL accumulation in Arabidopsis leaves. Plant Cell 17:2661–2675. doi: 10.1105/tpc.105.035766

Theißen G (2001) Development of floral organ identity: stories from the MADS house. Curr Opin Plant Biol 4:75–85. doi: 10.1016/S1369-5266(00)00139-4

Theißen G, Kim JT, Saedler H (1996) Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. J Mol Evol 43:484–516

Theißen G, Melzer R, Rümpler F (2016) MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. Development 143:3259–3271. doi: 10.1242/dev.134080

Theißen G, Saedler H (2001) Plant biology: Floral quartets. Nature 409:469–471. doi: 10.1038/35054172

Trigueros M, Navarrete-Gómez M, Sato S, et al (2009) The NGATHA genes direct style development in the arabidopsis gynoecium. Plant Cell 21:1394–1409. doi: 10.1105/tpc.109.065508

Unte US, Sorensen AM, Pesaresi P, et al (2003) SPL8, an SBP-box gene that affects pollen sac development in Arabidopsis. Plant Cell 15:1009–1019. doi: 10.1105/tpc.010678

Upson T, Andrews S (2004) The genus Lavandula. The Royal Botanic Gardens, Kew, Richmond, Surrey, UK

Urwin NAR, Horsnell J, Moon T (2007) Generation and characterisation of colchicine-induced autotetraploid Lavandula angustifolia. Euphytica 156:257–266. doi: 10.1007/s10681-007-9373-y

Vakili A, Sharifat S, Akhavan MM, Bandegi AR (2014) Effect of lavender oil (Lavandula angustifolia) on cerebral edema and its possible mechanisms in an experimental model of stroke. Brain Res 1548:56–62. doi: 10.1016/j.brainres.2013.12.019

Wang Z, Hobson N, Galindo L, et al (2012) The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J 72:461–473. doi: 10.1111/j.1365-313X.2012.05093.x

Wang Z, Zhou Z, Liu Y, et al (2015) Functional Evolution of Phosphatidylethanolamine Binding Proteins in Soybean and Arabidopsis. 27:323–336. doi: 10.1105/tpc.114.135103

Weigel D, Meyerowitz EM (1994) The abcs of floral homeotic genes. Cell 78:203–209. doi: 10.1016/0092-8674(94)90291-7

Wells R, Truong F, Adal AM, et al (2018) Lavandula essential oils: a current review of applications in medicinal, food, and cosmetic industries of lavender. Nat Prod Commun 13:1403–1417. doi: 10.1177/1934578x1801301038

Wells RS, Adal AM, Bauer L, et al (2020) Cloning and functional characterization of a floral repressor gene from Lavandula angustifolia. Planta 251:1–11. doi: 10.1007/s00425-019-03333-w

Wigge PA, Kim MC, Jaeger KE, et al (2005) Integration of spatial and temporal information during floral induction in Arabidopsis. Science (80-) 309:1056–1059. doi: 10.1126/science.1114358

Wils CR, Kaufmann K (2017) Gene-regulatory networks controlling inflorescence and flower development in Arabidopsis thaliana. Biochim Biophys Acta 1860:95–105. doi: 10.1016/j.bbagrm.2016.07.014

Wilson RN, Heckman JW, Somerville CR (1992) Gibberellin is required for flowering in Arabidopsis thaliana under short days. Plant Physiol 100:403–408. doi: 10.1104/pp.100.1.403

Woronuk G, Demissie Z, Rheault M, Mahmoud S (2010) Biosynthesis and therapeutic properties of lavandula essential oil constituents. Planta Med 77:7–15. doi: 10.1055/s-0030-1250136

Wu F, Shi X, Lin X, et al (2017) The abcs of flower development: mutational analysis of AP1/FUL-like genes in rice provides evidence for a homeotic (a)-function in grasses. Plant J 89:310–324. doi: 10.1111/tpj.13386

Wu G, Lin WC, Huang T, et al (2008) KANADI1 regulates adaxial-abaxial polarity in Arabidopsis by directly repressing the transcription of Asymmetric Leaves2. Proc Natl Acad Sci U S A 105:16392–16397. doi: 10.1073/pnas.0803997105

Wu G, Poethig RS (2006) Temporal regulation of shoot development in Arabidopsis thaliana by miRr156 and its target SPL3. Development 133:3539–3547. doi: 10.1242/dev.02521

Wu RM, Walton EF, Richardson AC, et al (2012) Conservation and divergence of four kiwifruit SVP-like MADS-box genes suggest distinct roles in kiwifruit bud dormancy and flowering. J Exp Bot 63:797–807. doi: 10.1093/jxb/err304

Wu X, Shi J, Xi M, et al (2010) A B Functional Gene Cloned from Lily Encodes an Ortholog of Arabidopsis PISTILLATA (PI). Plant Mol Biol 28:684–691. doi: 10.1007/s11105-010-0193-1

Xu Y, Gao S, Yang Y, et al (2013) Transcriptome sequencing and whole genome expression profiling of Chrysanthemum under dehydration stress. BMC Genomics 14:1–15. doi: 10.1186/1471-2164-14-662

Yamabhai M (2009) Sticky PCR: A PCR-based protocol for targeted protein engineering. Biotechnol J 4:544–553. doi: 10.1002/biot.200800198

Yang H, Shi G, Li X, et al (2019) Overexpression of a soybean YABBY gene, GmFILa, causes leaf curling in Arabidopsis thaliana. BMC Plant Biol 19:1–16. doi: 10.1186/s12870-019-1810-2

Yang Y, Fanning L, Jack T (2003) The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. Plant J 33:47–59. doi: 10.1046/j.0960-7412.2003.01473.x

Yanofsky MF (1995) Floral meristems to floral organs: genes controlling early events in Arabidopsis flower development. Annu Rev Plant Physiol Plant Mol Biol 46:167–188. doi: 10.1146/annurev.pp.46.060195.001123

Yant L, Mathieu J, Dinh TT, et al (2010) Orchestration of the floral transition and floral development in arabidopsis by the bifunctional transcription factor APETALA2. Plant Cell 22:2156–2170. doi: 10.1105/tpc.110.075606

Yoo SK, Chung KS, Kim J, et al (2016) CONSTANS Activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to Promote Flowering in Arabidopsis. Plant Physiol 139:770–778. doi: 10.1104/pp.105.066928.promotion

Yoo SK, Wu X, Lee JS, Ahn JH (2011) AGAMOUS-LIKE 6 is a floral promoter that negatively regulates the FLC/MAF clade genes and positively regulates FT in Arabidopsis. Plant J 65:62–76. doi: 10.1111/j.1365-313X.2010.04402.x

Yu LH, Miao ZQ, Qi GF, et al (2014) MADS-Box transcription factor AGL21 regulates lateral root development and responds to multiple external and physiological signals. Mol Plant 7:1653–1669. doi: 10.1093/mp/ssu088

Yu LH, Wu J, Zhang ZS, et al (2017) Arabidopsis MADS-Box Transcription Factor AGL21 Acts as Environmental Surveillance of Seed Germination by Regulating ABI5 Expression. Mol Plant 10:834–845. doi: 10.1016/j.molp.2017.04.004

Zagotta MT, Hicks, A. K, et al (1996a) The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. Plant J 10:691–702. doi: 10.1046/j.1365-313X.1996.10040691.x

Zagotta MT, Hicks KA, Jacobs CI, et al (1996b) The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. Plant J 10:691–702. doi: 10.1046/j.1365-313x.1996.10040691.x

Zhang X, Henriques R, Lin S, et al (2006a) Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protoc 1:1–7. doi: 10.1038/nprot.2006.97

Zhang X, Henriques R, Lin SS, et al (2006b) Agrobacterium-mediated transformation of Arabidopsis thaliana using the floral dip method. Nat Protoc 1:641–646. doi: 10.1038/nprot.2006.97

Zhao M, Yang S, Liu X, Wu K (2015) Arabidopsis histone demethylases LDL1 and LDL2 control primary seed dormancy by regulating DELAY OF GERMINATION 1 and ABA signaling-related genes. Front Plant Sci 6:1–9. doi: 10.3389/fpls.2015.00159

Ziaee M, Khorrami A, Ebrahimi M, et al (2015) Cardioprotective effects of essential oil of Lavandula angustifolia on isoproterenol-induced acute myocardial infarction in rat. Iran J Pharm Res IJPR 14:279–89

Zuzarte M, Gonçalves MJ, Cavaleiro C, et al (2009) Chemical composition and antifungal activity of the essential oils of lavandula pedunculata (Miller) Cav. Chem Biodivers 6:1283–1292. doi: 10.1002/cbdv.200800170

Appendices

Appendix A – Supplementary Data

Table A.1 Nanodrop Spectrometer data of extracted RNA from *L. angustifolia* floral and bud tissues. *L. angustifolia* bud RNA was used in reverse transcription and in the amplification of floral genes.

Sample	µg/ µl	260/230	230/280
L. angustifolia BD	1058.8	2.17	2.35
L. angustifolia FL	941.9	2.15	2.29

Table A.2 Nanodrop Spectrometer data of extracted plasmids from *E. coli* JM109 cells.

Sample	µg/ µl	260/230	230/280
1	228.9	2.30	1.89
2	238.0	2.09	1.86

Table A.3 Nanodrop Spectrometer data of extracted *E. coli* JM109 plasmids.

Sample	µg/ µl	260/230	230/280
CC	100.1	1.88	0.96
CAL	74.9	1.93	1.06
SVP	70.3	1.88	1.96
PI	92.9	1.87	0.48
GB	80.1	1.83	0.59
FT	79.8	1.88	1.85
AP1	96.7	1.92	1.54
AP2	31.2	1.93	0.32

Table A.4 Culture-forming units per plate of heat-shock transformed *E. coli* cells on LB tetracycline medium.

Gene	Dilute Plate	Concentrated Plate
CC	9	>100
CAL	10	>100
SVP	73	>100
PI	48	>100
GB	41	>100
FT	60	>100
AP1	61	>100
Plasmid	0	0

Table A.5 RPKM values of floral development genes in *L. angustifolia* and *L. x intermedia* tissues based on Differential expression for RNA Seq analysis.

Gene	**Tissue**			
	L. x intermedia VM	*L. x angustifolia* VM	*L. x angustifolia* FL	*L. angustifolia* BD
CC	0.00	0.36	102.40	16.68
CAL	4.44	7.09	41.56	45.93
SVP	25.42	11.51	0.61	2.80
PI	0.68	1.23	145.75	234.96
FT	0	10.55	28.90	79.67
AP1	0	0	24.68	53.02
Not cloned:				
SM	2.39	16.82	73.94	54.64
AP2	7.63	1.84	0	3.05
AP3	2.62	3.16	197.12	193.87
SOC1	28.03	27.61	4.56	4.47
FLC	8.01	22.30	25.79	22.31
SEP3	0	0.99	198.45	198.96
AG	0	0.19	45.95	21.73
EMF2	10.62	16.29	13.32	17.90

Appendix B – Supplementary Figures

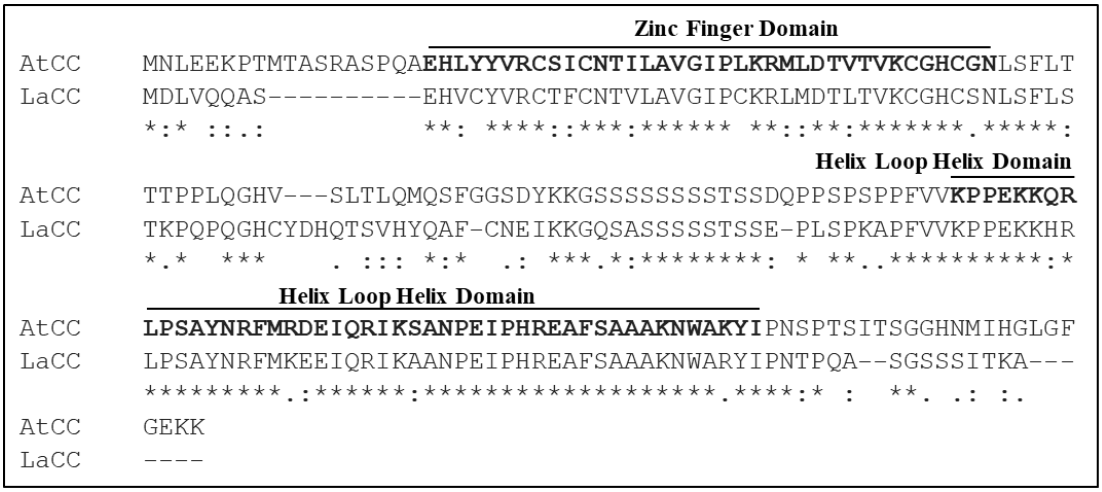

Figure B.1 Multiple Sequence Alignment of protein sequences of *L. angustifolia Crab's Claw (LaCC)* and *Arabidopsis thaliana Crab's Claw (AtCC)*. *AtCC* has zinc finger and helix loop helix domains and is a member of the YABBY family of transcription factors which is involved in carpel and nectary development (Bowman and Smyth 1999).

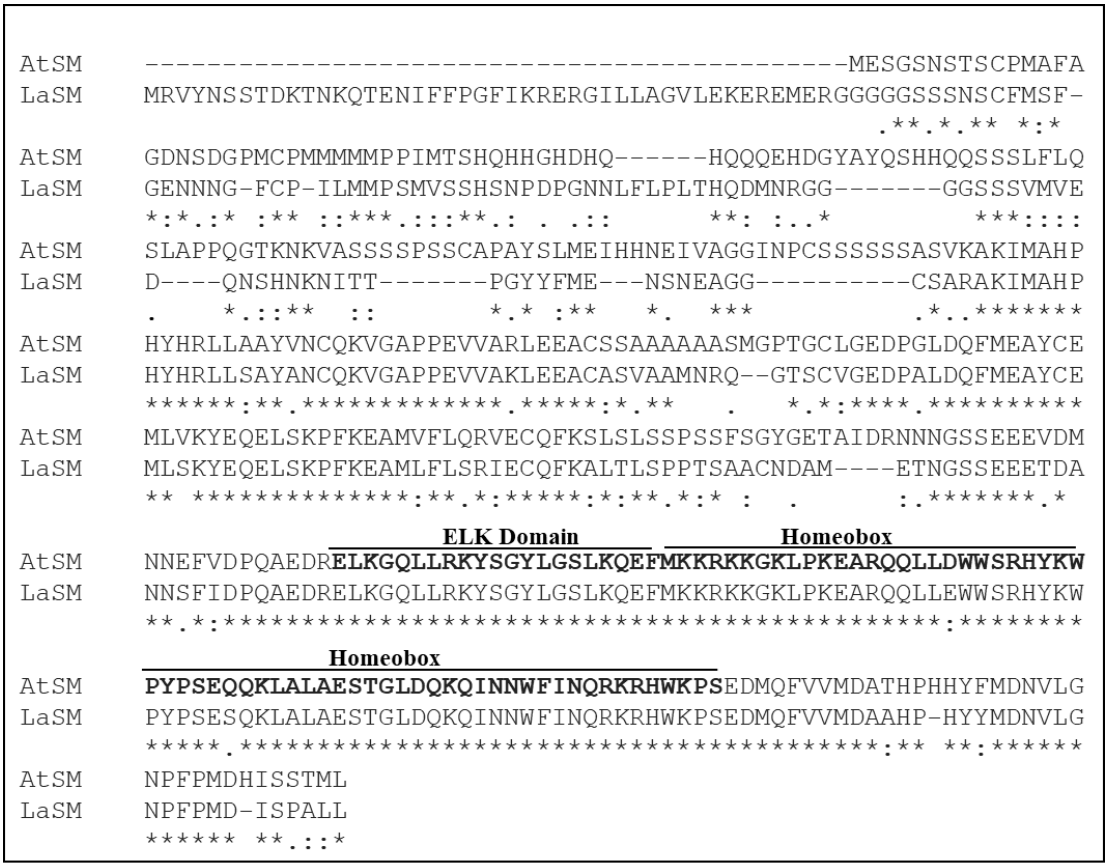

Figure B.2 Multiple Sequence Alignment of protein sequences of *L. angustifolia Shoot Meristemless* (*LaSM*) and *A. thaliana Shoot Meristemless* (*AtSM*). *AtSM* is a TALE/KMNOX Homeobox transcription factor involved in cell proliferation in the SAM (Aida et al. 1999; Byrne et al. 2002). The bolded portions of protein sequence indicate the sequence of the ELK Domain and homeobox.

Figure B.3 Multiple Sequence Alignment of protein sequences of *L. angustifolia Embryonic Flower 2* (*LaEMF2*) and *A. thaliana Embryonic Flower 2* (*AtEMF2*). *AtEMF2* is a transcription factor involved in repression of flowering in *A. thaliana* (Moon et al. 2003b). The Zinc finger and VEFS-Box domains are shown in bold.

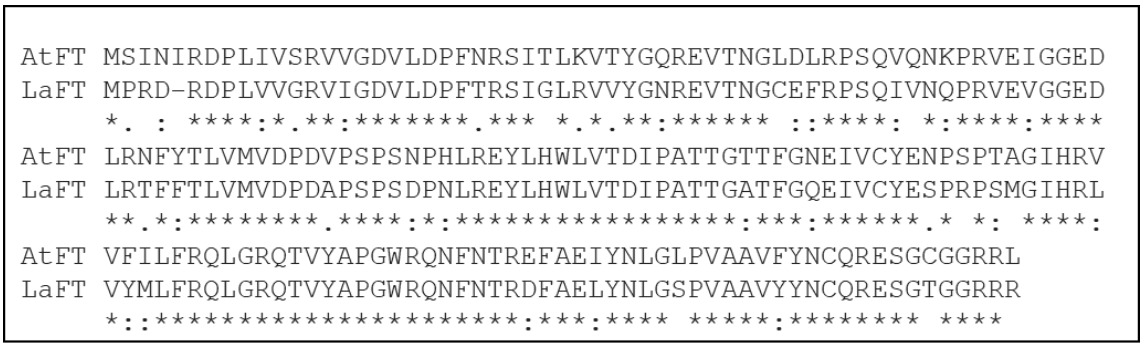

Figure B.4 Multiple Sequence Alignment of protein sequences of *L. angustifolia Flowering Locus T* (*LaFT*) and *A. thaliana Flowering Locus T* (*AtFT*). *AtFT* is a gene involved in floral initiation in response to photoperiod (Andrés et al. 2015). The FT protein acts as a phosphatidylethanolamine binding transcription factor (Wang et al. 2015).

Figure B.5 Multiple Sequence Alignment of protein sequences of *L. x intermedia Apetala2* (*LiAP2*) and *A. thaliana Apetala2* (*AtAP2*). *AtAP2* is a transcription factor involved in sepal and petal development in *A. thaliana* (Theißen et al. 2016)s. The bolded portion of the protein sequence indicates the sequence of the two DNA-binding domains.

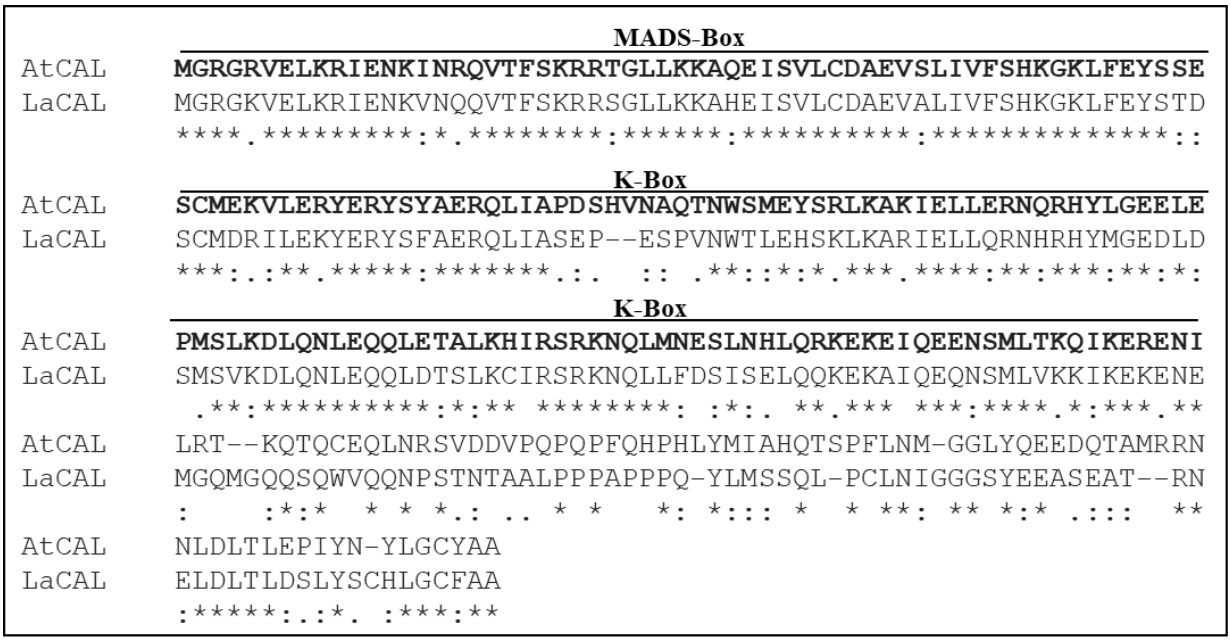

Figure B.6 Multiple Sequence Alignment of protein sequences of *Lavandula angustifolia Cauliflower* (*LaCAL*) and *Arabidopsis thaliana Cauliflower* (*AtCAL*). *AtCAL* is a MADS-Box TF involved in floral meristem identity specification in *A. thaliana* (Bowman et al. 1993). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

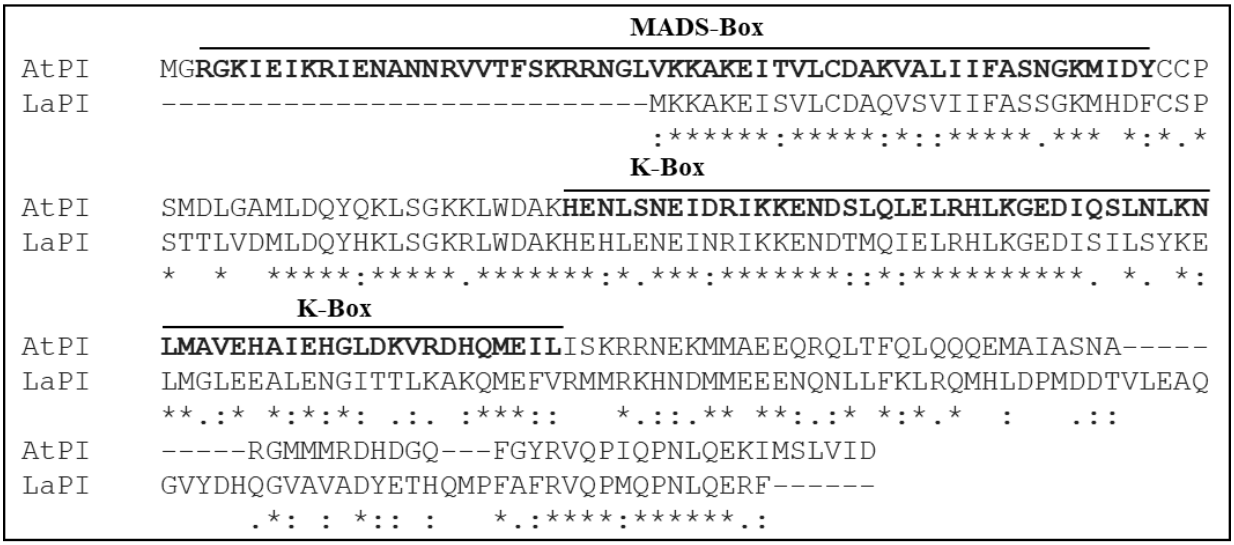

Figure B.7 Multiple Sequence Alignment of protein sequences of *L. angustifolia Pistillata* (*LaPI*) and *Arabidopsis thaliana Pistillata* (*AtPI*), which is required for petal and stamen specification. The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains (Krizek and Meyerowitz 1996; Theißen et al. 2016).

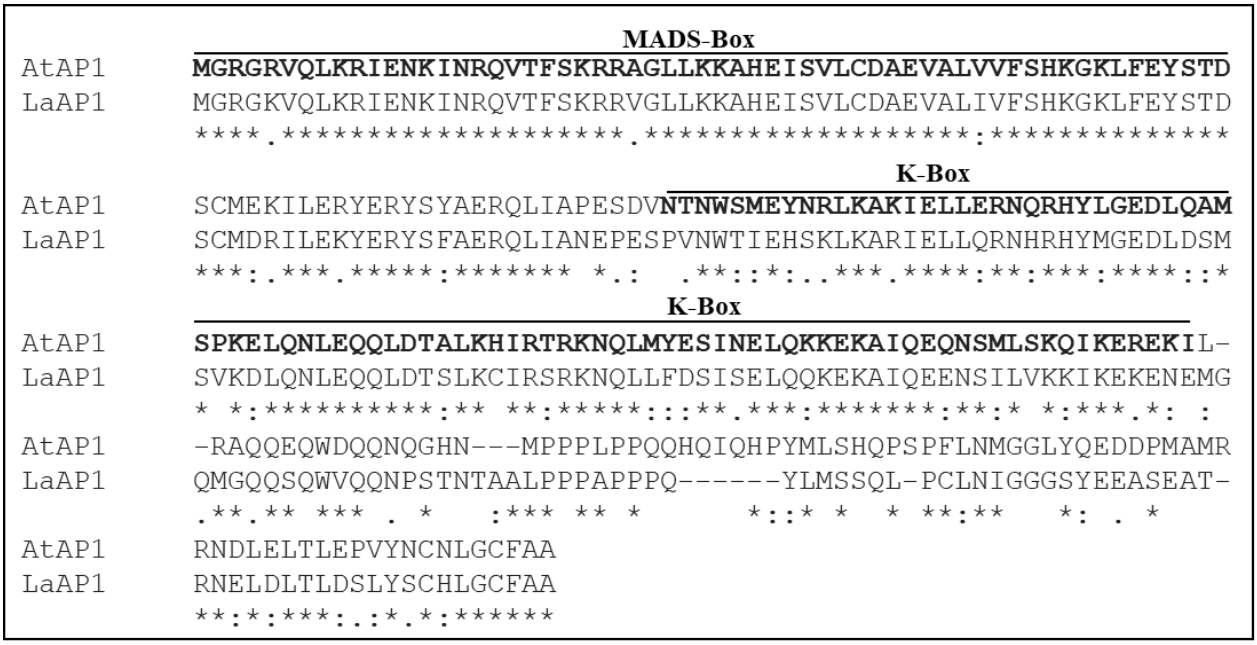

Figure B.8 Multiple Sequence Alignment of protein sequences of *L. angustifolia Apetala1* (*LaAP1*) and *A. thaliana Apetala1* (*AtAP1*). *AtAP1* is a MADS-Box TF involved in sepal and petal development (Theißen et al. 2016). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

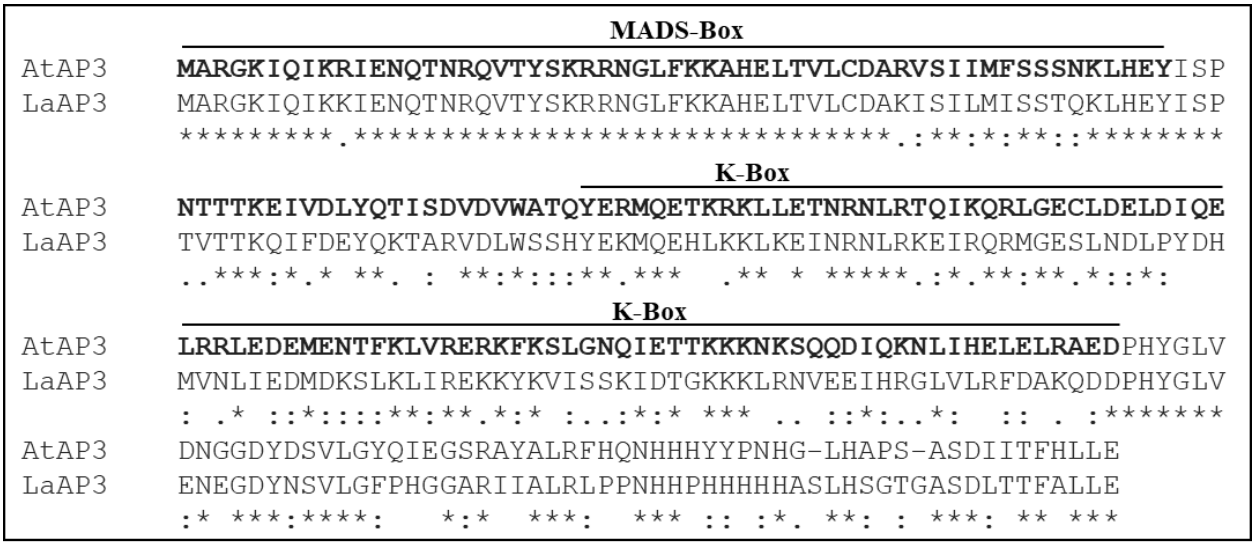

Figure B.9 Multiple Sequence Alignment of protein sequences of *L. angustifolia Apetala 3* (*LaAP3*) and *A. thaliana Apetala3* (*AtAP3*). *AtAP3* is a MADS-Box transcription factor involved in stamen development in *A. thaliana* (Theißen et al. 2016). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

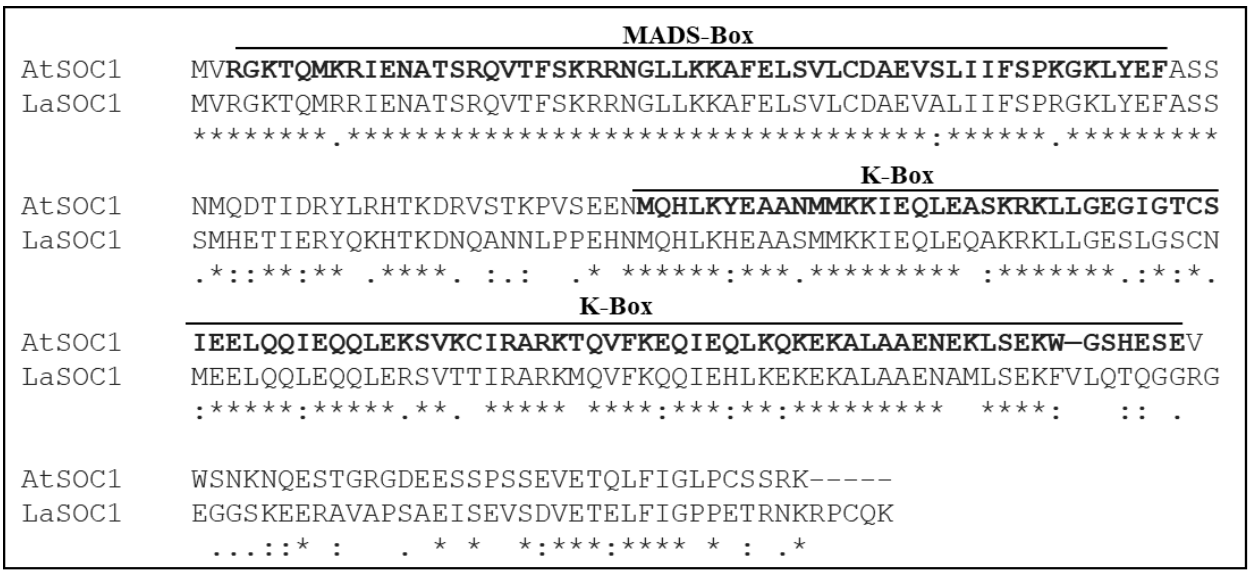

Figure B.10 Multiple Sequence Alignment of protein sequences of *L. angustifolia Suppressor of CONSTANS1* (*LaSOC1*) and *A. thaliana Suppressor of CONSTANS1* (*AtSOC1*). *AtSOC1* is a MADS-Box transcription factor involved in floral transition (Moon et al. 2003a). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

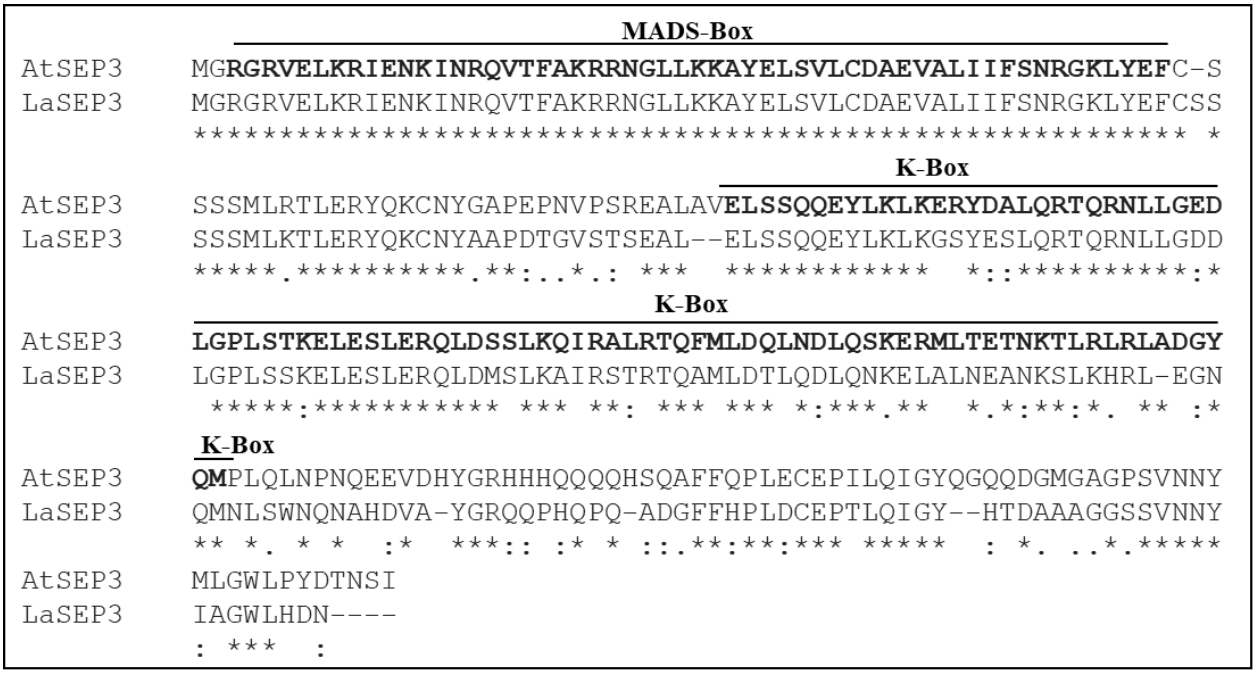

Figure B.11 Multiple Sequence Alignment of protein sequences of *L. angustifolia Sepallata 3* (*LaSEP3*) and *A. thaliana Sepallata 3* (*AtSEP3*). *AtSEP3* is a MADS-Box transcription factor required for floral organ development (Theißen et al. 2016). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and the K-Box domains.

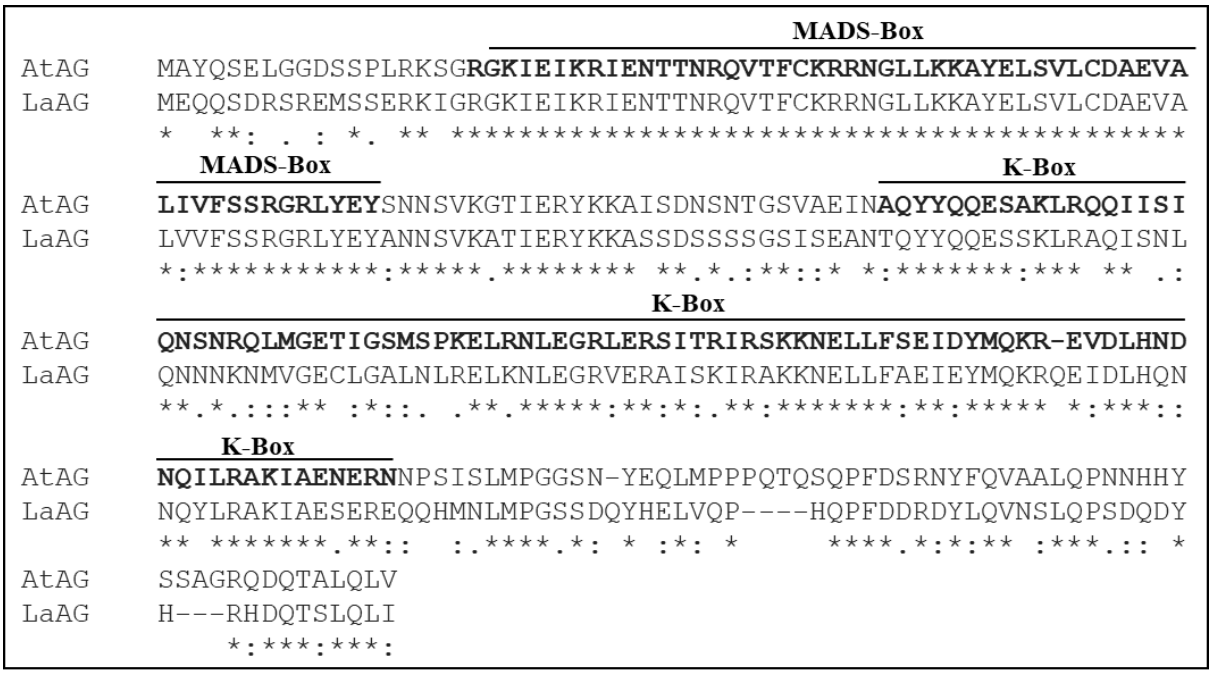

Figure B.12 Multiple Sequence Alignment of protein sequences of *L. angustifolia Agamous* (*LaAG*) and *A. thaliana Agamous* (*AtAG*). *AtAG* is a MADS-Box transcription factor involved in carpel development in *A. thaliana* (Theißen et al. 2016). The bolded portion of protein sequence indicates the sequence of two highly conserved domains of the MADS-Box transcription factors, the MADS-Box and K-Box domains.

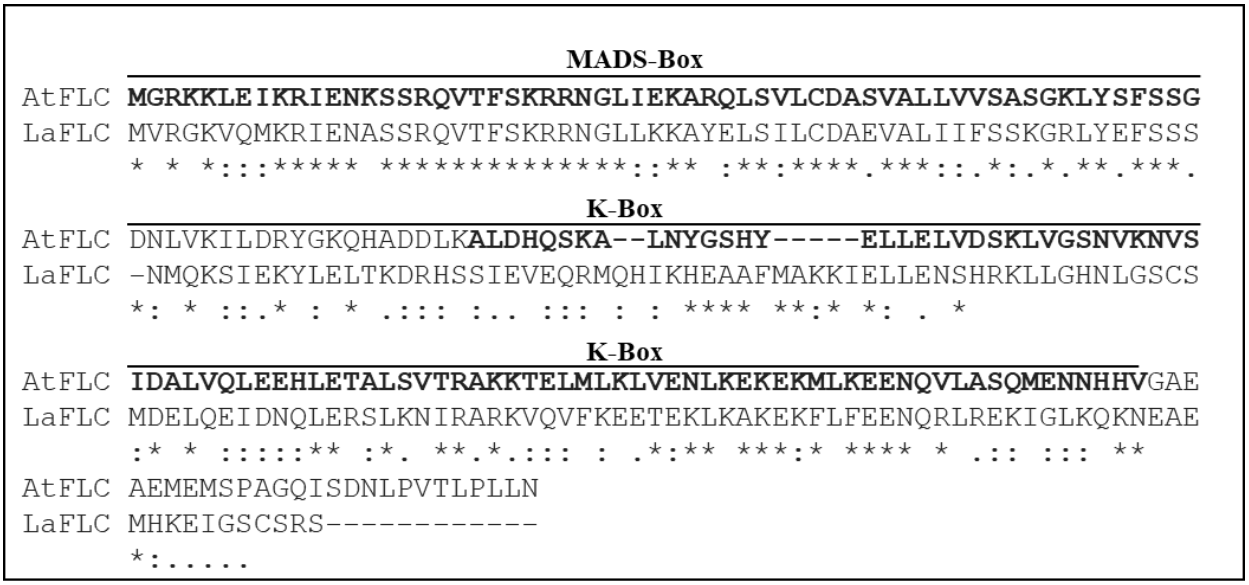

Figure B.13 Multiple Sequence Alignment of *L. angustifolia Flowering Locus C* (*LaFLC*) and *A. thaliana Flowering Locus C* (*AtFLC*). *AtFLC* is a floral repressor in *A. thaliana* (Henderson and Dean 2004).

Appendix C - Supplementary Tables

Table C.1 Genes involved in floral development in *A. thaliana.* Genes include ABCDE-type MADS-Box and other transcription factors including, corepressors, repressors, and floral initiation genes.

Gene	Gene Type	Function	Reference
ABCDE-Type Genes			
APETALA1 (AP1)	MADS-Box A-type function gene Transcription factor	Sepal development. Petal development when combined with B-class genes. Meristem identity gene	(Yanofsky 1995)
APETALA2 (AP2)	MADS-Box A-type function gene Transcription factor	Sepal development. Petal development when combined with B-class genes.	(Theißen et al. 2016)
APETALA3 (AP3)	MADS-Box B-type function transcription factor	Stamen development when combined with *AGAMOUS*	(Theißen et al. 2016)
PISTILLATA (PI)	MADS-Box B-type function gene Transcription factor	Stamen development when expressed with *AGAMOUS*	(Theißen et al. 2016)
AGAMOUS (AG)	MADS-Box C-type function gene Transcription factor	Carpel development	(Theißen et al. 2016)
SEEDSTICK (STK)	MADS-Box D-type function gene Transcription factor	Ovule identity	(Theißen et al. 2016)
SHATTERPROOF1 (SHP1)	MADS-Box D-type function gene Transcription factor	Ovule identity	(Theißen et al. 2016)
SEPALLATA1-3 (SEP1-3)	MADS-Box E-type function gene Transcription factor	Expressed in all whorls	(Theißen 2001)
Repressors			
SHORT VEGETATIVE PHASE (SVP)	Repressor of flowering-time genes	Represses flowering and during vegetative growth phase.	(Hartmann et al. 2000)
Additional Genes			
CLAVATA 1	Floral meristem identity gene	Establishes and maintains floral meristem Maintains floral meristem size	(Clark et al. 1993)
CAULIFLOWER (CAL)	Floral initiation gene	Specification of floral meristem identity	(Yanofsky 1995)
LEAFY (LFY)	Floral initiation gene	Meristem identity gene Aids in stamen and petal identity	(Yanofsky 1995)
FCA	Flowering time controlling gene	Promotion of flowering	(Page et al. 1999)
CONSTANS (CO)	Circadian rhythm gene	Initiation of flowering Flowering time control gene	(Searle and Coupland 2004a)
FLOWERING LOCUS T (FT)	Circadian rhythm gene	Flowering time gene and initiation of flowering in response to day length. Overexpression leads to severe early flowering. Mutation leads to late flowering under long days. Activation of *AP1, AP3, SOC1, FUL, CAL*	(Searle and Coupland 2004a; Mizoguchi et al. 2005)
SUPPRESSOR OF OVEREXPRESSIOELF3N OF CO 1 (SOC1)	Circadian rhythm gene	Initiation of flowering in response to day length. Overexpression leads to severe early flowering,	(Searle and Coupland 2004a)
The early flowering 3 (ELF3)	Flowering initiation in response to photoperiod	Flowering initiation in response to photoperiod. Mutants cause extreme early flowering Regulation of floral identity genes and floral meristem genes	(Zagotta et al. 1996a)

Table C.2 Full-length sequences of *Lavandula* genes identified within the *Lavandula* transcriptome. Source tissue is the tissue from which the gene was originally amplified during construction of the comprehensive *Lavandula* transcriptome, where VM = vegetative meristem, FL = flower, and BD = floral bud and ORF = open reading frame.

Gene name	**Source tissue**	***Lavandula* ORF sequence**
Shoot Meristemless	*L. angustifolia* SAM	ATGCGTGTGTACAACAGTAGCACAGATAAAACAAACAA ACAAACAGAAAACATTTTTTTTCCTGGGTTTATAAAGAG AGAGAGAGGCATATTATTAGCTGGAGTATTAGAGAAAG AGAGAGAAATGGAGCGCGGCGGTGGCGGTGGCTCGAGT AGCAATTCTTGTTTTATGTCTTTCGGAGAAAACAATAAT GGGTTTTGCCCTATACTCATGATGCCTTCGATGGTCTCTT CTCACTCTAATCCGGATCCGGGAAATAACCTATTTCTCC CTCTTACTCATCAAGACATGAACCGCGGCGGCGGCGGC AGTTCGTCGGTGATGGTTGAAGATCAGAATAGCCATAA CAAGAATATTACGACCCCAGGGTATTATTTTATGGAGA ACAGTAATGAAGCCGGTGGTTGTTCTGCTAGGGCTAAA ATCATGGCTCATCCTCACTACCACCGCCTCTTGTCAGCT TACGCCAACTGCCAAAAGGTAGGAGCGCCGCCGGAAGT AGTGGCGAAGCTGGAGGAGGCCTGCGCTTCCGTGGCGG CGATGAACCGGCAAGGCACAAGCTGCGTGGGCGAAGAT CCTGCGCTGGATCAATTCATGGAGGCCTACTGTGAGATG CTGAGCAAGTACGAGCAAGAACTGTCGAAACCCTTCAA AGAAGCCATGCTTTTCCTTTCCAGAATTGAGTGCCAGTT CAAAGCCCTCACCCTCTCGCCTCCCACCTCCGCTGCTTG TAATGATGCAATGGAGACGAATGGGTCATCCGAAGAAG AGACCGATGCCAACAACAGTTTCATAGACCCTCAAGCA GAAGACAGAGAACTCAAAGGCCAGCTCCTACGCAAATA CAGCGGATACTTAGGCAGCCTCAAGCAAGAATTCATGA AGAAGCGAAAGAAAGGCAAGCTGCCTAAGGAAGCTCG GCAGCAGCTGCTCGAATGGTGGAGCCGCCACTACAAAT GGCCTTATCCATCAGAATCACAGAAGCTGGCGTTGGCT GAGTCGACAGGGCTTGACCAGAAGCAGATAAACAACTG GTTTATCAACCAAAGGAAGAGGCATTGGAAGCCGTCTG AGGATATGCAGTTTGTAGTGATGGATGCTGCTCATCCTC ATTACTACATGGATAATGTGTTGGGTAACCCTTTTCCGA TGGATATTTCGCCCGCGCTGCTCTAA
Crab's claw	*L. angustifolia* FL	ATGGATTTGGTTCAACAAGCTTCCGAACATGTTTGCTAC GTCCGCTGCACCTTCTGCAACACCGTTCTCGCGGTTGGG ATTCCATGCAAGAGGCTGATGGACACTCTGACAGTGAA ATGTGGGCATTGCAGCAATCTCTCATTTCTTAGCACAAA ACCTCAGCCTCAAGGACACTGCTATGATCATCAAACAA GCGTTCATTATCAGGCTTTCTGTAATGAGATCAAGAAGG GGCAGTCTGCATCATCTTCTTCATCCACTTCGAGCGAAC CCTTGTCCCCAAAAGCTCCATTTGTGGTCAAACCTCCTG AGAAGAAGCATAGGCTACCATCTGCATACAATCGGTTT ATGAAAGAAGAGATTCAGCGCATCAAGGCGGCTAATCC GGAGATACCACATAGAGAGGCTTTCAGTGCAGCTGCAA AAAATTGGGCTAGGTATATTCCAAACACCCCTCAGGCG TCTGGTTCTAGCAGCATTACTAAGGCCTGA
Cauliflower	*L. angustifolia* BD	ATGGGGAGGGGGAAGGTGGAGCTGAAGAGAATCGAGA ACAAGGTGAACCAGCAGGTGACTTTCTCGAAGAGGCGA

		TCGGGGCTTCTGAAGAAAGCCCATGAGATCTCTGTGTTG TGTGATGCTGAAGTTGCTTTGATTGTTTTCTCTCATAAG GGGAAGTTGTTTGAGTACTCCACTGATTCTTGCATGGAT AGGATCCTTGAGAAGTATGAAAGATACTCATTTGCAGA AAGGCAGTTAATTGCTAGTGAGCCTGAGTCACCTGTTAA TTGGACCCTCGAACACAGCAAACTTAAGGCTAGGATTG AGCTCTTGCAAAGAAATCATAGGCACTACATGGGCGAA GATTTGGACTCGATGAGCGTGAAAGATCTGCAGAATCT GGAGCAGCAGCTTGACACTTCGCTTAAATGCATAAGAT CAAGAAAAAATCAGCTCTTGTTCGATTCGATCTCCGAGT TGCAGCAAAAGGAGAAAGCTATACAAGAGCAGAACAG CATGCTGGTAAAGAAGATCAAAGAGAAGGAGAATGAA ATGGGACAAATGGGACAACAGTCACAGTGGGTGCAGCA AAACCCTAGCACCAACACAGCAGCACTTCCACCACCAG CACCACCACCTCAATATCTCATGTCATCTCAACTCCCTT GCTTAAACATAGGAGGGGGATCGTACGAGGAAGCATCC GAGGCCACAAGGAACGAGCTAGACCTCACTCTTGACTC ACTCTATTCTTGCCATCTTGGATGCTTTGCCGCATGA
Short Vegetative Phase	*L. angustifolia* SAM	ATGGCTAGAGAGAAGATTCAGATCAAGAAAATCGACAA CGTCACCGCCCGGCAGGTCACTTTCTCCAAGAGGAGAA GAGGGCTCTTCAAGAAGGCCCAAGAACTCGCCGTTCTC TGCGACGCCGATGTGGGACTCATCATCTTCTCCTCCACC GGCAAACTCTTCGAATATGCCAGTACAAGTATGGATGA TATAGTGGGAAGGCATAATTTGCACTCCAAGAACCTTA GCAAGTTGGATCAGCCATGCCTCGAGCTGCAGCTAGTG GAGGATAGCAACCACTCCAGGCTGAGCAAAGAAGTGGT TGAGAGAAGCCATCAGCTAAGGCACATGAGAGGAGAA GAACTTCACGAATTAAGTGTTGAAGAGCTGCATCATCTG GAGAAGTCTCTTGAAGTTGGACTAACTCGCGTCATAGA GAAGAAGGGAGAGAAAATCATGACCGAGATTGGTCAG CTTCAAGAAAAGGGGATGGAACTCATGGAAGAGAACA AACGACTAAGGATGCAGGTGAGAGATTTATCGAATGGG GCGAGGAGAATGGGCACAATTACTAGTGAGTCAGAGAT AGTAATGTACGAAGAAGGGCAGTCGTCGGAGTCCGTCA AAAACGCCTGCAACTCCACCGGCCCGCCGCAAGACTAC GACAGCTCATATACATCTCTCAAGCTCGGGCTTCCTTAC ACTGGCTGA
Pistillata	*L. angustifolia* FL	ATGAAAAAGGCTAAGGAGATCAGCGTCTTGTGTGATGC TCAAGTCTCCGTTATCATATTCGCTAGTTCTGGGAAGAT GCACGACTTCTGCAGCCCTTCTACTACGCTGGTTGACAT GTTGGATCAGTACCATAAACTCTCTGGGAAGAGGCTCT GGGATGCAAAACATGAGCACTTGGAGAATGAAATCAAC AGAATCAAGAAAGAGAATGACACCATGCAGATTGAGCT GAGGCACCTGAAAGGAGAAGATATATCGATTTTGAGTT ACAAGGAGCTGATGGGGTTGGAGGAAGCTCTTGAGAAT GGAATCACAACACTCAAAGCCAAGCAGATGGAGTTCGT CCGGATGATGAGGAAACATAATGACATGATGGAGGAGG AGAACCAGAACCTCCTATTTAAGCTGAGGCAGATGCAT TTGGATCCAATGGATGACACCGTACTGGAAGCGCAAGG TGTGTATGATCATCAGGGTGTTGCTGTTGCAGACTACGA AACACACCAGATGCCGTTCGCCTTCCGTGTCCAGCCAAT GCAGCCTAATCTGCAGGAGCGGTTCTAG

Flowering Locus T	*L. angustifolia* BD	ATGCCAAGAGATCGAGACCCTCTAGTGGTAGGTAGAGT TATAGGAGATGTTTTGGATCCATTCACAAGATCCATTGG ACTGAGAGTTGTTTATGGGAACAGAGAGGTTACTAATG GTTGCGAGTTTAGGCCCTCTCAAATTGTCAACCAGCCTA GGGTTGAAGTTGGAGGGGAAGATCTTCGCACTTTCTTCA CTCTGGTCATGGTGGACCCTGATGCTCCAAGCCCCAGTG ATCCTAATCTTAGAGAATACTTGCACTGGCTAGTGACTG ATATTCCAGCAACTACTGGAGCAACCTTTGGACAAGAG ATCGTGTGTTACGAGAGTCCACGGCCTTCGATGGGGATC CACCGCCTGGTCTACATGTTGTTCCGGCAGTTGGGGCGG CAGACGGTCTACGCCCCCGGTTGGCGCCAAAACTTCAA CACGAGGGACTTTGCGGAGCTCTACAACCTTGGCTCCCC GGTGGCTGCGGTGTATTATAACTGTCAGAGAGAGAGTG GAACAGGCGGGAGAAGACGGTGA
Apetala1	*L. angustifolia* FL	ATGGGTAGAGGGAAAGTGCAACTGAAGAGGATAGAGA ACAAAATCAACAGACAAGTAACCTTCTCTAAGAGGAGA GTTGGCTTGTTGAAGAAAGCCCATGAGATCTCTGTGTTG TGTGATGCTGAAGTTGCTTTGATTGTTTTCTCTCATAAG GGGAAGTTGTTTGAGTACTCCACTGATTCTTGCATGGAT AGGATCCTTGAGAAGTATGAAAGATACTCATTTGCAGA AAGGCAGTTAATTGCTAATGAGCCTGAGTCACCTGTTAA CTGGACCATCGAACACAGCAAACTTAAGGCTAGGATTG AGCTCTTGCAAAGAAATCATAGGCACTACATGGGCGAA GATTTGGACTCGATGAGCGTGAAAGATCTGCAGAATCT GGAGCAGCAGCTTGACACTTCGCTTAAATGCATAAGAT CAAGAAAAAATCAGCTCTTGTTCGATTCGATCTCCGAGT TGCAGCAAAAGGAGAAAGCTATACAAGAGGAGAACAG CATTCTGGTAAAGAAGATCAAAGAGAAGGAGAATGAA ATGGGACAAATGGGACAACAGTCACAGTGGGTGCAGCA AAACCCTAGCACCAACACAGCAGCACTTCCACCACCAG CACCACCACCTCAATATCTCATGTCATCTCAACTCCCTT GCTTAAACATAGGAGGGGGATCGTACGAGGAAGCATCC GAGGCCACAAGGAACGAGCTAGACCTAACTCTTGACTC ACTGTATTCTTGCCATCTTGGATGCTTTGCCGCATGA
Apetala2	*L. x intermedia* SAM	ATGGAATATTCCGGCAGCTCTAGTTCCTCCATTGTTAAT GTGGAGACTTCGAGCACCGCCGGCGATGAGAACTCATG CTCCGATCATTTCGTCGGTTTCGATATTTGGAAGAGCAA CTGCGATTACGAGGAGAGTGAGAAATCGAGCTGCGGTT ACGTGACGAAGGAGTTTTTCCCGGGGACCGGAGGGGGG CCGACGCAGTGTTTGGATCTCTCTGAGAACCGGCATGA GCTGATGGAGCAGAGGATTATTTCTCGGGAGGAGCAGC AGCAGAAGCGGCAGCTAGTGAAGAAGAGTCGGCGGGG ACCTAGGTCTCGCAGCTCTCAGTATCGCGGCGTCACATT CTATCGCCGCACCGGACGCTGGGAATCGCACATTTGGG ACTGTGGTAAACAAGTTTATTTAGGGGGATTTGACACTG CACATGCTGCGGCACGGGCATACGACCGTGCTGCAATT AAATTTCGTGGACTGGATGCAGACATCAATTTTAATGTT GATGATTATTATGAAGATTTTAAGCAGATGAAGAATTTA ACCAAGGAAGAATTTGTGCACATACTTCGGCGCAAGAG TACAGGATTCTCAAGGGGAAGTTCGAAGTATAGGGGGG TGACGTTGCACAAATGTGGACGGTGGGAAGCTCGGAT

		GGGGCAGTTTCTTGGCAAGAAGTATATATATCTTGGATT GTTTGACACTGAGATAGAAGCTGCAAGGGCATATGATA AAGCTGCTATCAAGTGCAATGGAAGAGAAGCAGTTACC AATTTTGAGCCGAGTACATATGGTGAAGAGCTGAGCTC TGAGGGTGTCATTGGAGCTGGCGAGAACCACAATCTCA ACCTGAATTTGAGCATTGCTCCCCCTGAAGGAGCTCATT ATGAAATTCATAACATGGAGATGGGAAGCTCTCAGATG CACAGTGGCTTATGTAACCGGCCTGAAGGCGGACTAAA GAGCTCTGCTTCAGCCTCTGTGGCAGCTCAATCTTCCCA TGAACATGTAAATGTATGTGAATACCCTTTTGTTTGGAG GGGAGTCGATTCTTGTCTTCCTCCCATTTATAAGGGAAC AACAACAGAGAAGGTTGCGGAACTTGATTCTAATGAAC ATTGGACTAGAAACTCCCCGAACAATCTTCATGGGGATT ATCCTTTACCATTCTTCTCTACTGCAGCATCATCAGGATT CTCTACTTCATCAGCTGGTATCCATCAGTTGCCCTTCAG CTCGAGATCCCCCAACCATTATCTTCCACCTGTCACCAA TTTCAACAATGGATCTCATTACTACTGGAGGTGA
Apetala3	*L. angustifolia* BD	ATGGCTCGTGGGAAGAT**C**CAGATCAAGAAAATAGAGAA CCAAACAAATAGGCAGGTAACTTACTCCAAGAGAAGAA ATGGTTTGTTCAAGAAAGCTCATGAGCTTACTGTTCTTT GTGACGCAAAGATATCTATCCTTATGATCTCCAGTACTC AGAAACTTCATGAATATATCAGCCCCACAGTCACGACA AAGCAGATTTTTGATGAATACCAGAAGACTGCTAGGGT TGATCTTTGGAGCTCCCACTACGAGAAAATGCAAGAGC ACTTGAAGAAGCTGAAGGAGATTAACAGGAATCTGAGA AAGGAGATTAGGCAACGGATGGGAGAAAGCTTGAACG ACCTACCGTACGACCATATGGTTAATCTCATCGAAGATA TGGACAAGTCTCTCAAGTTGATTCGCGAAAAAAAGTAC AAAGTCATCAGCAGCAAAATTGATACCGGCAAGAAAAA GCTGAGGAATGTTGAAGAAATACATAGAGGCCTTGTGC TTCGATTTGATGCAAAGCAGGATGATCCACACTACGGCT TAGTTGAGAATGAAGGTGATTATAATTCCGTTCTGGGAT TTCCACACGGCGGGGCAAGGATCATCGCGCTTCGCCTG CCGCCTAATCACCACCCTCACCACCACCACCATGCTAGC CTTCACAGTGGAACCGGAGCCTCAGACCTCACTACATTT GCCTTGCTTGAGTGA
Suppressor of CONSTANS1	*L. x intermedia* SAM	ATGGTGAGGGGGAAGACTCAGATGAGGCGTATAGAGA ACGCCACAAGCAGGCAGGTGACCTTCTCAAAGAGGAGA AATGGGCTGCTCAAGAAGGCCTTTGAGCTCTCAGTTCTT TGTGATGCTGAGGTGGCACTCATCATCTTCTCTCCTAGA GGCAAGCTCTATGAATTTGCAAGCTCAAGCATGCATGA AACGATCGAGCGCTACCAGAAGCACACGAAAGATAATC AAGCCAACAACCTTCCTCCCGAGCACAACATGCAGCAC TTGAAGCATGAAGCAGCAAGCATGATGAAAAAAATCGA GCAGCTAGAACAAGCAAAGCGGAAGCTGCTTGGGGAA AGCTTAGGCTCATGCAACATGGAGGAGCTGCAGCAGCT AGAACAGCAGCTCGAGCGGAGTGTCACCACCATTCGTG CAAGGAAGATGCAAGTGTTCAAGCAGCAGATCGAGCAC TTGAAGGAGAAGGAGAAAGCCCTAGCTGCTGAGAATGC AATGCTATCTGAGAAGTTTGTGCTTCAAACGCAAGGAG GGAGAGGGGAAGGAGGGTCGAAGGAGGAGAGGGCGGT GGCGCCGTCGGCGGAGATAAGTGAGGTATCGGATGTGG

		AGACGGAGTTGTTTATTGGGCCGCCTGAGACGAGGAAT AAACGGCCGTGCCAAAAATAG
Sepallata 3	*L. angustifolia* FL	ATGGGAAGAGGGAGAGTTGAGCTGAAGAGGATAGAGA ACAAGATCAACAGGCAAGTGACCTTTGCTAAGCGAAGG AATGGGCTTTTGAAGAAAGCTTACGAGCTTTCGGTCCTT TGTGACGCTGAGGTTGCTCTGATTATCTTCTCCAATAGA GGAAAGCTGTACGAGTTTTGCAGCAGCTCTAGCAGCAT GCTCAAGACACTGGAAAGGTATCAGAAATGCAACTATG CAGCTCCAGATACCGGTGTATCGACAAGCGAGGCTCTG GAACTTAGTAGCCAGCAGGAATATTTGAAGCTAAAAGG AAGCTATGAATCCCTGCAACGCACTCAAAGGAACCTTC TTGGTGATGATCTTGGTCCTTTGAGCAGCAAGGAGCTCG AATCACTCGAGAGGCAACTCGATATGTCACTCAAGGCG ATCCGATCAACACGGACTCAAGCAATGCTGGATACACT TCAAGATCTTCAGAATAAGGAGCTTGCCCTCAATGAAG CTAACAAGAGTTTGAAACATCGGTTGGAGGGGAATCAA ATGAATCTATCATGGAATCAGAATGCTCATGATGTGGCC TATGGAAGACAGCAACCACACCAGCCTCAAGCTGATGG CTTTTTTCATCCTTTGGACTGTGAACCCACTCTGCAAATT GGGTACCATACTGATGCAGCAGCAGGAGGGTCTAGTGT GAATAACTACATAGCAGGGTGGCTACATGATAACTAA
ARF5	*L. angustifolia* FL	ATGGGTCTTCTCAGCATAATCCGGAAAATTAAGCGCAA AGAAAAGGAGATGCGCATTCTCATGGTGGGTCTTGATA ACTCTGGAAAGACGACAATAGTGATGAAGATTAACGGT GAAGACACTAGTATAATCAGTCCTACACTTGGATTCAAC ATCAAAACCATGGCGTACCAAAAGTATACTTTGAATAT ATGGGATGTTGGAGGGCAGAAAACCATAAGGTCTTATT GGAGGAATTATTTCGAGCAAACTGATGGGATGGTCTGG GTGGTGGATAGTTCAGACCTGCGAAGGCTGGATGATTG CAAATATGAGTTAGATAATCTTCTCAAAGAGGAGAGAC TATCAGGAGCATCATTGTTGATTTTTGCAAATAAGCAAG ACATACAAGGTTCCCTGTCGCCTGATGAGATAGCTAGA GTGCTCAACTTAGAATCCATGGATGGAAGCAGACACTG GAAAATAGTTGGGTGTAGTGCAGTGACTGGGGAAGGCC TGCTCCAAGGCTTCGATTGGTTGGTTCAAGACATTGCCT CCAGAATTTACGTGCTTGACTAA
Agamous	*L. angustifolia* FL	ATGGAACAGCAAAGTGATCGATCAAGGGAGATGTCATC AGAGAGGAAGATTGGGAGGGGGAAGATCGAGATCAAA CGGATCGAGAACACGACGAACCGGCAGGTAACTTTCTG CAAAAGGCGCAATGGTTTGCTGAAGAAGGCGTATGAAT TATCAGTGCTTTGTGATGCTGAGGTTGCTCTAGTTGTTTT CTCAAGCAGAGGCCGCCTCTATGAGTATGCTAACAACA GTGTTAAAGCAACCATTGAGAGGTACAAGAAAGCATCC TCAGATTCCTCAAGTAGTGGCTCTATTTCTGAAGCCAAT ACTCAGTATTACCAGCAGGAGTCTTCAAAACTTCGTGCA CAGATCAGTAATTTGCAGAATAATAACAAAAACATGGT TGGCGAATGTTTAGGGGCTTTAAATCTGAGAGAGCTCA AGAATCTGGAAGGCAGAGTTGAAAGAGCAATTAGCAAA ATCCGAGCCAAGAAGAATGAACTATTGTTTGCTGAGAT CGAGTACATGCAGAAGAGGCAGGAGATTGACTTGCATC AGAACAATCAGTACCTTCGAGCAAAGATAGCTGAGAGT

GAGAGAGAGCAGCAGCACATGAACTTGATGCCTGGTAG
TTCTGATCAGTATCATGAGCTCGTGCAACCTCATCAGCC
GTTCGATGATCGGGATTATCTTCAAGTCAATAGCCTCCA
ACCCTCAGATCAGGATTACCATAGGCACGACCAGACAT
CTCTTCAGTTAATCTGA

Embryonic flower 2 *L. angustifolia* SAM ATGCCTGGCATTCCTCTCGTGGCTCGTGAGACTGCGAAT
TGTGTTTGCAATTGCAGTTACTCTAGAGGTGCAGACCAT
ATGTGTCGCCAAGATCCCCGTGCGCATTTGTCCCCTGAA
GAGCAAATTGCAGCTGAGGAAAGCCTATCAATTTACTG
CAAGCCTGTTGAATTGTATAACATTCTTCAGCGCCGTGC
TGTTAGAAATCCATCCTTTCTTCAGAGATGTTTGCGGTA
CAAAATACAAGCAAAGCACAAGAAAAGGATTCAAATG
ACAATCTCCCTACCTTCAACTGTAAATGATGAATCTCGA
ATTCAAAGTCTATTTCCTTTGTATATAATGTTGGCAAGA
CCCGTTTATAGTCCTTCAGTTGTAGAGGAATCTGCAGTT
TATCGATTCAGTCGAGAATGTATATTGACTAAATGTACG
GGAGATGATGCAACGAACCAAGATGAAGCAAATTTTAT
TCTTCCTGAGATTAATAAACTATCAGCTGAAGTCAAATC
TGGCTCTCTTTCGATTTTGTTTGTTAGCTGCGCTGAATTG
TTTAAGGAGCATCGGGATGAGCTGTTATTTCCTGCTAAT
GCTGGAGGTCACTGCTTTCTGGGTAAGATTCCAATGGAG
TTACTTCATTTGTCATGGGAGAAGTCTCTGAACTTGAGT
TTCGGCAAGAGAGCCGAGATGCTGTCTACTGTTGACTTA
CATTCCTGCTTCCTGAAGACGGGCTGTTTTGGAGAAGAC
AAATGTTTTTCATTTCATAGTCCTCATGCTGGAGCTATG
CCTATGTCTCAACAACTGCAAGTCCGCATTGCCGCTGAA
GAACTTGGCTTCAGGGAGAGATCTGCTTATGATTCGTTT
GCATTCAGTGACATTCCTTCCTCTACTTTGCCTCATATTA
TCAGGTTGAGAACTGGAAATGTTGTCTTCAACTACAGAT
ACTATGATAACAAGTTGCAGAGGACCGAAGTGACGGAG
GACTTTGCATGCCCTTTCTGTTTGGTCAAATGTGCAAGC
TTTAAGGGTTTGAGATATCATTTGCCCTCCAGTCATGAC
TTATTCAGCTTTGAGTTTTGGGTGACAGAGGAATTTCAA
GCTGTCAATGTATCTGTCAAGACTGATACGTGGAGATCT
GAGATCGTTGCGGCTGGTGTTGACCCTAAGCAACAAAC
ATTTTTCTTTTGTTCTAAGGCACCACGACGTAGAAAAAT
GAAGAGTCCGCTGCAAAAATCAAAGTTTGTGCATCCTCT
TGTCTTAGATTCAGAAATGCTGGGGTCAGCCAATGAGCT
TAGAGAAAAGACCATTGGCGTGGAAGAGAGTGCGGACT
GTGATGCATCAAGTCCCATAGTTTCATCTGCCACAGCTC
ATTCATATGCAGACCCTGAATGTGTCCAGTCATTACCTG
GAAGCAATCTTGCACCTACAGCATTGCTACAGTTTGCAA
AGACAAGGAAATTGTCAGTTGAGCGCTCTGATCCTAGG
AATCGTGCACTCTTAAAGAAGCGGCAGTTCTTCCACTCT
CATAGAGCACAACCCATGGCCATGGAGCAAGTCCTATC
AGATCGCGATAGTGAGGATGAAGTTGATGATGATGTTG
CAGATCTTGAAGACCGAAGGATGCTGGATGACTTTGTG
GATGTCACAAAAGATGAGAAGCAATTGATGCATCTTTG
GAACTCTTTTGTTAGGAAGCAAAGGGTACTGGCAGATG
GTCATATCTCTTGGGCATGCGAGGCTTTCTCCAAATTGC
ACAGACAGGATCTTGTTCAGACACCTGCCTTACTTTGGT
GCTGGAGATTATTTATGATCAAGCTGTGGAACCATGGCC
TTCTGGACGCTCGTACGATGAACAGCTGTAACGTAATTC
TCGAGCAATCGGACCCACAGAGC

Flowering locus C	*L. angustifolia* FL	ATGGTGAGGGGGAAGGTGCAAATGAAGAGAATTGAGA ATGCATCAAGCCGACAAGTAACCTTTTCGAAGAGGCGA AATGGCCTGCTGAAGAAGGCGTACGAGCTCTCGATTCT CTGCGATGCAGAAGTGGCTTTGATCATCTTCTCTTCTAA AGGGAGACTCTATGAATTCTCAAGCTCCAACATGCAGA AGTCAATAGAGAAATATCTGGAACTTACAAAAGATCGC CATTCTAGCATTGAAGTTGAACAACGCATGCAGCATAT AAAACATGAAGCAGCATTCATGGCAAAGAAGATAGAGC TATTGGAGAATTCACACCGGAAGCTTTTAGGGCATAATC TGGGATCATGCTCCATGGATGAGCTTCAAGAGATTGAT AACCAGTTGGAGAGAAGCCTTAAGAACATAAGGGCCAG AAAGGTTCAAGTATTCAAGGAGGAAACAGAGAAACTAA AAGCAAAGGAGAAATTTCTCTTCGAAGAAAACCAAAGA CTCAGAGAAAAGATTGGGTTGAAGCAAAAGAATGAAGC AGAAATGCATAAAGAAATTGGGAGTTGCAGCCGGAGCA CGGTGAGCGCGGAGGCGGTGACG
Late elongated hypocotyl	*L. angustifolia* BD	ATGGATCCTTATTCTTCTGGAGAAGACCCAGTCATCAAG ACAAGAAAACCTTATACAATTACTAAGCAGCGTGAGAG ATGGACTGAAGAGGAGCATAATAGGTTTCTAGAAGCTC TCAAGCTCTATGGTCGTGCTTGGCAGCGAATAGAAGAA CATATAGGAACCAAGACTGCTGTGCAGATAAGAAGTCA TGCACAAAAGTTTTTTACAAAGTTGGAAAAGGAGGCTA TTGTTAAAGGAGTTCCAATAGGACAATCTCTTGATATCG ATATCCCTCCACCACGGCCTAAAAGAAAACCAAACAAT CCTTATCCTCGAAAGACAAGTGTGGGAACTCCCGCTTCC CAGGTTGAAGTGAAGGATGGACTATCAAATCCTATTTCT TCCTCATGTCAGAGTAAAGTGACCATGAATGCGGAGAG AGAACCTGAGCAATCTTGCGACTATGCGAAATTCGGAA ATAGTAAATACCATGACAAGGATAATTTTTCTGAGGCTT TCACTACTCTCAAGGAAGCTTCCTGTGAATCCCCATCGT CAGCATACAAAACTGCCTCCTTATCTGATGCTCCTAGAA AATCCTGTACATTTCGGCAATTTGCCCCACTCTCTAGAG AAGCAATGACTCAAGATGTAACCACTGCAATCCATGTT ACTACCGAAGCCAAAGGGATGAAGAAGTCTGATATTGA ACAGCCTTTCCAGGATAATCTCACTTTGAGGGTCTCAAA TGTTGGTGATCCATACCTTTTACAAGAGAAATCTGTCCA TACTAGAAAAGTAGATGAGTTGAAACAGACTGAAAATA ATGGAAGTGAAACTACTGCTGTGCAGGCCTCTCAGAAC TATCCTCGCCATGTTCCTGTGCACATTCTTGATGGGAGC TTAGGAATGAGGGCATACAATATCTCCCCCGATGTGTCA TATGCAGATTTTCGCCATATGGGTGGAGTCCATGCTCAT CAAAATCTATTCATGAACCCAGCTTCATCTGCTTCTTCT GAGCATTATAGCAATGCATCAAAATCTTCTGTTCACCAG GCATTTCCTAGTTATCATCCTGTCTTCGCTCCTATCCAAA GCCAAGACGATCCCCAGTCGTGTCTGCATGTATCACCGG CGTTTTCCAGTCTAATTGTGTCTTCTCTCTCCCAAAATCC TGCAGCCTATGCCGCAGCTAGTTTTGCAGCAACTTTGTG GCCGTCTCTGAACGTGGAAGCTTCAGTTGAGCCATCTAC AACTCCCAGTATGGCTGCAATTGCTGCAGCAACTGTAGC AGCAGCAACTGCATGGTGGGCAGCCCATGGCCTACTTC CTTTATGTGCTCCATATCATCCTGGCTTTACCTGTTCTCC TGCATCAGCATCTGCAGCTCCAATACCTAGCAGCCGAG GTATAACTGTTAACAATGAAAGAAGGGAGAGTCCTGAT

CCTACCTTAGGGGGTCAGCAACTGGAACCGGAATGTTC
GGAGGCTCTGCAAGAACAGCATTCAGTTTCAAAGTCAC
CATTGTTATCCTCATCAGATTCTGAAGAAAGTGAAGGTG
CAAAAAAGAATGTTGAGTCAACCGCTCCTGAAACAACA
AAAGATGAGGTTACTACGACTGAGCCGCATGATGCAGA
CAAGTCAACTAACAAAAAGCTTGTGGACCGTTCTTCATG
TGGATCAAATACACCTTCCAGCAGTGAAGTAGAAGCTG
AGGCTTTGGAAAAGGATACAAAGGAGAAAGACGAAAA
ACATATGAACGAGAAAGAGAAGCAAGAAGAACTTGAT
GCTATGCTTAATGTTAGTGATCCATTTGGTCGTCGTTGT
AGAAGCACAAGCAATGCTGGTGCCGGTGACTCTTGGAA
GGAAGTATCAGAGGAGGGCCGACTGGCATTTGACGCGC
TCTTCTCTAGAGAAATACTACCACAGAGTTTTTCACCTC
CACATAAGGGCAAGAAAACCTGTAACAAGGAAAGTGA
GAAGGATGAAAGCAGAATGCAGTTAGACCTTAATGAGA
AGACATGGGATACATGTTCACCCTTGATTAGTGAAACC
ATGAAAGAGGGAGCCCTACATACAGGACTAGAACTAGG
ACAAGGTAAGCTAAAAGCTCGACGGACAGGATTTAAAC
CTTACAAGAGATGTTCAGTGGAAGCCAAGGAGTTCAAG
GTGTCAGCCAACGGCCAAGATGAAGAGAAATGTCCGAA
ACGATTACGCTTGGAAGGAGAGGCTTCTACATGA