2016

TRANSCRIPTIONAL REGULATION OF SPECIALIZED METABOLITES IN *ARABIDOPSIS THALIANA* AND *CATHARANTHUS ROSEUS*

Craig M. Schluttenhofer
University of Kentucky, craig.schluttenhofer@uky.edu
Digital Object Identifier: https://doi.org/10.13023/ETD.2016.451

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Craig M. Schluttenhofer, Student

Dr. Ling Yuan, Major Professor

Dr. Arthur Hunt, Director of Graduate Studies

TRANSCRIPTIONAL REGULATION OF SPECIALIZED METABOLITES IN *ARABIDOPSIS THALIANA* AND *CATHARANTHUS ROSEUS*

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Agriculture
at the University of Kentucky

By

Craig Michael Schluttenhofer

Lexington, Kentucky

Director: Ling Yuan, Professor of Plant Physiology

Lexington, Kentucky

2016

ABSTRACT OF DISSERTATION

TRANSCRIPTIONAL REGULATION OF SPECIALIZED METABOLITES IN *ARABIDOPSIS THALIANA* AND *CATHARANTHUS ROSEUS*

For millennia humans have utilized plant specialized metabolites for health benefits, fragrances, poisons, spices, and medicine. Valued metabolites are often produced in small quantities and may command high prices. Understanding when and how the plant synthesizes these compounds is important for improving their production. Phytohormone signaling cascades, such as jasmonate (JA) activate or repress transcription factors (TF) controlling expression of metabolite biosynthetic genes. TFs regulating specialized metabolite biosynthetic genes can be manipulated to engineer plants with increased metabolite production.

WRKY transcription factor are known components of both JA signaling cascades and regulation of specialized metabolism. The presence of WRKY binding sites in promoters of several terpene indole alkaloids suggested their involvement in regulating biosynthesis of these compounds. A phylogenetic analysis was used to compare Arabidopsis and Catharanthus WRKY TFs families. Gene expression analysis identified WRKY TFs induced by JA in both Arabidopsis and Catharanthus, providing candidates for future characterization. WRKY TFs suggest a possible conserved regulatory network of TFs downstream of JA signaling cascades.

The origin and conservation of JA signaling in plants remain ambiguous. Identification of the first algal TIFY factor helped determine when JA signaling appeared. The charophyte, *Klebsormidium flaccidum* does not possess genes encoding key green-plant JA signaling components, including CORONATINE INSENSTIVE1, JASMONATE-ZIM DOMAIN, NOVEL INTERACTOR OF JAZ, and the JAZ-interacting bHLH factors, yet their orthologs are present in the moss. A molecular clock analysis dated the evolution of JA signaling evolution to during the early Ediacaran to late Cambrian periods 628 to 491 million years ago – a time corresponding to rapid diversification of animal predators.

The plant Mediator complex is a core component of gene expression. Conservation of the MED25 subunit in plants, and its known involvement in JA signaling implicates this factor in regulation of specialized metabolism. MED25 is involved in anthocyanin accumulation, but how it functions remains unknown. Characterization of MED25 in Arabidopsis revealed it interacts with the transcription factor GL3 as well as the JAZ1 repressor. Importantly, the interaction of

JAZ1 with MED25 reveals a new mechanism by which JAZ proteins regulate gene expression, improving our understanding of JA signaling.

KEYWORDS: *Arabidopsis thaliana*, *Catharanthus roseus*, jasmonic acid, MEDIATOR, phenylpropanoids, terpene indole alkaloids

Craig Michael Schluttenhofer

December 2nd, 2016

Date

TRANSCRIPTIONAL REGULATION OF SPECIALIZED METABOLITES IN *ARABIDOPSIS THALIANA* AND *CATHARANTHUS ROSEUS*

By

Craig Michael Schluttenhofer

Ling Yuan, Ph.D.
Director of Dissertation

Arthur Hunt, Ph.D.
Director of Graduate Studies

December 2nd, 2016
Date

ACKNOWLEDGEMENTS

I would like to give a huge heartfelt thanks to all the people that helped make my time at the University of Kentucky a success. First off, I would like to specially thank all the member of the Dr. Ling Yuan laboratory. Without their help and guidance, I would not have been able to successfully learn and perform the many experiments these projects required. Additionally, they have all become such great friends and I cherish the times we have spent together. I cannot begin to express all my gratitude to Dr. Barunava Patra and Dr. Sitakanta Pattanaik for the time they took out of their schedules to spend with me at the hospital when my health was poor. I also want to express a sincere thanks to Dr. Ling Yuan for allowing me the opportunity to work in his laboratory. I have learned more than I ever imagined under his guidance.

I would also like to thank God and acknowledge all my friends at the Holy Spirit Parish on campus. The Newman Center has provided much needed comfort and relaxation during stressful times. I would like to especially thank Fr. Steve Roberts, Fr. Albert DeGiacomo, and Fr. Alan Carter for their inspiration and advice over the years.

Finally, I would like to thank all my family. Without their support this work would not have been possible. In particular, the love, patience, and wisdom of my mom, Roberta Schluttenhofer, and dad, Michael Schluttenhofer, has help guide me through my time at Kentucky. Word cannot express my love for you, mom and dad. I love you so much. I also want to thank my sisters for their love and support. Most importantly, I would like to acknowledge my fiancée Hannah Wiedemann. Without her love and support much of this work would not have been possible. Thank you for always standing by my side and enabling me to do my best. I cannot express how much I look forward to spending our lives together. I would also like to thank your parents, Randal and Martha Wiedemann, for being so welcoming to me. They have truly become like second parents to me. I look forward to becoming part of your family and getting to know you all better over the coming years.

This dissertation is dedicated to God, Hannah Wiedemann, and all my family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

Use and Value of Medicinal Plants

Often considered natures pharmacy, plants are excellent metabolic factories. As sedentary organisms, plants produce many compounds to defend themselves against various stresses. These compounds, termed specialized or secondary metabolites, while not essential for growth, are critical to plant competitiveness in their natural habitats. Specialized metabolites may serve a number of roles within the plant. Plant produce three primary classes of specialized metabolites, alkaloids, phenylpropanoids, and terpenes. The plant kingdom synthesizes over 12,000 alkaloids, 8,000 phenylpropanoids, and 25,000 terpenes (Zwenger, 2008). Furthermore, plants can fuse compounds from different classes to make more complex molecules such as terpene indole alkaloids (TIA), the combination of a terpene moiety with an alkaloid compound. Many phenolic compounds possess UV absorptive functions that protect the plant from photooxidative damage to DNA and proteins (Li et al., 1993; Landry et al., 1995). Specialized metabolites serve as insecticidal, antibacterial, or antifungal compounds to resist abiotic stresses (Roepke et al., 2010; Schweizer et al., 2013). Some are protectants against drought or cold (Nakabayashi et al., 2014; Schulz et al., 2016). A few specialized metabolites serve as a source of inter-plant communication (Muroi et al., 2011). Still others function to prevent competition from nearby unrelated plant species (Rietveld, 1983; Jose and Gillespie, 1998). Moreover, specialized metabolites serve to attract pollinators or seed dispersers (Harborne, 2001; Tsahar et al., 2002; Sheehan et al., 2016). Any given specialized metabolite may possess one or more of these properties to aid the plant's survival and success.

In addition to benefiting the plant, specialized metabolites are valuable to humans. As components of our diet they provide essential nutrients or antioxidant properties beneficial to human health (Zafra-Stone et al., 2007; Hounsome et al., 2008). Specialized metabolites serve as spices for flavoring food and fragrances for perfumes (Schwab et al., 2008; Schilling et al., 2010). Additionally, some specialized metabolites possess pharmaceutical properties and can be used as supplements or drugs to improve human or animal health. Indeed, 27% of pharmaceutical drugs are directly derived from specialized metabolites, albeit not just from plants, but also fungi and bacteria (Newman and Cragg, 2012) (Figure 1.1). Another 45% of pharmaceutical drugs are inspired by specialized metabolites (Newman and Cragg, 2012) (Figure 1.1). Particularly for anti-cancer drugs, naturally occurring compounds remain a key source of chemical diversity. Many anti-cancer drugs are derived, or partially derived, from plants including paclitaxel, vinblastine, and, vincristine (Newman and Cragg, 2012). Thus, plants have been and remain a key source of chemical diversity to exploit for pharmaceutical and other human needs. Furthermore, due to the high degree of chemical complexity many of these metabolites (e.g. vinblastine) cannot be completely synthesized *de novo*. Thus, the metabolites, or their precursors, must be derived from plant material.

Utilizing plants for drug production typically leads to two critical problems: 1) the high cost of handling and processing plant materials increases final drug costs to patients and 2) inconsistent supplies of plant materials can lead to reduce production and drug shortages. Indeed, several plant-derived anti-cancer drugs including etoposide, paclitaxel, vinblastine, and vincristine, pharmaceuticals have experienced drug shortages in recent years (Chabner, 2011; Ventola, 2011). Efforts are needed to reduce costs and drug shortages to maintain consistent drug availability for patients. Both problems can partially be addressed by increasing metabolite concentration with the raw plant material. Higher metabolite concentrations in raw materials can

potentially increase ease of product extraction, thereby reducing processing costs to obtained the purified chemical. Similarly, higher metabolite concentrations allow less material to be processed. A need for less plant material can be valuable if drug shortages occur as growers can fulfill the required volume of raw product sooner.

In plants, increasing metabolite concentration can be accomplished by either targeting steps in the biosynthetic pathway (e.g. increasing expression of the rate limiting enzyme) or through manipulation of regulatory factors controlling the biosynthetic pathway. Alternatively, a combination of approaches can be used. Each method has demonstrated success. In *Artemisia annua*, overexpression of the mevalonic acid rate limiting enzyme 3-hydroxy-3-methylglutaryl CoA reductase (HMGR) increased artemisinin production 22% (Aquil et al., 2009). When HMGR and the artemisinin rate limiting step were combined a 7.65-fold increase was observed (Alam and Abdin, 2011). In *Artemisia annua*, overexpression of AP2-ERF, NAC, or WRKY transcription factor (TF) increases anthocyanin accumulation 67%, 79%, or 1.9-fold, respectively (Yu et al., 2012; Han et al., 2014; Lv et al., 2016). Overexpressing both a rate limiting step and a TF in *Catharanthus roseus* lead to higher levels of TIAs (Pan et al., 2012). Effective engineering to increase metabolite accumulation by targeting specific enzymes benefits, perhaps even requires, characterization of the entire biosynthetic pathway. The major advantage of this method is the potential to precisely manipulate a biosynthetic pathway towards a desired product. In contrast, TF can regulate expression of both known and unknown biosynthetic steps simultaneously. Therefore, the entire biosynthetic pathway does not need to be characterized. However, this may come with the expense of pleiotropic effects on plant growth, development, or non-target metabolites. Nevertheless, the later method may be particularly useful for species, which harbor long or complex biosynthetic pathways leading to the desired metabolite.

Identifying candidate TFs for metabolic engineering projects can take two approaches: top-down and bottom-up. The first approach, bottom-up, is to start with enzymes. Promoters taken from these enzyme biosynthetic genes are then used to find the corresponding regulators. Promoters of these regulators are used to find higher level regulators. Eventually, one factor controlling the entire, or majority, (master regulator) of the pathway is found. The top-down approach starts near the anticipated top of a regulatory cascade by using a factor that controls one or more major response networks. This factor can be used to identify various levels of subordinate factors until a master regulator of the desired pathway is found. A hallmark feature of many specialized metabolic pathways is their regulation by the phytohormone jasmonate (Aerts et al., 1994; Singh et al., 1998; Qi et al., 2011; Lenka et al., 2012), thus providing a potential starting point for the top-down approach.

The Jasmonate Signaling Network

Phytohormones are important signaling regulators that mediate plant responses to stress. Jasmonate (JA) is a key stress signaling phytohormone which prioritizes plant defense over plant growth by modulating levels of other hormones (Lorenzo et al., 2004; Sun et al., 2009; Yang et al., 2012). Specifically, JA regulates specialized metabolism and symbiosis with microorganisms, as well as mediates tolerance to wounding, pathogens, and insect herbivory (Feys et al., 1994; Thomma et al., 1998; Lorenzo et al., 2004; Dombrecht et al., 2007). JA also regulates developmental programs for male fertility, root growth, seed germination and senescence (Staswick et al., 1992; Xie et al., 1998; Li et al., 2004; Miao and Zentgraf, 2007). The JA signaling cascade has been well characterized in the model plant *Arabidopsis thaliana*. The

core JA signaling cascade is comprised of the conjugated signal molecule (JA-isoleucine), the CORONITINE INSENSITIVE1 (COI1) protein, JASMONATE ZIM DOMAIN (JAZ) proteins, and the bHLH transcription factor (TF) MYC2. COI1 is a leucine-rich repeat containing F-box protein that recognizes and interacts with JAZ proteins(Chini et al., 2007; Thines et al., 2007; Sheard et al., 2010). Along with S-PHASE KINASE-ASSOCIATED PROTEIN1, CULLIN, and RING-BOX PROTEIN1, COI1 forms a complex which polyubiquitinates JAZ proteins, marking them for degradation via the 26S proteasome (Chini et al., 2007; Thines et al., 2007). JAZ proteins belong to a larger TIFY family and are characterized by TIFY and JAS domains (Bai et al., 2011). MYC2, along with its homologs MYC3 and MYC4 in Arabidopsis(Niu et al., 2011), is a master regulator activating JA responsive genes (Dombrecht et al., 2007). JAZ represses activity of the bHLH factors belonging to sub-group IIIb (ICE1 and ICE2), IIIe (MYC2, MYC3, and MYC4) and IIIf (GL3, EGL3 and TT8), the AP2 TF TARGET OF EAT1 and 2 (TOE1 and TOE2), as well as the MYB factors PAP1, GL1 MYB21, and MYB24 (Niu et al., 2011; Qi et al., 2011; Song et al., 2011; Zhai et al., 2015). Recent work illustrates that MYC2, upon interaction with JAZ, undergoes a structural shift that hinders its activation activity (Zhang et al., 2015). JAZ interacts with NINJA via the TIFY domain and subsequently recruits the corepressor TOPLESS (TPL) via its ethylene-responsive element binding factor-associated amphiphilic (EAR) motif (Pauwels et al., 2010). Moreover, TPL can directly interact with JAZ proteins containing the EAR motif (Shyu et al., 2012). TPL recruits histone deacetylases to modify chromatin architecture to repress target gene transcription (Krogan et al., 2012; Wang et al., 2013). The multisubunit Mediator complex functions as both a coactivator and corepressor and recruits RNA polymerase II to the promoter of activated genes (Kelleher et al., 1990; Flanagan et al., 1991). Recent work shows that TPL interacts with MED13 to repress auxin signaling (Ito et al., 2016). Interaction of MED13 with TPL also likely functions to repress JAZ mediated repression of JA signaling. MYC2 interacts with MEDIATOR SUBUNIT OF RNA POLYMERASE 25 (MED25) to activate transcription of its JA responsive target genes (Çevik et al., 2012; Chen et al., 2012). The JAZ binding of MYC2 competes with MED25, further suppressing target gene activation (Zhang et al., 2015). In the presence of JA, COI1 and JAZ interact to form a coreceptor complex (Sheard et al., 2010) that polyubiquitinates JAZ for degradation via the 26S proteasome (Chini et al., 2007). Degradation of JAZ frees MYC2 and other TFs to activate downstream JA responsive genes (Chini et al., 2007; Hou et al., 2010; Qi et al., 2011; Song et al., 2011; Zhu et al., 2011).

Origin and Evolution of JA Signaling

While the JA signaling cascade has been well characterized using the model plant Arabidopsis, the origin of the pathway and individual genes remains ambiguous. The moss, *Physcomitrella patens*, responds to JA treatment but is not known to synthesize JA (Stumpe et al., 2010). In contrast, the Klebsormidiophyceae charophyte, *Klebsormidium flaccidum* synthesizes JA, but not JA-isoleucine (Hori et al., 2014; Koeduka et al., 2015). However, contradictory evidence exists for the presence of JA signaling in *K. flaccidum*. A previous study suggests the presence of *KfCOI1* and *KfMYC2* (Wang et al., 2015), whereas a prior report found no evidence for these genes (Hori et al., 2014). The uncertainty necessitates further examination to authenticate the presence of each component in the JA pathway and to address how, when, where, and why the pathway arose. The prediction of COI1 and MYC2 in charophytes suggests the impairment of JA recognition is due to the lack of JAZ proteins. To date, the TIFY family has only been shown to occur in embryophytes (Bai et al., 2011). The TIFY family is divided into four major subfamilies: JAZ, peapod (PPD), TIFY, and ZIM factors, based on the presence of CCT/JAS, TIFY/ZIM, and GATA Zinc finger (ZnF-GATA) domains (Bai et al., 2011). Which family member of TIFY factors arose first

in plants is not known, nor is its initial function. Understanding how JA signaling arose will provide a basis for interpreting its role in medicinal species, in turn leading to improved designs for defense and metabolic engineering projects.

The Terpene Indole Alkaloid Pathway in Catharanthus

Vinblastine and vincristine are two TIAs listed on the World Health Organizations List of Essential Medicines – that is medicines which are the indispensable for a fundamental health care system. TIAs are found in a limited number of plant species belonging to the families Apocynaceae, Loganiaceae, Nyssaceae and Rubiaceae. *Catharanthus roseus,* a member of the Apocynaceae family, also commonly known as Madagascar periwinkle or Annual Vinca, has become a model plant for understanding TIA biosynthesis and regulation (Memelink and Gantet, 2007; Facchini and De Luca, 2008). Catharanthus synthesizes over 130 TIAs (Heijden et al., 2004). Catharanthus TIAs can be characterized into two principle groups: monoterpene indole alkaloids (MIA), and the bisindole alkaloids (BIA), which are composed of two joined MIAs. The primary TIAs of interest in *Catharanthus* include ajmalicine, serpentine, vinblastine, and vincristine. The MIAs, ajmalicine and serpentine, also present in Indian Serpentwood (*Rauwolfia serpentina*), are utilized pharmaceutically for the treatment of hypertension (Vakil, 1949). The BIAs, vinblastine and vincristine, are species specific metabolites which have proven invaluable in the treatment of cancers (Holland et al., 1973).

The synthesis of BIAs is complex and requires at least 30 enzymes (Figure 1.2). Recent publications of Catharanthus transcriptomes and draft genome enable studies into TIA metabolism (Góngora-Castillo et al., 2012; Van Moerkercke et al., 2013; Dugé de Bernonville et al., 2015; Kellner et al., 2015). Indeed, the last several years have observed a rapid characterization of missing enzymatic steps of TIA biosynthesis (Asada et al., 2013; Besseau et al., 2013; Rai et al., 2013; Salim et al., 2013; Miettinen et al., 2014; Salim et al., 2014; Dugé de Bernonville et al., 2015; Kumar et al., 2015; Qu et al., 2015). In Catharanthus, TIA are formed by joining a terpene, secologanin, with the alkaloid tryptamine to form the first TIA, strictosidine. The shikimic acid pathway leads to tryptophan production. Tryptophan is then decarboxylated by TRYPTOPHAN DECARBOXYLASE (TDC) to produce tryptamine (De Luca et al., 1989).

Secologanin is formed through series of two connected pathways. The first pathway is the evolutionarily conserved 2-C-methyl-D-erythritol 4-phosphate (MEP) pathway which leads to the formation of isopententyl pyrophosphate (IPP). GERANYL DIPHOSPHATE SYNTHASE (GPPS) converts IPP and its isomer dimethylallyl into geranyl diphosphate (GPP) (Rai et al., 2013). The enzyme GERANIOL SYNTHASE (GES) catalyzes the formation of geraniol from GPP (Kumar et al., 2015). The second portion of the secologanin pathway is the conversion of geraniol to secologanin. Recent work has made significant strides to characterize all the missing steps in secologanin biosynthesis. The first step of the pathway requires GERANIOL 10-HYDROXYLASE (G10H), which is believed to a major limiting step in TIA biosynthesis (Kumar et al., 2015). G10H catalyzes the conversion of geraniol into geraniol 10-hydroxygeraniol (Collu et al., 2001). Activity of the cytochrome P450 requires the NADPH:CYTOCHROME P450 REDUCTASE (CPR2). While two other CPR-like genes occur in Catharanthus, their involvement in TIA biosynthesis is likely minimal (Parage et al., 2016). The next step is the conversion of 10-hydroxygeraniol into 10-oxogeranial by 10-HYDROXYGERANIOL OXIDOREDUCTASE (10HGO) (Miettinen et al., 2014). IRIDOID SYNTHASE (IS) cyclizes 10-oxogeranial into *cis*–*trans*-nepetalactol or *cis*-*trans*-iridodial (Geu-Flores et al., 2012). IS belongs to a small gene family which includes three additional paralogs

capable of utilizing 10-oxogeranial (Munkert et al., 2015). Next, IRIDOID OXIDASE (IO), also called 7-DEOXYLOGANETIC ACID SYNTHASE, leads to the formation of 7-deoxyloganetic acid (Miettinen et al., 2014; Salim et al., 2014). This 7-deoxyloganetic acid is glycosylated with glucose by 7-DEOXYLOGANETIC ACID GLUCOSYLTRANSFERASE (7DLGT) to produce 7-deoxyloganic acid (Asada et al., 2013; Miettinen et al., 2014). In the next reaction, 7-DEOXYLOGANIC ACID 7-HYDROXYLASE (7DLH) catalyzes the production of loganic acid (Salim et al., 2013; Miettinen et al., 2014). Loganic acid is methylated by LOGANIC ACID O-METHYLTRANSFERASE (LAMT) to form loganin (Murata et al., 2008). The last step, catalyzed by SECOLOGANIN SYNTHASE (SLS), produces secologanin (Irmler et al., 2000). Recent work identified two SLS genes in Catharanthus (Dugé de Bernonville et al., 2015). The first identified SLS isolated from cell culture expresses primarily in the root, but is barely detectable in most tissues. In contrast, SLS2 expression occurs at high levels in root and leaf tissues, where secologanin is synthesized (Dugé de Bernonville et al., 2015).

STRICTOSIDINE SYNTHASE (STR) joins secologanin and tryptamine to form strictosidine (Mizukami et al., 1979; Treimer and Zenk, 1979). The glucose moiety of strictosidine is then removed by the enzyme STRICTOSIDINE β-D-GLUCOSIDASE (SDG) to form a reactive aglycone (Geerlings et al., 2000). At this key point, the strictosidine aglycone metabolite diverges into multiple pathways including ajmalicine, catharanthine, and vindoline (O'Connor and Maresh, 2006). The enzymes responsible for catharanthine biosynthesis remain unidentified. Similarly, an unknown enzyme(s) catalyzes the conversion of an strictosidine aglycone intermediate into tabersonine. TABERSONINE 16-HYDROXYLASE (T16H) hydroxylates tabersonine into 16-hydroxytabersonine (Schröder et al., 1999). Two genes of T16H occur in Catharanthus (Besseau et al., 2013). The expression of T16H1 primarily localizes to flower, whereas TH16H2 mostly occurs in young leaves (Besseau et al., 2013). This newly added hydroxyl group is methylated by 16-HYDROXYTABERSONINE O-METHYLTRANSFERASE (16OMT) to produce 16-methoxytabersonine (Levac et al., 2008). Next, TABERSONINE 3-OXYGENASE (T3O) and TABEROSINE 3-REDUCTASE (T3R) converts 16-methoxytabersonine to 3-hydroxy-16-methoxy-2,3-dihydrotabersonine. Next, the nitrogen molecule of the indole moiety is methylated by N METHYLTRANSFERASE (NMT) forming desacetoxyvindoline (Liscombe et al., 2010). Desacetoxyvindole is hydroxylated by DESACETOXYVINDOLINE-4-HYDROXYLASE (D4H) to produce deacetylvindoline (Vazquez-Flota et al., 1997). This hydroxyl group is then methylated by DEACETYLEVINDOLINE 4-O-ACETYLETRANSFERASE (DAT) to produce vindoline (St-Pierre et al., 1998). The first BIA, α-3',4'-anhydrovinblastine, is produced by PEROXIDASE1 enzymatic coupling a catharanthine molecule with a vindoline molecule (Sottomayor et al., 1998; Costa et al., 2008)

The biosynthesis of BIAs also requires extensive intercellular metabolite transport (Burlat et al., 2004; Murata et al., 2008; Roepke et al., 2010). Initial studies suggest loganic acid moves from the internal phloem associate parenchyma (IPAP) cells into the epidermis (Courdavault et al., 2014). However, imagining of metabolites in single cells shows loganic acid in IPAP cells and loganin in both IPAP and epidermal cells (Yamamoto et al., 2016). This suggests that loganin, not loganic acid, is the mobile metabolite. Once formed, desacetoxyvindoline presumably moves from the epidermis to laticifer and idioblast cells (Guirimand et al., 2011). Notably, desacetoxyvindoline appears in laticifer and idioblast cells (Yamamoto et al., 2016). Thus, the mobile metabolite must occur prior to desacetoxyvindoline. Indeed, 16-methoxytabersonin as well as cathenamine, one of the strictosidine aglycone intermediates, occur in laticifer and idioblast cells (Yamamoto et al., 2016). Despite the known requirement for TIA intercellular movement, only one transporter is known. TPT2 transports catharanthine from epidermal cells to the leaf surface (Yu and De Luca, 2013). In contrast vindoline presumably localizes to laticifer and idioblast cells (St-Pierre et al., 1999). Spatial separation of catharanthine and vindoline

prevents accumulation of vinblastine (Roepke et al., 2010). Upon wounding, vindoline is released to the surface and couples with catharanthine to form vinblastine. Vinblastine inhibits microtubule assembly, thus disrupting cell division and results in cell death. The mechanism of vinblastine action suggests it likely affects both herbivores and pathogens of Catharanthus.

Transcriptional Regulation of Catharanthus TIAs

A principal elicitor of TIA production in Catharanthus, as well as natural products in many other medicinal species, is the phytohormone jasmonate which functions in plant defense signaling to protect the plant from biotic stresses (Roepke et al., 2010; Wei, 2010). However, the biosynthesis of TIAs is also regulated by UV light (Ramani and Chelliah, 2007; Binder et al., 2009), fungal elicitors (Menke et al., 1999), wounding (Frischknecht et al., 1987; Vázquez-Flota et al., 2004), drought (Jaleel et al., 2007), cold (Dutta et al., 2007, 2013) and salt stress (Dutta et al., 2013). Unfortunately, knowledge of JA and other hormone signaling cascades in Arabidopsis has been inadequately applied to Catharanthus. This problem is further aggravated by the fact that Arabidopsis cannot be used as a model for understanding TIA production as it does not produce these compounds. Thus, our understanding of transcriptional regulation of TIA production in Catharanthus remains limited. Despite these hurdles, important TF regulating JA induction of specialized metabolism have been identified. These approaches have focused on understanding transcriptional regulation of key biosynthetic genes.

Successfully improving metabolite yield requires understanding key limitations to the biosynthetic pathway. Biosynthetic steps which either commit precursors to the metabolic pathway or limit conversion of one metabolite to another remain excellent targets for manipulation to improve chemical production. Several key committed steps in TIA biosynthesis have been the primary targets of studies for transcriptional regulation. In TIA biosynthesis, the committed steps to MIAs include the formation of the tryptamine precursor, formation of the secologanin, and the coupling of these two precursors into strictosidine. Formation of tryptamine and strictosidine are catalyzed by the enzymes TDC and STR (De Luca and St Pierre, 2000). Secologanin is formed by a series of enzymatic reactions from geraniol, starting with G10H. Increased TIA accumulation by feeding cell cultures with loganin suggests that G10H is a rate limiting step in TIA production (van der Fits and Memelink, 2000). Studies investigating committed or rate-limiting steps of alkaloid producing species reveal several key families of TFs: the AP2-ERF, bHLH, and WRKY families.

STR is the first committed step to TIA metabolism. Expression of *STR* increases in response to JA and fungal elicitor treatment (Menke et al., 1999). Deletion analysis of *the STR* promoters revealed the presence of several *cis*-regulatory elements including the jasmonate- and elicitor-responsive elements (JERE). The JERE in the *STR* promoter has been used to isolate the AP2/ERF family members, OCTADECANOID-RESPONSIVE CATHARANTHUS AP2-DOMAIN 1 (ORCA1) and ORCA2. Both ORCA1 and ORCA2 bound a GCC motif *cis*-element in the *STR* promoter (Menke et al., 1999). However, only ORCA2 expression increases upon JA elicitation. Moreover, overexpression of ORCA2 in *Catharanthus* cells, strongly activates the *STR* promoter, whereas ORCA1 has minimal effect. These findings suggest that ORCA2 plays a crucial role in JA and elicitor responsive expression of *STR* (Menke et al., 1999). Indeed, ORCA2 has an extensive role in regulating TIA pathway genes, particularly later steps in the pathway, i.e. vindoline biosynthesis (Li et al., 2013). Another AP2/ERF TF, ORCA3, has been isolated from T-DNA activation tagged cell lines of *Catharanthus*. *ORCA3* expression is rapidly induced by JA and its overexpression in

Catharanthus cell lines induced genes from the tryptamine, secologanin, and vindoline pathways. In cell cultures, ORCA3 binds the *CPR*, *STR* and *TDC* promoters and upregulates their expression (van der Fits and Memelink, 2000). To identify a higher-tier regulator of the TIA pathway, studies analyzed the ORCA3 promoter. The *ORCA3* promoter has a 74-bp region containing a bipartite JA responsive element (Vom Endt et al., 2007). The JA responsive element is composed of an A/T-rich quantitative sequence and a qualitative component. The qualitative element contains a G-box motif which functions to turn ORCA3 expression on or off in response to JA. A yeast-one hybrid system was used to isolate factors binding the quantitative element. Several At-HOOK type TFs bind this region (Vom Endt et al., 2007). Two (2D7 and 2D173) bind to the A/T-rich quantitative sequence and activate the *ORCA3* promoter. Based on the role of MYC2 in Arabidopsis, the Catharanthus MYC2 ortholog was isolated (Zhang et al., 2011). MYC2 binds the G-box within the qualitative sequence of the *ORCA3* promoter. In cell culture, MYC2 activated the *ORCA3* promoter. In contrast, RNAi-mediated knock-down of *MYC2* significantly decreased *ORCA3* and *ORCA2* expression. (Zhang et al., 2011). These findings suggest that JA-responsive expression of TIA pathway genes is controlled by a TF cascade where MYC2 acts upstream of ORCAs. In addition to AP2/ERF and WRKY TFs, the *STR* promoters are regulated by several other TFs including BOX P-BINDING FACTOR 1 (BPF1) and G-BOX BINDING FACTORS (GBF1 and GBF2). The MYB-like protein BPF1 was isolated from *Catharanthus* using a JA and elicitor responsive region of the *STR* promoter in a yeast one-hybrid assay (van der Fits et al., 2000). Unlike other TIA regulators, expression of *BPF1* is induced by fungal elicitors but remains unchanged in response to JA. This study indicates that BPF1 is possibly involved in an elicitor-responsive but JA-independent signal transduction pathway in *Catharanthus* (van der Fits et al., 2000). Using a β-estradiol inducible system in hairy roots, overexpression of *BPF1* induces tryptamine, secologanin, and TIA pathway genes (Li et al., 2015). Similarly, induction of BPF1 overexpression leads to increased *BIS1*, *GBF1*, *GBF2*, *MYC2*, *ORCA2*, *ORCA3*, *WRKY1*, *ZCT1* (ZINC-FINGER *C. roseus* TRANSCRIPTION FACTOR 1), *ZCT2*, and *ZCT3* expression. Consistent with the activation of both activator and repressor TFs, alkaloid accumulation was very mild in *BPF1* overexpression induced hairy root cultures (Li et al., 2015). The TFs GBF1 and GBF2 belong to the basic leucine zipper (bZIP) family and bind the G-box (CACGTG) motif in the *STR* promoter. Transient bombardment assays showed that GBF1 and GBF2 act as repressors of the *STR* promoter (Siberil et al., 2001).

The enzyme TDC commits the indole moiety to the TIA pathway. Analysis of *Catharanthus* TIA pathway promoters such as *TDC*, *G10H*, *CPR*, and *DAT*, revealed the presence of multiple W-box *cis*-elements, a canonical DNA-binding motif for WRKY TFs (Schluttenhofer et al., 2014). However, the biological significance of those *cis*-elements in TIA pathway promoters remains. WRKY1, isolated from *Catharanthus* seedlings, has been shown to play a crucial role in TIA biosynthesis. WRKY TFs regulate numerous specialized metabolic pathways (Schluttenhofer and Yuan, 2015) *WRKY1* is induced by JA, ethylene, and gibberellic acid (Suttipanta et al., 2011). Overexpression of *WRKY1* in *Catharanthus* hairy roots results in up-regulation of several TIA pathway genes, most notably *TDC*. WRKY1 binds the W-box motif in the *TDC* promoter and trans-activates the *TDC* promoter in *Catharanthus* cells. Consistent with TIA pathway activation, the *WRKY1* overexpressing hairy roots accumulate higher amounts of serpentine compared to the control. Surprisingly, in *WRKY* overexpression lines, the transcription repressors, *ZCT1*, *ZCT2* and *ZCT3* are upregulated. Furthermore, expression of transcriptional activators, *ORCA2*, *ORCA3* and *MYC2*, are reduced. This suggests CrWRKY1 is functions both as an activator and repressor of the TIA pathway. Consistent with this hypothesis is the decreases levels of catharanthine in overexpression hair roots (Suttipanta et al., 2011). ZCT1, ZCT2 and ZCT3 belong to the Cys_2/His_2 (TF IIIA-type) zinc finger protein family, and were isolated by yeast one-hybrid using the elicitor

responsive region of the *TDC* promoter. Expression of *ZCT* genes is induced by yeast extract and JA. The ZCT TFs bind to multiple regions of both *TDC* and *STR* promoters. The binding site of ZCTs in the *STR* promoter is distinct but overlaps with binding sites for ORCAs. In a transient assay, ZCT proteins repress the activities of *TDC* and *STR* promoters, suggesting their role as potential transcriptional repressors in the TIA pathway (Pauw et al., 2004). Additionally, ZCT1 and ZCT2 bind and repress activity of the HYDROXYMETHYLBUTENYL 4-DIPHOSPHATE SYNTHASE (HDS) promoter (Chebbi et al., 2014). HDS is a step in the MEP pathway which forms the IPP precursor for geraniol biosynthesis. Recently, in contrast to low concentrations (50-250 μM), high levels (1000 μM) of JA were shown to suppress TIA production (Goklany et al., 2013). Under low JA concentrations, expression of ORCA TFs exceeds those of ZCTs. When 1000 μM JA is used, the inverse is true and ZCT expression exceeds ORCA expression. Thus, activation or repression of TIA biosynthesis may depend upon the ratio of ORCA to ZCT TFs (Goklany et al., 2013). In addition to JERE, UV-light responsive *cis*-regulatory sequences have been identified in both *TDC* promoter. The GT-1 and 3AF1 TFs bind multiple elements in the *TDC* promoters to enhance their expression in response to UV-light (Ouwerkerk et al., 1999).

Secologanin forms the second moiety necessary for formation of TIAs. Secologanin starts with the formation of 10-hydroxygeraniol from geraniol, catalyzed by the enzyme G10H, in the first committed step to secologanin synthesis. Analysis of the *G10H* promoter indicates the presence of several putative DOF, GBF, MYB, and WRKY TF binding sites (Suttipanta et al., 2007). Recent works illustrates the role of two bHLH factors in regulating the secologanin terpene pathway. A clustering analysis integrating TFs with TIA biosynthetic genes identified bHLH IRIDOID SYNTHASE 1 (BIS1). In cell cultures, the BIS1 activates the GES, G10H, 10HGO, IRS, 7DLGT, and 7DLH promoters (Van Moerkercke et al., 2015). Later, based on similarity, BIS2 was identified. As with BIS1, BIS2 binds GES, G10H, 10HGO, IRS, 7DLGT, and 7DLH promoters (Moerkercke et al., 2016).

The formation of vindoline, a precursor to vinblastine and vincristine, requires DAT. DAT transfers an acetyl group onto deacetylvindoline to produce vindoline. JA and light both regulate DAT expression (Aerts and De Luca, 1992; Wang et al., 2010). Analysis of the *DAT* promoter revealed ABA, auxin, JA, light- and defense-regulated *cis*-elements (Wang et al., 2010; Makhzoum et al., 2011). Wang et al. (2010) identified three TGACG motif within DAT promoter which are involved in JA responsive expression of this gene. The observations that *DAT* is not regulated by ORCAs, implicate additional unidentified TFs involved in regulation of the TIA pathway.

Summary

Vinblastine and vincristine are two essential anti-cancer drugs which must be extracted from Catharanthus. Catharanthus plants synthesize these compounds in very low quantities. Enhancing TIA production could reduce time and expenditure for drug production. The phytohormone JA is a key elicitor of specialized metabolism. JA signaling has been well studied in models plants such as Arabidopsis. However, limited knowledge of Arabidopsis JA signaling has been applied to Catharanthus. Furthermore, unlike Arabidopsis, Catharanthus produces TIA, some of which are specific to this species. JA activates TFs which induce expression of TIA biosynthetic genes and eventually result in metabolite accumulation. A limited number of TFs have been identified which regulate Catharanthus TIA production. However, gaps remain in our understanding of specialized metabolism regulation by JA. These knowledge gaps occur throughout the JA signaling cascade.

Areas Needing Further Research

To improve our understanding of JA regulation of specialized metabolism, I identified several key questions which need addressed. First, how do large TF families contribute to JA response and what portion of factors are involved? Additionally, is their induction to JA conserved between distantly related plant species? Second, how, when, and why did JA signaling arise in plants? Currently, little is known about the evolutionary history of components involved in JA signaling. Furthermore, what research exists on this topic is contradictory. Finally, besides the COI1-JAZ coreceptors, the Mediator subunit MED25 plays a key function in JA signaling. I wanted to know where does this factor fall in the hierarchy of transcriptional regulation of specialized metabolism? And by what mechanism does it use? My work addresses these questions to further fill knowledge gaps in our understanding of JA signaling. This knowledge can be used in the future to help improve production of valuable plant metabolites.

Objectives

Objective 1

I wanted to address how do large TF families contribute to JA response? What portion of factors are involved? And is induction to JA conserved across divergent species? To address these questions, I analyzed the WRKY TF family. WRKY1 was previously demonstrated by our lab to regulate TIA production (Suttipanta et al., 2011). WRKYs form one of the largest TF families in plants. Thus, additional WRKY factors should be present in Catharanthus. As WRKYs are important components of plant defense, JA responsive factors may contribute to TIA regulation. As such, I sought to characterize the Catharanthus WRKY family and identify those which are JA responsive. I also aimed to determine if induction of WRKY TFs to JA was conserved across species and to utilize this information to identify candidate regulators of the TIA pathway.

Objective 2

The JA signaling pathway is essential to multiple aspects of plant biology, including specialized metabolite production. However, how, when, and why the JA signaling cascade arose remains unknown. I sought to understand how, when, and why JA signaling evolved in plants? An increasing number of fully sequenced genomes permits gene identification across a diverse range of species. Phylogenetic analyses can then be applied to determine the evolutionary history of key genes involved in the JA signaling cascade. Therefore, I sought to determine the origin of when and how JA signaling arose in plants.

Objective 3

Mediator is essential for activation of gene transcription. The MED25 subunit functions as a key hub for regulating plant defense and response to JA. Prior reports indicate a role of MED25 in regulating anthocyanin accumulation. How does MED25 function to regulate

specialized metabolism? Where does it fall in the transcriptional hierarchy of specialized metabolism? And what mechanism does it use? To better understand these roles, I sought to characterize the role of MED25 in anthocyanin regulation using the model plant Arabidopsis.

Relationship of Projects

These projects are connected by a unifying theme: improve our understanding of jasmonate signaling. Despite being a conserved hormone signaling cascade, much remains to be learned about JA. Specifically, how JA regulates specialized metabolism remains understudied. Without appropriate control of metabolite production and accumulation, a given product may not provide benefits or may be toxic to the host plant. While each project is distinct, all three take a top-down approach to understanding JA responsive regulators. Objective one starts with identifying the entire Catharanthus family to elucidate potential candidate regulators based on response to JA. WRKY TFs function downstream of JA signaling and other phytohormone cascades. They are also highly diverse between plant species, but whether JA response is conserved across species remains unknown. Furthermore, if JA responsive WRKY TFs function through or independently of the COI1-JAZ coreceptor complex also remains ambiguous. Objective two begins at the very origin of the pathway, by investigating the evolutionary history of JA. I investigate how each signaling component arose, when this occurred, and when a functional network established in plants. Understanding when the earliest components arose improves our understanding of how the JA network is wired and developed those connections between components. Objective three characterizes MED25, one of the earliest JA signaling components to evolve. While MED25 has a known role in JA signaling and predicted role in anthocyanin development, how it functions to regulate specialize metabolism is unknown. Characterizing Mediator component will allow us to identify interacting TFs (upper level regulators) which can modulate known anthocyanin regulators. Collectively, these projects will improve our understanding of how the upper tiers of JA signaling work. These findings can be used in the future to improve genetic engineering of specialized metabolism in Catharanthus and other medical species.

Notice of Publication

The “Transcriptional Regulation of Catharanthus TIAs” section has been previously published. Changes were made to update information since publication. This section, and others relating to regulation of alkaloid and terpene biosynthesis, was my contribution to the co-authored review cited below.

Citation: B. Patra, C. Schluttenhofer, Y. Wu, S. Pattanaik, and L. Yuan. 2013. Transcriptional regulation of secondary metabolite biosynthesis in plants. Biochimica et Biophysica Acta - Gene Regulatory Mechanisms 1829 (11): 1236-1247.

Table 1.1. List of Catharanthus TFs regulating TIA biosynthesis. The name, genome identifier number and transcription factor family of Catharanthus regulators of TIA biosynthesis.

Locus	Name	Transcription Factor Family
CRO_T003206	BIS1	bHLH
CRO_T009037	BIS2	bHLH
CRO_T021525	BPF-1	MYB-HB Like
CRO_T026312/CRO_T001597	GBF1	bZIP
CRO_T020924/CRO_T003332	GBF2	bZIP
CRO_T014132	ORCA1	AP2-ERF
CRO_T030992	ORCA2	AP2-ERF
CRO_T030273	ORCA3	AP2-ERF
CRO_T018065	ZCT1	C2H2
CRO_T007669	ZCT2	C2H2
CRO_T015507	ZCT3	C2H2
CRO_T021622	WRKY1	WRKY
CRO_T016859	MYC2	bHLH
CRO_T026362	2D7	DUF296
CRO_T027183	2D173	DUF296

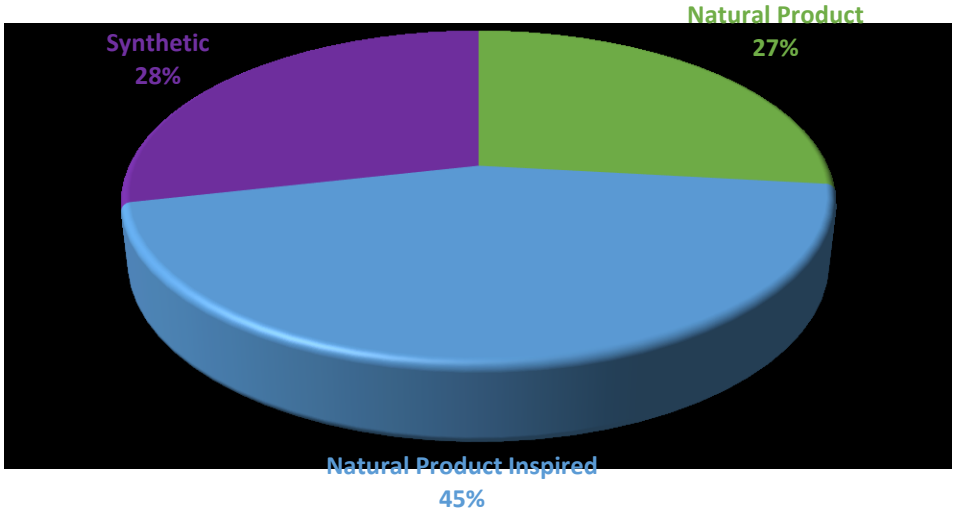


Figure 1.1. The origin of all pharmaceutical drugs marketed since 1981. Data modified from Newman and Cragg, 2012.

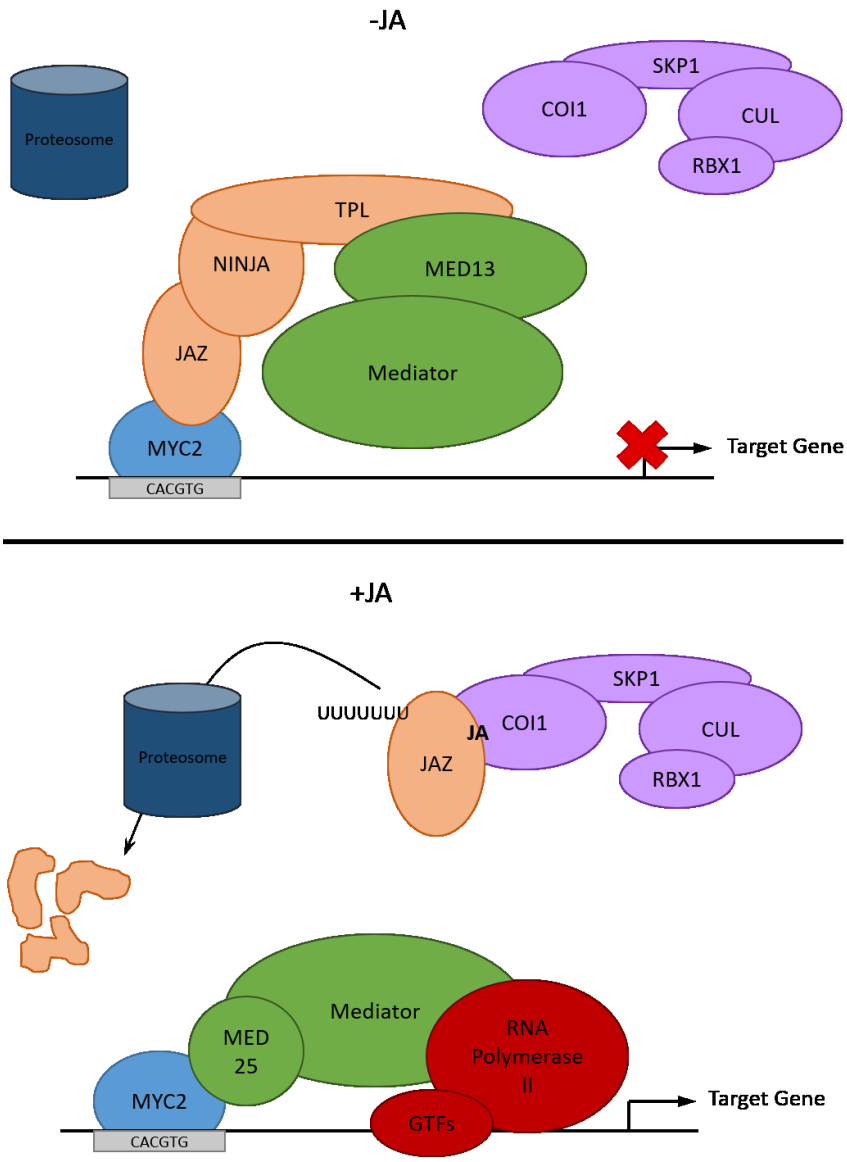


Figure 1.2. Model of JA Signaling in plants. In absence of JA (top panel), JASMONATE ZIM DOMAIN (JAZ), NOVEL INTERACTOR OF JAZ (NINJA), TOPLESS and (TPL) form a corepressor complex. JAZ binds MYC2 and other JA-responsive TFs to repress activation of their target genes. TPL interacts with MED13, part of the repressive CYCLIN-DEPENDENT KINASE8 module, to inhibit Mediator recruitment of RNA Polymerase II for gene transcription. In presence of JA (bottom panel), JA permits the interaction of JAZ with the F-box protein CORONITINE INSENSITIVE1 (COI1). Interaction of JAZ and COI1 leads to polyubiquitination (U) of JAZ, targeting it for degradation by the 26S proteasome. Degradation of JAZ frees MYC2 and other JA responsive TFs. Free MYC2 can recognize the G-box *cis*-element (CACGTG). As part of Mediator, Mediator subunit 25 (MED25) can interact with DNA-bound MYC2. Mediator then can recruit RNA Polymerase II and general transcription factors (GTF) needed to activate transcription of target genes.

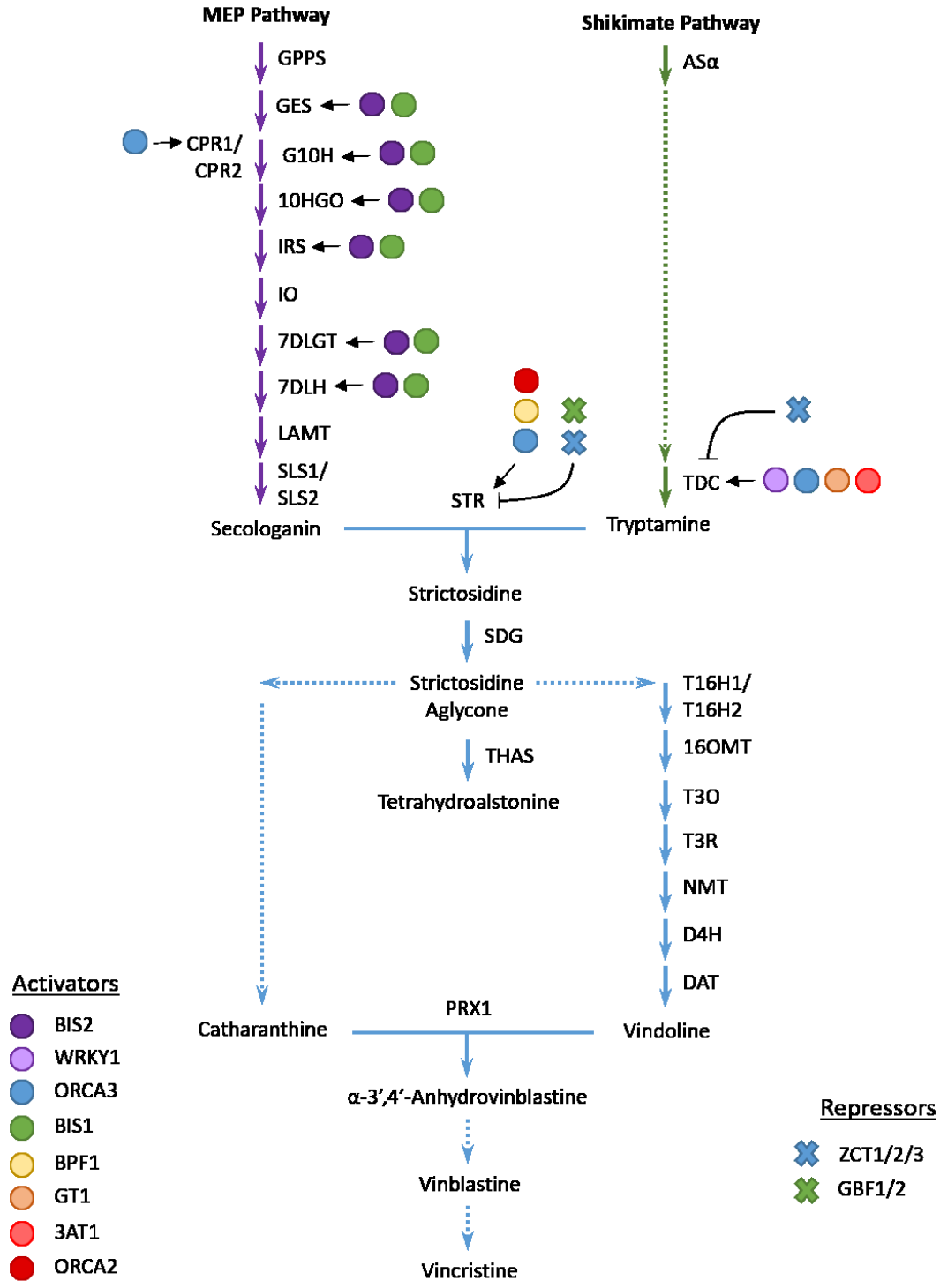


Figure 1.3. The TIA pathway in *Catharanthus roseus*. Enzymes are indicated by arrows and abbreviated names. Activators and repressors of the TIA pathway are depicted as circles or crosses, respectively. Transcription factor markers are indicated next to gene promoters they bind. Enzymes of the secologanin, tryptamine, and TIA pathway are indicated by purple, green,

and blue arrows, respectively. Enzymes are ANTHRANILATE SYNTHASE (ASα), TRYPTOPHAN DECARBOXYLASE (TDC), GERANYL DIPHOSPHATE SYNTHASE (GPPS), GERANIOL SYNTHASE (GES), GERANIOL 10-HYDROXYLASE (G10H), NADPH:CYTOCHROME P450 REDUCTASE (CPR), 10-HYDROXYGERANIOL OXIDOREDUCTASE (10HGO), IRIDOID SYNTHASE (IS), IRIDOID OXIDASE/7-DEOXYLOGANETIC ACID SYNTHASE (IO), 7-DEOXYLOGANETIC ACID GLUCOSYLTRANSFERASE (7DLGT), 7-DEOXYLOGANIC ACID 7-HYDROXYLASE (7DLH), LOGANIC ACID O-METHYLTRANSFERASE (LAMT), SECOLOGANIN SYNTHASE1 (SLS1), SECOLOGANIN SYNTHASE2 (SLS2), STRICTOSIDINE SYNTHASE (STR), STRICTOSIDINE β-D-GLUCOSIDASE (SDG), TABERSONINE 16-HYDROXYLASE1 (T16H1), TABERSONINE 16-HYDROXYLASE2 (T16H2), 16-HYDROXYTABERSONINE O-METHYLTRANSFERASE (16OMT), TABERSONINE 3-OXYGENASE (T3O), TABERSONINE 3-REDUCTASE/ALCOHOL DEHYDROGENASE1 (T3R), 16-METHOXY-2,3-DIHYDRO-3-HYDROXY-TABERSONINE-N-METHYLTRANSFERASE (NMT), DESACETOXYVINDOLINE-4-HYDROXYLASE (D4H), DEACETYLEVINDOLINE 4-O-ACETYLETRANSFERASE (DAT), PEROXIDASE1/ANHYDROVINBLASTINE SYNTHASE (PRX1), and TETRAHYDROALSTONINE SYNTHASE (THAS). TF genes are 3AT1, bHLH IRIDOID SYNTHASE1 (BIS1) and 2 (BIS2), BOX P-BINDING FACTOR1 (BPF1), G-BOX BINDING FACTOR 1 (GBF1) and 2 (GBF2), GT1, OCTADECANOID-RESPONSIVE CATHARANTHUS AP2-DOMAIN 2 (ORCA2) and 3 (ORCA3), WRKY1, ZINC FINGER-BINDING PROTEIN 1 (ZCT1), 2 (ZCT2), and 3 (ZCT3).

CHAPTER TWO: ANALYSES OF *CATHARANTHUS ROSEUS* AND *ARABIDOPSIS THALIANA* WRKY TRANSCRIPTION FACTORS REVEAL INVOLVEMENT IN JASMONATE SIGNALING

Abstract

To combat infection to biotic stress, plants elicit the biosynthesis of numerous natural products, many of which are valuable pharmaceutical compounds. Jasmonate is a central regulator of defense response to pathogens and accumulation of specialized metabolites. *Catharanthus roseus* produces a large number of terpenoid indole alkaloids (TIAs) and is an excellent model for understanding the regulation of this class of valuable compounds. Recent work illustrates a possible role for the Catharanthus WRKY transcription factors (TFs) in regulating TIA biosynthesis. In Arabidopsis and other plants, the WRKY TF family is also shown to play important role in controlling tolerance to biotic and abiotic stresses, as well as secondary metabolism. Here, I describe the WRKY TF families in response to jasmonate in Arabidopsis and Catharanthus. Publicly available Arabidopsis microarrays revealed at least 30% (22 of 72) of WRKY TFs respond to jasmonate treatments. Microarray analysis identified at least six jasmonate responsive Arabidopsis WRKY genes (AtWRKY7, AtWRKY20, AtWRKY26, AtWRKY45, AtWRKY48, and AtWRKY72) that have not been previously reported. The Catharanthus WRKY TF family is comprised of at least 48 members. Phylogenetic clustering reveals 11 group I, 32 group II, and 5 group III WRKY TFs. Furthermore, I found that at least 25% (12 of 48) were jasmonate responsive, and 75% (9 of 12) of the jasmonate responsive CrWRKYs are orthologs of AtWRKYs known to be regulated by jasmonate. Overall, the CrWRKY family, ascertained from transcriptome sequences, contains approximately 75% of the number of WRKYs found in other sequenced asterid species (pepper, tomato, potato, and bladderwort). Microarray and transcriptomic data indicate that expression of WRKY TFs in Arabidopsis and Catharanthus are under tight spatio-temporal and developmental control, and potentially have a significant role in jasmonate signaling. Profiling of CrWRKY expression in response to jasmonate treatment revealed potential associations with secondary metabolism. This study provides a foundation for further characterization of WRKY TFs in jasmonate responses and regulation of natural product biosynthesis.

The following section has been published:

Citation: C. Schluttenhofer, S. Pattanaik, B. Patra, and L. Yuan. 2014. Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. BMC Genomics. 15(1): 502-522.

Introduction

Transcription factors (TFs) play a critical role in responding to jasmonate to elicit the synthesis of TIAs in Catharanthus (Menke et al., 1999; Van Der Fits and Memelink, 2001; Zhang et al., 2011). Negative regulators also mediate jasmonate signaling of the TIA pathway in Catharanthus (Sibéril et al., 2001; Pauw et al., 2004). In plants, WRKY TFs are critical regulators of response to biotic and abiotic stress. WRKY TFs have been attributed to tolerance of drought (Ren et al., 2010), salt (Jiang and Deyholos, 2009), nutrient deficiency (Devaiah et al., 2007), osmotic (Chen et al., 2010), cold (Zou et al., 2010), heat (Li et al., 2010), oxidative (Li et al., 2011), wounding (Skibbe et al., 2008), pathogens (Zheng et al., 2006), and UV-B stresses (Wang et al., 2007). The WRKY TF family is primarily a plant specific family with the exception of several examples in protozoa (Zhang and Wang, 2005). The WRKY domain is a 60 to 70 amino acid long DNA binding domain that recognizes the W-box (TTGACC/T); however, recent studies suggest this cis-element may be more degenerate and other components are involved for WRKY binding to DNA in response to a specific stimulus (Brand et al., 2013; Machens et al., 2014). The N-terminal portion of the WRKY domain is characterized by a highly conserved WRKYGQK motif whereas the C-terminal region of the domain contains either a Cys2-His2 or Cys2-His-Cys zinc-finger (Eulgem et al., 2000). WRKY TFs are distinguished by the presence of one or two WRKY domains. Group I WRKYs typically contain two WRKY domains whereas group II and group III members only contain one WRKY domain (Eulgem et al., 2000). Up to five subgroups (IIa, IIb, IIc, IId, and IIe) are recognized in the group II WRKY TFs (Eulgem et al., 2000).

In Arabidopsis, WRKY TFs are well established in salicylic acid (SA) and defense signaling pathways (Yu et al., 2001; Dong et al., 2003; Kalde et al., 2003). The majority of Arabidopsis WRKY TFs are induced by treatment with SA (Dong et al., 2003). However, the importance of WRKY TFs in JA signaling network is relatively less studied. Li et al. (Li et al., 2004) identified AtWRKY70 as a positive regulator of SA signaling and negative regulator of jasmonate signaling. Mutation of AtWRKY33 enhances susceptibility to necrotrophic pathogens by up-regulating JAZ proteins, repressors of jasmonate signaling (Zheng et al., 2006; Birkenbihl et al., 2012). Jasmonate positively regulates AtWRKY18 and AtWRKY40 which also negatively regulate abscisic acid (ABA) response (Pandey et al., 2010). AtWRKY6, AtWRKY8, AtWRKY11, AtWRKY17, AtWRKY25, AtWRKY28, AtWRKY38, AtWRKY60, AtWRKY62, and AtWRKY70 are also differentially expressed by jasmonates to regulate plant defense (Robatzek and Somssich, 2001; Zhou et al., 2005; Journot-Catalino et al., 2006; Zheng et al., 2007; Kim et al., 2008; Chen et al., 2010; Wu et al., 2011; Hu et al., 2012).

Over the last several years, WRKY TFs have emerged as a key family in the induction of natural product biosynthesis (Suttipanta et al., 2011; Li et al., 2013). CjWRKY1, from *Coptis japonica*, regulates the production of the benzylisoquinoline alkaloid, berberine (Kato et al., 2007). Cotton (*Gossypium arboreum*) GaWRKY1 affects the biosynthesis of the sesquiterpene, gossypol (Xu et al., 2004). Multiple biosynthetic genes for the sesquiterpene lactone, artemisinin, valuable as an anti-malaria drug, are regulated by *Artemisia annua* WRKY1 (Ma et al., 2009). *Hevea brasiliensis* WRKY1 is present in the latex of mechanically wounded (tapped) trees suggesting involvement in rubber latex synthesis (Zhang et al., 2012). The *Taxus chinensis* WRKY1 was found to regulate the expression of 10-DEACETYLBACCATIN III-10 B-O-ACETYL TRANSFERASE (DBAT), a gene encoding a key enzyme catalyzing a rate limiting step in the biosynthesis of the anticancer terpene, paclitaxel (Li et al., 2013). In Arabidopsis, camalexin biosynthesis is mis-regulated in *wrky33* mutant (Mao et al., 2011). Over-expression in Arabidopsis of *Panax quinquefolius WRKY1*, a jasmonate responsive WRKY from American ginseng, is found to enhance

expression of genes related to drought, salt, and disease resistance, leading to improvement of seedling survival to drought and salt stress, in addition to regulating the expression of genes related to triterpene biosynthesis (Sun et al., 2013). In Catharanthus, *CrWRKY1* has been demonstrated to respond to jasmonate, ethylene, and gibberellin signaling to regulate TIA production (Suttipanta et al., 2011). Over-expression of *CrWRKY1* increased the production of serpentine while simultaneously decreasing catharanthine accumulation, suggesting this WRKY may function in governing gene expression that specifically directs the flow of metabolites to synthesize TIAs in Catharanthus roots.

Identification of jasmonate responsive WRKYs will provide useful information on plant defense and natural product regulatory networks. Understanding the number and types of WRKY TFs present in Catharanthus will provide a clearer picture on the regulation of TIAs by this important TF family. Arabidopsis microarray data was analyzed to help identify the involvement of the WRKY TFs in jasmonate signaling. Then, medicinal plant transcriptome sequence data was used to identify the Catharanthus family of WRKY TFs. Expression data from Catharanthus revealed the induction of multiple WRKY transcripts by methyl jasmonate (JA) treatment. Seventy-five percent of the jasmonate responsive CrWRKYs are orthologs of AtWRKYs known to be regulated by jasmonate. JA-induced WRKYs provide potential candidates for further regulation of TIA accumulation in Catharanthus. The identification of orthologs for WRKY TFs known to be involved in specialized metabolism in other plant species indicates the possible involvement of additional WRKY TFs in regulation of TIA production in Catharanthus.

Results and Discussion

WRKY TFs Are Involved in Jasmonate Signaling

The role of WRKY TF family in SA signaling and plant defense is well established and has been systematically analyzed in Arabidopsis, but remains less clear for jasmonate signaling. Jasmonate is a key phytohormone regulating the production of specialized metabolites in many plant species, including Catharanthus. While Arabidopsis does not synthesize TIAs as found in Catharanthus, studying AtWRKYs can answer several important questions. First I wanted to determine whether the WRKY family is important for regulating jasmonate signaling in a model species, such as Arabidopsis. Second, I wanted to elucidate jasmonate responsive AtWRKYs to aid the identification of CrWRKY orthologs with potentially conserved regulatory functions. Comparison of orthologous jasmonate responsive WRKYs from Arabidopsis and Catharanthus will identify WRKYs that are potentially involved in modulating jasmonate signaling, and in turn identify candidates that regulate TIA production.

To clearly establish the role of WRKY TFs in jasmonate signaling I first identified jasmonate responsive WRKYs in the model plant Arabidopsis. To ascertain jasmonate responsive Arabidopsis WRKY TFs I used publicly available microarray datasets (Table 2.1). The ATH1 Affymetrix arrays used contain probes identifying 85% (61 of the 72) of Arabidopsis WRKY TFs (Table 2.2). From five datasets, I identified 39 AtWRKY TFs that significantly change in response to jasmonate treatment (Table 2.3). Of the 39 jasmonate responsive AtWRKY genes, 22 were differentially expressed in at least two jasmonate treated datasets. *AtWRKY6*, *AtWRKY18*, *AtWRKY45*, and AtWRKY53 were differentially expressed in three jasmonate treated datasets. Expression of *AtWRKY7*, *AtWRKY69* and *AtWRKY75* were significantly changed in response to jasmonate in four datasets. *AtWRKY40* and *AtWRKY47* expression were significantly changed in all five datasets. Notably, *AtWRKY47*,

AtWRKY69, and *AtWRKY75* expression were significantly differentially regulated to jasmonate treatment in four or more microarray experiments, but did not survive application of the Benjamini-Hochberg false discovery (B-H FDR) in any dataset. Arabidopsis *WRKY6*, *WRKY11*, *WRKY17*, *WRKY25*, *WRKY46*, and *WRKY53*, which have known roles in jasmonate signaling, were identified as being differentially expressed, but did not survive the B-H FDR in any dataset (Table 2.3). While the B-H FDR is less conservative than other procedures (e.g. Bonferroni correction), genes on the upper end of significant p-values (genes with small fold changes) may not be easily detected even if the response is consistent. This becomes apparent by the large reduction in significant genes after B-H FDR (Table 2.1). In total, eleven WRKYs survived the B-H FDR in at least one dataset indicating these are jasmonate responsive. I identified five WRKY TFs previously reported to be jasmonate responsive (Li et al., 2004; Zheng et al., 2006; Kim et al., 2008; Pandey et al., 2010). Additionally, I identified at least six Arabidopsis WRKY TFs (*AtWRKY7*, *AtWRKY20*, *AtWRKY26*, *AtWRKY45*, *AtWRKY48*, and *AtWRKY72*) previously unreported to have a role in jasmonate response (Table 2.1 and Table 2.3). Expression of *AtWRKY7*, *AtWRKY40*, and *AtWRKY45* changed significantly and survived the B-H FDR in two datasets. For these eleven AtWRKYs, the change in expression in response to jasmonate treatment was small, around 2.5 fold (Table 2.4A-E). AtWRKY40 displayed the greatest change in expression, with a 9-fold induction of transcripts after 1 hr of treatment with JA. Collectively, at least 18% of AtWRKY (11 of 61), up to 64% (39 of 61) or more, were jasmonate responsive WRKYs based on the microarrays analyzed. Overall, at least 30% (22 of 72) of AtWRKY TFs play a role in jasmonate response indicating the importance of this family in the jasmonate signaling network. Further experiments may eventually reveal that upwards of 50% (39 of 72) of WRKYs are involved in Arabidopsis jasmonate response (Table 2.3).

The small overlap between WRKY genes differentially expressed in response to jasmonate treatment in the microarray experiments suggested a tight developmental and/or spatiotemporal regulation in Arabidopsis. Two-way ANOVAs analyzing expression in response to time, as well as its combined effect with jasmonate treatment, further indicated AtWRKY regulation is temporally dependent (Table 2.5A-C). The expression of 20 AtWRKY genes was time dependent (Table 2.5A). Jasmonate treatment was found to regulate the expression of AtWRKY38 and AtWRKY70 in a time dependent manner. The genetic background of jasmonate signaling pathway mutants had less effect on WRKY gene expression. CORONATINE INSENSITIVE 1 (COI1) has been established as a jasmonate receptor (Xie et al., 1998; Sheard et al., 2010). AtWRKY72 was the only WRKY family member found to be regulated in a COI1-dependent manner when the B-H FDR was applied (Table 2.5B). No WRKYs were found to be dependent on MYC2 (Table 2.5C), a major transcriptional regulator of jasmonate signaling (Lorenzo et al., 2004; Dombrecht et al., 2007). These results indicate that response of Arabidopsis WRKY TFs to jasmonate treatment is highly dependent upon tissue, timing and culture conditions, and likely occurs through several major pathways. Furthermore, WRKY TFs may be important COI1-independent regulators of jasmonate response.

Unsupervised agglomerative hierarchical clustering analysis of AtWRKY TFs was performed to identify similar patterns of gene expression which may indicate related functions (Eisen et al., 1998). Gene expression of AtWRKYs formed two major clusters. Clustering of experiments revealed more similarities within an experiment than by jasmonate treatment (Figure 2.1). Additionally, the two major clusters separated those experiments in which the sampled tissues were from either plants or cell cultures. These findings further support AtWRKY gene expression in response to jasmonate as highly dependent on culture conditions and environment. The two major clusters were further subdivided into two or three clusters. Jasmonate responsive WRKYs previously annotated or identified by the microarray analysis

primarily occurred in cluster one and all of the three sub-clusters. The distribution of jasmonate responsive AtWRKYs indicates at least two major pathways for the regulation of AtWRKY gene expression. Interestingly, the only WRKY identified by microarray analysis to be COI1-depenedent, AtWRKY72, occurred in cluster 2b, distinct from expression patterns of other jasmonate responsive WRKYs. These data further suggest that there are complex tissue and environmental controls over jasmonate responses that likely occurs through several major pathways. These findings from Arabidopsis provide foundational information about the involvement of WRKY TFs in jasmonate response and for exploiting these factors in genetic engineering of transcriptional regulatory networks for natural product production.

Identification of Catharanthus WRKY TFs

Previously the Yuan laboratory identified a JA responsive group III type WRKY TF, CrWRKY1, as important for the regulation of TIA in Catharanthus (Suttipanta et al., 2011). Furthermore, the Arabidopsis data indicates the AtWRKY family as important for regulating jasmonate response. Jasmonate responsive CrWRKY TF may be important for regulating the production of valuable TIAs. Elucidation of CrWRKYs regulating specialized metabolite production will be valuable for future genetic engineering projects to increase production of pharmaceutically valuable TIAs. As the first step to identify important CrWRKY regulators of specialized metabolism I sought to identify all WRKY family members in Catharanthus. The recent release of 14 medicinal plants Illumina sequenced transcriptomes by the Medicinal Plant Genomic Resource (MPGR), including Catharanthus, provides the opportunity to identify many CrWRKY TFs (Góngora-Castillo et al., 2012). To identify CrWRKYs, I first downloaded all protein sequences from MPGR and isolated a single protein sequence for each locus. I assumed the individual copy of each locus reflects the total number of functional genes within the genome. While this method may include potential errors, such as RNA-sequencing artifacts, establishing single copies of genes allows the identification of WRKY family members and approximate family size. Due to possible variations in splicing and/or incomplete splicing of introns I searched contig assemblies with the longest predicted protein sequence for each gene in the MPGR database. All CrWRKY proteins identified from the MPGR database, described below, were manually verified to contain a WRKY domain. In several cases (*CrWRKY8*, *CrWRKY13*, *CrWRKY17*, *CrWRKY21*, *CrWRKY34*, *CrWRKY37*, and *CrWRKY47*), alignment results among Catharanthus contigs for a locus and the closest matching AtWRKY TF, were utilized to remove a conserved intron following the WRKYGQK consensus sequence or correct a frame shift, to generate a full WKRY domain sequence. As single base pair insertion in CrWRKY8 was not clear by aligning other copies of this contig, the region spanning the insertion was cloned for verification.

Searching for the established invariant consensus sequences WRKYGQK and known alternative WRKYGKK, WRKYGEK, and WRKYGSK consensus sequences from the list of proteins, duplicate results were eliminated and 46 putative WRKY TFs were identified (Table 2.6). Comparatively, MPGR annotated 47 potential CrWRKY TF encoding genes. However, only 35 WRKYs overlapped between manual searches for the consensus motifs and the MPGR annotated datasets.

To further validate the number of WRKY TFs a list of the single longest predicted proteins for each locus was submitted to the National Center for Biotechnology Information Conserved Domain Database (NCBI CDD) and the Samuel Roberts Nobel Foundation PlantTFcat (PlantTFcat) server, for protein domain identification (Marchler-Bauer et al., 2011). The NCBI CDD identified

52 WRKY domain-containing proteins (Table 2.6). Similarly, PlantTFcat (http://plantgrn.noble.org/PlantTFcat/) also identified 52 WRKY domain-containing proteins. The majority of additional proteins identified by NCBI CDD and PlantTFcat as WRKY TFs, had incomplete N-terminal ends of the WRKY domain (CrWRKY11, CrWRKY15, CrWRKY48, and CrWRKY49). One additional predicted WRKY TF, CrWRKY32, contained a WRKYGRK motif. CrWRKY9, which was identified by NCBI CDD, but not PlantTFcat, had an incomplete C-terminal portion of the WRKY domain. Contig Cra15757 was predicted by PlantTFcat to be a WRKY TF. Inspection of this protein sequence did not reveal the presence of a WRKY consensus or zinc finger binding motif. Of the 47 proteins annotated as WRKYs by MPGR, only 40 were found to be true WRKY TFs as identified by NCBI CDD and PlantTFcat. In total, 52 proteins in Catharanthus were predicted as WRKY TFs (Table 2.2).

Of the 52 possible WRKY TFs from Catharanthus, at least 48 appear to be authentic (Table 2.7). The MPGR database contained full WRKY domain sequences for 52 domains from 43 TFs. Partial domain sequences were found for nine WRKYs. 3' rapid amplification of cDNA ends (RACE) or 5' RACE was performed to obtain the necessary domain sequence for 5 WRKYs. 3' RACE was performed on *CrWRKY9*. For *CrWRKY11*, *CrWRKY12*, *CrWRKY15*, and *CrWRKY48*, 5' RACE was used to obtain the rest of the WRKY domain sequence. Clones could not be found for four genes (*CrWRKY49*, *CrWRKY50*, *CrWRKY51*, and *CrWRKY52*). Expression data, available from MPGR, revealed these four WRKYs are not present in any of the 23 samples sequenced. To validate the MPGR expression data, quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to measure the transcript levels of *CrWRKY49*, *CrWRKY50*, *CrWRKY51*, and *CrWRKY52*. Gene specific transcripts for *CrWRKY49*, *CrWRKY50*, *CrWRKY51*, and *CrWRKY52* could not be detected in root, stem, leaf, or whole plant samples. Transcripts for the same four WRKYs could also not be found in 0, 1, 2, and 4 hour JA-treated samples. This data suggests that these predicted partial WRKY sequences are not in any of my samples, and that they may be either artifacts of RNA-sequencing, temporally regulated, or induced by a factor not present in these growing conditions. WRKY TFs are known to play key roles in plant senescence (Robatzek and Somssich, 2001). However, senescing medicinal plant tissues were not utilized for sequencing in the MPGR. Inclusion of senescing tissues may slightly increase the total WRKY number to more closely reflect fully sequenced plant species. Future investigations with different treatment conditions may detect the expression of *CrWRKY49*, *CrWRKY50*, *CrWRKY51*, and *CrWRKY52*.

The Catharanthus WRKY family appears to be one of the smallest reported WRKY TF families to date. Only the moss *Physcomitrella patens*, the lycophyte *Selaginella moellendorffii*, and Castor bean (*Ricinus communis*), with 37, 35, and 47 WRKYs respectively, are reported to have fewer WRKY TFs (Rensing et al., 2008; Rushton et al., 2010; Li et al., 2012). My results suggest that the Catharanthus WRKY family is similar in size to *Cucumis sativus*, *Fragaria vesca*, *Jatropha curcas*, and *Carica papaya* with 55, 56, 58 and 66 WRKYs, respectively (Pandey and Somssich, 2009; Ling et al., 2011; Miao et al., 2012; Xiong et al., 2013). To further investigate the size of the Catharanthus WRKY family, I identified the WRKY TFs from serpentwood (*Rauvolfia serpentina*) transcriptome sequences (Table 2.8) (Góngora-Castillo et al., 2012). Serpentwood is closely related to Catharanthus and also produces pharmaceutically valuable TIAs (Figure 2.2). I found 54 serpentwood TFs, a number close to the 52 WRKYs identified in Catharanthus. The number of WRKYs belonging to each subgroup was also similar between these two species (Figure 2.3). However, as both serpentwood and Catharanthus WRKYs were identified from transcriptome data, the actual size of the families may be larger. To address this possibility, I identified WRKY families from tomato (*Solanum lycopersicum*) (Consortium, 2012), potato (*Solanum tuberosum*) (Consortium, 2011), pepper (*Capsicum annuum*) (Kim et al., 2014), and bladderwort (*Utricularia*

gibba) (Ibarra-Laclette et al., 2013), all species of which have complete genome sequence available (Table 2.8). I identified 81 WRKY TFs in tomato (Table 2.8, Figure 2.4), as previously reported (Huang et al., 2012). Bladderwort, pepper, and potato each contained 65, 66, and 75 complete WRKY TFs, respectively (Table 2.8, Figure 2.3, and Figure 2.4). These data suggest the ancestor of the Gentianales (Catharanthus and serpentwood), Lamiales (bladderwort), and Solanales (pepper, potato and tomato) likely contained around 65 WRKY TFs. Therefore, I conclude that greater than 75% (52 out of 65) of Catharanthus and serpentwood WRKY TFs were identified from transcriptome data. Together, the six asterid species contained a similar number of WRKY TFs as found in the Brassicales, Arabidopsis and papaya (Wu et al., 2005; Pandey and Somssich, 2009). These data, combined with that from other WRKY families (Pandey and Somssich, 2009; Ling et al., 2011; Miao et al., 2012; Xiong et al., 2013), suggests that *Brachypodium distachyon* (86 WRKYs), *Oryza sativa* ssp. *japonica* (105 WRKYs), *Populus trichocarpa* (104 WRKYs), and *Zea mays* (119 WRKYs) may contain atypically large WRKY families (Dong et al., 2003; Zhang and Wang, 2005; Pandey and Somssich, 2009; Tripathi et al., 2012; Wei et al., 2012) compared to other angiosperms. Arabidopsis and rice both contain expansions in the group III WRKY subfamily (Wu et al., 2005; Zhang and Wang, 2005), whereas an expansion of group IIe occurs in potato and tomato (Huang et al., 2012) (Table 8). I did not find any evidence for subfamily expansions in Catharanthus or serpentwood (Table 2.8; Figure 2.5).

Phylogenetic Analysis of Catharanthus WRKY TFs

To determine the relationship among Catharanthus WRKY TFs, a phylogenetic tree was constructed with 282 WRKY domains from 243 TFs from Catharanthus, *Amborella trichopoda*, Arabidopsis and rice (*Oryza sativa* ssp. *japonica*) (Figure 2.5). WRKY sequences from *Chlamydomonas reinhardtii* (XP_001692342), *Dictyostelium discoideum* (XP_643786), and *Giardia lamblia* (XP_001708807) were included as an outgroup. Additional, outgroup sequences include human *GCMa* (BAA13651) and *FLYWCH CRAa* (EAW85450). I used 84 and 105 WRKY domains from 72 and 94 TFs, from Arabidopsis and rice respectively, to construct the phylogenetic tree. Thirty-five domains from 29 Amborella WRKY TFs were also included in the phylogenetic analysis (Project, 2013). I incorporated the WRKY sequences of Amborella, an evolutionary basal angiosperm, to reduce long-branch attractions during phylogenetic tree construction. Amborella was selected over *Physcomitrella patens* (moss) and *Selaginella moellendorffii* (spikemoss) since the WRKY sequences from this phylogenetically important species remains unreported yet provides valuable insights about WRKY evolution. The phylogenetic tree contained 58 domains from 48 CrWRKY TFs (Table 2.9). To ascertain potential functions, I compared Catharanthus and Arabidopsis WRKY TFs by identifying orthologs. I identified 11 group I, 32 group II, and five group III WRKY TFs in Catharanthus. Group II WRKY TFs can be classified into groups IIa, IIb, IIc, IId, or IIe (Eulgem et al., 2000). In Catharanthus, I identified three group IIa, five group IIb, thirteen group IIc, four group IId and seven group IIe WRKY TFs.

Evolutionarily, group I WRKY TFs, such as those found in algae, are some of the most ancient of WRKYs (Wu et al., 2005; Zhang and Wang, 2005). Recent evidence suggests that the group I WRKYs, and other WRKY TFs, originated from an ancestral group IIc-like domain (Brand et al., 2013). As previously reported for this group, ten group I CrWRKYs contained two WRKY domains with the N-terminal domain forming a separate clade and the C-terminal WRKY domains forming part of the group IIc clade (Wu et al., 2005; Zhang and Wang, 2005). To identify orthologs and paralogs in Amborella, Arabidopsis, Catharanthus and rice, I used OrthoMCL (Li et al., 2003).

I found Catharanthus contains six coorthologs (CrWRKY2, CrWRKY3, CrWRKY4, CrWRKY5, CrWRKY8 and CrWRKY51) to AtWRKY33 (Table 2.10). According to the phylogenetic tree, CrWRKY5 is most closely related to AtWRKY33 (Figure 2.5).

Group IIa was the only group of WRKYs that had similar numbers between Catharanthus and Arabidopsis. Rice had four group IIa WRKY TFs whereas both Arabidopsis and Catharanthus each contained three. The three group IIa WRKYs from Catharanthus are coorthologs to AtWRKY40 (Table 2.10).

Previous reports indicate some plants contain variants of the highly conserved WRKYGQK domain, such as WRKYGKK, WRKYGEK, WRKYGSK, among others (Zhang and Wang, 2005). Variation in this region can reduce, eliminate, or alter DNA binding activity (Duan et al., 2007). WRKY TFs with variants of the consensus sequence may recognize different cis-elements. I found AtWRKY50, AtWRKY51, and AtWRKY59 belong to the group IIc WRKY subfamily and possess a WRKYGKK motif as previously reported (Eulgem et al., 2000). Two CrWRKYs were identified that contain variants of the highly conserved WRKYGQK motif. CrWRKY23 and CrWRKY32 contain WRKYGKK and WRKYGRK sequence motifs, respectively. Mutagenesis of the conserved glutamine was previously demonstrated to reduce, but not eliminate, DNA binding (Duan et al., 2007). More recently, AtWRKY50 was found to generally bind the GAC core of the W-box with less preference for 5' or 3' bases (Brand et al., 2013). Therefore, both WRKYGKK and WRKYGRK variants are expected to still bind DNA. *Nictotiana tabacum* WRKY12, containing a WRKYGKK motif, has been found to bind the WK-box cis-element (TTTTCCAC), but not the W-box, which regulates expression of the plant defense gene PATHOGENESIS RELATED1 (van Verk et al., 2008). However, *Capsicum annuum* WRKY1, a WRKYGKK motif WRKY TF involved in plant defense, can still recognize the W-box (Oh et al., 2008). The phylogeny and protein alignment of the DNA-binding WRKY domains of CrWRKY23, CaWRKY1 and NtWRKY12 revealed more similarity of CrWRKY23 to CaWRKY1 (Figure 6A and B), suggesting that despite the variant WRKYGKK motif, CrWRKY23 likely still recognizes the W-box element or at least the GAC core. Although *Hordeum vulgare* WRKY46 (SUSIBA2) contains a WRKYGQK motif, HvWRKY46 recognizes the Sugar Responsive (AATAGAAAA) and W-box cis-elements to regulate barley genes involved in starch metabolism (Sun et al., 2003). This leaves open the possibility that some Catharanthus WRKY TFs may not recognize W-box cis-elements.

Group III WRKY TFs are believed to have dramatically expanded during the evolution of angiosperms and can be classified into different subgroups depending on the species (Kalde et al., 2003; Huang et al., 2012). Arabidopsis contains fourteen group III WRKY TFs which are further divided into eight group IIIa and six group IIIb. In Arabidopsis, most group III WRKY transcription are induced by plant pathogens (Kalde et al., 2003). I identified only five group III WRKY TFs in Catharanthus. Similarly, I identified 5 group III WRKYs in serpentwood and bladderwort (Table 8). Proportionally, the number of group III CrWRKY TFs is the smallest compared to Arabidopsis. The low number of group III CrWRKYs, and similar number from serpentwood and bladderwort, suggests this group has not undergone significant expansion such as occurred in rice or Arabidopsis (Huang et al., 2012). CrWRKY1 and CrWRKY48 were found to be coorthologs to AtWRKY70 and AtWRKY54 (Table 2.10). AtWRKY70 modulates SA and jasmonate signaling (Li et al., 2004). Interestingly, CrWRKY1 differentially directs the flow of unknown precursors into TIA products (Suttipanta et al., 2011), a feature possibly governed by its jasmonate responsive gene expression. CrWRKY45, CrWRKY46, and CrWRKY47 are coortholgs of AtWRKY41, AtWRKY46, and AtWRKY53. AtWRKY46 and AtWRKY53 are partially functionally redundant in regulating plant defense (Hu et al., 2012).

The Yuan lab previously reported the role of CrWRKY1 in regulating gene expression and TIA accumulation in Catharanthus (Suttipanta et al., 2011). CrWRKY1 is a group III WRKY with overall protein sequence homology closest to AtWRKY70, and corresponds to MPGR contig number Cra16284. Phylogenetically, CrWRKY1 is located towards the base of the group III clade and does not clearly group with its Arabidopsis or rice orthologs (Figure 2.7A). To identify the unique feature of CrWRKY1, I analyzed the protein sequence alignment. The invariant tryptophan starting the WRKYGQK motif was used as the reference point for comparing alignments. Alignment of CrWRKY1 to other group III WRKY TFs revealed that CrWRKY1 lacks an amino acid between the two conserved cysteine residues at positions 21 and 29 (Figure 2.7B). The closest rice WRKY TFs, OsWRKY21, OsWRKY61 and OsWRKY47, all have altered spacing within the WRKY domain sequence. OsWRKY47 possesses an additional proline residue between the WRKYGQK sequence and the conserved arginine residue at position 16 (Eulgem et al., 2000). The conserved arginine at position 16 was changed to threonine followed by a TQS motif in OsWRKY61. OsWRKY47 contains an extra DDP sequence between positions 41 and 42 compared to all other Arabidopsis, rice, and Catharanthus group III WRKY TFs. The altered spacing in the WRKY domain may give these WRKY TFs unique structural properties important for target gene regulation.

Expression Profiling Reveals Multiple Jasmonate Responsive CrWRKYs

MPGR provides RNA-sequencing based expression data from different tissues for all sequenced medicinal plants. For Catharanthus, RNA-sequencing data is also available for different tissues, seedlings, cell suspension cultures, and hairy root cultures. These data provide an opportunity to understand the induction of WRKY genes in response to conditions that induce the TIA pathway. Furthermore, several treatments allow for comparison of induction to the same hormone in varying tissues. In response to JA, a potent and important elicitor of natural product formation, including TIAs, in Catharanthus and other medicinal species (Aerts et al., 1994; Zabetakis et al., 1999; Biondi et al., 2000; Yu et al., 2000), MPGR expression data indicates multiple WRKY TFs are either up or down regulated in Catharanthus.

To identify or validate WRKY TFs that are up- or down regulated by JA, I performed qRT-PCR on whole plant samples (root, stem, and leaves) that were collected from one month old soil grown plants at 0, 1, 2, and 4 hours after JA treatment. Successful induction with JA was verified by measuring JAZ2 expression (Figure 2.8A). To determine which WRKY TFs were possible regulators of TIA production, I sought to measure the expression of multiple pathway genes such as *G10H*, *TDC*, and *STR* (Figure 2.8B). These genes were selected to represent early (*G10H* and *TDC*) and middle portions (*STR*) of the TIA pathway. Any WRKY induced prior or simultaneously to these genes possibly could regulate that corresponding portion of the pathway and any subsequent segments.

I selected at least two genes from each CrWRKY subgroup. Four genes (*CrWRKY5*, *CrWRKY8*, *CrWRKY13*, and *CrWRKY28*) were selected based on involvement of their orthologs in regulating secondary metabolism genes in other species. Analysis of sixteen CrWRKYs identified twelve which displayed significant changes in expression in response to jasmonate (Figure 2.8C-D). The fold change for most CrWRKYs with a significant response to JA was 2 fold or less (Figure 2.8C-D), similar to my microarray findings for jasmonate responsive Arabidopsis WRKYs (Table 2.4). *CrWRKY8* was up-regulated 1 hour after JA treatment then decreased by 4 hours after treatment. *CrWRKY5* was down-regulated significantly at both 2 and 4 hours after JA treatment. *CrWRKY13*, similar to the ABA responsive *AtWRKY40*, was significantly up-regulated 1 and 2 hours

after JA treatment. *CrWRKY38* was up-regulated by 2 hours after JA treatment. *CrWRKY18*, *CrWRKY21*, *CrWRKY41*, *CrWRKY45*, and *CrWRKY48* were all significantly down-regulated at all time points after JA treatment. Two WRKYs, *CrWRKY26* and *CrWRKY36*, had a bimodal expression pattern that was down-regulated at 1 and 4 hours, but not 2 hours, after JA treatment. A bimodal expression pattern has been observed for some regulators of the TIA pathway (Menke et al., 1999). *CrWRKY35* was down-regulated 2 and 4 hours after treatment. Nine of 12 CrWRKYs analyzed were down-regulated to JA treatment. In total at least 25% (12 of 48), and probably more, of CrWRKY TFs are regulated by jasmonate. Of the twelve jasmonate responsive CrWRKYs, nine have an AtWRKYs ortholog which were either previously reported and/or identified here by microarray analysis ($p<0.05$ and survived B-H FDR) as differentially regulated by jasmonate (Figure 2.8C-D and Table 2.3). When compared to the less stringent list of AtWRKYs which had expression significantly changed ($p<0.05$) in response to jasmonate, but did not survive the B-H FDR, all twelve CrWRKYs have orthologs to jasmonate responsive AtWRKYs.

To identify potential WRKYs regulating TIA biosynthesis through the jasmonate signaling pathway, I compared the induction times of WRKY genes (Figure 2.8C-D) to early and mid-biosynthetic genes of the pathway (Figure 2.8B). Similar to previous reports in cell cultures (Menke et al., 1999), induction of *STR* by JA began approximately 2 hours after treatment. However, expression of *G10H* decreased starting at 1 hour after JA treatment, and was further down-regulated 4 hours after treatment (Figure 2.8B). TDC transcript levels remained unchanged to JA treatment in mature Catharanthus plants. Prior reports of *TDC* (Menke et al., 1999; Van der Fits and Memelink, 2000; Collu et al., 2001; Peebles et al., 2009; Wei, 2010) and *G10H* (Van der Fits and Memelink, 2000; Collu et al., 2001; Peebles et al., 2009; Wei, 2010) transcript induction by jasmonate treatment was identified in seedlings, hairy roots, or cell cultures; however, my experiments were performed in intact mature Catharanthus plants. Expressions of WRKY TFs possibly contributing to TIA regulation are predicted to be altered before early and mid steps of the TIA pathway. Expression of *CrWRKY8*, *CrWRKY13*, *CrWRKY18*, *CrWRKY21*, *CrWRKY26*, *CrWRKY36*, *CrWRKY41*, *CrWRKY45* and *CrWRKY48* changed by 1 hour after JA treatment indicating these WRKYs could possibly regulate the expression of early TIA pathway genes. Altered expression of all twelve CrWRKYs responding to JA occurred by 2 hours after treatment, the same time at which significant induction of *STR* occurred. Contrary to *TDC* and *G10H*, which contain four and one W-boxes in their promoters respectively (Suttipanta et al., 2007; Suttipanta et al., 2011), the characterized *STR* promoter does not contain any W-box elements for WRKYs to bind, but this does not exclude the possibility that WRKY regulate other TFs controlling STR expression. The spatio-temporal regulation of CrWRKYs, by reducing TDC responsiveness and down-regulating G10H, is one possible reason why mature Catharanthus plants do not accumulate TIA in response to jasmonate treatment (Pan et al., 2010). These findings suggest that all CrWRKYs I ascertained as differentially expressed in response to jasmonate are possible regulators of early and middle steps of TIA biosynthesis. Presumably, these CrWRKY could also regulate downstream steps of the pathways which are temporally expressed later.

The Jasmonate Response of Catharanthus WRKY Varies Among Plant Culture Conditions

I sought to determine the similarities between my qRT-PCR results and the transcriptome data published by MPGR. As I used one month-old plants to quantify gene expression, and no data on JA treated mature plants are provided by MPGR, I correlated my data to three different datasets each representing one aspect of my samples (5 day JA treated seedlings, 6 hour JA

treated cell suspension cultures, and 24 hour JA treated hairy root cultures). Seedlings treated with JA most closely represent my samples in physiology as both are whole plant tissues; however, the MPGR dataset used seedlings rather than mature plants, which may respond to JA differently (Pan et al., 2010). While cell cultures are considerably different in physiology from whole plants, the earliest time sample (6 hours after JA treatment) was closest to my sample times of 1, 2, and 4 hours after JA treatment. Hairy root cultures require several weeks to develop to sufficient size; therefore, the age of this tissue most likely represents a similar age as my plant samples, despite my shorter JA treatment time. The Pearson correlation coefficient was calculated to measure the relationship between the datasets. CrWRKY11 and CrWRKY21 were excluded from the correlations as they appear two times in the MPGR datasets without expression values. The correlation between MPGR seedling and cell culture datasets, as well as between cell culture and hairy root datasets, was quite low (r=0.179 and r=0.227 respectively), indicating considerable difference in CrWRKY response to JA in cell culture systems. However, there was a high correlation between seedling and hairy root MPGR datasets (r=0.892) for JA treated CrWRKY genes. My qRT-PCR data showed that the correlations ranged from 0.061, between 1 hour JA treated plants and 24 hour JA treated hairy roots, to 0.790, between 1 hour JA treated plants and 6 hour JA treated cell suspension cultures (Figure 2.9A). Overall, the three MPGR datasets correlated well with all three time points of qRT-PCR data (median value of the 9 correlations = 0.555) indicating similar expression changes in response to jasmonate treatment. Increasing time after JA treatment in whole plants increased the correlation with seedling and hairy root cultures. Cell cultures, despite higher similarities in the 4 hour JA treated plant and 6 hour JA treated cell culture time frame, showed a lower correlation between the 1 hour JA treated plant and 6 hour JA treated cell cultures. Similar to Arabidopsis, these findings in Catharanthus suggest significant differences exist between jasmonate response in various cultural conditions, including intact seedlings, adult plants, cell cultures, and hairy root cultures.

Gene expression clusters often contain genes with related functions (Eisen et al., 1998), including those in natural product formation (Vanderauwera et al., 2005). Recently, clustering of MPGR expression data has aided the identification of Catharanthus IRIDOID SYNTHASE (Geu-Flores et al., 2012). To identify potential clusters of CrWRKY TFs with similar expression pattern which may indicate WRKY functions, I performed a hierarchical clustering. Unsupervised agglomerative hierarchical clustering of 23 transcriptome gene expression datasets from MPGR revealed three primary clusters: a plant tissue cluster, a hairy root cluster, and a protoplast cluster (Figure 2.9B). Clusters of plant culture type indicate a greater difference between cultural conditions than between JA treatment. However, clear differences exist between JA treated and untreated samples within subgroups of each primary cluster. Clustering of CrWRKY gene expression revealed three primary clusters. CrWRKYs of cluster one were most up-regulated in different plant tissues, suggesting a role in plant development. The second cluster of WRKY genes is up-regulated in hairy root cultures. Members of cluster two may be important for regulating metabolism and resource direction into primarily root produced alkaloids, such as ajmalicine and serpentine. Identification of CrWRKY1, which plays a role in serpentine production, in cluster two supports this idea. CrWRKY46, ortholog to AaWRKY1, a trichome expressed WRKY in *Artemisia annua*, was also found in this cluster. CrWRKY34 and CrWRKY35 orthologous to PqWRKY1 which is suggested to regulate terpene biosynthesis in *Panax quinquefolius* (American ginseng) roots also occur in this cluster. The third cluster consisted of CrWRKY that were up-regulated in response to JA or yeast extract (YE), an elicitor of TIA biosynthesis, in protoplasts. Most of this cluster was also up-regulated in hairy root cultures. Most CrWRKYs orthologous to WRKYs regulating natural product formation in other species were identified in this cluster. Importantly,

four CrWRKYs (CrWRKY2, CrWRKY5, CrWRKY13, and CrWRKY28), similar to those with known roles in secondary metabolism (Table 2.10), were identified as part of the same sub-cluster in cluster three. The four members of this cluster may play key roles in regulation of natural product formation in Catharanthus. A second sub-cluster of cluster three, composed of six members, contained five CrWRKYs (CrWRKY4, CrWRKY14, CrWRKY22, CrWRKY26, and CrWRKY27) which are orthologs to WRKY TFs regulating natural products in other species. The sixth member of this sub-cluster, CrWRKY41, is also a jasmonate responsive WRKY. Members of this sub-cluster may also be important for regulation of natural products in Catharanthus. CrWRKY TFs, determined by qRT-PCR to be JA responsive, were distributed across all three clusters, indicating jasmonate broadly regulates WRKYs from each cluster.

Predicted Role of CrWRKY Orthologs in Secondary Metabolism

Catharanthus produces alkaloids, terpenes and latex, all classes of compounds that contain biosynthetic genes involved in their production which have been implicated to be regulated by WRKY TFs in other species (Xu et al., 2004; Kato et al., 2007; Zhang et al., 2012). To associate biosynthesis of natural compounds with jasmonate-responsive CrWRKYs, I compared Catharanthus WRKY TFs to those known to control secondary metabolism in other plant species (Table 10). CjWRKY1 is involved in the regulation of the benzylisoquinoline alkaloid berberine (Kato et al., 2007). In Catharanthus, CrWRKY28 grouped closely with CjWRKY1 (Figure 5). The ortholog of CjWRKY1 in Arabidopsis is AtWRKY75.

AtWRKY33 plays a role in regulating biosynthesis of camalexin, an indole ring and N-containing defense molecule, and functions downstream of MITOGEN-ACTIVATE PROTEIN KINASE 3 and 6 (Mao et al., 2011). Recently, CrMPK3 was shown to regulate TIA accumulation (Raina et al., 2012). As Catharanthus produces over 130 different TIA metabolites, the multiple coorthologs to AtWRKY33 may be important for regulating diverse products of this pathway. However, further experiments are needed to demonstrate whether the orthologous TFs in Catharanthus act downstream of CrMPK3 and are involved in TIA biosynthesis. Interestingly, CmWRKY1, from Chlamydomonas, was also an ortholog to AtWRKY33, suggesting a possible early function of this TF in defense and regulating secondary metabolism. Chlamydomonas produces gelatin structures and quarum sensing mimetic compounds to deter rotifer predators and bacterial competitors (Teplitski et al., 2004; Lurling and Beekman, 2006).

AaWRKY1, which is involved in regulating the accumulation of artemisinin, has three coortholgs in Catharanthus, CrWRKY45, CrWRKY46, and CrWRKY47 (Table 2.10). In Arabidopsis three group IIIa WRKYs, AtWKRY41, AtWRKY46, and AtWRKY53 are coortholgs to AaWRKY1. In Catharanthus, increased production of HMGR and terpenes have negatives effect on the accumulation of certain TIAs (Ayora-Talavera et al., 2002). In *A. annua*, AaWRKY1 affects the expression of 3-HYDROXY-3-METHYLGLUTARYL-COA REDUCTASE (HMGR) (Ma et al., 2009), a rate limiting enzyme in the mevalonate pathway. Both AaWRKY1 and CrWRKY46 are jasmonate responsive genes; therefore, at least CrWRKY46 may have an evolutionarily conserved function in regulating the flux of carbon into Catharanthus terpenes.

The rate limiting enzyme in the production of paclitaxel, DBAT, is regulated by TcWRKY1 (Li et al., 2013). Phylogenetically, TcWRKY1 is basal to the group IIa and IIb clades (Figure 2.1). Catharanthus CrWRKY16 and CrWRKY20 were identified as coorthologs to TcWRKY1 (Table 2.10).

In Arabidopsis, the group IIb AtWRKY6, AtWRKY31, and AtWRKY42 were found to be coortholgs to TcWRKY1 (Table 10).

GaWRKY1, from cotton, regulates a sesquiterpene cyclase leading to the production of gossypol (Xu et al., 2004). AtWRKY40 was found to be the Arabidopsis ortholog to GaWRKY1 (Table 2.10). AtWRKY18 and AtWRKY60 formed their own ortholog group independent of AtWRKY40. Catharanthus, however, contains three coortholgs to GaWRKY1, CrWRKY13, CrWRKY14, and CrWRKY15. As with AtWRKY40, these CrWRKYs may have a role in negative regulation of ABA response (Chen et al., 2010) and positive regulation of jasmonate responses (Pandey et al., 2010). Supporting this idea, like AtWRKY40 and GaWRKY1, I found expression of CrWRKY13 was induced by jasmonate treatment. Drought, salinity, and cold all affect TIA accumulation in Catharanthus (Jaleel et al., 2008; Dutta et al., 2013), thus at least CrWRKY13 may function in regulating the accumulation of TIAs in response to abiotic stress and plant defense.

HbWKRY1 expression is related to latex production in rubber trees (Zhang et al., 2012). In Catharanthus, CrWRKY21, CrWRKY22, CrWRKY26, CrWRKY27, CrWRKY29, and CrWRKY33 are coorthologs to HbWRKY1. Six coothrologs (AtWRKY8, AtWRKY12, AtWRKY23, AtWRKY28, AtWRKY48, AtWRKY51, and AtWRKY71) to HbWRKY1 exist in Arabidopsis (Table 4). Like HbWRKY1, at least four Arabidopsis (AtWRKY8, AtWRKY28, AtWRKY48, and AtWRKY51) and two Catharanthus WRKYs (CrWRKY21 and CrWRKY26) are regulated by jasmonate. Phylogenetically, the group IIc AtWRKY23 and CrWRKY26 are the WRKYs most similar to HbWRKY1 in Arabidopsis and Catharanthus, respectively (Figure 2.5). As Catharanthus also produces a latex compound, the jasmonate regulated CrWRKY26 or one of the other paralogs, may function in the regulation of latex or terpene production in Catharanthus.

Heterologous over-expression, in Arabidopsis, of the JA responsive American ginseng WRKY TF, PqWRKY1, increased drought and salt stress tolerance, in addition to regulating terpene biosynthetic genes (Sun et al., 2013). AtWRKY7 in Arabidopsis and CrWRKY34 and CrWRKY35 in Catharanthus are orthologs to PqWRKY1 (Table 2.10). Contrary to the report by Sun et al. (2013), which classifies PqWRKY1 as a group IIc WRKY, I found PqWRKY1 actually falls within the IId subgroup when compared to the entire Arabidopsis WRKY family (Figure 2.5). AtWRKY7, CrWRKY35, and PqWRKY1 are each regulated by jasmonate supporting the possible conserved evolutionary function of these proteins in regulating terpene biosynthesis.

Kalde et al. (2003) reported a role of most Arabidopsis group III WRKY TFs in plant defense. Overall, no clear trend was observed for WRKY TFs possibly involved in secondary metabolism belonging to a specific group or subgroup. Further work is needed to verify the predicted roles of these CrWRKYs in the regulation of secondary metabolism.

Comparative genetics across species has provided invaluable information that lead to the isolation and functional understanding of several key regulators in natural product formation. In Arabidopsis, the bHLH factor AtMYC2 is known as a central regulator of jasmonate signaling pathway. The orthologs of AtMYC2, CrMYC2 and NtMYC2, from Catharanthus and tobacco, respectively, have thus been isolated and characterized. While Arabidopsis does not produce TIAs or nicotine, CrMYC2 and NtMYC2 act in the jasmonate signaling pathway to regulate biosynthesis of these metabolites (Shoji and Hashimoto, 2011; Zhang et al., 2011). Moreover, AtMYC2 can bind the jasmonate-responsive elements present in the promoter of Catharanthus ORCA3, an AP2/EFR TF gene, and activates its expression, illustrating the conserved nature of these orthologous regulators (Montiel et al., 2011). These reports further strengthen my reasoning for cross-species comparison of WRKY TFs from Catharanthus, Arabidopsis, and other medicinal plant

to identify regulators conserved in jasmonate response and possibly secondary metabolite production.

Comparison of CrWRKYs with orthologs from other species, that are known to regulate natural products or respond to jasmonate treatment, helped us develop a model for WRKY regulation of TIA biosynthesis in Catharanthus (Figure 2.10). In this model, jasmonate acts as a central regulator of the TIA pathway with both positive and negative effects on WRKYs. Phytohormones, including ABA, ethylene and gibberellin (GA), are also likely involved in CrWRKYs regulation. Overall, this work provides a fundamental base for which future experiments can be designed to help elucidate the molecular mechanisms controlling the biosynthesis of highly valuable TIAs.

Conclusion

Taken together, my results illustrate a role for the Arabidopsis WRKY family in mediating the jasmonate response pathway. These findings strengthened my reasoning for investigating Catharanthus jasmonate responsive WRKY TF which are potentially involved in regulation of TIA biosynthesis. Results from Arabidopsis and Catharanthus suggest that the regulation of WRKY gene expression in response to jasmonate is dependent upon environmental and spatio-temporal context. Such information can be important in designing metabolic engineering projects. Furthermore, I identified numerous jasmonate responsive orthologs between AtWRKY and CrWRKY TFs that may be functionally conserved or partially conserved. The jasmonate responsive CrWRKYs are potential candidate TFs for having key roles in modulating jasmonate signaling and regulating TIA biosynthesis. Information on how AtWRKYs response to various phytohormones and stresses may also apply to Catharanthus. This information may be useful for understanding how other phytohormones also contribute to the regulation of TIA production. Moreover, elucidation of CrWRKY functions may provide valuable insights into the regulation of natural product biosynthesis in other medicinal plants.

Materials and Methods

Plant Growth Conditions

Catharanthus 'Little Bright Eyes' seeds were surface sterilized and were germinated in the dark at 30°C for 3 days on MS plates, before being then transferred to an ambient temperature 24 h light photoperiod tissue culture room for an additional 4 days. Seedlings were transferred to soil and grown at ambient temperature under 24 h light. Samples were collected from 1 month-old Catharanthus plants treated with JA for 0, 1, 2, or 4 hours. The JA experiment was performed once with each time having three replicates. Three plants were combined from each replicate time sample. JA treatment consisted of spray application of 100 µM JA then placing plants under a clear plastic dome sealed with tape. Whole plants were harvested, roots quickly washed, and then frozen in liquid nitrogen.

WRKY TF Identification

Contigs translated into protein sequence were downloaded from the MPGR. The single longest copy of each contig translated into protein sequence was identified using Microsoft Excel. Each unique contig number, which translated into some protein sequence, was determined to represent a unique gene distinguished by locus number. To differentiate potential WRKY genes each distinct contig locus number, but not different length variants of the same locus number, were considered as a unique product. As observed in Arabidopsis and other species, the multiple contig copies that comprise many of the loci may represent splice variants, not fully sequenced transcripts or different alleles. A Microsoft Excel file containing all protein encoding contigs was searched to manually identify WRKY and WRKYGQK invariant motif containing proteins. A FASTA file of the single longest protein encoding contig for the entire genome was submitted to the NCBI CDD and PlantTFcat servers to identify whole and partial WRKY domains containing contigs. The process was performed for *Amborella trichopoda*, *Arabidopsis thaliana*, *Capsicum annuum*, *Catharanthus roseus*, *Oryza sativa* ssp. *japonica*, *Rauvolfia serpentina*, *Solanum lycopersicum*, *Solanum tuberosum*, and *Urticularia gibba*. A file containing WRKY TFs from Catharanthus, Amborella, Arabidopsis, and rice was submitted to OrthoMCL (Li et al., 2003) to identify orthologs and paralogs. WRKY TFs involved in regulating secondary metabolism from other species were also included. GenBank accession numbers for medicinal plant WRKY TFs included are: AaWRKY1 (FJ390842), CjWRKY1 (AB267401), TcWRKY1 (JQ250831), GaWRKY1 (AY507929), HbWRKY1 (GU372969), and PqWRKY1 (AEQ29014).

Phylogenetic Tree Construction

The unrooted phylogenetic trees for Catharanthus, Amborella, Arabidopsis, and *Oryza sativa* ssp. *japonica* and medicinal plant WRKY TFs were constructed using the MEGA5 software. The neighborhood joining method, with bootstrap values of 2000, was utilized to conduct the phylogeny test. The analysis used p-distance of amino acid sequence to determine substitution rate. Gaps or missing data were excluded as needed, according to the pairwise deletion option. Phylogenetic trees analyzing the bladderwort, pepper, potato, serpentwood, and tomato WRKY families were constructed in the same way.

RNA Extraction

RNA was extracted using an extraction buffer composed of 1% 1,5-naphthalenedisulfonic acid and 4% p-aminosalicylic acid prepared in diethylpryocarbonate (DEPC) treated water. A 5 M sodium hydroxide solution was added until the extraction buffer was fully dissolved. For each RNA sample 5 mL of extraction buffer solution was mixed with 5 mL of liquefied phenol. Ground samples were added to the extraction buffer/phenol solution, vortexed 1 minute, 5 mL of chloroform added, then vortexed again. Samples were spun down for 10 minutes at 6000 rpm at 4°C. The aqueous phase was transferred to a 50 mL centrifuge tube and 1/10th the volume of 3 M sodium acetate (pH 5.3) was added along with 2 times the volume of chilled 100% ethanol. Samples were incubated on ice 1 hr prior to centrifugation. The supernatant was discarded and the pellet dried for 30 minutes. The dried pellet was resuspended in 4 mL of autoclaved DEPC treated water and 2.5 mL of 8 M lithium chloride and incubated at 4°C overnight. The RNA was then precipitated by centrifuging at the above. The pellet was rinsed with chilled DEPC treated

80% ethanol. The ethanol was decanted and the RNA allowed to dry for 30 minutes before resuspending in sterile DEPC treated water.

cDNA Synthesis and Gene Expression

Synthesis of first strand cDNA from total RNA isolated from plant tissue and quantitative reverse transcription polymerase chain reaction (qRT-PCR) were performed as previously reported (Suttipanta et al., 2007). Samples for the JA treatment consisted of 3 biological replicates each with 3 technical replicates. The comparative cycle threshold method was used to measure the transcript levels. All primers used for qRT-PCR can be found in Supplemental Table 7. Significant differences in gene expression were calculated using the Student's T-test. P-values of 0.05 and 0.01 were considered significant and highly significant, respectively.

5' and 3' Rapid Amplification of cDNA Ends (RACE) and Cloning

5' and 3' RACE was performed using the RACE kit (Invitrogen) as directed by the manufacturer. A nested set of PCRs was performed to isolate the target sequence. The first PCR reaction used the AAP primer and a gene specific primer; whereas the second nested reaction used the AUAP primer and a second gene specific primer. 3' RACE was performed as for gene expression cDNA synthesis with the modification of using the 3' AP primer to create a 3' adapter. The 3' target sequence was amplified through PCR using a nested set of gene specific primers and the adapter specific 3' AUAP primer. All 3' and 5' RACE primers are listed in Supplemental Table 8.

Cloning and Sequencing of Partial WRKY Domains

Partial WRKY domain sequences to be cloned were amplified using 5' or 3' RACE. The sequence was then separated on an agarose gel and the DNA purified using a Wizard® SV Gel and PCR Clean-Up System (Promega). The purified DNA was ligated into pGEM-T Easy vector (Promega). Plasmid isolation was performed using a Wizard ® Plus SV Minipreps DNA Purification System (Promega). 250-300 ng plasmid DNA were sequenced with either the T7 or SP6 primer using DTCS Quick Start (Beckmann) according to the manufacturer's protocol. Sequencing was performed with a CEQ™ 8000 Genetic Analyzer System (Beckman Coulter) Sanger sequencer.

Arabidopsis Microarray Analysis and Gene Expression

Jasmonate treated microarray datasets were collected from NCBI, EMBL, and TAIR. RMA Express was used for array normalization of each experiment (Bolstad et al., 2003). Background adjustment, quantile normalization, and median polish were applied. Data was exported as log transformed data then analyzed by two-way ANOVA using the MEV software (Saeed et al., 2003). Two-way ANOVAs were performed on each dataset to determine response to jasmonate treatment and another variable (genotype or time). Controls and probes not linked to a gene were eliminated post-ANOVA prior to application of the false discovery rate. The B-H FDR was calculated in Microsoft Excel according to Thissen et al. (2002). Significant differences ($p<0.05$) were determined before and after application of the Benjamini-Hochberg false discovery rate (B-

H FDR). Significant differences before the B-H FDR was applied were both included because qRT-PCR for CrWRKY TFs indicated small (less than 2 fold), yet significant, changes to jasmonate treatment.

Hierarchical Clustering and Correlations

The Pearson correlation coefficient was calculated to measure the relationship between qRT-PCR and MPGR datasets. First, the fold change for Catharanthus WRKY genes from the MPGR dataset was calculated in reference to 0 hour control treatments. Differences between control and JA treated datasets were then adjusted using the Catharanthus reference gene EF1α (Cra3894) as an internal control (Wei, 2010), as this gene was used as the internal control for qRT-PCR expression measurements. The correlation coefficient between fold changes in expression was calculated using Microsoft Excel. As a control, the expression of TIA biosynthetic genes was analyzed for JA responsive induction in seedling, protoplast, and hairy root datasets. As anticipated expression of TIA biosynthetic genes increased upon JA treatments, although the magnitude of response differed between datasets.

Unsupervised agglomerative hierarchical clustering was performed using the GenePattern (Reich et al., 2006) website (http://genepattern.broadinstitute.org). The Pearson correlation was used as a distance measure for both row and column clustering. The clustering method was pairwise-average linkage with a row centering. Row centering was performed by subtracting the median value of each row. A global color scheme using a color gradient was applied for visualization. Purple, black, and green indicate increased gene expression, no change and decreased gene expression respectively. For Arabidopsis microarray data expression values were first analyzed with RMA Express and log transformed values exported for analysis with GenePattern (Bolstad et al., 2003; Reich et al., 2006).

Table 2.1. WRKY Differentially Expressed to Jasmonate. The Arabidopsis WRKY transcription factors differentially expressed in response to jasmonate treatment in five experiments before and after the application of the Benjamini-Hochberg false discovery rate.

			Genes after Two-Way ANOVA (P=0.05)			Genes after Two-Way ANOVA (P=0.05) and B-H FDR		
Dataset	**Source**	**Samples Used**	**No. Genes**	**No. WRKY Genes**	**WRKYs**	**No. Genes**	**No. WRKY Genes**	**WRKYs**
E-ATMX-13	EMBL	JA treated timecourse (0.5, 2, and 6 hr) in cell suspension cultures	2819	11	WRKY6, WRKY9, WRKY15, WRKY18, WRKY25, WRKY26, WRKY39, WRKY40, WRKY47, WRKY54, WRKY69	116	none	none
E-GEOD-28600	EMBL	JA and JA+ABA treated (3 and 24 hr) T87 cell cultures	5279	21	WRKY1, WRKY6, WRKY7, WRKY16, WRKY21, WRKY35, WRKY36, WRKY38, WRKY40, WRKY43, WRKY45, WRKY47, WRKY52, WRKY53, WRKY54, WRKY67, WRKY69, WRKY70, WRKY71, WRKY72, WRKY75	533	3	WRKY7, WRKY38, WRKY70
E-MEXP-883	EMBL	JA treated (6 hr) WT and myc2 plants	3348	14	WRKY6, WRKY7, WRKY11, WRKY18, WRKY20, WRKY23, WRKY26, WRKY33, WRKY39, WRKY40, WRKY45, WRKY47, WRKY69, WRKY75	568	4	WRKY26, WRKY33, WRKY40, WRKY45
GSE21762	NCBI	JA treated WT and coi1 seedlings	3743	16	WRKY3, WRKY7, WRKY17, WRKY22, WRKY25, WRKY31, WRKY40, WRKY46, WRKY47, WRKY52, WRKY53, WRKY60, WRKY70, WRKY72, WRKY74, WRKY75	175	1	WRKY72
ME00337	TIAR	JA treated (0.5, 1, 3 hr) time course on WT seedlings	4796	15	WRKY3, WRKY7, WRKY18, WRKY20, WRKY21, WRKY23, WRKY38, WRKY40, WRKY45, WRKY47, WRKY48, WRKY53, WRKY60, WRKY69, WRKY75	950	6	WRKY7, WRKY18, WRKY20, WRKY40, WRKY45, WRKY48

Table 2.2. Arabidopsis WRKY TFs present on the Affymetrix microarray. The list of WRKY TFs present or absent from the Affymetrix arrays used in this study. The Arabidopsis Affymetrix array contains probes to identify the expression of 61 WRKY TFs. Eleven of the 72 WRKY TFs in Arabidopsis are not represented on the array.

Present on AFFY Arrays	Absent from AFFY Arrays
WRKY1, WRKY2, WRKY3, WRKY4, WRKY6, WRKY7, WRKY8, WRKY9, WRKY10, WRKY11, WRKY12, WRKY13, WRKY14, WRKY15, WRKY16, WRKY17, WRKY18, WRKY19, WRKY20, WRKY21, WRKY22, WRKY23, WRKY25, WRKY26, WRKY27, WRKY28, WRKY30, WRKY31, WRKY32, WRKY33, WRKY34, WRKY35, WRKY36, WRKY38, WRKY39, WRKY40, WRKY42, WRKY43, WRKY44, WRKY45, WRKY46, WRKY47, WRKY48, WRKY52, WRKY53, WRKY54, WRKY55, WRKY56, WRKY57, WRKY58, WRKY60, WRKY61, WRKY65, WRKY66, WRKY67, WRKY69, WRKY70, WRKY71, WRKY72, WRKY74, WRKY75	WRKY24, WRKY29, WRKY41, WRKY49, WRKY50, WRKY51, WRKY59, WRKY62, WRKY63, WRKY64, WRKY68

Table 2.3. JA responsive Arabidopsis WRKY TFs. The WRKY TFs identified as having significantly altered gene expression in at least one jasmonate treated dataset. WRKYs cited as identified in this study are those which were had significantly altered gene expression and survived the B-H FDR in at least one dataset. N/A indicates a probe to identify that WRKY is not available on the Affymetrix array but has been reported to be involved in jasmonate response. References are shown for WRKYs with reported function in jasmonate response.

JA Responsive WRKY	No. Datasets Expression	Reference
WRKY1	1	
WRKY3	2	
WRKY6	3	[1]
WRKY7	4	This study
WRKY8	0	[2]
WRKY9	1	
WRKY11	1	[3]
WRKY15	1	
WRKY16	1	
WRKY17	1	[3]
WRKY18	3	This study, [4]
WRKY20	2	This study
WRKY21	2	
WRKY22	1	
WRKY23	2	
WRKY25	2	[5]
WRKY26	2	This study
WRKY28	0	[6]
WRKY31	1	
WRKY33	1	This study, [7,8]
WRKY35	1	
WRKY36	1	
WRKY38	2	This study, [9]
WRKY39	2	
WRKY40	5	This study, [4]
WRKY43	1	
WRKY45	3	This study
WRKY46	1	
WRKY47	5	
WRKY48	1	This study
WRKY50	N/A	[10]
WRKY51	N/A	[10]
WRKY52	2	
WRKY53	3	[11]
WRKY54	2	
WRKY60	2	
WRKY62	N/A	[12]

Table 2.3. cont.

JA Responsive WRKY	No. Datasets Expression	Reference
WRKY67	1	
WRKY69	4	
WRKY70	2	This study, [13,14]
WRKY71	1	
WRKY72	2	This study
WRKY74	1	
WRKY75	4	

***** References**:**

1. Robatzek S, Somssich IE: *The Plant Journal* 2001, 28(2):123-133.
2. Chen L, Zhang L, Yu D: *Molecular Plant-Microbe Interactions* 2010, 23(5):558-565.
3. Journot-Catalino N, Somssich IE, Roby D, Kroj T: *Plant Cell* 2006, 18(11):3289-3302.
4. Pandey SP, Roccaro M, Schön M, Logemann E, Somssich IE: *The Plant Journal* 2010, 64(6):912-923.
5. Zheng Z, Mosher S, Fan B, Klessig D, Chen Z: *BMC Plant Biology* 2007, 7(1):2.
6. Wu L-t, Zhong G-m, Wang J-m, Li X-f, Song X, Yang Y: *African Journal of Microbiology Research* 2011, 5(30):5481-5488.
7. Birkenbihl RP, Diezel C, Somssich IE: *Plant Physiology* 2012, 159(1):266-285.
8. Zheng Z, Qamar SA, Chen Z, Mengiste T: *The Plant Journal* 2006, 48(4):592-605.
9. Wang Z, Cao G, Wang X, Miao J, Liu X, Chen Z, Qu L-J, Gu H: *Plant Cell Reports* 2008, 27(1):125-135.
10. Gao Q-M, Venugopal S, Navarre D, Kachroo A: *Plant Physiology* 2011, 155(1):464-476.
11. Miao Y, Zentgraf U: *The Plant Cell Online* 2007, 19(3):819-830.
12. Mao P, Duan M, Wei C, Li Y: *Plant and Cell Physiology* 2007, 48(6):833-842.
13. Ren C-M, Zhu Q, Gao B-D, Ke S-Y, Yu W-C, Xie D-X, Peng W: *Journal of Integrative Plant Biology* 2008, 50(5):630-637.
14. Li J, Brader G, Palva ET: The WRKY70 Transcription Factor: *Plant Cell* 2004, 16(2):319-331.

Table 2.4. Identification of JA responsive Arabidopsis WRKY TFs using microarray data. The fold change of jasmonate responsive Arabidopsis WRKY TFs from five microarray datasets. Only those jasmonate responsive AtWRKYs which has a significant p-value ≤0.05 and survived application of the B-H FDR are included.

A.

Dataset	WRKY	Fold Change after 0.5 h JA treatment	Fold Change after 2 h JA treatment	Fold Change after 6 h JA treatment
E-ATMX-13	none	---	---	---

B.

Dataset	WRKY	Fold Change after 3 h JA treatment	Fold Change after 24 h JA treatment
E-GEOD-28600	WRKY7	0.92	0.60
E-GEOD-28600	WRKY38	0.69	0.40
E-GEOD-28600	WRKY70	0.91	1.25

C.

Dataset	WRKY	Fold Change in COL-0 after JA treatment	Fold Change in *myc2* after JA treatment
E-MEXP-883	WRKY26	1.47	1.77
E-MEXP-883	WRKY33	1.78	2.49
E-MEXP-883	WRKY40	2.21	2.07
E-MEXP-883	WRKY45	2.49	3.76

D.

Dataset	WRKY	Fold Change in COL-0 after JA treatment	Fold Change in *coi1* after JA treatment
GSE21762	WRKY72	0.80	0.73

E.

Dataset	WRKY	Fold Change after 0.5 h JA treatment	Fold Change after 1 h JA treatment	Fold Change after 3 h JA treatment
ME00337	WRKY7	0.80	0.51	0.62
ME00337	WRKY18	4.99	6.39	2.65
ME00337	WRKY20	0.88	0.84	0.92
ME00337	WRKY40	6.25	9.06	5.37
ME00337	WRKY45	1.09	2.52	3.81
ME00337	WRKY48	0.90	0.88	0.62

Table 2.5. Arabidopsis WRKYs with altered expression in JA mutants. Arabidopsis WRKY TFs were analyzed for differential expression by genotype in A) *coi1* and B) *myc2* mutants or C) by time. Analysis was performed using a two-way ANOVA. WRKYs before and after the application of the B-H FDR are presented.

A.

Time-Dependent			
Genes after Two-Way ANOVA (P=0.05)		Genes after Two-Way ANOVA (P=0.05) and B-H FDR	
Time-Dependent	**Time + JA Interaction**	**Time-Dependent**	**Time + JA Interaction**
WRKY1, WRKY6, WRKY7, WRKY10, WRKY13, WRKY16, WRKY17, WRKY18, WRKY20, WRKY21, WRKY22, WRKY25, WRKY27, WRKY32, WRKY36, WRKY38, WRKY39, WRKY43, WRKY45, WRKY47, WRKY48, WRKY54, WRKY58, WRKY66, WRKY70, WRKY75	WRKY7, WRKY18, WRKY25, WRKY38, WRKY70	WRKY3, WRKY7, WRKY13, WRKY16, WRKY18, WRKY21, WRKY22, WRKY25, WRKY28, WRKY38, WRKY39, WRKY40, WRKY43, WRKY44, WRKY45, WRKY47, WRKY48, WRKY52, WRKY54, WRKY70	WRKY38, WRKY70

B.

COI1-Dependent			
Genes after Two-Way ANOVA (P=0.05)		Genes after Two-Way ANOVA (P=0.05) and B-H FDR	
Genotype	**Genotype + JA Interaction**	**Genotype**	**Genotype + JA Interaction**
WRKY6, WRKY8, WRKY11, WRKY15, WRKY17, WRKY18, WRKY22, WRKY23, WRKY25, WRKY33, WRKY38, WRKY46, WRKY48, WRKY52, WRKY53, WRKY69, WRKY70, WRKY72, WRKY74, WRKY75	WRKY8, WRKY17, WRKY18, WRKY40, WRKY46, WRKY53, WRKY70, WRKY72	WRKY72	none

Table 2.5. cont.

C.

MYC2-Dependent			
Genes after Two-Way ANOVA (P=0.05)		Genes after Two-Way ANOVA (P=0.05) and B-H FDR	
Genotype	**Genotype + JA Interaction**	**Genotype**	**Genotype + JA Interaction**
WRKY16, WRKY26, WRKY33, WRKY36, WRKY45, WRKY53	WRKY28	none	none

Table 2.6. Bioinformatics methods identifying each WRKY gene. WRKY domain containing proteins were identified using 4 sources: manual searching (i.e. searching proteins for a WRKY motif), PlantTFcat, NCBI CDD, and MPGR. Rows indicate overlap in genes identified from the different sources. The bottom row provides a total number of genes identified by each method.

Gene	Manual WRKY Search	PlantTFcat	NCBI CDD	MPGR
Cra549	X	X	X	
Cra1311	X	X	X	X
Cra1702	X	X	X	X
Cra2068	X	X	X	X
Cra2271	X	X	X	
Cra2950	X	X	X	X
Cra3503	X	X	X	X
Cra3760	X	X	X	X
Cra3799	X	X	X	X
Cra4234	X	X	X	X
Cra5093	X	X	X	X
Cra5497	X	X	X	X
Cra6088	X	X	X	X
Cra6519	X	X	X	X
Cra7867	X	X	X	X
Cra8145	X	X	X	X
Cra8670	X	X	X	X
Cra9152	X	X	X	X
Cra9369	X	X	X	X
Cra10348	X	X	X	X
Cra11684	X	X	X	
Cra13263	X	X	X	
Cra13321	X	X	X	X
Cra16284	X	X	X	
Cra16307	X	X	X	
Cra17347	X	X	X	
Cra18915	X	X	X	X
Cra18989	X	X	X	X
Cra19330	X	X	X	X
Cra19395	X	X	X	X
Cra19580	X	X	X	X
Cra20290	X	X	X	X
Cra21821	X	X	X	X
Cra22395	X	X	X	
Cra22725	X	X	X	X
Cra23742	X	X	X	
Cra24943	X	X	X	X
Cra28262	X	X	X	X
Cra30069	X	X	X	X
Cra37309	X	X	X	

Table 2.6. cont.

Gene	Manual WRKY Search	PlantTFcat	NCBI CDD	MPGR
Cra43671	X	X	X	
Cra43896	X	X	X	X
Cra56567	X	X	X	X
Cra65443	X	X	X	X
Cra70197	X	X	X	X
Cra105225	X	X	X	X
Cra22691		X	X	X
Cra24719		X	X	X
Cra54213		X	X	X
Cra55720		X	X	X
Cra102390		X	X	X
Cra11128			X	
Cra5637				X
Cra10341				X
Cra15757		X		
Cra16285				X
Cra53604				X
Cra72531				X
Cra76953				X
Cra82407				X
Total	46	52	52	47

Table 2.7. A list of Catharanthus WRKY domain containing proteins. A list of Catharanthus WRKY domain containing proteins along with locus number and group number are presented.

Catharanthus WRKY	Locus	Group
CrWRKY1	Cra16284	III
CrWRKY2	Cra549	I
CrWRKY3	Cra4234	I
CrWRKY4	Cra5497	I
CrWRKY5	Cra6088	I
CrWRKY6	Cra8145	I
CrWRKY7	Cra9152	I
CrWRKY8	Cra10348	I
CrWRKY9	Cra11128	I
CrWRKY10	Cra13321	I
CrWRKY11	Cra22691	I
CrWRKY12	Cra43671	I
CrWRKY13	Cra1311	IIa
CrWRKY14	Cra13263	IIa
CrWRKY15	Cra54213	IIa
CrWRKY16	Cra2068	IIb
CrWRKY17	Cra3503	IIb
CrWRKY18	Cra18915	IIb
CrWRKY19	Cra19580	IIb
CrWRKY20	Cra22725	IIb
CrWRKY21	Cra2271	IIc
CrWRKY22	Cra2950	IIc
CrWRKY23	Cra6519	IIc
CrWRKY24	Cra8670	IIc
CrWRKY25	Cra9369	IIc
CrWRKY26	Cra19330	IIc
CrWRKY27	Cra22395	IIc
CrWRKY28	Cra24943	IIc
CrWRKY29	Cra28262	IIc
CrWRKY30	Cra37309	IIc
CrWRKY31	Cra43896	IIc
CrWRKY32	Cra102390	IIc
CrWRKY33	Cra105225	IIc
CrWRKY34	Cra1702	IId
CrWRKY35	Cra3760	IId
CrWRKY36	Cra7867	IId
CrWRKY37	Cra17347	IId
CrWRKY38	Cra11684	IIe
CrWRKY39	Cra16307	IIe
CrWRKY40	Cra19395	IIe
CrWRKY41	Cra20290	IIe

Table 2.7. cont.

CrWRKY42	Cra21821	IIe
CrWRKY43	Cra23742	IIe
CrWRKY44	Cra30069	IIe
CrWRKY45	Cra3799	III
CrWRKY46	Cra5093	III
CrWRKY47	Cra18989	III
CrWRKY48	Cra24719	III
CrWRKY49	Cra55720	
CrWRKY50	Cra56567	
CrWRKY51	Cra65443	
CrWRKY52	Cra70197	

Table 2.8. The distribution of WRKY TFs from nine plant species. WRKY TFs identified from the sequenced genomes of *A. trichopoda*, *A. thaliana*, *C. annuum*, *O. sativa*, *S. lycopersicum*, *S. tuberosum*, and *U. gibba*. *C. roseus* and *R. serpentina* WRKY TFs were identified from transcriptome sequences in the MPGR database. Complete and partial WRKY domain containing proteins were identified using the NCBI Conserved Domain Database. The presence of WRKY domains were manually verified and phylogenetic analyses were conducted to determine WRKY subgroups for each of the species.

Species	Complete WRKY TFs	Partial WRKY TFs	Group I	Group IIa	Group IIb	Group IIc	Group IId	Group IIe	Group III	Unassigned
Amborella trichopoda	29	3	7	2	4	5	2	4	5	3
Arabidopsis thaliana	72	2	14	3	8	18	7	8	14	2
Capsicum annuum	66	4	16	4	6	13	11	7	9	4
Catharanthus roseus	48	4	11	3	5	13	4	7	5	4
Oryza sativa ssp. japonica	93	6	15	4	8	16	7	10	34	5
Rauvolfia serpentina	49	5	10	2	4	12	5	5	5	11
Solanum lycopersicum	78	3	15	5	8	17	6	17	11	2
Solanum tuberosum	75	9	14	5	6	14	7	15	14	9
Urticularia gibba	65	7	16	4	4	18	7	11	5	7

Table 2.9. WRKY domains of Catharanthus WRKY TFs. Catharanthus contained at least 48 WRKY TFs with 56 WRKY domains. The 70 amino acid sequence of each WRKY domain is provided.

WRKY Name	WRKY Domain
CrWRKY1	ETETKYSSTMEDEYAWRKYGQKDILRSNFPRCYFRCTHKNEGCKAT KQVQIVTKNPLMYQTTYFGQHTCN
CrWRKY10-C	VHAAGDVGISGDGYRWRKYGQKMVKGNPHPRNYYRCTSAGCTVR KHIEMAKDNSNGVIITYKGRHDHDMP
CrWRKY10-N	FSVPAQKTPYPDGYNWRKYGQKQVKSPQGSRSYYRCTYSKCSAKKI ECSDNSNRVIEIVYRSCHNHDPPE
CrWRKY11	LQIESEIDVLDDGYRWRKYGQKVVKGNPNPRSYYKCTSAGCPVRKH VERASEDIKSVITTYEGKHNHEVP
CrWRKY12-C	VQTMSEVDVINDCYRWRKYGQKLVKGNPNPRSYYRCSNSGCPVKK HVERSSHDPKIVITTYEGKHDHEMP
CrWRKY12-N	VSPRIQEKALDDGYNWRKYGQKLVKGNVFVRSYYKCTYASCTSKKQ VERSYDGRLTDIKYIGKHEHPKPQ
CrWRKY13	TEASDTSLIVKDGYQWRKYGQKVTRDNPSPRAYFKCSFAPSCPVKKK VQRSIEDQSILVATYEGEHNHPH
CrWRKY14	TDPDDKSLVVKDGYHWRKYGQKVTKDNPSPRAYFKCSFAPTCQVK KKVQRSVGNAAILVATYEGEHNHQP
CrWRKY15	TEASDTSLIVKDGYQWRKYGQKVTRDNPSPRAYFKCSFAPSCPVKKK VQRSIEDQSIVVATYEGEHNHSK
CrWRKY16	VRARSEAPMISDGCQWRKYGQKMAKGNPCPRAYYRCTMGVGCP VRKQVQRCAEDRSILITTYEGHHNHPL
CrWRKY17	VRARCETATMNDGCQWRKYGQKIAKGNPCPRAYYRCTVAPTCPV RKQVQRCAEDTSILITTYEGTHNHSL
CrWRKY18	VRVRCDTPTMNDGCQWRKYGQKIAKGNPCPRAYYRCTVAPNCPV RKQVQRCAEDMSILITTYEGTHNHTL
CrWRKY19	VRARCETATMNDGCQWRKYGQKIAKGNPCPRAYYRCTVAPGCPV RKQVQRCLEDMSILITTYEGTHNHPL
CrWRKY20	VRARSEAPMITDGCQWRKYGQKMAKGNPCPRAYYRCTMAAGCP VRKQVQRCADDRTILITTYEGNHNHPL
CrWRKY21	FQTRSDVDVLDDGYKWRKYGQKVVKNSLHPRSYYRCTHNNCRVKK RVERLSEDCRMVITTYEGRHNHTPC
CrWRKY22	FTTKSEIDHLEDGYRWRKYGQKAVKNSPFPRSYYRCTSQKCSVKKRV ERSFQDPSIVITTYEGQHNHHCP
CrWRKY23	FRTKSQVEILDDGYKWRKYGKKMVKNSPNPRITTDAQLKDAPVKKR VERDKEDPKYVITAYEGIHNHQGP
CrWRKY24	FKTLSDVDVLDDGYKWRKYGQKVVKNTQHPRSYYRCTQDNCRVKK RVERLAEDPRMVITTYEGRHIHSPS
CrWRKY25	RIKSCDSAMTDDGYKWRKYGQKSIKNSPNPRSYYRCTNPRCAAKKQ VERSSDDPDTLIITYEGLHLHFAY
CrWRKY26	FMTKSDVDHLEDGYRWRKYGQKAVKNSPFPRSYYRCTSASCNVKK RVERCLNDPSLVITTYEGQHNHQTP
CrWRKY27	FMTKSEVDHLEDGYRWRKYGQKAVKNSPFPRSYYRCTNTKCTVKK RVERSSEDPTIVITTYEGQHCHHTV

Table 2.9. cont.

CrWRKY28	FQTRSQVDILDDGYRWRKYGQKAVKNNKFPRSYYRCTYQGCNVKK QVQRLSKDEGIVVTTYEGMHSHPIE
CrWRKY29	FMTKSEVDHLEDGYRWRKYGQKAVKNSPYPRSYYRCTTQKCPVKK RVERSFQDPSIVITTYEGTHNHHVP
CrWRKY2-C	VQTTSEVDLLDDGYRWRKYGQKVVKGNPYPRSYYKCTSPGCNVRK HVERAATDPKAVITTYEGKHNHDVP
CrWRKY2-N	QASILVDKPADDGYNWRKYGQKQVKGSEYPRSYYKCTHQNCPVKK KVERSQDGQVTEIIYKGQHNHPPPQ
CrWRKY30	FHTRSTEDILDDGYKWRKYGQKSVKNSSHPRSYYRCTHHTCNVKKQ IQRLSKDTSVVVTTYEGIHSHPCE
CrWRKY31	FQTRSADDVLDDGYRWRKYGQKSVKNSKYPRSYYRCSQHTCNVKK QVQRLSKDTGIVVTTYEGIHNHPCE
CrWRKY32	FKTKSDVEILDDGFKWRKYGRKMVKNSINPRNYYKCSVEGCPVKKR VERDNNDSRYVVTTYEGIHNHQGP
CrWRKY33	FMTKSEIDQLDDGFRWRKYGQKAVKNSPFPRSYYRCTTAGCGVKK RVERSSEDATIVITTYEGMHNHCSP
CrWRKY34	AISMKMADIPPDDYSWRKYGQKPIKGSPHPRGYYKCSSVRGCPARK HVERALDDPSMLIVTYEGEHNHSL
CrWRKY35	AISMKMADIPPDDYSWRKYGQKPIKGSPHPRGYYKCSSVRGCPARK HVERALDDPTMLIVTYEGEHNHSH
CrWRKY36	AISNKLADIPPDEYSWRKYGQKPIKGSPHPRGYYKCSSMRGCPARK HVERCLEDPSMLIVTYEGEHNHPR
CrWRKY37	AISSKIADIPADEYSWRKYGQKPIKGSPYPRGYYKCSTVRGCPARKHV ERATDDPKMLIVTYEGEHRHVQ
CrWRKY38	SRNRTEVYPPPDSWSWRKYGQKPIKGSPYPRGYYRCSSSKGCPARK QVERSRLDPTKLLITYSSEHNHSL
CrWRKY39	RLKGEMGAPPSDSWAWRKYGQKPIKGSPYPRGYYRCSSSKGCPAR KQVERSRIDPTMLMVTYTCEHNHPW
CrWRKY3-C	VQTVSEVDILDDGYRWRKYGQKVVRGNPNPRSYYKCTNAGCPVRK HVERASHDPKAVITTYEGKHNHNVP
CrWRKY3-N	TSSITSDRSSDDGYNWRKYGQKLVKGSEFPRSYYKCTYPNCEVKKIFE RSPDGQITEIVYKGSHDHPKPQ
CrWRKY40	DKKQKKEGPPLDCWSWRKYGQKPIKGSPYPRGYYRCSTSKGCSAKK QVERCRTDPTVLIVTYTSTHNHAT
CrWRKY41	VCQVPAEALSSDTWSWRKYGQKPIKGSPYPRGYYRCSTSKGCLARK QVERNRSDPGMFIVTYTAEHNHPM
CrWRKY42	SRPSSGEVVPSDLWAWRKYGQKPIKGSPYPRGYYRCSSSKGCSARK QVERSRNDPNMLVITYTSEHNHPW
CrWRKY43	VHQMTQEELSGDSWAWRKYGQKPIKGSPYPRNYYRCSTSKGCSAR KQVERCPTDPNIFVVSYSGEHTHPR
CrWRKY44	VIQVTAEDLSSDKWAWRKYGQKPIKGSPYPRSYYRCSSSKGCLARK QVEQSCKDPSIFIVTYTAEHSHSQ
CrWRKY45	CSGIGQEGPVDDGYNWRKYGQKDILGAIFPRSYYRCTHRYTQGCLA TKQVQKSEEDSSIFEVTYKGRHSC

Table 2.9. cont.

CrWRKY46	SSDNGLEGPSDDGYSWRKYGQKHILGAKYPRSYYRCTYRHIQNCW VTKQVQRSDEDPTIFEITYRGAHTC
CrWRKY47	SPGTGLEGPLEDGYSWRKYGQKDILGAKYPRGYYRCTHRPVQGCLA TKQVQRSDDDPTIFQITYRGRHTC
CrWRKY48	TWTQNSSTLIDDGYAWRKYGQKVILNADYPRNYFRCTHKFDQECQ ATKQVQMIQENPPLYRTTYHGHHTC
CrWRKY4-C	VQTRSEVDLLDDGYKWRKYGQKVVKGNPHPRSYYRCTYAGCNVRK HVERASTDAKAVVTTYEGKHNHDIP
CrWRKY4-N	VAAVALDKPADDGYNWRKYGQKLVKAKEHPRSYYKCTHLNCPVKK KVERATDGHVAEITYKGQHNHEMPQ
CrWRKY5-C	VQTTSDIDILDDGYRWRKYGQKVVKGNPNPRSYYKCTYAGCPVRKH VERASHDLRAVITTYEGKHNHDVP
CrWRKY5-N	SQYLREQRKSEDGYNWRKYGQKQVKGSENPRSYYKCTFPSCPTKKK VERNLEGHITEIVYKGNHNHAKPQ
CrWRKY6-C	VQNTVDSEIIRDGFRWRKYGQKVVKGNPYPRSYYRCTSLKCNVRKY VERTSEDPTAFITTYEGKHNHEMP
CrWRKY6-N	SHSTLGDRPSYDGYNWRKYGQKQVKGSEYPRSYYKCTHPNCPVKK KVERSLDGQIAEIVYKGEHNHPKPQ
CrWRKY7-C	VQTTSEVDILDDGYRWRKYGQKVVKGNPNPRSYYKCTSAGCTVRK HVERASHDLKSVITTYEGKHNHDVP
CrWRKY7-N	GDPNIGGAPAEDGYNWRKYGQKQVKGSEYPRSYYKCTHQNCQVK KKVERSQEGHITEIIYKGAHNHPKPP
CrWRKY8-C	VQTTSDIDILDDGYRWRKYGQKVVKGNPNPRSYYKCTSPGCPVRKH VERASHDLRSVITTYEGKHNHDVP
CrWRKY8-N	QQTMSERRRAEDGYNWRKYGQKNVKGSENPRSYYKCTFPSCPTKK KVERSVDGQITEIVYKGNHNHAKPQ
CrWRKY9-C	VQTMSEVDVINDCYRWRKYGQKLVKGNPNPRSYYRCSNSGCPVKK HVERSSHDPKIVITTYEGKHDHEMP
CrWRKY9-N	VSPRIQEKALDDGYNWRKYGQKLVKGNVFVRSYYKCTYASCTSKKQ VERSYDGRLTDIKYIGKHEHPKPQ

Table 2.10. Orthologs and paralogs for Arabidopsis and Catharanthus WRKY. Orthologs were identified using OrthoMCL. WRKYs in 'red font' are jasmonate responsive, either according to the literature or by my findings in this study. WRKY highlighted in light gray are TFs known to regulate secondary metabolism in *A. thaliana*, *Artemisia annua*, *C. rosues*, *Coptis japonica*, *Gossypium arboreum*, *Hevea brasiliensis*, *Panax quinquefolius*, and *Taxus chinensis*.

Catharanthus roseus	***Arabidopsis thaliana***	**Other**
CrWRKY1, CrWRKY48, CrWRKY52	AtWRKY54, AtWRKY70	
CrWRKY2, CrWRKY3, CrWRKY4, CrWRKY5, CrWRKY8, CrWRKY51	AtWRKY20, AtWRKY33	CmWRKY1
CrWRKY6	AtWRKY44	
CrWRKY7	AtWRKY2	
CrWRKY9, CrWRKY12	AtWRKY1	
CrWRKY10	AtWRKY32	
CrWRKY13, CrWRKY14, CrWRKY15	AtWRKY40	GaWRKY1
CrWRKY16, CrWRKY20	AtWRKY6, AtWRKY31, AtWRKY42	TcWRKY1
CrWRKY17, CrWRKY18	AtWRKY72	
CrWRKY19	AtWRKY9	
CrWRKY21, CrWRKY22, CrWRKY26, CrWRKY27, CrWRKY29, CrWRKY33	AtWRKY8, AtWRKY12, AtWRKY23, AtWRKY28, AtWRKY48, AtWRKY51, AtWRKY57, AtWRKY71	HbWRKY1
CrWRKY23, CrWRKY32	AtWRKY50	
CrWRKY24	AtWRKY13	
CrWRKY25	AtWRKY49	
CrWRKY28	AtWRKY75	CjWRKY1
CrWRKY30, CrWRKY31	AtWRKY24, AtWRKY43, AtWRKY56	
CrWRKY34, CrWRKY35	AtWRKY7	PqWRKY1
CrWRKY36	AtWRKY21	
CrWRKY37	AtWRKY11, AtWRKY17	
CrWRKY38	AtWRKY69	
CrWRKY39	AtWRKY65	
CrWRKY41, CrWRKY43	AtWRKY22, AtWRKY27	
CrWRKY42	AtWRKY14, AtWRKY35	
CrWRKY45, CrWRKY46, CrWRKY47	AtWRKY41, AtWRKY46, AtWRKY53	AaWRKY1

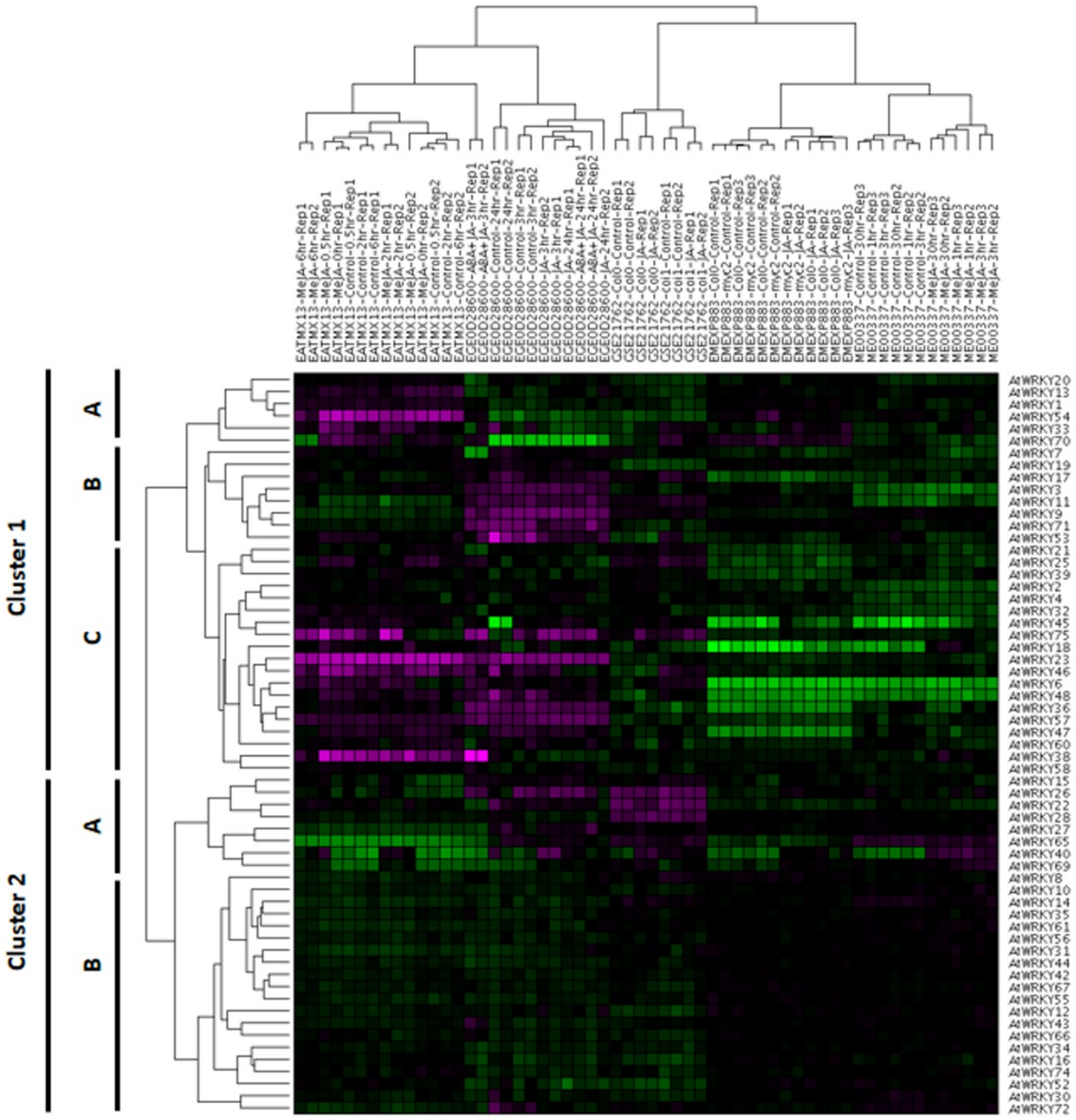


Figure 2.1. Clustering analysis of Arabidopsis WRKY TF. Hierarchical cluster analysis of the Arabidopsis WRKY TF family was performed using GenePattern. The clustering method was a pairwise average linkage with distance measured using the Pearson correlation coefficient. Data was log transformed. The median value was subtracted from each row. Color is based on global expression with purple being up-regulated and green down-regulated.

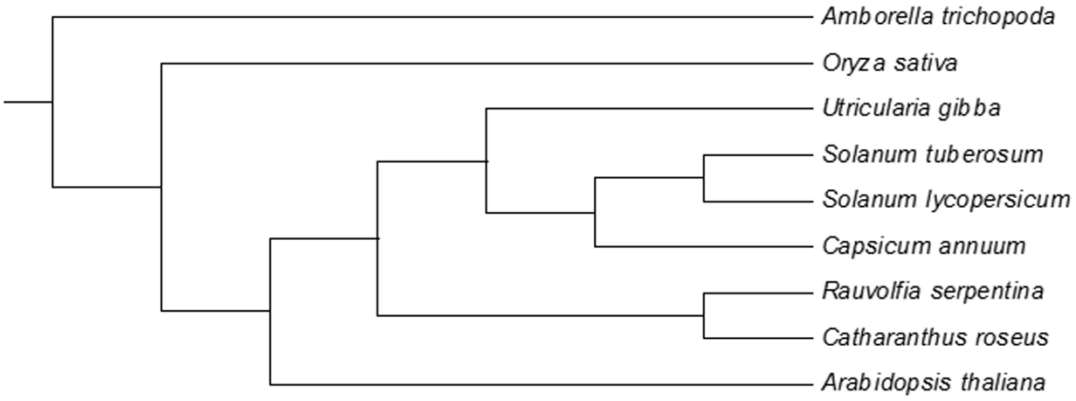


Figure 2.2. Evolutionary relationship of Asterids used in this study. A phylogenetic tree showing the taxonomic relationship of nine plant species. The species tree was computed using the NCBI Common Tree then visualized with MEGA5 software.

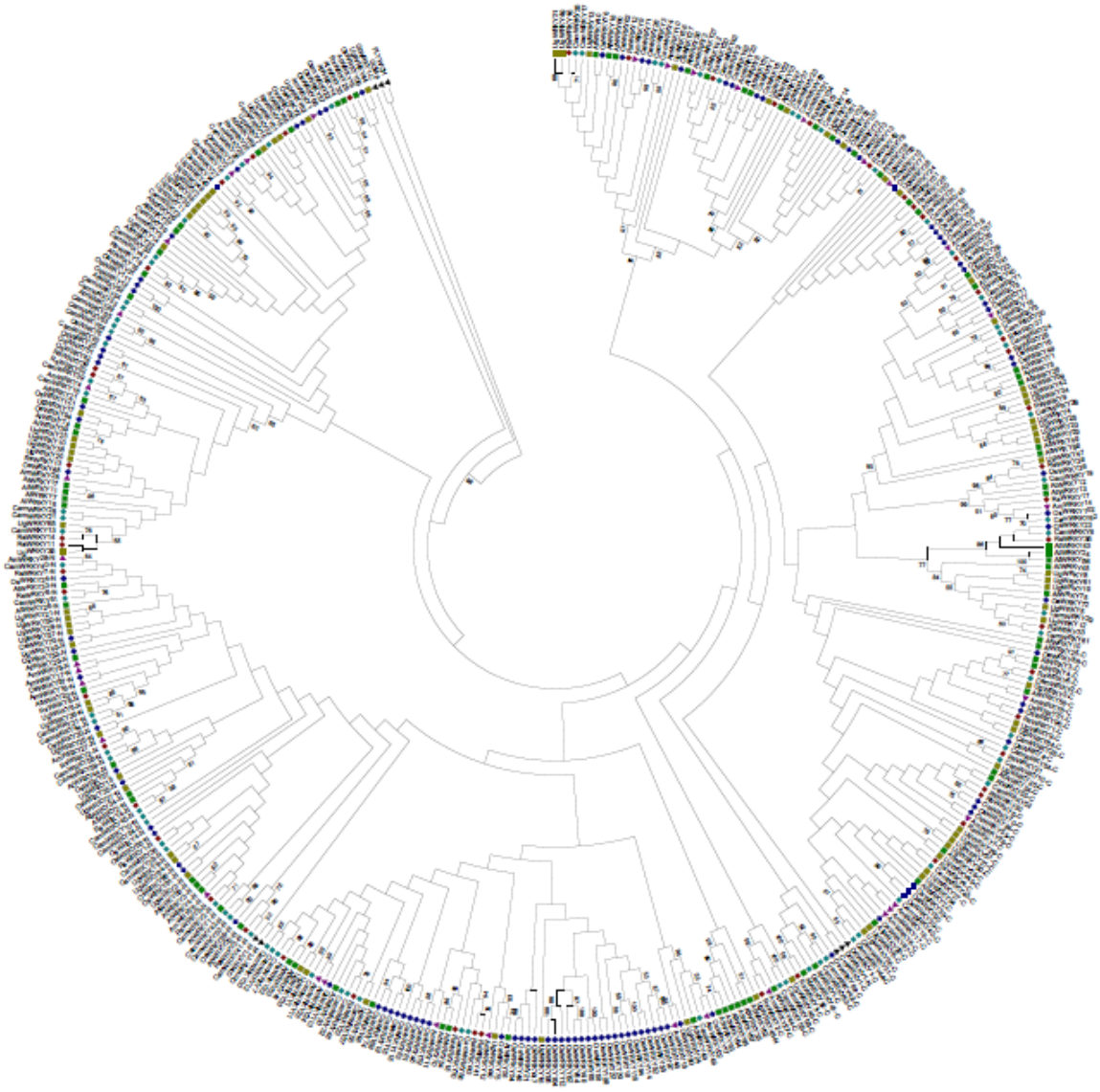


Figure 2.3. Phylogenetic analysis of bladderwort, pepper, and serpentwood WRKY families. The phylogenetic tree with *A. thaliana* (green square), *A. trichopoda* (purple triangle), *C. annuum* (teal dot), *O. sativa* (blue diamond), *R. serpentina* (red dot), and *U. gibba* (gold square) was constructed in MEGA5 using the Neighbor-Joining method with P-distance substitution model and 2000 bootstraps. Proteins used as an outgroup are indicated by a black triangle. WRKY domain alignment was performed with ClustalW.

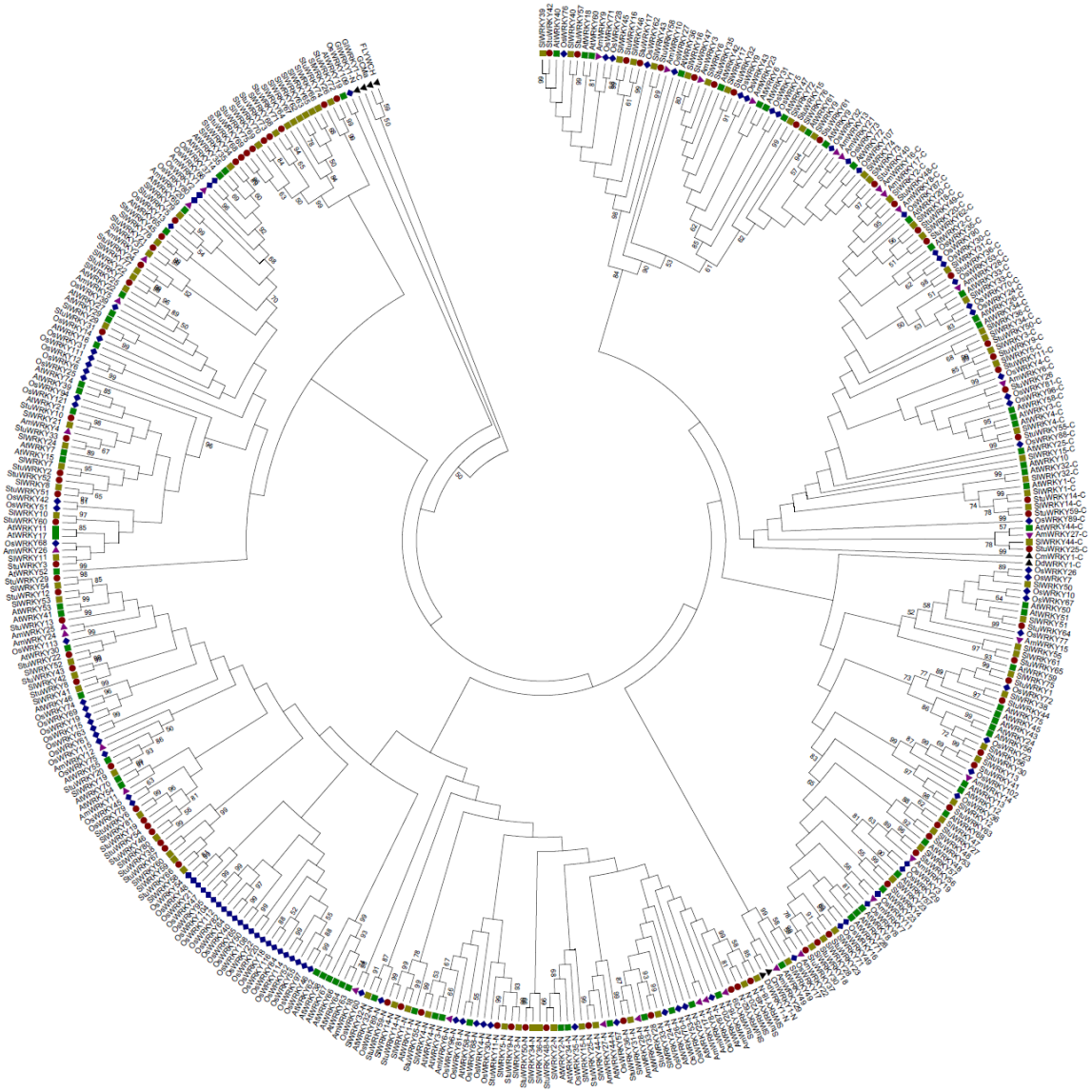


Figure 2.4. Phylogenetic analysis of potato and tomato WRKY families. The phylogenetic tree with *A. thaliana* (green square), *A. trichopoda* (purple triangle), *O. sativa* (blue diamond), *S. lycopersicum* (gold square), and S. tuberosum (red dot) was constructed in MEGA5 using the Neighbor-Joining method with P-distance substitution model and 2000 bootstraps. Proteins used as an outgroup are indicated by a black triangle. WRKY domain alignment was performed with ClustalW.

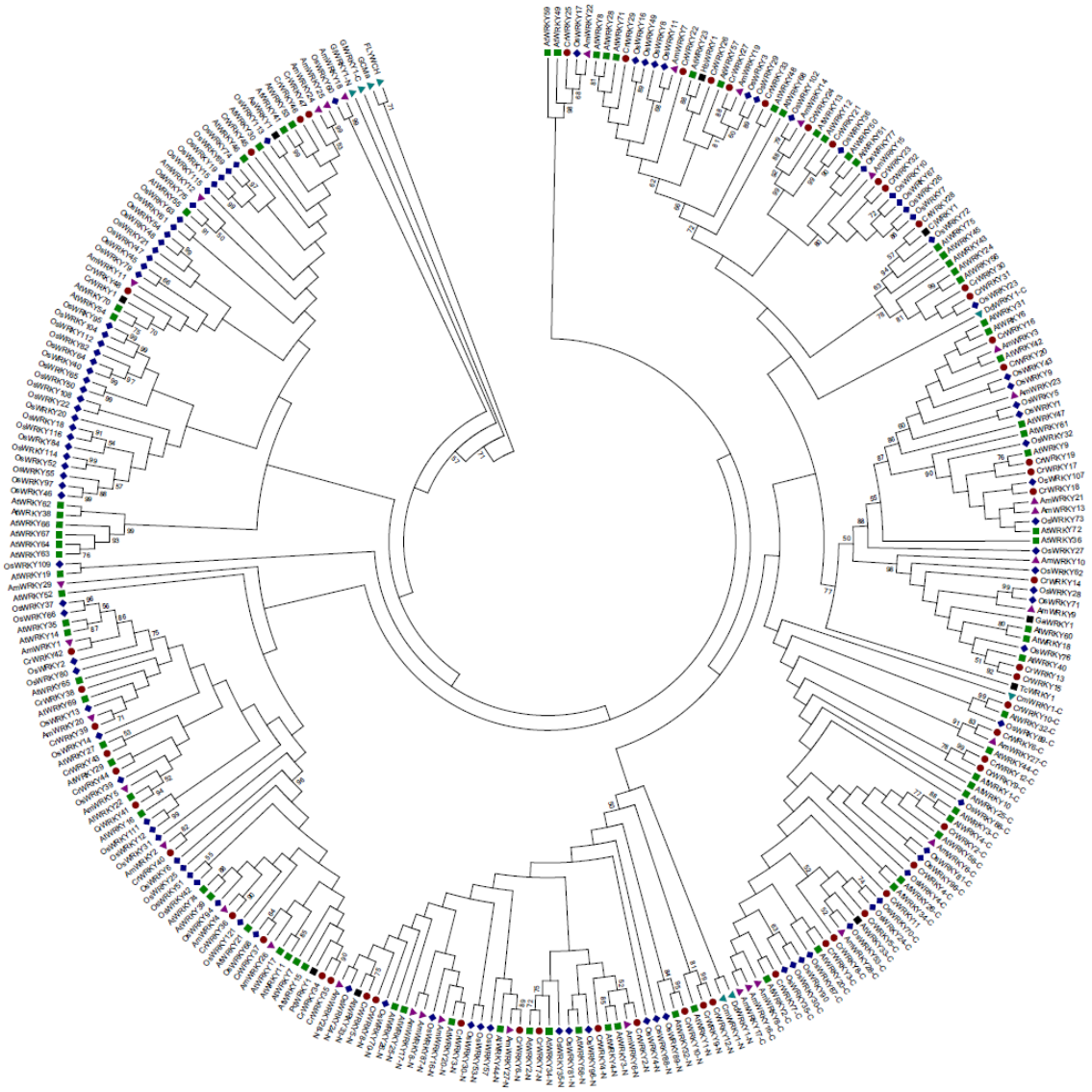


Figure 2.5. Phylogenetic tree of the Catharanthus WRKY family. The phylogenetic tree of *C. roseus* (red dot), *A. trichopoda* (inverted purple triangle), *A. thaliana* (green square), and *O. sativa* (blue diamond) was constructed using the Neighbor-Joining method with P-distance substitution model, pairwise-deletion, and a bootstrap value of 2000. WRKY domain alignment was performed with ClustalW. Proteins used as an outgroup are indicated by a teal triangle. WRKY TFs from six additional medicinal species, *A. annua* (AaWRKY1), *C. japonica* (CjWRKY1), *G. arboreum* (GaWRKY1), *H. brasiliensis* (HbWRKY1), *P. quinquefolius* (PqWRKY1), and *T. chinensis* (TcWRKY1), were included. WRKY TFs involved in regulating secondary metabolism are indicated with a black dot.

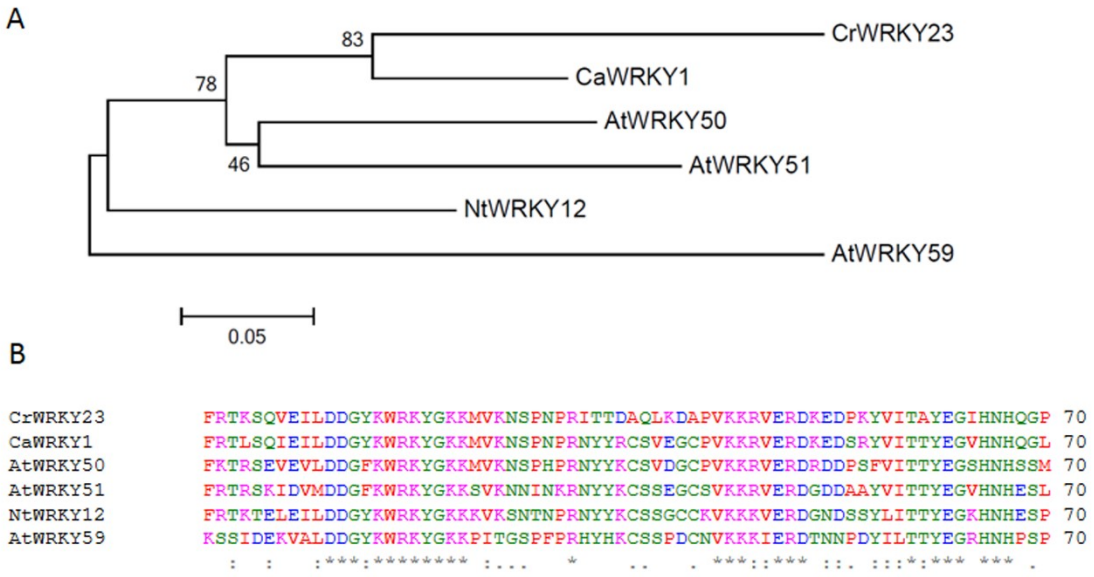


Figure 2.6. Phylogenetic relationship of WRKYGKK WRKYs. The phylogenetic relationship and alignment of CrWRKY23 to other WRKYGKK containing WRKY transcription factors. A. The phylogenetic tree was constructed in MEGA5 using the Neighbor-Joining method with P-distance substitution model, pairwise-deletion, and a bootstrap value of 2000. The tree is unrooted and branch lengths drawn to scale to evolutionary distances. Alignment was performed using ClustalW. B. Alignment of WRKY domain sequence was performed using ClustalW.

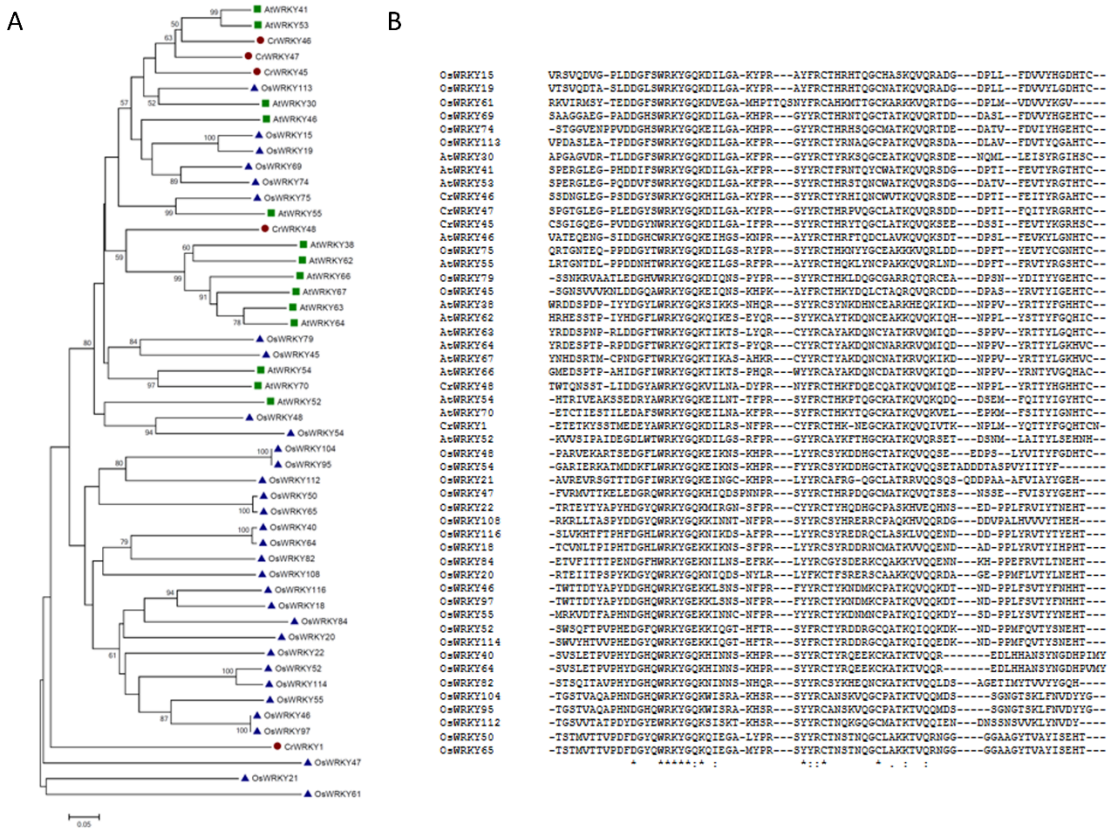


Figure 2.7. Analysis of the Catharanthus group III WRKY members. A) The phylogenetic relationship and B) alignment of CrWRKY1 to other group III WRKY TFs. The phylogenetic tree was constructed in MEGA5 using the Neighbor-Joining method with P-distance substitution model and a bootstrap value of 2000. WRKY domain alignment was performed with ClustalW. B. Alignment of the closest related rice WRKY genes and Catharanthus group III WRKYs was performed using ClustalW.

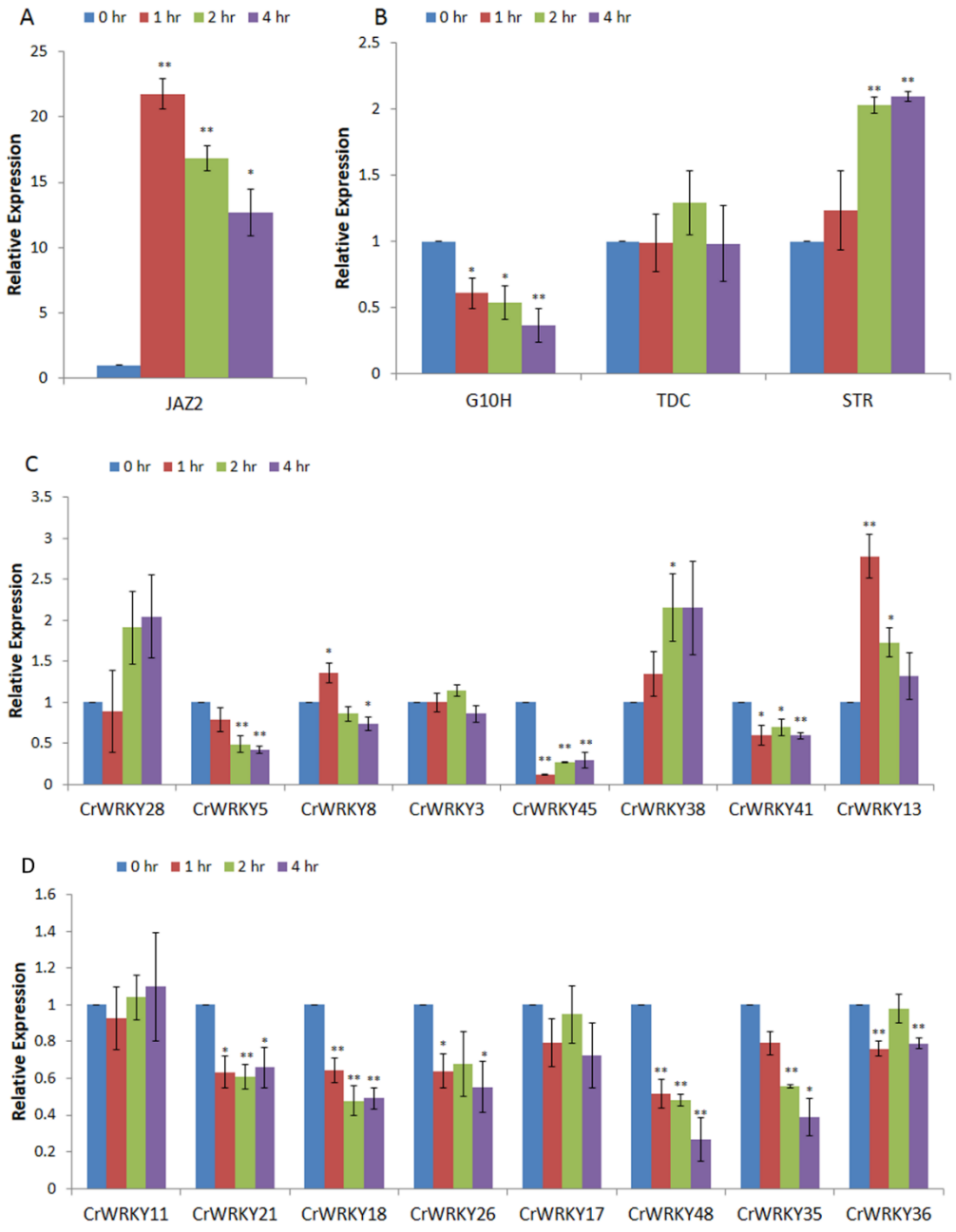


Figure 2.8. Gene expression of Catharanthus WRKY TFs in response to JA. Quantitative reverse-transcription PCR (qRT-PCR) for quantification of gene expression was performed on mature Catharanthus plants. Three whole plants were combined for each biological replicate. Each time point consisted of three biological replicates. Three technical replicates were measured per time point sample. A. Expression of Catharanthus *JAZ2* transcripts were determined in after 0, 1, 2, or 4 hours of JA treatment. B. Expression of the TIA biosynthetic genes *G10H*, *TDC*, and *STR*. C-D. Expression of 16 Catharanthus WRKY transcription factors in response to 0, 1, 2, and 4 hours of JA treatment. Significant and highly significant, p-value<0.05 or 0.01 respectively, changes in gene expression were determined using a Student's T-test.

A

	Seedling 5 day MeJA	Cell Culture 6 hr. MeJA	Hairy Root 24 hr. MeJA
Plant 1 hr. MeJA	0.103	0.790	0.061
Plant 2 hr. MeJA	0.615	0.413	0.555
Plant 4 hr. MeJA	0.690	0.307	0.660

B

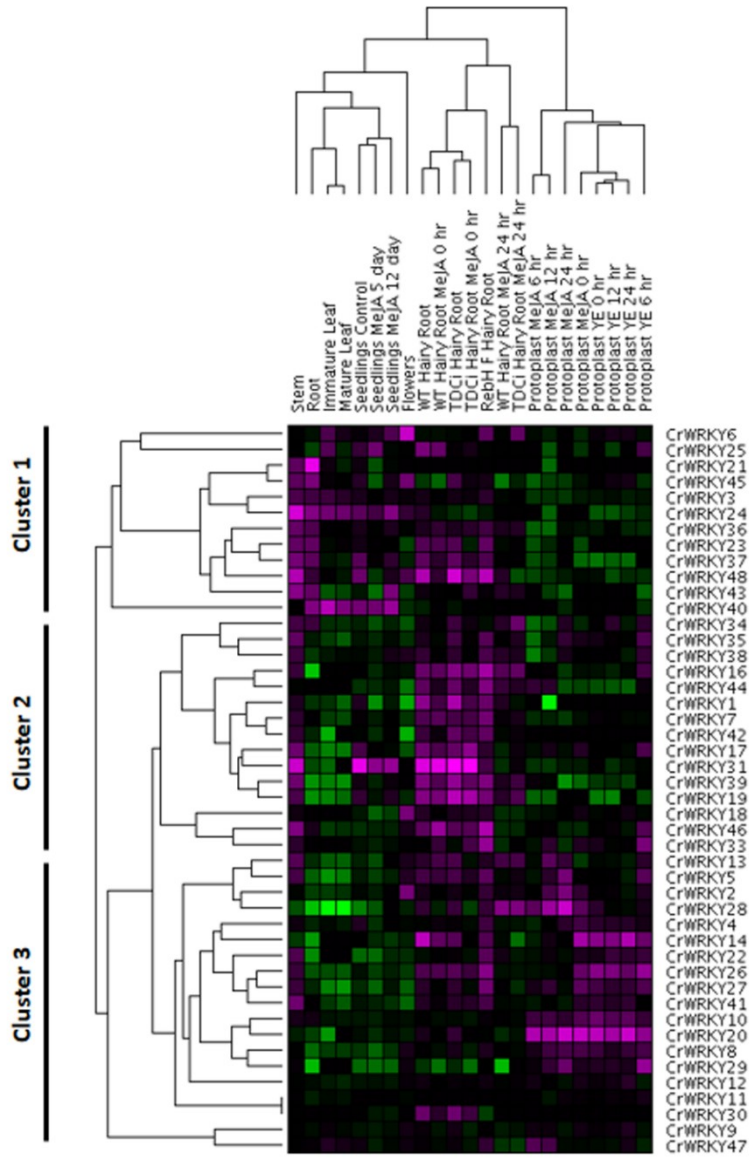


Figure 2.9. Clustering analysis of Catharanthus WRKYs. A. Pearson correlation between fold change of qRT-PCR expression and MPGR datasets for JA treated samples. Fold change for both sets was calculated using the reference gene EF1α as an internal control. B. Hierarchical cluster analysis of MPGR transcriptome data for the CrWRKY TF family was performed using GenePattern. The clustering method was a pairwise average linkage with distance measured using the Pearson correlation coefficient. Data was log transformed. The median value was subtracted from each row. Color is based on global expression with purple being up-regulated and green down-regulated.

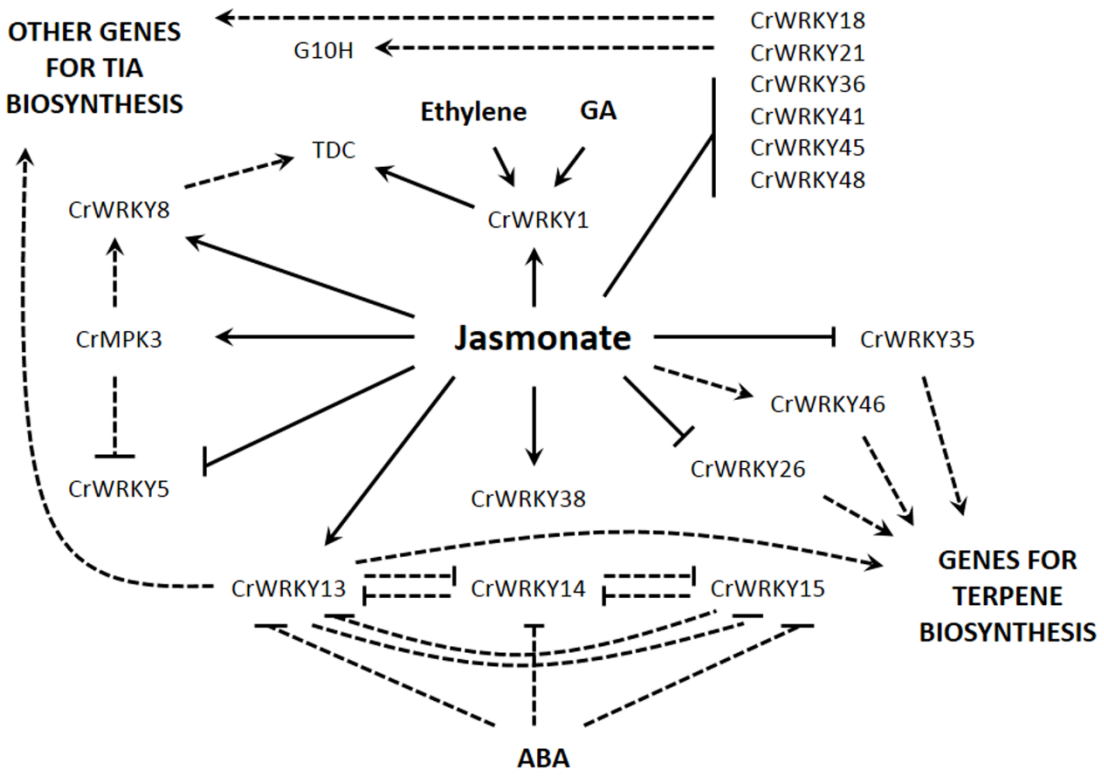


Figure 2.10. Model for Catharanthus WRKY TFs regulation of TIAs. A model of WRKY TFs function in Catharanthus based on expression data and known roles in *A. thaliana*. The model depicts CrWRKYs which were either similar to a WRKY with a known role in regulating natural product formation in another species or had transcript level differentially expressed in response to JA treatment. Jasmonate has both positive and negative effects on CrWRKY transcript accumulation which is possibly important for fine-tuning TIA and terpene biosynthetic gene expression. The hormones abscisic acid (ABA), ethylene, and gibberellin (GA) also are likely important for regulation of Catharanthus WRKYs. Solid lines depict known regulations and dashed lines indicate hypothetical regulatory functions.

CHAPTER THREE: ORIGIN AND EVOLUTION OF JASMONATE SIGNALING

Abstract

Adaption from an aqueous to terrestrial environment was a milestone in plant evolution. Jasmonate (JA) signaling is a key mediator of plant development and defense which arose during this time. The charophyte, *Klebsormidium flaccidum* does not possess genes encoding key green-plant JA signaling components, including COI1, JAZ, NINJA, MYC2, and the JAZ-interacting bHLH factors, yet their orthologs are present in the moss, *Physcomitrella patens*. TIFY family genes were found in charophyta and chlorophya algae, the first identified outside the embryophyte linage. Early TIFYs possibly contributed to male reproductive development and later to food storage for the embryo. JAZs evolved from ZIM genes of the TIFY family through changes to several key amino acids. A molecular clock analysis suggested that the JA signaling cascade arose in the last common ancestor of embryophytes during the early Ediacaran to late Cambrian periods, between 628 and 491 million years ago. This Cambrian Explosion era is known for rapid animal diversification. I propose that increased predation from the explosion of fauna drove improvements in male reproduction. The subsequent evolution of ZIM into JAZ repressors, along with acquisitions of other JA signaling components, additional hormone networks, and defensive mechanisms, potentially allowed plants to explore and colonize a new terrestrial niche.

Introduction

Aquatic and terrestrial plants are periodically exposed to a variety of environmental stress. In the oceans, tidal fluctuations expose certain algal species to daily desiccation and predation. Algal species in inland freshwater bodies are also subject to similar stresses due to water fluctuations as a result of seasonal and annual variations in rainfall. Furthermore, meteorological phenomenon (e.g. wind) on freshwater bodies can drift algae ashore, exposing them to additional stress, such as drought, extreme sunlight, and new diseases. How Viridiplantae, also known as the green lineage or plants, evolved to tolerate stress is poorly understood. The kingdom Viridiplantae is comprised of highly diverse photosynthetic members belonging to the phyla Chlorophyta, Charophyta, and Embryophyta (land plants). Charophytes are mostly freshwater algae species that form six classes: Chlorokybophyceae, Charophyceae, Coleochaetophyceae, Klebsormidiophyceae, Mesostigmatophyceae, and Zygnematophyceae. Contrary to early findings (Karol et al., 2001; Finet et al., 2010), recent evidence suggests that the Zygnematophyceae are the direct ancestors of embryophytes (Wickett et al., 2014). Zygnematophyceae are unicellular, flagella-less, mostly freshwater algae that form short chain filamentous or small colonies and reproduce by conjugation (McCourt et al., 2004). During transition from water to land, early embryophyte ancestors developed adoptive features that include resistance to stress, such as desiccation, and support structures (Karol et al., 2001). Elucidating the development of stress tolerance is, therefore, a key component of understanding land plant evolution.

The objective of this study was to initiate the process to clarify the evolutionary path of JA signaling. I sought to provide initial evidence into the origin and early functions of the biologically important JA pathway, including how, when, where, and why it evolved. I determined that *K. flaccidum* has orthologs of MED25 and TPL, but not COI1, JAZ, MYC2, or NINJA. However, these key JA pathway constituents are present in moss. Importantly, I identified the ZIM ZnF-GATA factors as the first TIFY family member which appeared outside the Embryophyta. Based on gene expression across plant genera I found early ZIM TFs probably regulated male reproductive structure development. Comparison of ZIM TFs to JAZ repressors revealed key sequence changes required for formation of the JAS domain. My results, combined with a molecular clock analysis suggest that core JA signaling occurred in the last common ancestor of embryophytes, which arose in the early Ediacaran to late Cambrian, between 628 and 491 million years ago. I propose that the JA signaling cascade initially evolved to fine-tune regulation of male reproductive development. Continued predation from herbivores during the Cambrian Explosion likely drafted JA signaling into regulating stress tolerance and may have driven aquatic plants to search for new predator free niches, including land.

Results

COI1 and TIR1 Originate from a Single F-Box Protein during Land Plant Colonization

Recently, Wang et al. (2015) proposed that two F-box proteins, TIR1 and COI1, involved in auxin and JA signaling, respectively, derived from the same gene during the charophyte to embryophyte transition. They suggested kfl00028_0560 and kfl00434_0030 as a pair of genes ancestral to COI1 and TIR1, respectively. BLAST results to the *K. flaccidum* genome returned kfl00028_0560 as the top hit for COI1 and TIR1 (E-value: 7e-12 and 6e-10, respectively). Further inspection with the NCBI conserved domain database (CDD) showed that kfl00028_0560 lacks an

F-box sequence. Alternatively, kfl00434_0030 contains a partial F-box sequence but was the second best BLAST hit for COI1 (E-value: 3e-08) and fourth for TIR1 (E-value: 8e-08; proceeded by kfl00290_0040 and kfl00508_0080, E-values: 8e-09 and 1e-08, respectively). Furthermore, a BLAST search in the NCBI database for the COI1 gene in the charophyte lineages Charophyceae, Chlorokybophyceae, Coleochaetophyceae, Klebsormidiophyceae, Mesostigmatophyceae, and Zygnematophyceae did not result in any hits with more than 5% sequence coverage and an E-value <0.1. Therefore, I suspected that kfl00028_0560 and kfl00434_0030 were mistakenly identified as ancestral sequences due to solely relying on the BLAST hits and the exclusion of additional *K. flaccidum* family members from the phylogenetic analyses.

COI1 contain both an F-box and AMN1 superfamily domains. To further investigate the origin of the AtCOI1 receptor I constructed phylogenetic trees using the AMN1 superfamily. The AMN1 superfamily was utilized as the F-box family is excessively large for multi-species comparisons. Additionally, the previously proposed COI1, kfl00028_0560, does not contain an F-box domain. I used AMN1 members from the charophyte *K. flaccidum*, the bryophyte *P. patens*, the lycophyte *Selaginella moellendorffii* (spikemoss), the gymnosperm *Picea abies* (spruce), the basal angiosperm *Amborella trichopoda* (Amborella), and the model dicot *Arabidopsis thaliana* (Arabidopsis) to reconstruct the evolutionary history of COI1. Topology of maximum likelihood (ML), maximum parsimony (MP), minimum evolution (ME), and neighbor-joining (NJ) consistently grouped AtCOI1 and AtTIR1 in sister clades (Figure 3.1, Figure 3.2, Figure 3.3, and Figure 3.4). Identification of an ancestral *K. flaccidum* COI1 was inconsistent between trees. ME and NJ trees placed kfl00020_0460 and kfl00101_0280 as ancestral to AtCOI1 and AtTIR1; however, MP identified Kfl00834_0040 as the potential ancestral sequence (Figure 3.2, Figure 3.3, and Figure 3.4). The ML tree did not identify a discernible COI1 ancestral gene (Figure 3.1). These results indicate a true ancestral COI1 sequence is not present in *K. flaccidum*. In contrast, kfl00028_0560 consistently clustered basal to Arabidopsis At5g27920 and At5g01720, one moss, six *Spikemoss*, four spruce, and two Amborella AMN1 proteins (Figure 3.1A, Figure 3.2, Figure 3.3, and Figure 3.4). Grouping of the AMN1 genes using OrthoMCL, a program for predicting orthologs, did not cluster any *K. flaccidum* sequence with COI1 or TIR1 (Table 3.1). Importantly, proteins for both COI1 and TIR1 are present in the moss (Figure 3.1; Wang et al., 2015), suggesting the origin of these receptors occurred late during the evolution of charophytes.

To further eliminate the possibility of a COI1 in *K. flaccidum* I aligned the sequences proposed by the phylogenetic trees and the BLAST search with the authentic AtCOI1. As the auxin receptor TIR1 is closely related to and could potentially be the progenitor of COI1, it was included in the alignment. Alignment of AtCOI1, AtTIR1, At5g01720, kfl00020_0460, kfl00028_0560, kfl00101_0280, and kfl00834_0040 revealed few conserved JA binding amino acid residues in the *K. flaccidum* sequences (Figure 3.5). AtCOI1 contains three key loops necessary for jasmonate and JAZ binding (Sheard et al., 2010). In *K. flaccidum,* all loop regions, JA-Isoleucine binding sites, or JAZ interacting residues display limited, if any, similarity to the AtCOI1 sequence.

Structural comparisons of protein may indicate similarities not noticeable by amino acid sequence alignments. Proteins structures were modeled using the PHYRE2. PHYRE2 uses a homology based approach to devise the structure of a query sequence using structural data of known proteins and domains. The TM-Score algorithm calculates the structural similarity between proteins; scores < 0.17 indicate random structural similarity whereas values >0.5 are considered to have similar folds (Xu and Zhang, 2010). Structural comparisons using indicated AtCOI1 and AtTIR1 possess a similar fold, TM-score of 0.36 (Figure 3.1B). Supporting the phylogenetic relationship, kfl00028_0560 was more structurally similar to At5g01720 than

AtCOI1. Kfl00101_0280 was also more similar structurally to At5g01720. In contrast, kfl00020_0460 showed lower structural similarity to the three Arabidopsis proteins. Despite the close phylogenetic relationship in ME and NJ trees, kfl00020_0460 and kfl00101_0280 were structurally different. Kfl00834_0040 showed reasonable structural similarity to AtTIR1. Collectively, the incorrect BLAST hits, inconsistent phylogenetic relationships, and structural differences suggest a functional COI1 ortholog does not exist in *K. flaccidum*.

Identification of Charophyte Algae TIFY Proteins

The TIFY family contains the key jasmonate signaling regulators JAZ. To date, TIFY proteins have not been identified in any algal species, and thus the origin of JAZ factors remains elusive. I identified a TIFY protein in *K. flaccidum*, kfl00064_0080. Using the *K. flaccidum* DNA sequence, a search for other charophyte TIFYs identified one from *Chaetosphaeridium globosum* (Coleochaetophyceae), another from *Spirogyra pratensis* (Zygnematophyceae), and two closely related genes from *Nitella mirabilis* (Charophyceae). The two *N. mirab*ilis proteins (JV811897.1 and JV811898.1), distinguished by a WMGCDLQVV insertion in the C-terminal end of JV811897.1, are likely encoded by alleles of the same gene. No hits were retrieved for the Chlorokybophyceae or Mesostigmatophyceae. The only hit in the Klebsormidiophyceae belonged to the *K. flaccidum* gene. Interestingly, a BLAST search of Chlorophyta in the NCBI database identified a *Tetraselmis subcordiformis* and two other *Tetraselmis sp.* TIFY proteins. No TIFY proteins could be identified in the genomes of *Bathycoccus prasinos*, *Chlamydomonas reinhardtii*, *Chlorella variabilis*, *Coccomyxa subellipsoidea*, *Micromonas pusilla* CCMP1545, *Micromonas pusilla* RCC299, *Ostreococcus lucimarinus*, *Ostreococcus tauri*, or *Volvox carteri*. These sequences represent the first algal TIFY factors which can be further utilized to understand the evolution of the TIFY family in plants.

The Charophyte TIFY proteins Are ZIM Factors

The TIFY family is divided into four subfamilies, ZIM, TIFY, PPD, and JAZ (Bai et al., 2011). I surmised the algae TIFYs were the founding members of the family. To determine which subfamily originated first I investigated the domains present in the charophyte TIFY proteins. I identified complete TIFY, CCT, and Zn-GATA domains in the *K. flaccidum*, *M. polymorpha*, *N. mirabilis*, and *S. pratensis* proteins (Figure 3.6). The Zn-GATA domain identifies the charophyte TIFY proteins as ZINC-FINGER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM - LIKE (ZIM) factors, which I subsequently named KfZIM, NmZIM (JV811898.1), and SpZIM. The TIFY protein from *C. globosum* also contained a CCT domain, but as it was incomplete I cannot exclude the possibility of a missing C-terminal Zn-GATA domain. KfZIM was BLAST searched in NCBI for related sequences. I identified a similar TIFY protein from the liverwort *Marchantia polymorpha* (AY603044.1, E-value: 4e-70) and the moss *Pohlia nutans* (E-value: 3e-23). Analysis of these proteins also revealed TIFY, CCT, and Zn-GATA domains, indicating they are also ZIM factors which I named MpZIM and PnZIM. Four, four, two, and three ZIM factors were found in moss, spikemoss, Amborella, and Arabidopsis, respectively (Figure 3.7). No ZIMs occurred in the spruce; however, it contained numerous putative JAZ factors. A total of nine moss, six spikemoss, 46 spruce, six Amborella, and 12 Arabidopsis JAZ proteins were identified. Possible evidence for diversification after divergence was found as separate mono-species clades for multiple moss, spikemoss, and spruce JAZ proteins. Furthermore, multiple spruce factors only containing a TIFY

domain clustered with JAZ members. Moss, spikemoss, spruce, Amborella, and Arabidopsis possess three, five, two, one and one TIFY subfamily members, respectively. There were only two Arabidopsis and four spikemoss PPD subfamily members.

Structural comparisons indicated that AtPPD1 folding resembled that of AtZML2 and AtJAZ1. AtTIFY8 folding was also similar to AtJAZ1. A comparison of algal KfZIM, NmZIM, and SpZIM proteins to Arabidopsis TIFY factors indicated greater structural similarity with AtZML2 (Table 3.2). In contrast, the Chlorophyte alga TsZIM displayed greater structural similarity to JAZ1 and MpZIM. The best match for MpZIM to Arabidopsis factors was AtTIFY8, albeit the structural similarity is low. MpZIM was most similar to the modeled TsZIM structure. SpZIM and NmZIM had a similar folding structure to the earlier charophyte factor, KfZIM. Collectively, the results indicate charophyte TIFYs are ZIM proteins.

Origin of the CCT and ZnF-GATA Domains of ZIM Factors

It is possible the ZIM family originated in plants to fulfill the function of these factors. To identify the potential original function of ZIM proteins I conducted a BLAST search excluding the Viridiplantae. The top 10 hits resulted in a blue light receptor from *Trichoderma reesei*, cutinases from *Cryptococcus gattii* and *Mucor ambiguus*, and WHITE COLLAR 2 (WC2) proteins from *Cryptococcus gattii*, *Phycomyces blakesleeanus*, and *Phaeosphaeria avenaria* (Figure 3.8A). As ZIM functions in plant light signaling (Shikata et al., 2004) I considered the WC2 proteins as a possible ancestral gene. White collar proteins, involved in fungal light signaling (Linden and Macino, 1997; Talora et al., 1999), are Zn-GATA factors akin to ZIMs (Teakle and Gilmartin, 1998). Alignment of KfZIM with *C. gattii* WC2 revealed some homology between the sequences (13.1% identity and 20.5% similarity), most notably between domains and within the Zn-GATA DNA binding motif (Figure 3.8B). The unrooted phylogenetic tree placed WC2 proteins basal to ZIM factors (Figure 3.8). Modeled KfZIM and CgWC2 structures were compared, but displayed limited similarity (TM-score = 0.1651). To determine if WC2 factors were the authentic progenitors of ZIM factors I constructed the evolutionary history of the ZnF-GATA family with a phylogenetic tree of 16 moss, 14 *K. flaccidum*, 9 *B. prasinos*, 17 *C. reinhardtii*, 11 *C. variabilis*, 7 *C. subellipsoidea*, 9 *M. pusilla* CCMP1545, 8 *M. pusilla* RCC299, 10 *O. lucimarinus*, 5 *O. tauri*, 10 *V. carteri*, 7 red alga *Cyanidioschyzon merolae*, and 1 glaucophyte *Cyanophora paradoxa* GATA TFs along with 5 WC2 and 4 WC1 factors. The one ZnF-GATA TF identified in the ancient *C. paradoxa* lineage, and was used to root the tree (Figure 3.9 and Figure 3.10). Interestingly, the *C. paradox* protein had characteristics intermediate to type IVa and type IVb GATA TFs as it contains a $Cx_2Cx_{17}Cx_2C$ motif, but shares more conserved residues with fungal type IVb factors. ZIMs contain a $Cx_2Cx_{20}Cx_2C$ domain and appear to have separated early from other type IVb factors (Shikata et al., 2004). Consistent with these finding, algal ZnF-GATAs contained anywhere from 16 (e.g. Bathy16g00780) to 47 amino acids (CmaCMK203C) between the conserved Cysteine pairs, although most contained 18 residues. Several algae were also found to contain several ZnF-GATA factors with a $Cx_2Cx_{17}Cx_2C$ domain (Bathy16g02130 and kfl01232_0020) found in animals (Teakle and Gilmartin, 1998), which was hitherto unreported in plants. Analysis of the ZnF-GATA family indicated that WC2 were not the ancestor to ZIM TF. The low similarity of ZIM and WC2 proteins was due to a shared, albeit distinctly different, ZnF-GATA domain. Despite inconsistent topologies of trees due to limited sequences for early green algae, ZIMs cluster with the prasinophyte genes Bathy02g02870, Mic59665, and Olu25644 (Figure 3.9 and Figure 3.10). These prasinophyte sequences contain a $Cx_2Cx_{20}Cx_2C$ motif, consistent with the charophyte ZIM TFs. Collectively,

these data indicate that ZnF-GATA domain of ZIM TFs diverged from traditional GATA sequences early during plant evolution, at least prior to the Klebsormidiophyceae lineage, but potentially as far back as the last common ancestor of all Viridiplantae.

To determine the origin of the ZIM TFs CCT domain I constructed a phylogenetic tree of the CCT superfamilies from 39 Arabidopsis, 29 moss, 7 *K. flaccidum*, 5 *B. prasinos*, 8 *C. reinhardtii*, 4 *C. variabilis*, 5 *C. subellipsoidea*, 3 *M. pusilla* CCMP1545, 4 *M. pusilla* RCC299, 5 *O. lucimarinus*, 4 *O. tauri*, 5 *V. carteri*, and 3 *C. merolae* members. No proteins with a CCT domain were found in *C. paradoxa*. Both ML and ME trees clustered the same chlorophyte algae sequences with ZIM factors (Figure 3.11). The algae sequences present with ZIMs were a *B. prasinos* (Bathy10g03140), two C. reinhardtii (Cre01.g066750.t1.2 and Cre06.g278200.t1.2), a *M. pusilla* CCMP1545 (MicC56149), a *M. pusilla* RCC299 (Mic63995), an *O. lucimarinus* (Olu33434), an *O. tauri* (Ota34685), and three from *V. carteri* (Vc86975, Vc100325, and Vc102923). Another Arabidopsis protein, ASML2, also clustered with ZIM fators, suggesting it also originated from the same ancestral CCT domain. These results suggest that the CCT domain found in ZIM factors likely originated in the last common ancestor of the Viridiplantae.

To determine conserved changes involved in the evolution of the CCT into a JAS/CCT_2 domain I constructed a logo for each type of motif. CCT and JAS domains differ in several key residues; notably, at positions 1, 5, 12, 21, 22, 23, 24, and 25 (Figure 3.12). Position 1 in the CCT domain possessed several residues, but in JAS this location is an obligate proline. In ZIMs, position 5 was principally occupied by a leucine; however, in a bryophyte, a lycophyte, and a gymnosperm, this residue was a conserved lysine (Figure 3.12B and C). Only in angiosperms does position 5 also include a possible arginine residue at this site. A noticeable change in the transition from CCT to JAS domain was the conversion of an obligate arginine to a conserved leucine at position 12. The JAS domain lost conserved aspartate, lysine, lysine, and isoleucine residues at positions 21, 22, 23, and 24, respectively. An obligate tyrosine exists at position 26 in ZIMs and 27 in JAS. The residue prior to this tyrosine changed from an obligate arginine in ZIMs to a conserved proline in JAS domains. The pattern of residues for the JAS domain was similar across moss, spikemoss, spruce, and angiosperms (Figure 3.12C).

Only a remnant of the CCT lysine at position 22 could still be recognized in the moss and spikemoss JAS domain web logo. To further investigate if all features of JAS domains were present by the evolutionary arrival of embryophytes, I aligned moss and spikemoss JAZ proteins (Figure 3.12D). Three of the six moss JAZs lacked the C-terminal region; however, the other three factors (PpJAZ1, PpJAZ2, and PpJAZ3) contained all the unique JAS residues including the conserved proline and tyrosine. Notably, all of the spikemoss JAZ proteins lack the last two conserved residues. In contrast, only three of the 30 spruce JAZ factors had an incomplete JAS domain. Of all the 43 angiosperm sequences investigate, only StJAZ5 lacked a complete JAS domain. These results indicate key residue changes as the JAS domain probably evolved in the last common ancestor to all embryophytes.

Plant ZIMs Are Abundantly Expressed in Mature Male Reproductive Structures and Maturing Seeds

The CCT domain of ZIM factors was first identified in the photoperiodic light signaling genes TIMING OF CAB EXPRESSION1 (TOC1) and CONSTANS (CO). Intriguingly, AtZIM also has a known role in light signaling (Nishii et al., 2000; Robson et al., 2001). To investigate the possibility of a retained ZIM response to light I utilized expression data from eFP Browser (Winter et al.,

2007). Arabidopsis ZIM, ZML1, and ZML2 were slightly upregulated to four hours of blue, red, far-red, and white light compared to a dark treated control (Figure 3.13).

BLAST result annotates MpZIM factor as a male specific factor. To elucidate the possible evolutionarily conserved function of ZIM proteins in flowering I analyzed expression data for these factors in 10 plant species, including seven dicots [Arabidopsis, *Glycine max* (soybean), *Populus trichocarpa* (poplar), *Solanum lycopersicum* (tomato), *S. pennellii*, *S. pimpinellifolium*, *S. tuberosum* (potato)] and three monocots [*Hordeum vulgare* (barley), *Oryza sativa* (rice), and *Zea mays* (maize)]. Analyses of expression throughout plant development consistently revealed elevated levels of ZIM factors in plant leaves (Figure 3.13 and Figure 3.14). Six species (Arabidopsis, barley, maize, poplar, potato, and rice) had expression data publicly available for male reproductive tissues. Importantly, Arabidopsis, barley, maize, poplar, and rice displayed, for at least one ZIM factor, elevated levels of expression in male reproductive structures. Expression in Arabidopsis, illustrated that AtZIM was highly expressed in mature pollen (Figure 3.13). Monocots consistently showed the greatest expression of ZIMs in anthers (Figure 3.14). Arabidopsis, potato, tomato, *S. pennellii*, and *S. pimpinellifolium* all exhibited increased ZIM expression as seed or fruit maturation progressed. Similarly, maize and rice both displayed elevated expression of ZIMs in seed endosperm. ZIMs; therefore, may have evolved an early role in regulating male reproductive maturity and seed maturation.

It has been proposed that JA signaling, along with other hormones, evolved to allow plants to cope with new stresses from living on land (Wang et al., 2015). As the transition of plants from water to land and subsequent exposure to new stresses correlates with the expansion of the TIFY family I investigated the expression of ZIM factors in response to abiotic and biotic factors. Microarrays for Arabidopsis and potato include responses to various plant hormones and stress conditions. Both Arabidopsis and potato ZIMs responded to osmotic stress with mannitol (Figure 3.15). Potato ZIMs displayed some response to salt and heat treatment (Figure 3.15A). Arabidopsis ZIM TFs also showed some response to heat and UV stress. Neither Arabidopsis nor potato ZIM TFs responded strongly to hormone treatment. Collectively, my findings suggest that ZIMs may have had a limited role in hormone signaling and response to stress.

Group IIIb, IIId, IIIe, and IIIf bHLH Are Absent from the Genome of the Charophyte Alga K. flaccidum

In the study by Wang et al. (2015), BLAST results proposed orthologs of MYC2 in *K. flaccidum*. BLAST hits in large TF families may incorrectly identify putative orthologs due to similarity of conserved DNA-binding regions while sequentially and structurally diverse in non-binding regions differ. To clarify whether an ortholog of MYC2 exists within charophytes I constructed a phylogenetic tree of the bHLH family from *K. flaccidum*, moss, spikemoss, spruce, Amborella, and Arabidopsis. Members of the bHLH family were identified in each species using PlantTFcat (Dai et al., 2013). I ascertained one *C. merolae*, 12 *K. flaccidum*, 113 moss, 56 spikemoss, 195 spruce, 91 Amborella, and 162 Arabidopsis bHLH domain containing TFs. Putative orthologs groups were identified with OrthoMCL. No *K. flaccidum* orthologs were found for MYC2 or other group IIIe bHLH TF. Surprisingly, no MYC2 ortholog was detected in spikemoss. A phylogenetic tree was constructed to further investigate the possibility of a KfMYC2 factor (Figure 3.16). Supporting the OrthoMCL predictions, MYC2 orthologs occurred only in moss, spruce, Amborella, and Arabidopsis, but not *K. flaccidum* or spikemoss.

JAZ proteins directly interact with the Arabidopsis group IIIb, IIId, IIIf bHLH factors (Qi et al., 2011; Hu et al., 2013; Qi et al., 2015). No group IIIf TFs (GL3, EGL3, or TT8) were identified in *K. flaccidum*, but occur in moss, spikemoss, spruce, Amborella, and Arabidopsis. Similarly, orthologs of the group IIIb INDUCER OF CBF EXPRESSION 1 (ICE1) and ICE2, positive regulators of cold tolerance and stomatal development (Kanaoka et al., 2008; Hu et al., 2013), do not exist in *K. flaccidum*. In Arabidopsis, four JASMONATE-ASSOCIATED MYC2-LIKE (JAM) factors are group IIId bHLHs which repress JA signaling (Sasaki-Sekimoto et al., 2013; Qi et al., 2015). These analyses indicate orthologs of JAMs first appeared in the gymnosperm spruce. Overall, I did not find orthologs of repressive JAM bHLH TFs or group IIIb, IIIe, IIIf activators in *K. flaccidum*.

The Evolution of NINJA and the VICH Domain

JAZ proteins interact with the C-terminal end of NINJA which then recruits TPL via its N-terminal ethylene-responsive element binding factor-associated amphiphilic repression (EAR) motif (Pauwels et al., 2010). NINJA contains the C-terminal end of a domain of unknown function, DUF1675; henceforth I call VICH after the conserved residues, localized to the region interacting with JAZ (Figure 3.17). I identified 2 *K. flaccidum*, 2 moss, 4 spikemoss, 3 spruce, 3 Amborella, and 6 Arabidopsis VICH domain containing factors. No VICH domain proteins were found in chlorophyte algae *B. prasinos*, *C. reinhardtii*, *C. variabilis*, *C. subellipsoidea*, *M. pusilla* CCMP1545, *M. pusilla* RCC299, *O. lucimarinus*, *O. tauri*, *V. carteri*, the rhodophyte *C. merolae*, or the glaucophyte *C. paradoxa*. The *K. flaccidum* kfl00432_0070 and kfl00088_0230 both had incomplete domains lacking the C-terminal conserved VICH residues. Phylogenetic analyses grouped Pp1s336_28V6, Pp1s257_112V6, Smo402278, Smo439999, MA_10435900g0010, and Atr00056.20 as part of the AtNINJA clade, which I termed CI (Figure 3.17C). The remaining VICH proteins formed a second clade, called CII. Unrooted phylogenetic trees grouped both *K. flaccidum* basal to the CI and CII. As the key function of NINJA is to connect JAZ to TPL I searched for EAR motifs in each of the sequences. EAR motifs occur in two primary forms, as LxLxL or DLNxxP; only the former type was identified in VICH domain proteins. Except for Pp1s257_112V6, all CI NINJA orthologs had a single EAR motif in the N-terminal end of the protein. All CII members, excluding Atr00154.36, also possessed a single EAR motif towards the N-terminal end of the sequence. However, Smo439999 and Smo446487 each contained two EAR motifs both localized to the N-terminus of the proteins. Kfl00432_0070 contained three EAR motifs localized towards the center of the protein; in contrast, kfl00088_0230 did not possess any. These results suggest that NINJA proteins arose in an early embryophyte from a predecessor VICH domain that first appeared in a charophyte alga.

Identification of the Corepressor TOPLESS in Charophyte Algae

Arabidopsis TPL belongs to the Groucho/TUP1 transcription corepressor family (Liu and Karmarkar, 2008; Causier et al., 2012), and interacts with NINJA to repress JA response (Pauwels et al., 2010). TPL family members contain a conserved Lissencephaly type 1–like homology (LisH), a domain C-terminal adjacent to LisH (CTLH), and WD40 motif (Liu and Karmarkar, 2008). As the WD40 family is large, >200 members in each Arabidopsis, *C. paradoxa*, and *K. flaccidum* I focused on the LisH domain. To determine if a true ortholog of TPL exists in *K. flaccidum* I constructed a phylogenetic tree with 12 moss, 17 spikemoss, 14 spruce, 10 Amborella, 17 Arabidopsis, 6 *K. flaccidum*, 3 *C. merolae* and 2 *C. paradoxa* LisH families (Figure 3.18 and Figure 3.19). One clade

contained TPL factors from each embryophyte species. Towards the base of this clade was the *K. flaccidum* sequence kfl00881_0020, suggesting the presence of a TPL ortholog in this algal species. No TPL type factors occurred in *C. paradoxa*, *C. merolae* or any chlorophyte algae. These findings were supported by the OrthoMCL placement of kfl00881_0020 in the same group with TPL and TPR factors (Table 3.3). Additionally, I identified Kfl01154_0010 as the *K. flaccidum* ortholog of the LUENIG corepressor group. These findings indicate that TPL factors arose in streptophytes prior to *K. flaccidum* but post divergence from the chlorophytes.

In addition to NINJA, the repressor TPL interacts directly with JAZ proteins containing an EAR motif. I sought to determine when the EAR motif occurred in JAZ factors as an alternative mechanism for regulation of JA response. I identified EAR motifs in four Arabidopsis, one Amborella, six spruce, and eight spikemoss JAZ proteins (Table 3.4). Additionally, I found EAR motifs in two moss and three spruce TIFY factors, as well as one spikemoss peapod gene, SmPPD3. The only ZIM TF to contain an EAR motif was from *S. pratensis*. SmJAZ5, SmJAZ7, and SmPPD3 each possessed two EAR motifs. In angiosperms, the EAR motif was located towards the N- or C-terminal end of JAZ proteins. In contrast, moss, spikemoss, and spruce EAR motifs predominantly localized to the middle of proteins. All factors identified contained the LxLxL type EAR motif. PaJAZ1, PaJAZ4, and PaTIFY12 each contained a combination of the two EAR motifs. Most EAR motifs occurred prior to the TIFY domain, with only a few located after the CCT domain. In dual EAR factors SmJAZ5 and SmJAZ7, one motif occurred before the TIFY domain and the other after the CCT region. Both EAR motifs of SmPPD3 were present prior to the TIFY domain. In SmJAZ3 and SmJAZ4 the EAR motif was positioned between the TIFY and CCT domains. In the SpZIM TF, the EAR motif overlapped with the TIFY domain. Collectively, these results suggest that an LxLxL type EAR motif was preferentially located prior to the TIFY domain in the middle of proteins in early embryophytes, but later situated towards the ends of angiosperm JAZ factors.

MEDIATOR OF RNA POLYMERASE SUBUNIT 25 Is Present in a Charophyte Genome

The ACID domain of MED25 interacts with MYC2 to regulate JA signaling (Çevik et al., 2012; Chen et al., 2012). In plants, MED25 has only been identified in embryophytes (Mathur et al., 2011); thus, to further evaluate the origin of JA signaling I sought to determine if MED25 is present in a charophyte genome. A BLAST search in *K. flaccidum* of the characterized Arabidopsis MED25 identified one putative ortholog, kfl00621_0040 (E-value: 2e-38). A reciprocal BLAST search of kfl00621_0040 in Arabidopsis best matched to AtMED25 (E-value: 4e-21). Alignment of the KfMED25 and AtMED25 peptides found 31.4% similarity, 20.4% identity, and 40.3% gaps between the sequences (Figure 3.20A). As low similarity exists between plant, animal, and fungal Mediator subunits (Bäckström et al., 2007), I compared KfMED25 to the evolutionary closer moss MED25. Two genes have been reported to encode MED25 in moss (Mathur et al., 2011). BLAST of the moss genome with AtMED25 identified Pp1s233_16V6 and Phpat.025G010600/ Pp1s233_44V6 as potential orthologs (E-values: 3.5e-52 and 8.1e-31, respectively). Both genes contain a von Willebrand domain, but only Pp1s233_16V6 possessed the ACID domain. Alignment of the *K. flaccidum*, Arabidopsis, and moss MED25 peptides revealed that KfMED25 also possessed a von Willebrand (vW) and ACID domains (Figure 3.20). Notably, KfMED25 lacks the glutamine rich C-terminal region found in AtMED25, and to a lesser extent PpMED25.

I sought to further determine when MED25 was acquired by the Viridiplantae lineage. BLAST searches of KfMED25 and AtMED25 in NCBI returned a complete *Araucaria cunninghamii* (gymnosperm), as well as putative MED25 factors in the chlorophytes *Auxenochlorella*

protothecoides (XP_011401634.1), *C. subellipsoidea* (XP_005645244.1), *Monoraphidium neglectum* (XP_013895617.1), *Tetraselmis* sp. GSL018 (JAC70819.1), and *V. carteri* (XP_002953330.1). Searching sequence algal genomes for the conserved vW domain I identified putative MED25s from *C. reinhardtii*, *C. subellipsoidea*, and *V. carteri*. Identical sequences from *C. subellipsoidea* and *V. carteri* were identified by both methods. No MED25 sequences were found in *B. prasinos*, *C. variabilis*, *M. pusilla* CCMP1545, *M. pusilla* RCC299, *O. lucimarinus*, *O. tauri*, the red algae *C. merolae*, *Chondrus crispus*, *Galdieria sulphuraria*, or the glaucophyte *C. paradoxa*. Delineation of embryophyte vW domains revealed two moss, four spikemoss, one spruce, one Amborella, and the one annotated Arabidopsis MED25 proteins. Three amoebozoa species contained *Dictyostelium discoideum*, *Dictyostelium purpureum*, and *Polysphondylium pallidum*. A phylogenetic tree was constructed using the 18 identified MED25 sequences (Figure 3.20B). The tree was rooted with the opisthokont MED25 from *D. discoideum*. Sequences clustered similar to that of the species tree (Figure 3.20C), suggesting I identified *bona fide* algal MED25 orthologs. Therefore, the results suggest that MED25 probably occurred in the last common ancestor of all plants, but was subsequently lost from certain algae clades (i.e. Mamiellaceae).

Development of the JA Signaling Pathway during the Ediacaran and Cambrian Periods

Current fossils suggest the presence of early land plants 475 million years ago (MYA), but paleobotanical evidence for the origin of charophytes is severely limited (Wellman et al., 2003). Molecular clock analyses of gene sequences can be utilized to overcome limitations of the fossil record and identify times of divergence. To determine the evolutionary history of charophytes and when JA signaling was acquired I conducted a molecular clock analysis. Small rRNA subunits (16S) for 98 Unikont-Ophistokont, Glaucophyta, Chlorophyta, and Streptophyta sequences were used to reconstruct the Viridiplantae evolutionary history. Rhodophyta were excluded based on clustering with Unikont-Ophistokont sequences, inconsistent with the generally accepted history of plant evolution. The ML tree was rooted with the cluster of Unikont-Ophistokont sequences. A general time reversible model was applied to the ML tree to calculate divergence times. The TimeTree database indicates a divergence time of 150 MYA (median of 33 studies) between monocots and dicots, and this value was used to calibrate the molecular clock. With my molecular clock analysis, the appearance of glaucophytes was estimated at 1300 MYA (Figure 8). The divergence of Streptophyta and Chlorophya from the last common ancestor of all plants probably occurred approximately 1076 MYA. Approximately 746 MYA the Mesostigmatophyceae-Chlorokybophyceae groups diverged from other Charophyta during the middle Cryogenian period. In my tree, no clear charophyte sister group to embryophytes was identified. The Charophyceae grouped independently of a Klebsormidiophyceae- Coleochaetophyceae-Zygnematophyceae clade. Within the latter clade, the Klebsormidiophyceae, Coleochaetophyceae, and Zygnematophyceae split 441 MYA during the early Silurian. The Klebsormidiophyceae- Charophyceae -Coleochaetophyceae- Zygnematophyceae diverged from the Embryophya during the Ediacaran period 628 MYA. The bryophytes, lycophytes, monilophytes, and gymnosperms diverged from their ancestors 491 (late Cambrian), 469 (middle Ordovician), 433 (middle Silurian), and 326 (late Silurian) MYA, respectively. In this analysis, the important angiosperm crown group was calculated at 174 MYA. Around 118 MYA the dicots diverged into asterids and rosids groups. Known fossils were used to validate the accuracy of time estimates (Scott and Chaloner, 1983; Edwards et al., 1992; Hill et al., 1997; Friis et al., 2001; Wellman et al., 2003; Gerrienne et al., 2004). Based on the molecular clock data I concluded that

the JA signaling pathway evolved between early Ediacaran and late Cambrian periods 628 to 491 MYA (Figure 3.21).

Discussion

One major goal in studying hormone signaling networks is to understand how, when, where, and why a given pathway evolved. As most hormones regulate multiple responses, difficulty arises in determining which function arose first. JA contributes to regulation of male fertility, root growth, seed germination, production of specialized metabolite, and symbiosis with microorganisms, as well as arbitrates cold, wounding, pathogens and insect herbivory responses (Staswick et al., 1992; Feys et al., 1994; Thomma et al., 1998; Xie et al., 1998; Li et al., 2004; Lorenzo et al., 2004; Dombrecht et al., 2007; Zhang et al., 2007; Hu et al., 2013). These functions coincide with traits acquired during the transition of plants from water to land (Rensing et al., 2008). Oxylipin biosynthesis originated in the last common ancestor of plants and animal (Lee et al., 2008) and the charophyte *K. flaccidum* produces JA (Hori et al., 2014). However, when the JA signaling pathway arose in plants is less clear. A previous work suggested that JA signaling evolved in the last common ancestor of all embryophytes (Wang et al., 2015). Here I found: 1) in contrast to a previous report, no evidence for COI1, JAZ, MYC2, JAM or other bHLH targets of JAZ in the charophyte alga *K. flaccidum*, 2) evidence for the first TIFY factor in algae, 3) that ZIM TFs were the ancestral genes to JAZ factors and likely evolved to regulate male reproductive structure development and later seed food storage, and 4) presence of the MED25 in Charophyta and Chlorophyta algae. Furthermore, a molecular clock analysis of the plant lineage indicates JA signaling arose during the Ediacaran to late Cambrian, 628-491 MYA. Collectively, my findings refine when each JA signaling component arose and further support the hypothesis that the complete pathway was present in the last common ancestor of embryophytes (Figure 3.22).

Identification of the ancestral COI1 and JAZ is paramount for understanding the evolution of JA signaling. Consistent with Hori et al. (2014) and in contrast to Wang et al. (2015), I could not find evidence for an authentic COI1 ortholog in *K. flaccidum*. Neither was a clear TIR1 ortholog identified in the charophyte; although, phylogenetic analyses (Figure 3.1A, Figure 3.2, Figure 3.3, and Figure 3.4) suggested a common progenitor AMN1 factor for both the auxin and JA receptors (Wang et al., 2015). In contrast to *K. flaccidum*, both COI1 and TIR1 orthologs were present in moss. Consistent with previous reports (Sheard et al., 2010), the modeling indicated a high similarity between the Arabidopsis COI1 and TIR1 structures (Figure 3.1B), suggesting that auxin and JA signaling cascades evolved in parallel during the transition of plants from water to land.

Here, I report TIFY domain containing factors from charophyte and chlorophyte algae as ZIM TFs (Figure 3.6 and Figure 3.7). As other types of TIFY proteins are not present in algae, I consider ZIMs as the initial founding family members from which TIFY and JAZ derived. The ZIM CCT domain was first identified based on its similarity to the protein-protein interaction domain of CO which is involved in photoperiodic response (Robson et al., 2001). ZIMs contain a unique $Cx_2Cx_{20}Cx_2C$ motif. Other plant or animal ZnF-GATA TFs have $Cx_2Cx_{18}Cx_2C$ or $Cx_2Cx_{17}Cx_2C$, respectively (Teakle and Gilmartin, 1998). Congruent with previous reports, alignment and phylogenetic analysis of the Zn-GATA family in the Viridiplantae revealed that plant factors are most similar to fungal type IVb members (Figure 3.9 and Figure 3.10), including WC2 (Teakle and Gilmartin, 1998). Notably, I identified one ZnF-GATA in *C. paradoxa* having a $Cx_2Cx_{17}Cx_2C$ sequence, suggesting that the plant $Cx_2Cx_{18}Cx_2C$ motif is a derivative of the animal type domain (Figure 3.9 and Figure 3.10).

To determine the potential role of ZIMs in early land plant evolution I analyzed coexpression data from 10 species (Figure 3.13, Figure 3.14, and Figure 3.15). The investigated ZIMs did not strongly respond to JA or other hormone treatments (Figure 3.15). Additionally, except for osmotic stress, ZIMs displayed little response to abiotic or biotic stresses (Figure 3.15). As previously reported for Arabidopsis ZIMs (Shikata et al., 2004; Shaikhali et al., 2012), a mild up-regulation in response to light was observed (Figure 3.13 and Figure 3.14). Gene expression data consistently indicated a strong up-regulation of most ZIMs in leaves, male reproductive structures, and maturing fruits (Figure 3.13 and Figure 3.14). Elevated ZIM expression in maturing fruits, but not embryo, insinuates a role in controlling food storage. As JA regulates male fertility, it is tempting to speculate that the early functions of ZIM TFs and subsequently derived JAZ factors were to regulate male reproductive development. Importantly, evolution of JA signaling coincides with developments in male reproductive structures. The gametophyte is the primary structural body of both charophytes and bryophytes. However, compared to most algal species, bryophyte spores experience periods of exposure to air which can lead to desiccation. Additionally, bryophyte spores, particularly in liverworts and hornworts but less so in mosses, are surrounded and more protected by large multicellular structures (i.e. antheridiophore). Charophytes lack extensive multicellular protection of spores. As such charophyte spores are more susceptible to stress and predation. Thus, as a charophyte species evolved into the embryophyte linage, plants developed increased desiccation tolerance and protection of the male gamete.

I determined a JAS domain consensus motif similar to previous reports (Sheard et al., 2010). Knowing the JAZ evolutionary history, I identified three key amino acid changes which occurred as the ZIM CCT domain evolved into the JAS domain (Figure 3.12A and B). Truncation of the CCT domain and absence of a GATA DNA-binding region suggest JAZ factors arose from a loss of the C-terminal end of a ZIM TF. Position 1 of the ZIM CCT domain evolved from various amino acids to a conserved proline. The conserved leucine at position 5 mutated into a lysine or arginine in embryophyte JAZ factors. The third key residue change involved conversion of an arginine in ZIMs to a leucine at position 12 in JAZ. As only 21 amino acids are necessary for a functional JAZ degron, other changes between CCT and JAS domains may play alternative roles (Sheard et al., 2010). The C-terminal portion of the JAS domain, including a conserved proline, contributes to nuclear localization (Grunewald et al., 2009). In ZIM TFs the DKKIRY sequence, well suited for nuclear localization, was replaced by X_4PY in JAZ. Why JAZ require this conserved proline in the JAS domain and has evolved away from more favorable nuclear localization residues remains unclear.

My results indicate that orthologs of group IIIe bHLH TFs occurred in moss, spruce, and Amborella, but, surprisingly, not in spikemoss (Figure 3.16). Previously, Thaler et al. (2012) reported a putative spikemoss *MYC2*. This gene was present in my dataset but not identified by plantTFcat due to the lack of a bHLH domain. Orthologs of ICE1, ICE2, GL3, EGL3, TT8, and MYC1 were present in moss, spikemoss, spruce, and Amborella (Figure 3.16). ICE1 and ICE2 regulate cold tolerance and stomatal development for gas exchange across an impermeable cuticle (Kanaoka et al., 2008; Hu et al., 2013). GL3, EGL3, and TT8 regulate development of herbivore deterring trichomes, flavonoid UV protectants, and seed fatty acid accumulation. Importantly, both group IIIb and IIIf regulate traits that are acquired as plants transition from aqueous to terrestrial environments. In contrast, the group IIId bHLH JAMs are negative regulators of JA signaling (Sasaki-Sekimoto et al., 2013). Group IIId bHLHs interact with JAZ proteins (Song et al., 2013), to fine tune gene expression by binding the G-box *cis*-elements in promoters of MYC2/3/4 genes, physically obstructing activator-DNA interactions (Song et al., 2013; Qi et al., 2015). Here,

I identified the earliest orthologs of JAMs in gymnosperms (Figure 3.16), suggesting fine tuning of JA signaling occurred later in plant evolution.

Groucho/Tup1 family members are found in animals and fungi (Liu and Karmarkar, 2008). I identified proteins with LisH domain as far back as glaucophytes (Figure 3.18, Figure 3.19, and Table 3.3). In addition to their involvement in the JA signaling cascade, TPL and TPRs possess functions in other phytohormone pathways including auxin, abscisic acid, brassinosteroids, ethylene, gibberellin, salicylic acid, and strigolactones (Szemenyei et al., 2008; Pauwels et al., 2010; Zhu et al., 2010; Causier et al., 2012; Jiang et al., 2013; Fukazawa et al., 2014; Oh et al., 2014); furthermore, they mediate apical and axillary meristem formation, circadian rhythm regulation, embryogenesis, leaf blade formation and development, and sporogenesis (Long et al., 2006; Szemenyei et al., 2008; Kwon et al., 2012; Tao et al., 2013; Wang et al., 2013; Chen et al., 2014; Zhang et al., 2014). Given the multiple functions and the long evolutionary history of Groucho/Tup1 factors, the role of TPL in JA signaling was probably co-opted from another preexisting pathway.

I uncovered the occurrence of MED25 predating the JA signaling cascade and MYC2, thus its original function was likely to recognize other unknown TFs (Figure 3.22A). MED25 regulates JA signaling (Kidd et al., 2009; Çevik et al., 2012; Chen et al., 2012; Yang et al., 2014) by interacting with MYC2, ERF1, ORA59, and other TFs that function in propagating the JA response (Ou et al., 2011; Çevik et al., 2012; Chen et al., 2012; Yang et al., 2014). However, MED25 also regulates abscisic acid, auxin, ethylene, and salicylic acid signaling as well as response to light and iron deficiency, cell size, flowering time, lateral root and root hair development (Kidd et al., 2009; Elfving et al., 2011; Iñigo et al., 2011; Xu and Li, 2011; Chen et al., 2012; Klose et al., 2012; Sundaravelpandian et al., 2012; Koprivova et al., 2014; Raya-González et al., 2014; Yang et al., 2014). Here, I identified a MED25 ortholog in the charophyte alga *K. flaccidum* (Figure 3.20A and B). Additionally, MED25 orthologs were identified in the Chlorophyta algae *C. reinhardtii*, *C. subellipsoidea*, and *V. carteri* (Figure 3.7B). As all members of the Chorophyta Mamiellaceae family lack a MED25, the subunit was probably lost from an ancestor of these algal lineages.

The role of JA in signaling outside of angiosperms is enigmatic. Contrasting reports exist for the presence of JAs in moss (Oliver et al., 2009; Stumpe et al., 2010; De Leon et al., 2012). *P. patens* synthesizes the JA precursor 12-oxophytodecanoic acid (OPDA), yet both JA and OPDA are perceived (De León et al., 2015; Yamamoto et al., 2015). Various red and green algae respond to JA (Christov et al., 2001; Collén et al., 2006; Kováčik et al., 2011; Raman and Ravi, 2011). Furthermore, the Chlorophyta Dunaliella, Rhodophyta *Gelidium latifoliu*m, fungus *Lasiodiplodia theobromae*, protist Euglena and a Spirulina cyanobacterium synthesize JA (Aldridge et al., 1971; Krupina and Dathe, 1991; Ueda et al., 1991; Ueda et al., 1991; Fujii et al., 1997). Importantly, some algae (e.g. *C. vulgaris*) both synthesize JA and respond to exogenous JA treatment despite lacking JAZ proteins (Ueda et al., 1991; Czerpak et al., 2006). The function of JA in these algal species is intriguing. One possibility is the existence of a COI1-independent JA signaling cascade. Work with *coi1* mutants suggests that this mechanism occurs in land plants (Stintzi et al., 2001). I previously proposed WRKY TFs may be involved (Schluttenhofer et al., 2014), which is supported by their presence in algal genomes (Figure 3.22B). Alternatively or concurrently, JA and its derivatives may have direct antimicrobial properties independent of signaling. JA inhibits the growth of the fungus *Candida albicans* (Morris et al., 1979) and other oxylipins display antimicrobial activity (Prost et al., 2005). My establishment of the core JA signaling pathway around the time of the last common embryophyte ancestor, in combination with the reported

roles of its synthesis and responses in red and green algae, suggests a more ancient oxylipins signaling response network exists in the Archaeplastida.

Molecular clock analysis of the Viridiplantae indicated the divergence of proposed charophyte ancestors from embryophytes during the early Ediacaran period (Figure 8b). The predicted timelines are compatible with the geological record, as the earliest known fossil evidence occurs shortly after each estimate (Figure 3.21B). My analysis estimated the appearance of bryophytes at 491 MYA (Figure 3.21A), whereas the earliest fossils are spores dating from 475 MYA (Wellman et al., 2003). The earliest known vascular plant, Cooksonia, from 415 MYA (Edwards et al., 1992), is preceded by my estimate of lycophyte divergence at 469 MYA. The estimate for monilophyte divergence, 433 MYA, occurs prior to early fern fossils from the Upper Devonian (Hill et al., 1997). The molecular clock estimated that gymnosperm divergence occur post Runcaria fossils, the earliest plant bearing seed-like structures from 385 MYA (Gerrienne et al., 2004), but prior to the early conifer specimens from 314 MYA (Scott and Chaloner, 1983). A previous report indicated an angiosperm from the late Jurassic period (Sun et al., 1998), but this was later re-dated to the early Cretaceous (Swisher et al., 1999) along other basal angiosperm species (Friis et al., 2001; Gomez et al., 2015). Given the consistency of my molecular clock analysis with the fossil record, I used it to date the period when JA signaling evolved. Based on the identification of necessary pathway genes, JA signaling evolved in the last common ancestor of embryophytes, after their divergence from charophytes but prior to the first bryophyte, 628-491 MYA (Figure 3.21A).

My estimates put acquisition of JA signaling during the warm, inter-glacial, Ediacaran-Cambrian periods (Figure 3.21B). Notably, attainment of the JA signaling pathway overlaps with the “Cambrian explosion”, a time of great species radiation, particularly of fauna phya (Morris, 1993; Briggs, 2015). Fossils denoting the expansion of grazing metazoan species (Caron et al., 2006; Smith, 2012) precedes those of the earliest land plant specimens (Wellman et al., 2003) and is almost certainly equated to increased algal herbivory. Increased predation likely necessitated algal evolution of improved reproduction to enhance survival. Importantly, my molecular clock analysis indicated that the Cambrian explosion coincides with the evolutionary development of free swimming sperm and increasingly protected algal reproductive bodies found in the Charophyceae, Coleochaetophyceae, Zygnematophyceae (McCourt et al., 2004) and Embryophya. These results suggest ZIM TFs may have evolved to fill the roles in regulation of male reproductive structures (Figure 3.13 and Figure 3.14). As JA signaling also regulates male development, I hypothesize that a connection between ZIM and the improved reproduction systems, equating to enhanced algal survival, could have allowed evolution of ZIM into JAZ factors as continued pressure of predation drove the development of additional complex defense mechanisms that required extra regulators. My findings indicate that other regulators of JA signaling, the COI1 receptor, NINJA, and multiple bHLH TF regulators, evolved during this period of explosive diversification. In addition to driving development of JA signaling, predation and the avoidance thereof may have triggered grand consequences on land plant evolution. Competition between new algal species and high rates of predation during the Cambrian may have forced algal species to search for new predator-deficient niches, i.e. land, in which to grow. I thus propose that the Cambrian explosion of fauna species pressured an alga species, now well-equipped with JA signaling and other necessary stress-tolerant regulatory mechanisms, to colonize land to avoid predation and the increasing competition from other algae species. This agrees with a suggestion that defense signaling networks (abscisic acid, JA, and salicylic acid) evolved in the last common ancestor of land plants, prior to several hormones, such as brassinosteroids, ethylene, and gibberellin, with roles more oriented toward plant growth and development (Wang et al., 2015).

While increasing stress and predation could apply positive selection pressure for evolution of JA signaling, evolving genes were also likely subjected to purifying selection which removed deleterious variants which arose. It is probable during diversification of signaling cascades, particularly with TFs, that mutant genes encoding protein variants improperly regulating essential genes arose which resulted in detrimental consequences. Therefore, development of the JA signaling cascade likely resulted from both positive and negative selection pressures. To answer how rapidly and which selection pressure provided greater contributions to evolution of JA signaling requires further study. Sequencing of additional charophyte genomes will be important to address these questions.

As JA signaling is a major determinant of plant response to the environment, I sought to determine how it evolved in the Viridiplantae lineage. Here, I provide support for the hypothesis that JA signaling primarily evolved during the transition of waterborne charophytes into land plants (Wang et al., 2015). However, the multiple protein domains necessary to form JA signaling components existed as early as the glaucophytes (Figure 3.22). Nevertheless, to further dissect and pinpoint the origin of JA signaling components requires genome sequencing of additional charophytes and early embryophytes. The origin of JA signaling acquisition coincides with a time of increasing fauna species which, due to elevated levels of predation, may have driven plants onto land. The identification of ZIM TFs as the progenitor to JAZ factors suggests that plants may have evolved the JA pathway for protection of the male gamete. As phytohormone signaling underlies most plant responses, a thorough understanding of their evolution can provide information on how biological networks develop, synthesize or co-opt new components, and adapt to those changes. These insights will allow future studies to test ancestral gene functions and to integrate developmental and genetic features which occurred during the aqueous to terrestrial environment transition. In addition to understanding fundamental plant biology and evolution, elucidation of phytohormone regulation provides insights into what phenotypic consequences they may be elicited from perturbation for human uses. As JA regulates plant response to wounding and pathogens, as well as key targets of crop improvement, its evolutionary history may guide future genetic engineering. My findings advance our knowledge on plant evolution and the acquisition of phytohormone signaling cascades. My work begins to answer how, when, where, and why JA signaling arose; however, more work is needed to further define this key advancement in plant evolution.

Materials and Methods

Identification of Genes

Proteins were identified by BLAST result into National Center for Biotechnology Information (NCBI). Protein families were identified by their domains using NCBI Conserved Domain Database (Marchler-Bauer et al., 2015) according to Schluttenhofer et al. (2014). In brief genome sequences of interest were downloaded then converted into an Excel file; single copies of proteins were determined and then submitted to the NCBI Conserved Domain Database. Results were then used to identify proteins with a domain of interest. Orthologs were identified using phylogenetic trees and results from OrthoMCL (Li et al., 2003).

Phylogenetic Analyses

Multiple alignments of proteins were performed using the Clustal Omega program. The Needle algorithm was used to generate pairwise alignments. Alignments for phylogenetic trees used the ClustalW algorithm. ML, MP, ME, NJ, and UPGMA phylogenetic trees were constructed using MEGA5.2. Specific parameters for each tree are listed in the corresponding figure legend. Sequences forming incompatible alignment pairs were exclude from alignments and constructed trees.

The small subunit rRNA sequences used for the molecular clock were obtained from SILVA (http://www.arb-silva.de/). The divergence time between monocots and dicots (150 MYA) was based on the median of 33 studies published in TimeTree (http://www.timetree.org/).

Gene Expression

Expression of Arabidopsis, poplar, Medicago, soybean, rice and maize ZIM factors were obtained from eFP Browser, http://bar.utoronto.ca/ (Winter et al., 2007). To identify probes for poplar and Medicago I BLASTED ZIM sequences against the PLEXdb (http://www.plexdb.org/index.php). BLAST results were conducted in Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html#) to identify orthologs of *TUBULIN3* which was used as a control gene for comparison of fold change in all datasets.

Protein Structures and Motifs

The structures of proteins were generated by homology threading using the PHYRE2 server (Kelley et al., 2015). Structures were then compared using TM-Score Online (http://zhanglab.ccmb.med.umich.edu/TM-score/) to obtain a TM-score. TM-scores were evaluated according to Xu and Zhang (2010). Logos were created using WebLogo (http://weblogo.berkeley.edu/). ZIMs factors and embryophyte JAZ proteins were used to construct the CCT and CCT_2/JAS logos, respectively. Domains were aligned with Clustal Omega prior to logo generation.

Genome Databases

Genome sequences for *Bathycoccus prasinos*, *Chlorella variabilis*, *Coccomyxa subellipsoidea*, *Cyanidioschyzon merolae*, *Cyanophora paradoxa*, *Micromonas pusilla* CCMP1545, *Micromonas pusilla* RCC299, *Ostreococcus lucimarinus*, and *Ostreococcus tauri* were obtained from the Joint Genome Institute (http://genome.jgi.doe.gov/). *Arabidopsis thaliana*, *Chlamydomonas reinhardtii*, *Physcomitrella patens*, *Selaginella moellendorffii* and *Volvox carteri* sequences were downloaded from Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html). *Amborella trichopoda* and *Picea abies* sequences were obtained from the Amborella Genome Database (http://www.amborella.org/) and Spruce Genome Project (http://congenie.org/start), respectively. Sequences for *K. flaccidum* were downloaded directly from *Klebsormidium flaccidum* Genome Project website (http://www.plantmorphogenesis.bio.titech.ac.jp/~algae_genome_project/klebsormidium/).

Table 3.1. OrthoMCL ortholog groups of AMN1 family members.

Ortholog Group	*K. flaccidum*	*P. patens*	*S. moellendorffii*	*P. abies*	*A. trichopoda*	*A. thaliana*
OG5_126657	kfl00011_0110, kfl00013_0250, kfl01028_0010	Pp1s105_237V6, Pp1s107_101V6, Pp1s176_100V6, Pp1s287_25V6, Pp1s35_31V6	Smo108463, Smo162512, Smo419616	MA_136805g0010, MA_327255g0010, MA_414622g0010, MA_8025229g0010	---	AT1G15740
OG5_127136	kfl00304_0120	---	---	---	---	---
OG5_127678	kfl00028_0560, kfl00088_0180, kfl00175_0070, kfl00175_0270, kfl00285_0160, kfl00290_0040, kfl00513_0020, kfl00581_0090, kfl00595_0040, kfl00730_0080	Pp1s279_17V6, Pp1s58_121V6, Pp1s97_169V6	Smo118815, Smo142102, Smo184280, Smo437235, Smo445748, Smo75506, Smo87311, Smo88742	MA_10435873g0020, MA_114225g0010, MA_8692330g0010	Atr00003.99, Atr00025.126	AT4G15475, AT5G23340
OG5_127738	---	Pp1s26_228V6	---	---	---	---
OG5_129307/RPP	---	---	---	MA_160159g0010	---	AT1G63860, AT2G14080, AT3G04220, AT3G44400, AT3G44480/RPP1, AT3G44670, AT4G16920, AT4G16950/RPP5, AT4G16960, AT5G11250
OG5_131898	kfl00020_0460	---	---	---	---	---
OG5_133501	kfl00010_0620	---	---	---	---	AT2G06040, AT5G21900
OG5_134851	---	---	---	---	---	AT4G36150
OG5_135305	kfl00099_0260, kfl00434_0100	---	---	---	---	---
OG5_137188	---	Pp1s17_234V6	---	---	---	---
OG5_140019	---	---	---	---	---	AT5G27920
OG5_141134	---	---	---	MA_10426597g0010, MA_10428924g0010, MA_10428924g0020, MA_10436081g0010	Atr00023.52	AT5G01720
OG5_143138/TIR1	---	Pp1s137_148V6, Pp1s196_87V6	Smo168175, Smo170974, Smo178850, Smo179436	MA_15842g0010	Atr00003.426	AT3G62980/TIR1, AT4G03190/AFB1
OG5_144350	---	Pp1s37_126V6	---	---	---	---
OG5_145014	---	---	---	MA_140161g0010	---	AT1G80570
OG5_149464/SLOMO	kfl00352_0100	Pp1s180_48V6, Pp1s23_195V6, Pp1s284_41V6, Pp1s46_277V6	Smo133451, Smo137667	MA_33148g0010	Atr00024.244	AT4G33210
OG5_152427	---	---	Smo101488, Smo102329	---	---	---
OG5_153187	---	Pp1s79_129V6	---	MA_10429136g0010, MA_107074g0010	---	AT5G07670, AT5G51370, AT5G51380
OG5_153396	kfl00009_0200	Pp1s128_42V6	Smo133097, Smo133180	MA_107288g0010	---	---
OG5_156349/FBW2	---	---	---	---	---	AT4G08980
OG5_156629/COI1	---	Pp1s277_20V6		MA_108477g0010, MA_15205g0010	Atr00029.327	AT2G39940/COI1
OG5_159818/EBF1	kfl00325_0170	Pp1s188_33V6	Smo123441, Smo233861	MA_10436543g0020, MA_64990g0010	Atr00078.21	AT2G25490
OG5_160079	---	---	---	MA_10434210g0010, MA_97184g0010	Atr00010.351, Atr00010.352	---
OG5_163790	---	Pp1s283_5V6	---	MA_760407g0010	---	---
OG5_164744	---	---	---	MA_121693g0010	Atr00135.41	---
OG5_166919	kfl00582_0050	---	---	---	---	---
OG5_169541	kfl00128_0070	---	---	---	---	---
OG5_177872	---	---	---	---	---	AT3G48880
OG5_177965	kfl00834_0040	---	---	---	Atr00159.5	---

Table 3.1. cont.

OG5_178237	---	Pp1s197_61V6	Smo31890	---	Atr00001.157	AT2G17020
OG5_178250	---	---	---	---	---	AT2G36370
OG5_178271	kfl00508_0080	---	---	MA_57733g0010	Atr00066.158	AT3G07550
OG5_184711	---	Pp1s104_63V6	---	---	---	---
OG5_190642	---	---	Smo425843, Smo447499	---	---	---
OG5_190918	---	Pp1s148_40V6	---	---	---	---
OG5_190984	kfl00169_0150		Smo130925, Smo90468	---	Atr00133.40	AT3G58530
OG5_205662	---	---	Smo89604	---	---	---
OG5_212059	---	---	---	MA_46507g0010	---	AT1G55590
OG5_243917	kfl00101_0280	Pp1s81_23V6	---	MA_8336373g0010	---	---

Table 3.2. The structural relationship between TIFY factors. The structures of Arabidopsis TIFY factors and early ZIM TFs were modeled with PHYRE2. Structures were compared for similarity using the TM-score. TM-scores between combinations of factors are shown.

Gene	AtZML2	AtJAZ1	AtPPD1	AtTIFY8	MpZIM	KfZIM	NmZIM	SpZIM	TsZIM
AtZML2	1								
AtJAZ1	0.1525	1							
AtPPD1	0.1806	0.1779	1						
AtTIFY8	0.1646	0.1727	0.1594	1					
MpZIM	0.1363	0.1090	0.1405	0.1505	1				
KfZIM	0.2256	0.1392	0.1583	0.1445	0.1659	1			
NmZIM	0.201	0.1639	0.1606	0.1623	0.17	0.2035	1		
SpZIM	0.1681	0.159	0.157	0.1307	0.1325	0.1826	0.1516	1	
TsZIM	0.1707	0.1924	0.1515	0.1595	0.1833	0.1543	0.1782	0.1655	1

Table 3.3. OrthoMCL ortholog groups of LisH domain proteins.

Group	Gene
NO_GROUP	AT1G11110, AT2G25420, Mic63300
OG5_128341	CmaCMP137C, Pp1s208_168V6
OG5_128400	AT1G61150, AT4G09300, Atr00003.330, Cpx9788, Cre01.g029050.t1.1, Csu33734, Cva36609, kfl00052_0320, MA_10429281g0010, MA_10430298g0010, Mic72406, Smo437992, Smo439020, Pp1s268_9V6, Pp1s32_299V6, Pp1s379_20V6, Vc65822
OG5_128648	Pp1s14_43V6
OG5_128821	AT5G67320/HOS15, Bathy13g02710, CmaCML100C, Cre12.g527500.t1.1, Csu53824, Cva48736, kfl00125_0160, MicC20648, Mic103376, Olu26408, Ota8824, Pp1s182_92V6, Smo429030, Smo89330, Smo90375, Smo93124, Vc63837
OG5_128980	Cva133982, kfl00125_0180, MicC47037, Mic55886, Mic78967
OG5_129140	AT5G08560, AT5G43920, CmaCMS241C
OG5_129558	AT1G73720/SMU1, Atr00069.97, Bathy03g00550, Csu53714, kfl00262_0080, Pp1s136_176V6
OG5_130102	Cre10.g446400.t1.2, Mic58000
OG5_130687	Cpx37420
OG5_131556	AT4G31160/DCAF1, Atr00017.53
OG5_132761	Atr00007.236, MA_10436559g0010, Smo405761, Smo430305
OG5_136328	AT2G32700/LUH, AT4G32551/LUG, Atr00067.172, Atr00122.38, Atr00002.372, Atr00037.161, kfl01154_0010, MA_99040g0010, MA_3866g0010, MA_704500g0010, Pp1s45_38V6, Smo112245, Smo410492, Smo79275, Smo94317
OG5_136634	AT1G15750/TPL, AT1G80490/TPR1, AT3G15880/TPR4, AT3G16830/TPR2, AT5G27030/TPR3, Atr00048.113, Atr00051.29, kfl00881_0020, MA_10430083g0010, MA_366520g0010, MA_98943g0010, MA_33469g0010, MA_16979g0010, MA_10436445g0020, MA_97574g0010, MA_118509g0010, Pp1s316_34V6, Pp1s99_260V6, Smo115161, Smo163891, Smo439915, Smo79194, Smo88677
OG5_140055	Pp1s37_113V6
OG5_143134	AT5G57120
OG5_184276	Pp1s316_8V6

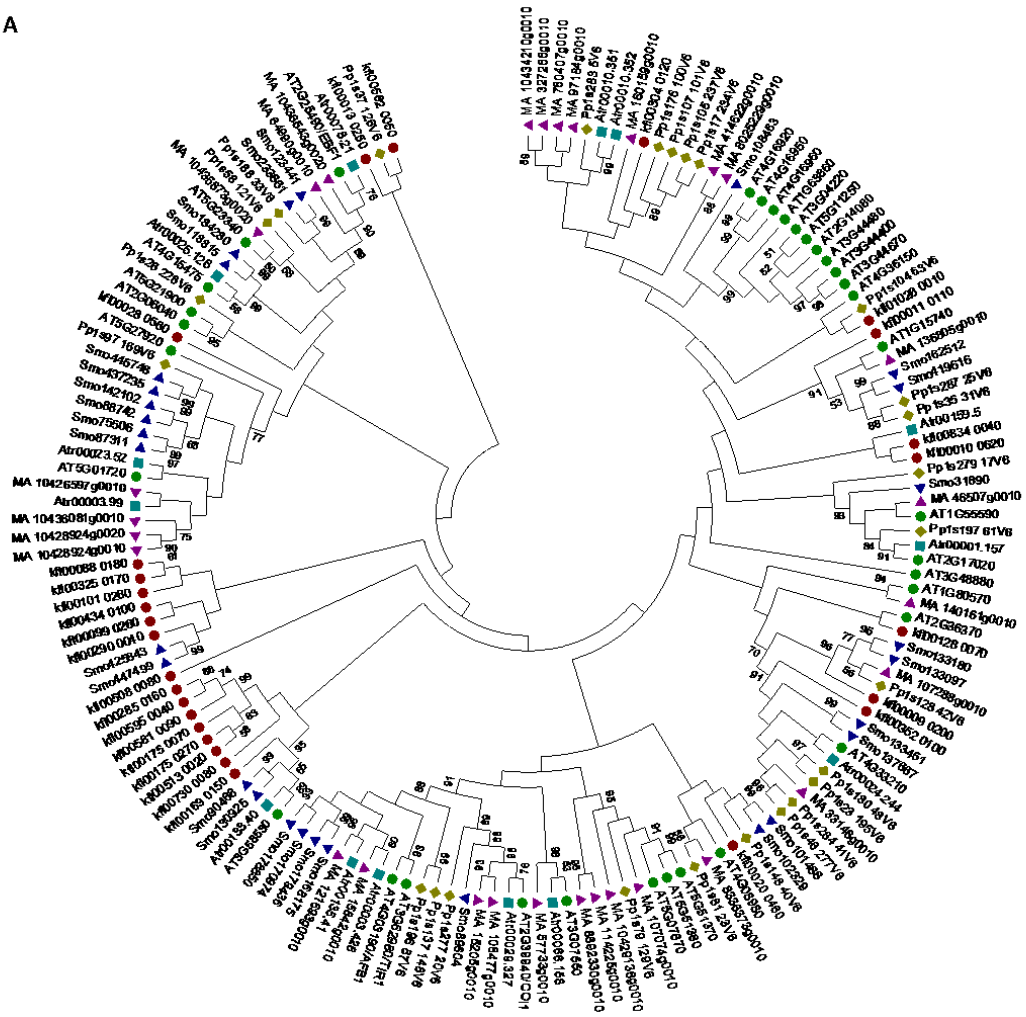
Table 3.4. TIFY family proteins with EAR motifs.

Factor	EAR Motif Start	Length of Protein	Ratio Ear Motif/Protein Length	Type of EAR Motif	Position TIFY	Position CCT
AtJAZ5	269	274	0.98	LDLRL	92-127	182-207
AtJAZ6	264	269	0.98	LELKL	100-135	186-211
AtJAZ7	25	148	0.17	LELRL	54-89	124-144
AtJAZ8	8	131	0.06	LELRL	39-74	107-127
AmJAZ1	228	233	0.98	LGLHL	107-142	177-202
PaJAZ1	131	289	0.45	LDLSLLP	155-197	259-283
PaJAZ4	86	260	0.33	LDLSLLP	138-172	230-254
PaJAZ7	304	328	0.93	LSLGL	138-172	260-284
PaJAZ19	97	322	0.30	LILYL	204-237	270-294
PaJAZ20	54	303	0.18	LNLNL	171-201	252-275
PaJAZ23	75	310	0.24	LRLGL	192-223	262-288
PaTIFY1	253	270	0.94	LHLRL	191-224	---
PaTIFY7	30	158	0.19	LVLKL	78-113	---
PaTIFY12	169	293	0.58	LDLSLLP	207-241	---
PpTIFY5	179	525	0.34	LILKL	315-344	---
PpTIFY6	177	545	0.32	LKLIL	318-347	---
SmJAZ1	102	281	0.36	LTLPL	153-186	221-243
SmJAZ2	101	280	0.36	LTLPL	152-185	220-242
SmJAZ3	138	235	0.59	LLLSL	62-91	172-190
SmJAZ4	138	235	0.59	LLLSL	62-91	172-190
SmJAZ5	211	480	0.44	LDLSLSV	252-283	347-367
	412		0.86	LKLQL		
SmJAZ6	1239	1346	0.92	LSLSLRP	1265-1293	1329-1346
SmJAZ7	211	480	0.44	LDLSLSV	252-283	347-367
	412		0.86	LKLQL		
SmJAZ10	123	441	0.28	LSLSLGP	149-177	213-229
SmPPD3	113	262	0.43	LSLLL	151-178	203-219
	126		0.48	LCLSL		
SpZIM*	172	435	0.40	LQLTL	168-203	236-278

At, Am, Pa, Sm, and Sp indicate *Arabidopsis thaliana*, *Amborella trichopoda*, *Picea abies*, *Selaginella moellendorffii*, and *Spirogyra pratensis*, respectively.

* From a charophyte alga

A



B

Gene	AtCOI1	AtTIR1	At5g27920	Kfl00020	Kfl00028	Kfl00101	Kfl00834
AtCOI1	1						
AtTIR1	0.3649	1					
At5g27920	0.3019	0.3827	1				
Kfl00020	0.2115	0.2072	0.1937	1			
Kfl00028	0.3276	0.2866	0.4218	0.1796	1		
Kfl00101	0.2644	0.2936	0.3825	0.1564	0.3815	1	
Kfl00834	0.2996	0.3567	0.2921	0.2023	0.2698	0.3101	1

Figure 3.1. Phylogenetic analysis of the charophyte and embryophyte AMN1 family. A) The phylogenetic analysis of the AMN1 family was performed using 26 moss, 29 spikemoss, 29 spruce, 14 Amborella, 34 Arabidopsis, and 27 *Klebsormidium flaccidum*. The unrooted maximum likelihood tree was constructed with MEGA5.2 using the ClustalW alignment, Dayhoff plus frequency model, uniform rates among sites, 100 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange heuristic method with a strong branch swap filter starting from an initial neighborhood joining tree. B) The structural similarity of AtCOI1 was compared to select Arabidopsis and *K. flaccidum* AMN1 domain proteins. TM-scores are indicated for each combination of modeled proteins.

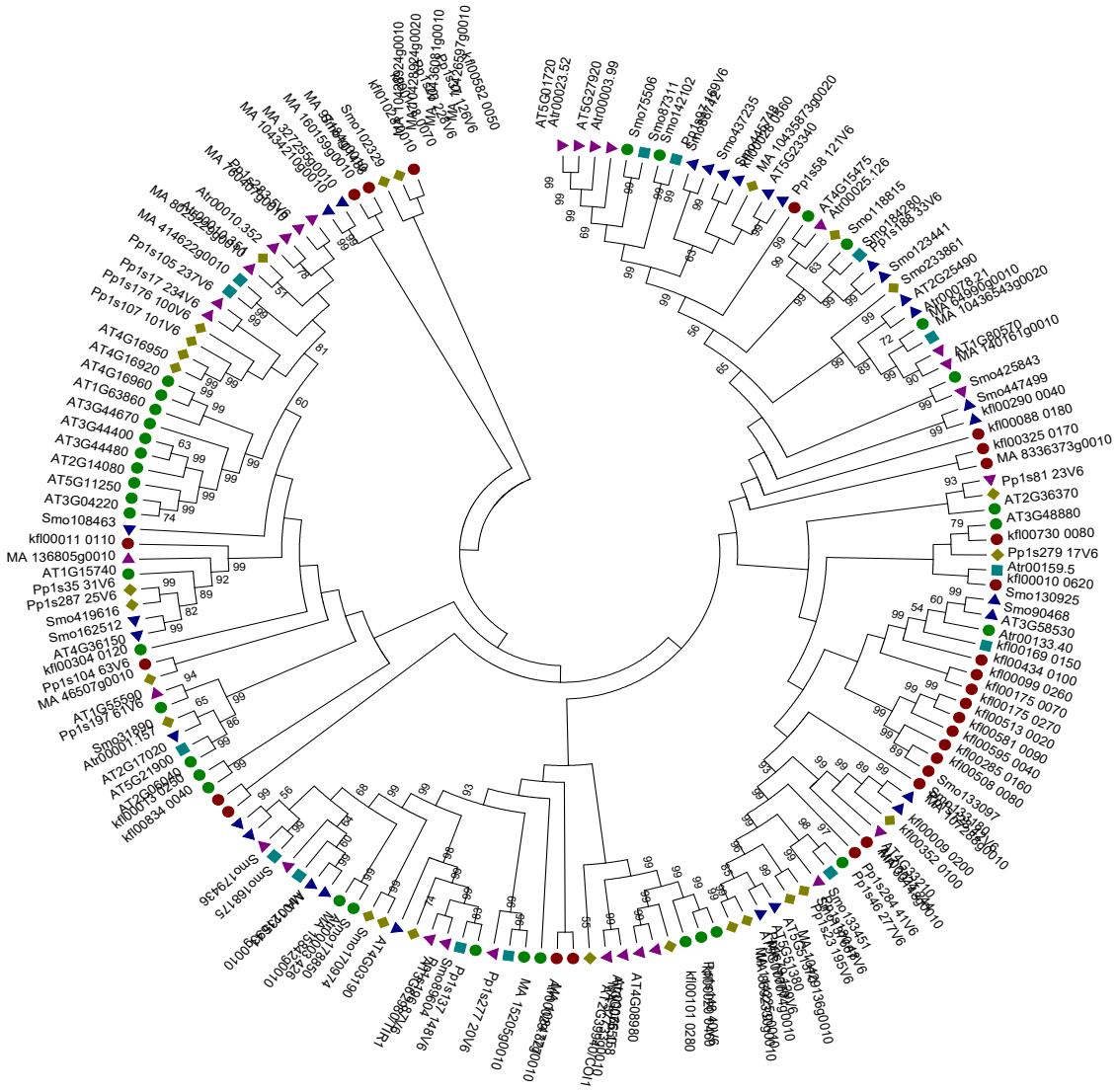


Figure 3.2. Minimum evolution tree of the AMN1 family. The phylogenetic analysis of the AMN1 family was performed using 14 Amborella, 34 Arabidopsis, 29 spruce, 26 moss, 29 spikemosee, and 27 *Klebsormidium flaccidum* members. Alignments were generated with the ClustalW algorithm. The unrooted minimum evolution was generated with MEGA5.2. The tree was constructed with the p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences.

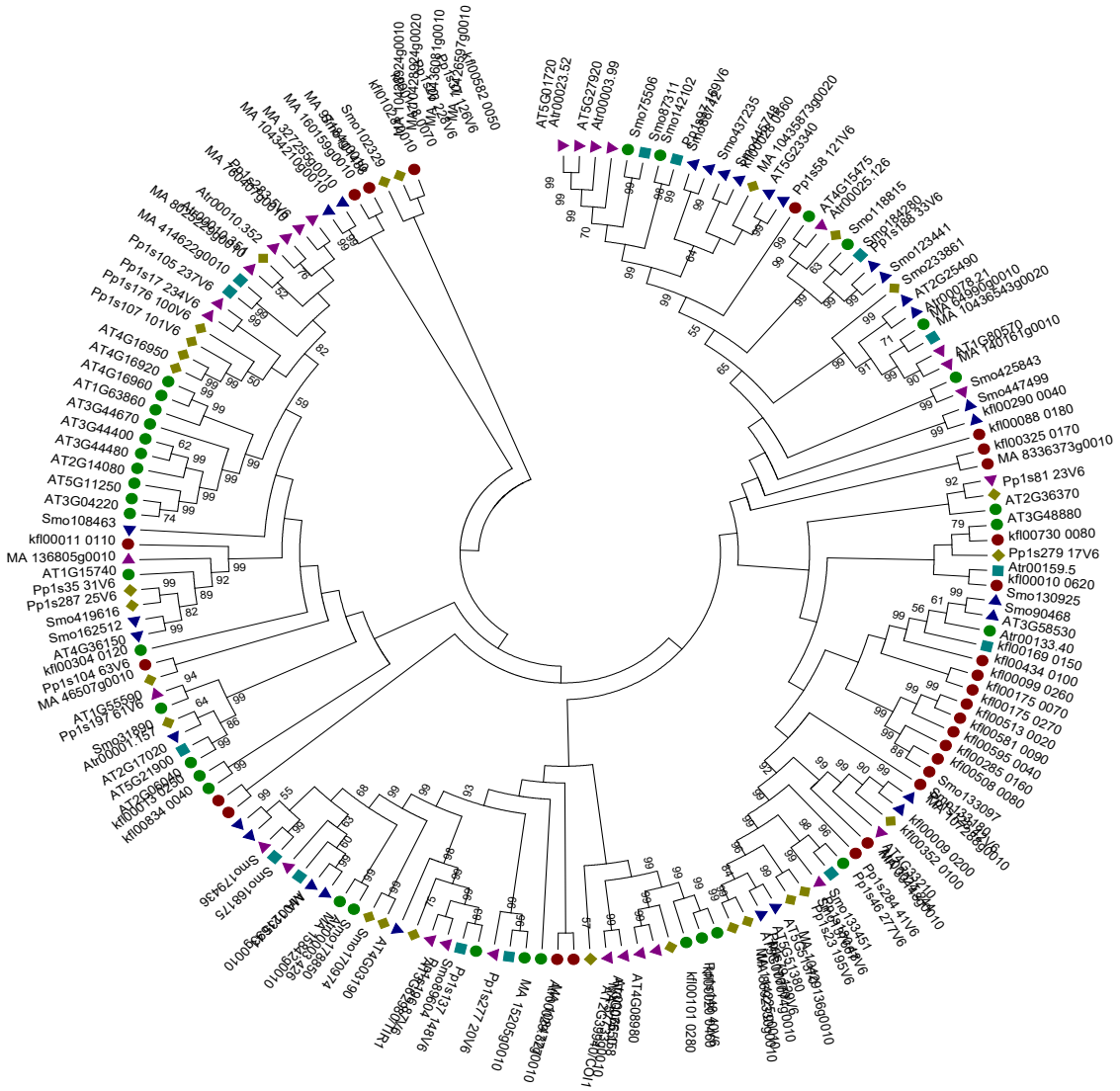


Figure 3.3. Neighbor-joining tree of the AMN1 family. The phylogenetic analysis of the AMN1 family was performed using 14 Amborella, 34 Arabidopsis, 29 spruce, 26 moss, 29 spikemosee, and 27 *Klebsormidium flaccidum* members. Alignments were generated with the ClustalW algorithm. The unrooted neighbor-joining tree was generated with MEGA5.2. The tree was constructed with the p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences.

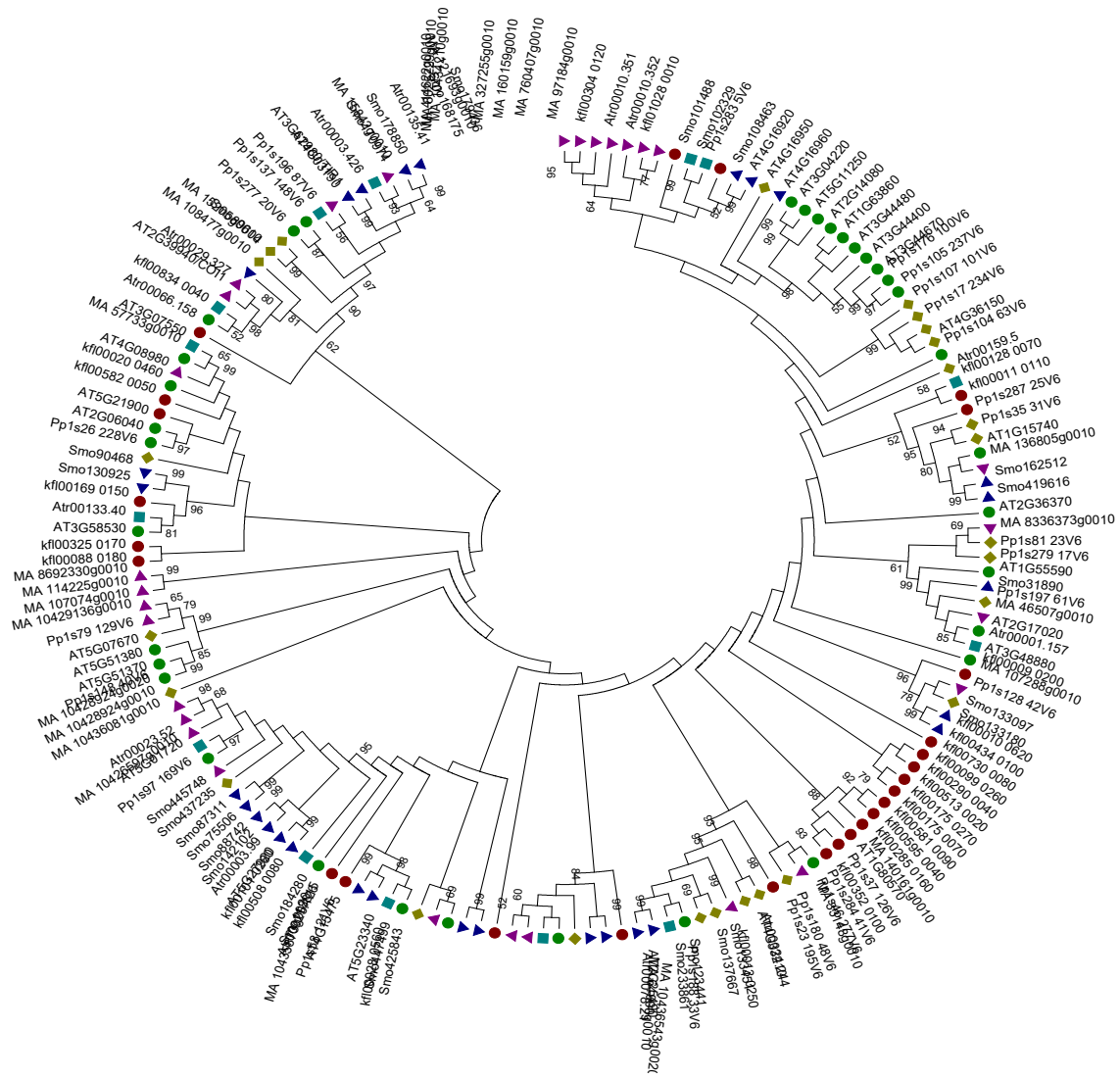


Figure 3.4. Maximum parsimony tree of the AMN1 family. The phylogenetic analysis of the AMN1 family was performed using 14 Amborella, 34 Arabidopsis, 29 spruce, 26 moss, 29 spikemosee, and 27 *Klebsormidium flaccidum* members. Alignments were generated with the ClustalW algorithm. The unrooted maximum parsimony was generated with MEGA5.2. The tree was constructed using the subtree-pruning-regrafting algorithm, 100 bootstrap replications, and pairwise deletion of gaps or missing sequences.

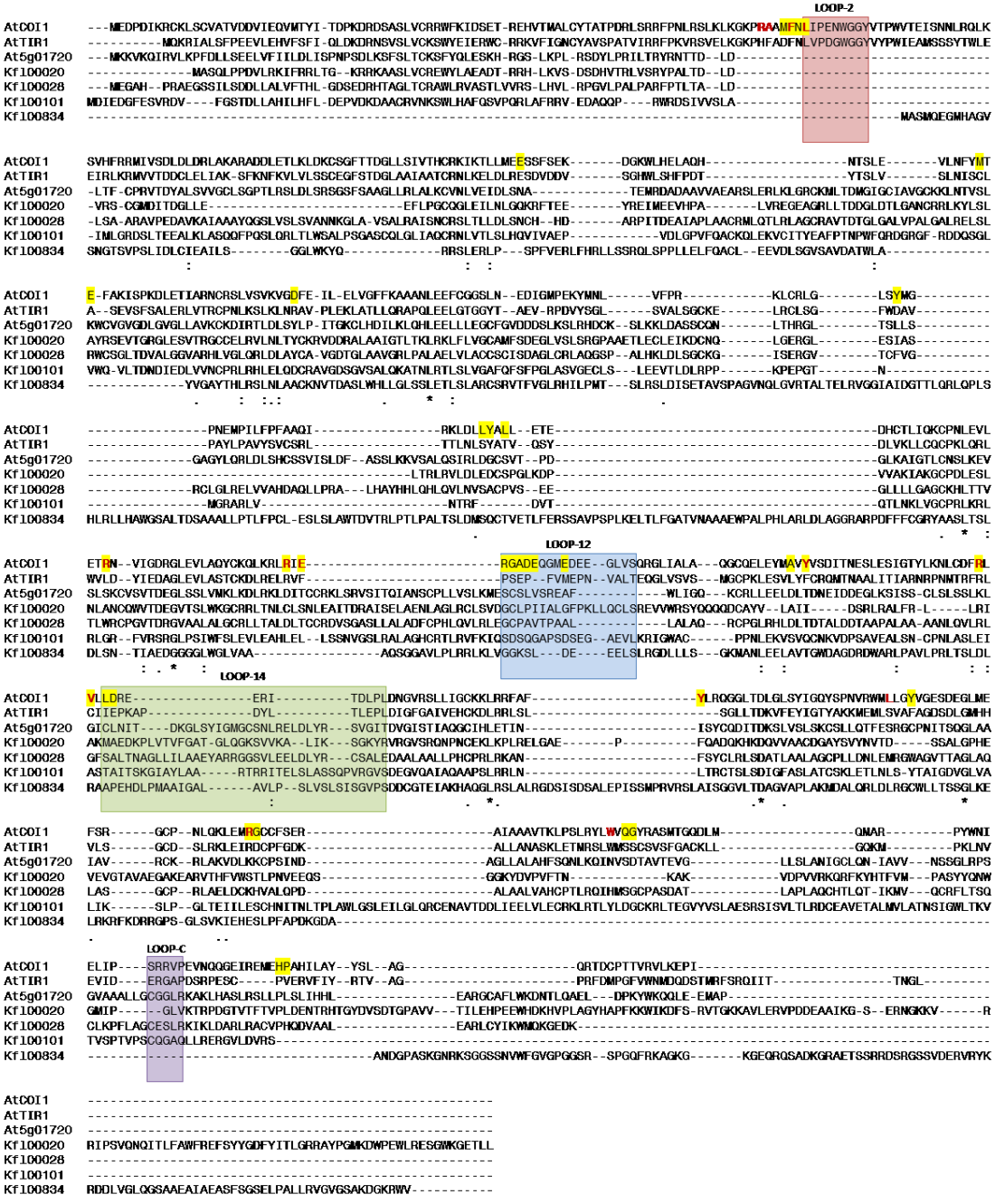


Figure 3.5. Alignment of putative KfCOI1 proteins. The amino acid sequence alignment of AtCOI1, AtTIR1, At5g01720, kfl100020, kfl00028, kfl00101, and kfl000834 is shown with conserved COI1 loop-2, loop-12, loop-14, and loop-C in red, blue, green and purple boxes, respectively. Residues interacting with JAZ proteins are highlighted and those binding jasmonate-isoleucine are indicated in bold red letters.

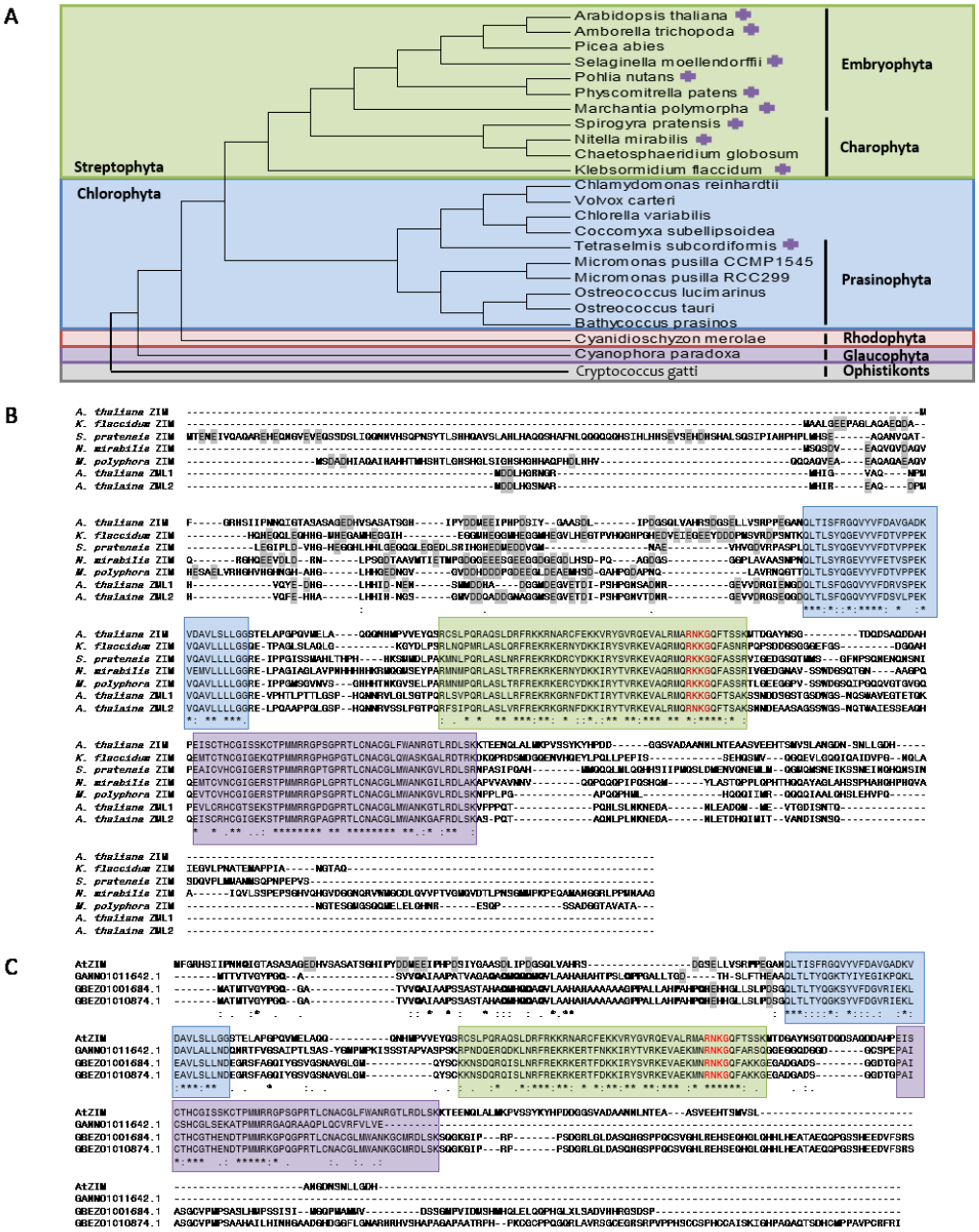


Figure 3.6. Relationship and alignment of charophyte and embryophyte TIFY factors. A) The species relationships of viridiplantae used in this study. Positive signs indicate species with a ZIM factor. B) The alignment of ZIM proteins from Arabidopsis, *Klebsormidium flaccidum*, *Marchantia polymorpha*, *Nitella mirabilis* and *Spirogyra pratensis* indicates the conserved TIFY (blue box), CCT (green box), and Zn-GATA domains (purple box). The putative nuclear localization signal sequences are in red. Acidic residues of the N-terminal acidic region are highlighted in gray. The ZIM from *Chaetosphaeridium globosum* was excluded because the sequence is incomplete. C) The alignment of *A. thaliana* ZIM and Chlorophyta proteins indicates the conserved TIFY (blue box), CCT (green box), and Zn-GATA domains (purple box). The putative nuclear localization signal sequences are in red. Acidic residues of the N-terminal acidic region residues are highlighted in gray. The proline-glutamine rich region residues are in bold.

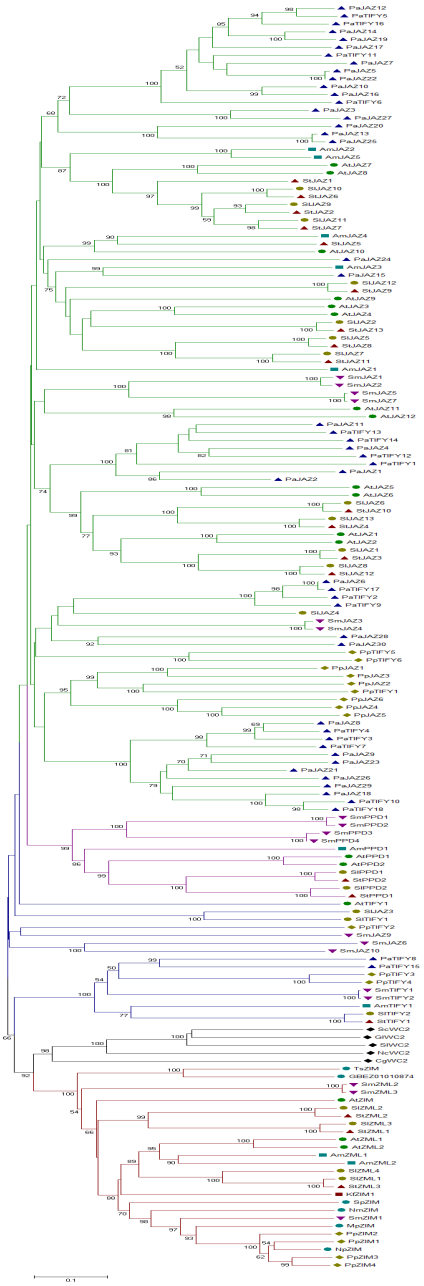


Figure 3.7. Phylogenetic identification of algal TIFY proteins as ZIM TFs. The phylogenetic analysis of the TIFY family members was performed using 9 Amborella, 18 Arabidopsis, 16 Oryza sativa, 48 spruce, 16 moss, 19 spikemoss, 21 *Solanum lycopersicum*, 19 *Solanum tuberosum*,*1 Klebsormidium flaccidum*, 1 *Chaetosphaeridium globosum*, 1 *Marchantia polymorpha*, 1 *Nitella mirabilis*, 1 *Pohlia nutans*, 1 *Spirogyra pratensis*, and 1 *Tetraselmis subcordiformis* sequences. Five WC2 proteins from fungi (*Cryptococcus gattii*, *Ganoderma lucidum*, *Neurospora crassa*, *Schizophyllum commune*, and *Serpula lacrymans*) were also included. The minimum evolution tree was constructed with MEGA5.2 using the ClustalW alignment, p-distance method, 1000 bootstrap replications, and pairwise deletion of gaps or missing sequences. Branches are colored according to factor type: green (JAZ), purple (PPD), blue (TIFY), red (ZIM) and black (non-ZIM ZnF-GATA).

A

Accession No.	Annotation	E-value
XP_006966680.1	blue light regulator 2 [Trichoderma reesei QM6a]	2.00E-07
KGB78050.1	cutinase palindrome-binding protein [Cryptococcus gattii R265]	2.00E-07
KIR34516.1	white collar 2 protein [Cryptococcus gattii MMRL2647]	2.00E-07
EPB82531.1	hypothetical protein [Mucor circinelloides f. circinelloides 1006PhL]	2.00E-07
KIR85749.1	white collar 2 protein [Cryptococcus gattii IND107]	3.00E-07
XP_003194137.1	hypothetical protein Cryptococcus gattii WM276]	3.00E-07
CAQ77079.1	putative white collar 2 protein [Phycomyces blakesleeanus]	4.00E-07
GAN04588.1	cutinase gene palindrome-binding protein [Mucor ambiguus]	4.00E-07
ACT46736.1	white collar-2 [Phaeosphaeria avenaria f. sp. tritici]	4.00E-07
ACT46737.1	white collar-2 [Phaeosphaeria avenaria f. sp. tritici]	4.00E-07

B

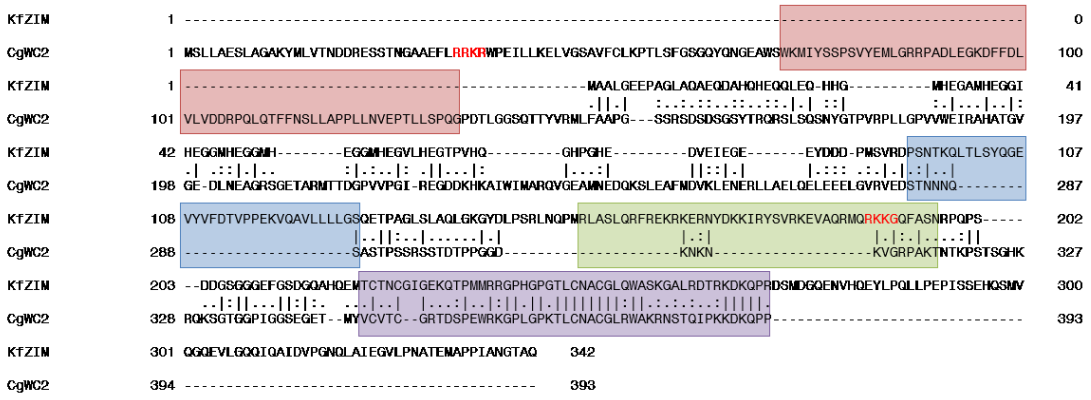


Figure 3.8. KfZIM non-plant BLAST hits and alignment. A) The top 10 BLAST hits, excluding Viridiplantae sequences, to the KfZIM protein. B) To identify conserved regions the KfZIM protein was compared to the fungus *Cryptococcus gattii* WC2 sequence. The pairwise sequence alignment was performed using EMBOSS stretcher. The PAS, TIFY, CCT, and GATA-ZnF domains are boxed in red, blue, green, and purple, respectively.

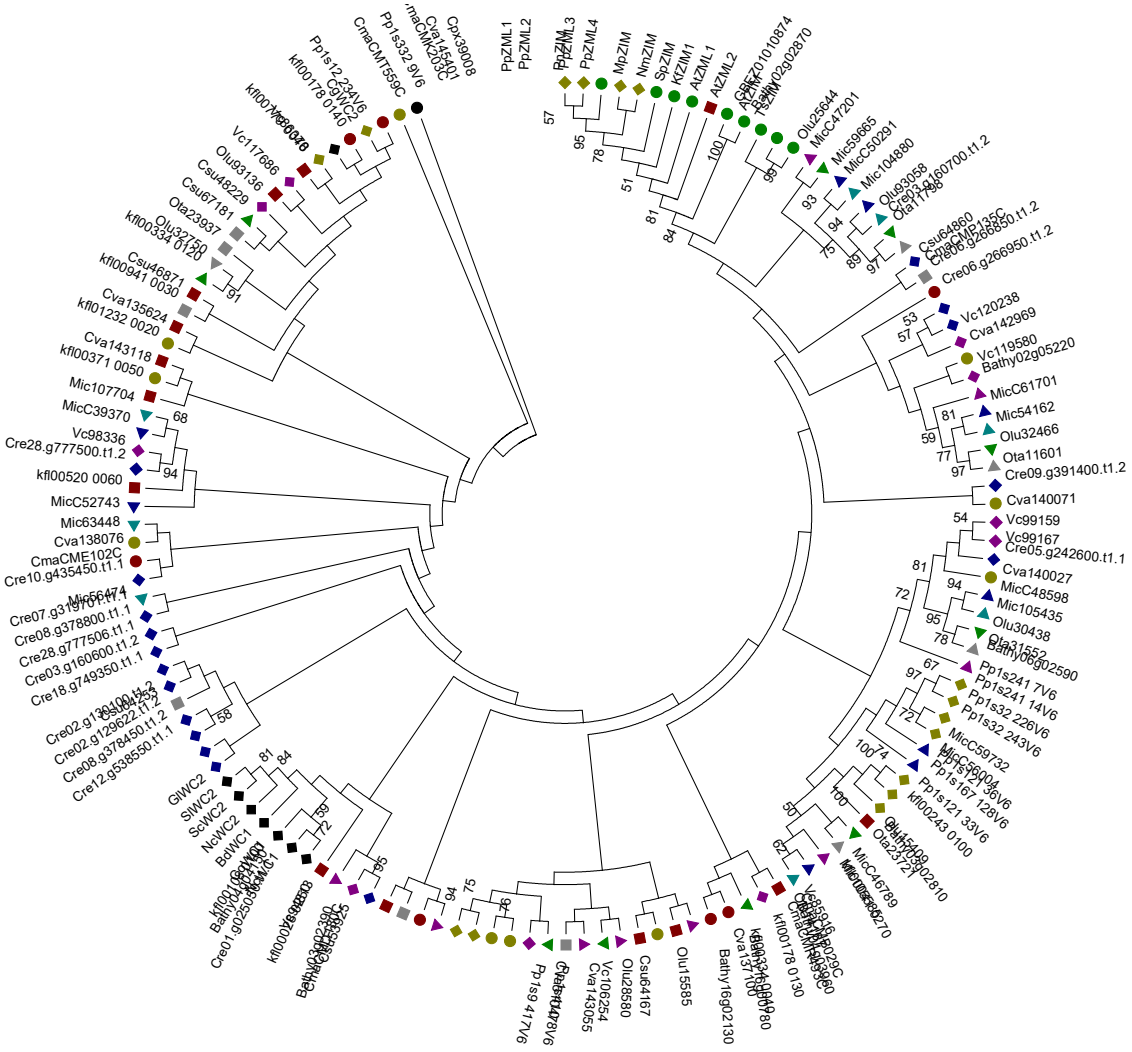


Figure 3.9. Ancestral algal TIFY factors belong to the ZIM GATA TF family. The phylogenetic analysis for the GATA TFs was performed using15 moss, 14 *K. flaccidum*, 9 *B. prasinos*, 17 *C. reinhardtii*, 11 *C. variabilis*, 7 *C. subellipsoidea*, 9 *M. pusilla* CCMP1545, 8 *M. pusilla* RCC299, 10 *O. lucimarinus*, 5 *O. tauri*, 10 *V. carteri*, 7 *C. merolae*, and 1 *C. paradoxa*, along with 5 WC2 and 3 WC1 factors. The maximum likelihood tree was constructed with MEGA5.2 using the ClustalW alignment, Equal input model, uniform rates among sites, 100 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange heuristic method with a strong branch swap filter starting from an initial neighborhood joining tree. The tree was rooted with the *C. paradoxa* sequence.

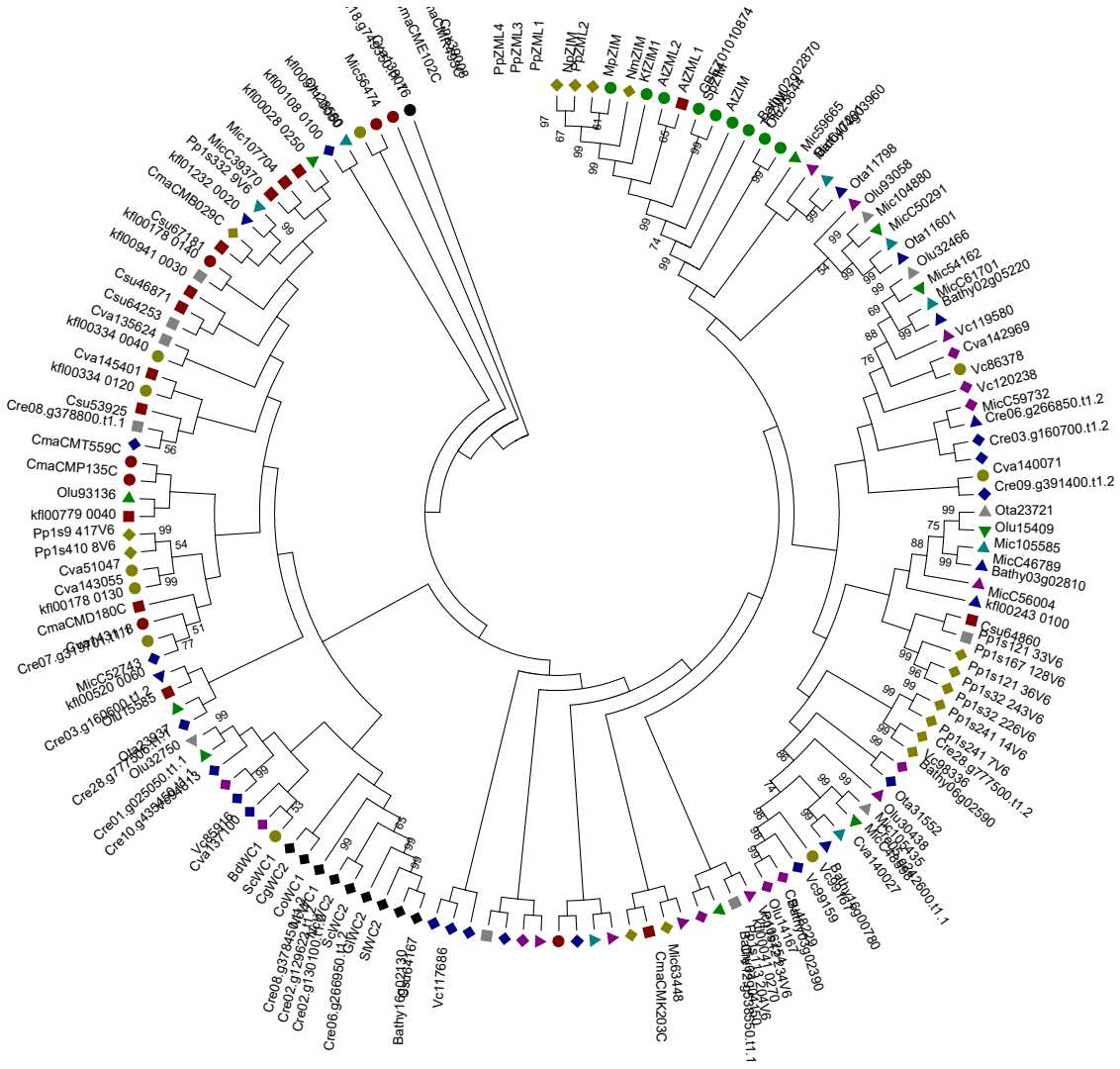


Figure 3.10. Minimum evolution tree of GATA TFs. The phylogenetic analysis of GATA TFs was performed using 16 moss, 14 *K. flaccidum*, 9 *B. prasinos*, 17 *C. reinhardtii*, 11 *C. variabilis*, 7 *C. subellipsoidea*, 9 *M. pusilla* CCMP1545, 8 *M. pusilla* RCC299, 10 *O. lucimarinus*, 5 *O. tauri*, 10 *V. carteri*, 7 *C. merolae*, and 1 *C. paradoxa,* along with 5 WC2 and 4 WC1 factors. The unrooted minimum evolution tree was constructed with MEGA5.2 using the ClustalW alignment, p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange algorithm starting from an initial neighborhood joining tree.

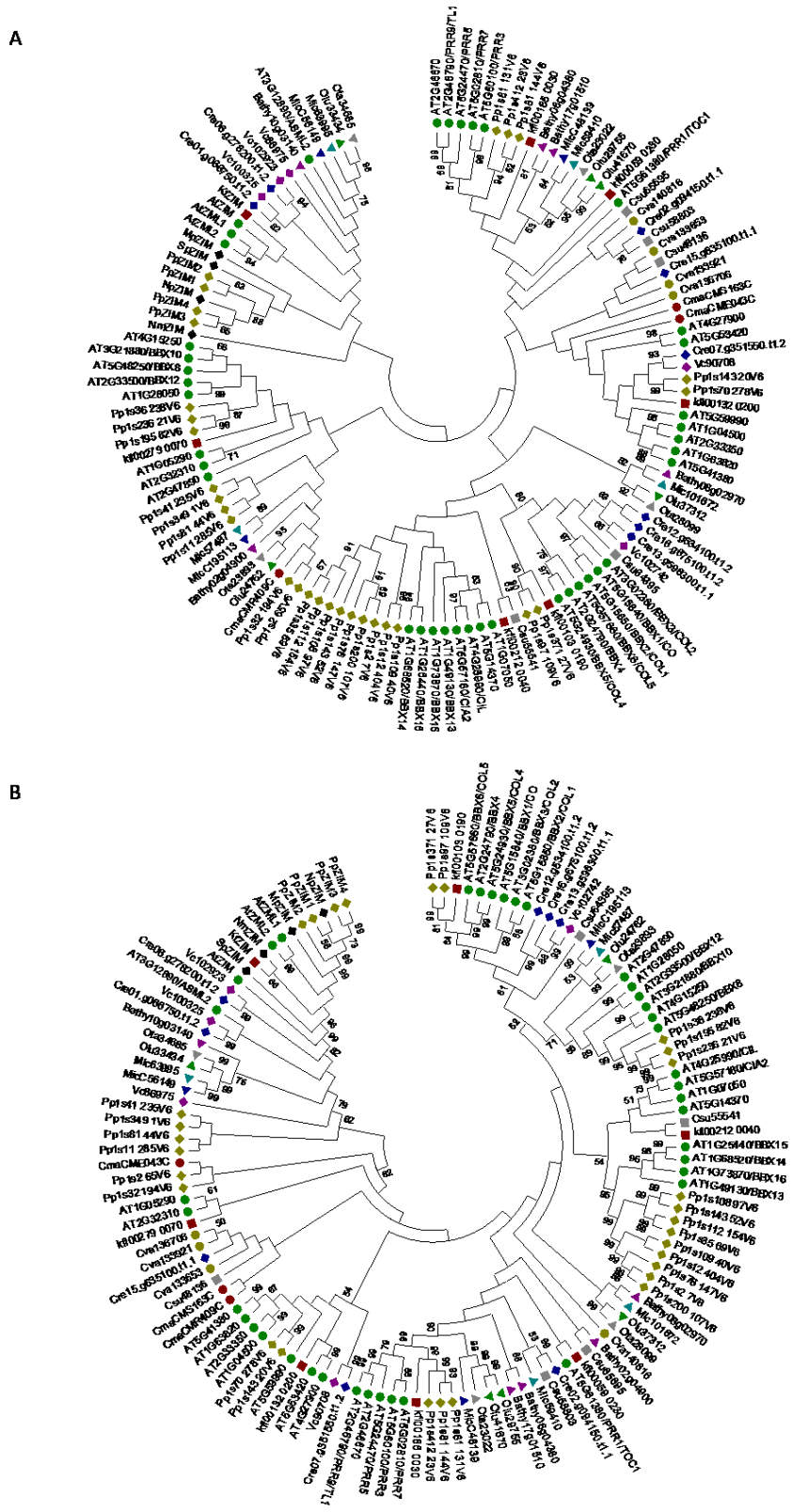


Figure 3.11. Phylogenetic analysis of CCT domain family members. The phylogenetic analysis of CCT members was conducted using 39 Arabidopsis, 29 moss, 7 *K. flaccidum*, 5 *B. prasinos*, 8 *C.*

reinhardtii, 4 *C. variabilis*, 5 *C. subellipsoidea*, 3 *M. pusilla* CCMP1545, 4 *M. pusilla* RCC299, 5 *O. lucimarinus*, 4 *O. tauri*, 5 *V. carteri*, and 3 *C. merolae* sequences. Alignments were performed with the ClustalW algorithm and trees assembled in MEGA5.2. A) The unrooted maximum likelihood tree was constructed with an Equal input model, uniform rates among sites, 100 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange heuristic method with a strong branch swap filter starting from an initial neighborhood joining tree. B) The unrooted minimum evolution tree was constructed with a p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange algorithm starting from an initial neighborhood joining tree.

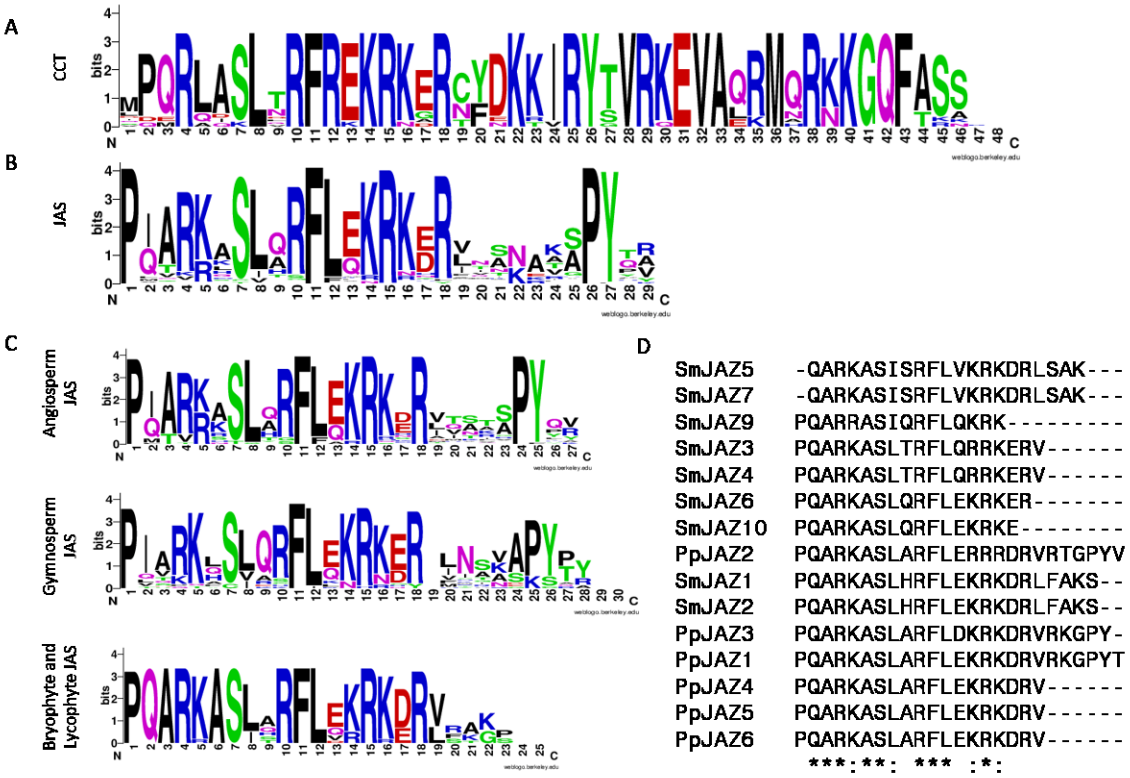


Figure 3.12. Amino acid differences between the CCT and JAS domains. The weblogos showing amino acid residues for A) CCT and B) JAS domains are indicated. C) Weblogos indicate amino acid residue occurrence in JAS domains in early land plants, gymnosperms, and angiosperms. D) The alignment of moss and spikemoss JAZ proteins was performed using Clustal Omega.

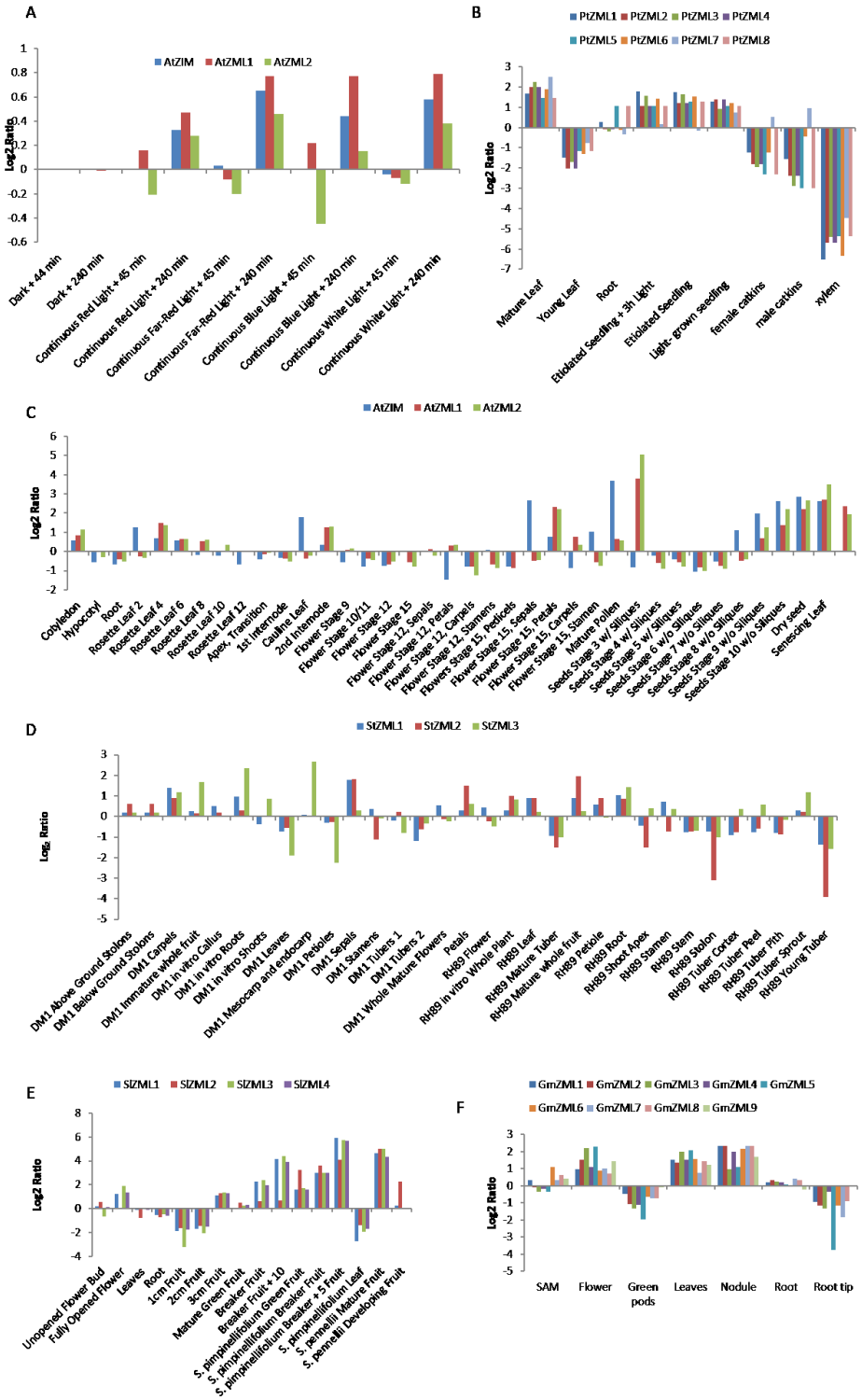


Figure 3.13. ZIM TF expression in response to light and during plant development. A) The expression of *Arabidopsis thaliana* ZIM TFs in response to light are depicted. Expression of ZIMs in B) *Poplar trichopoda*, C) *Arabidopsis thaliana*, D) *Solanum tuberosum*, E) *Solanum lycopersicum*, and F) *Glycine max* in various developmental stages are presented. All expression values were obtained from eFP Browser.

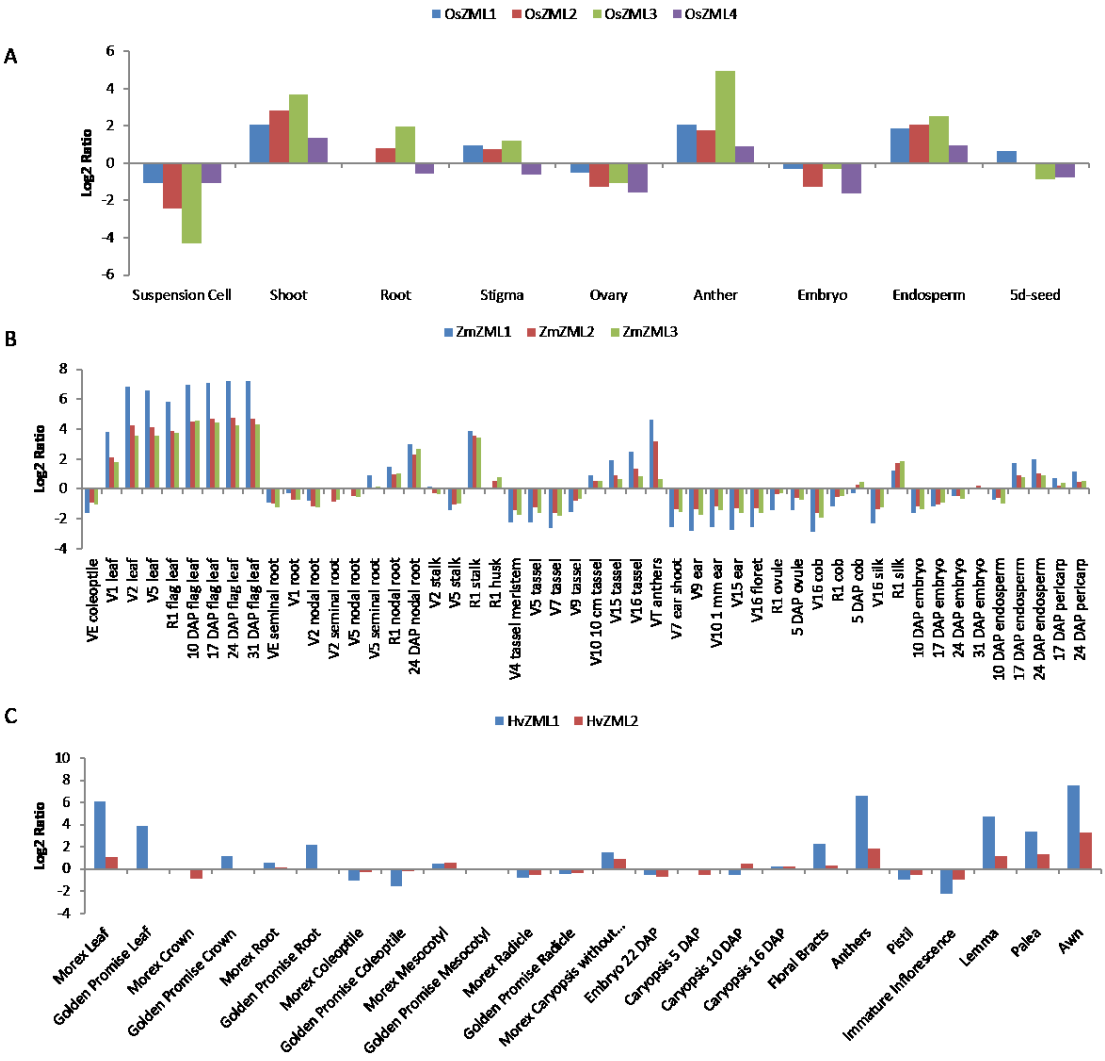


Figure 3.14. ZIM TF expression during monocot plant development. Expression of ZIMs in A) *Zea mays* and B) *Hordeum vulgare* in various developmental stages are presented. All expression values were obtained from eFP Browser.

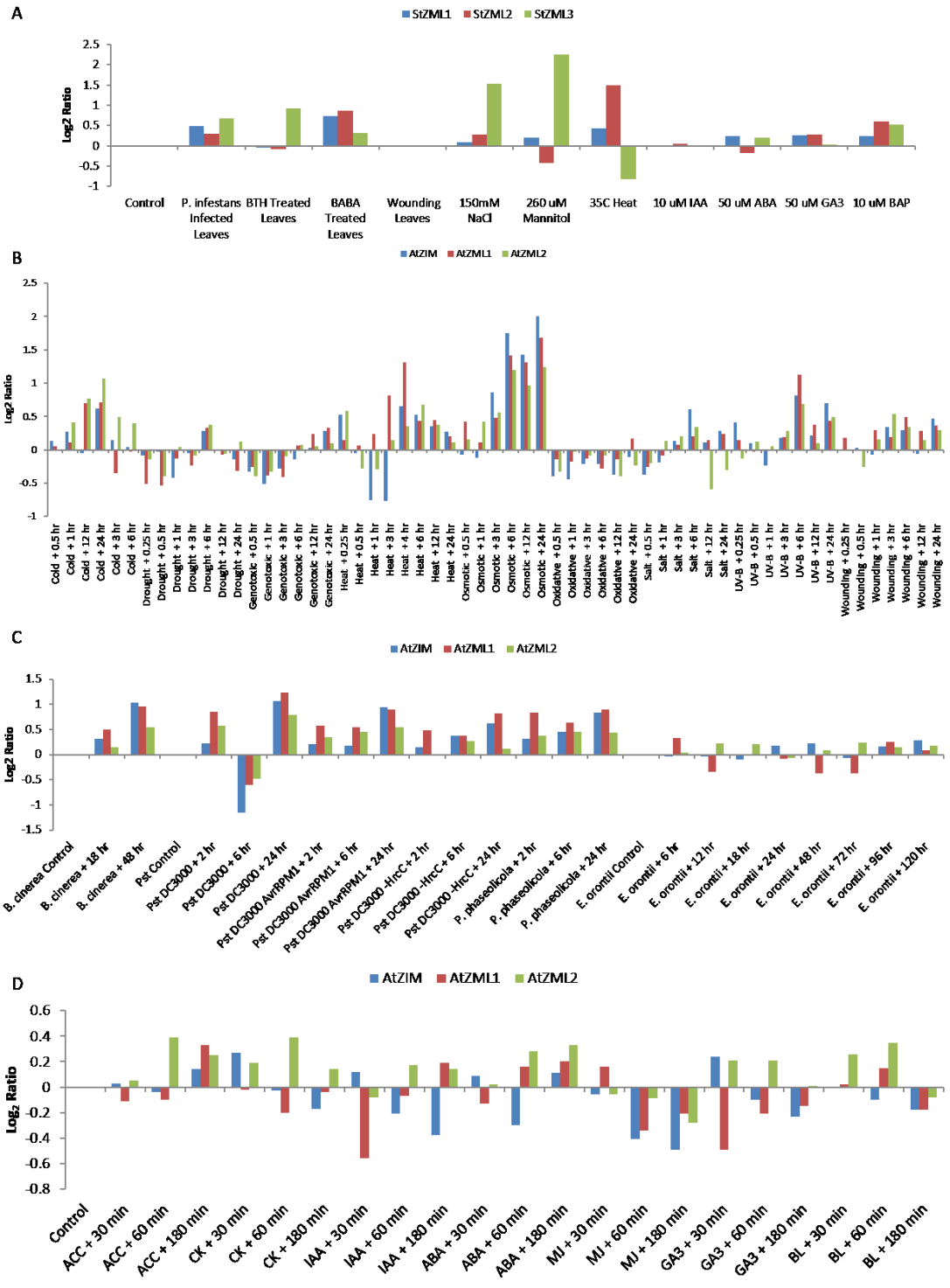


Figure 3.15. Expression of ZIM TFs in response to hormone treatment and stress. A) Gene expression data were obtained for potato in response to biotic and abiotic stress. The expression of Arabidopsis ZIM factors in response to B) abiotic stress, C) hormone treatment, and D) and pathogen infection are shown. All gene expression data were obtained from eFP Browser.

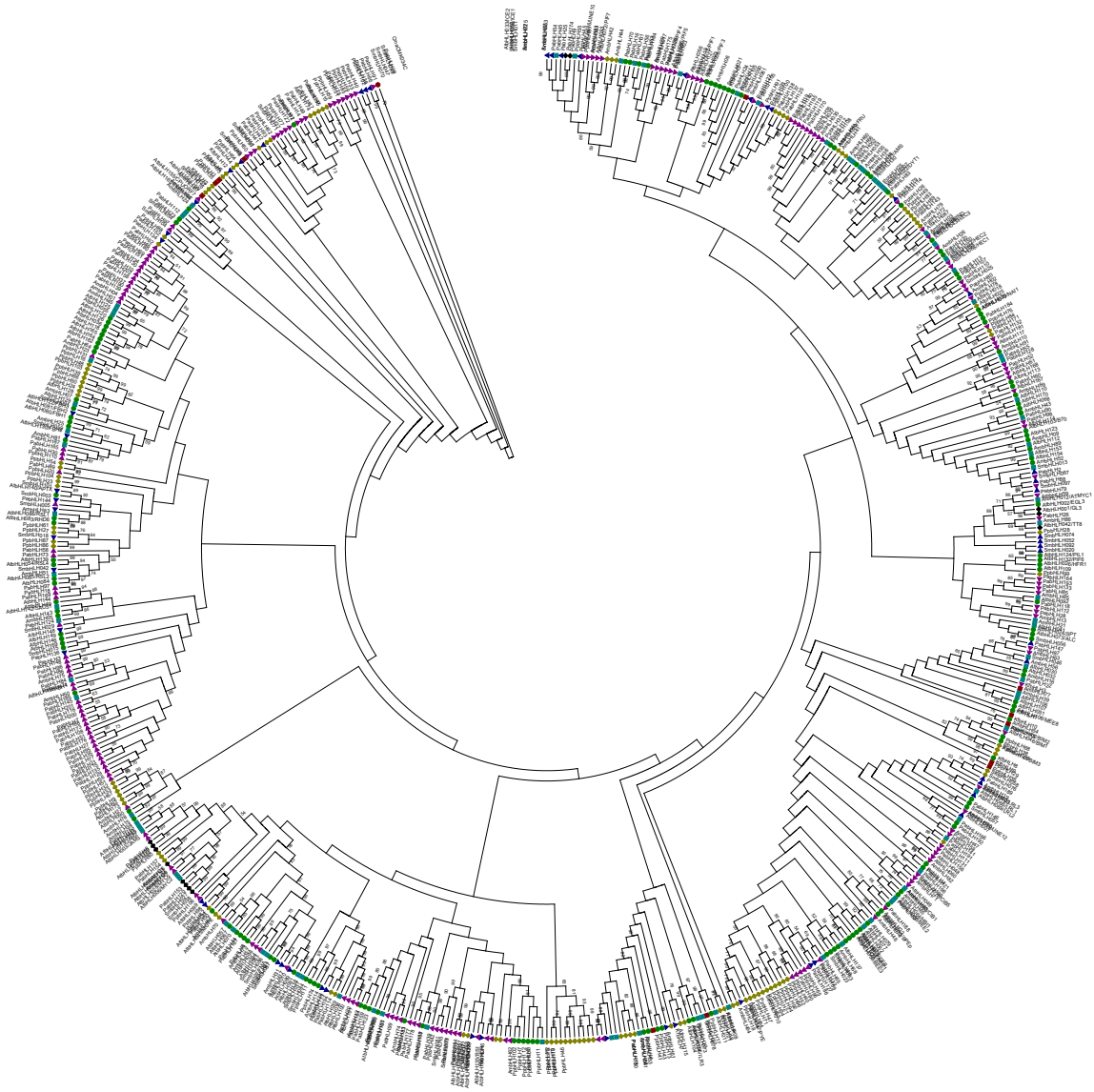


Figure 3.16. Minimum evolution tree of the Viridiplantae bHLH family. The bHLH TF minimum evolution phylogenetic tree was constructed using 1 *C. merolae*, 12 *K. flaccidum*, 113 moss, 56 spikemoss, 195 spruce, 91 Amborella, and 162 Arabidopsis members. The bHLH family tree was generated with MEGA5.2 using the ClustalW alignment, p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange algorithm starting from an initial neighborhood joining tree. The tree was rooted with the outgroup *C. merolae* sequence**.**

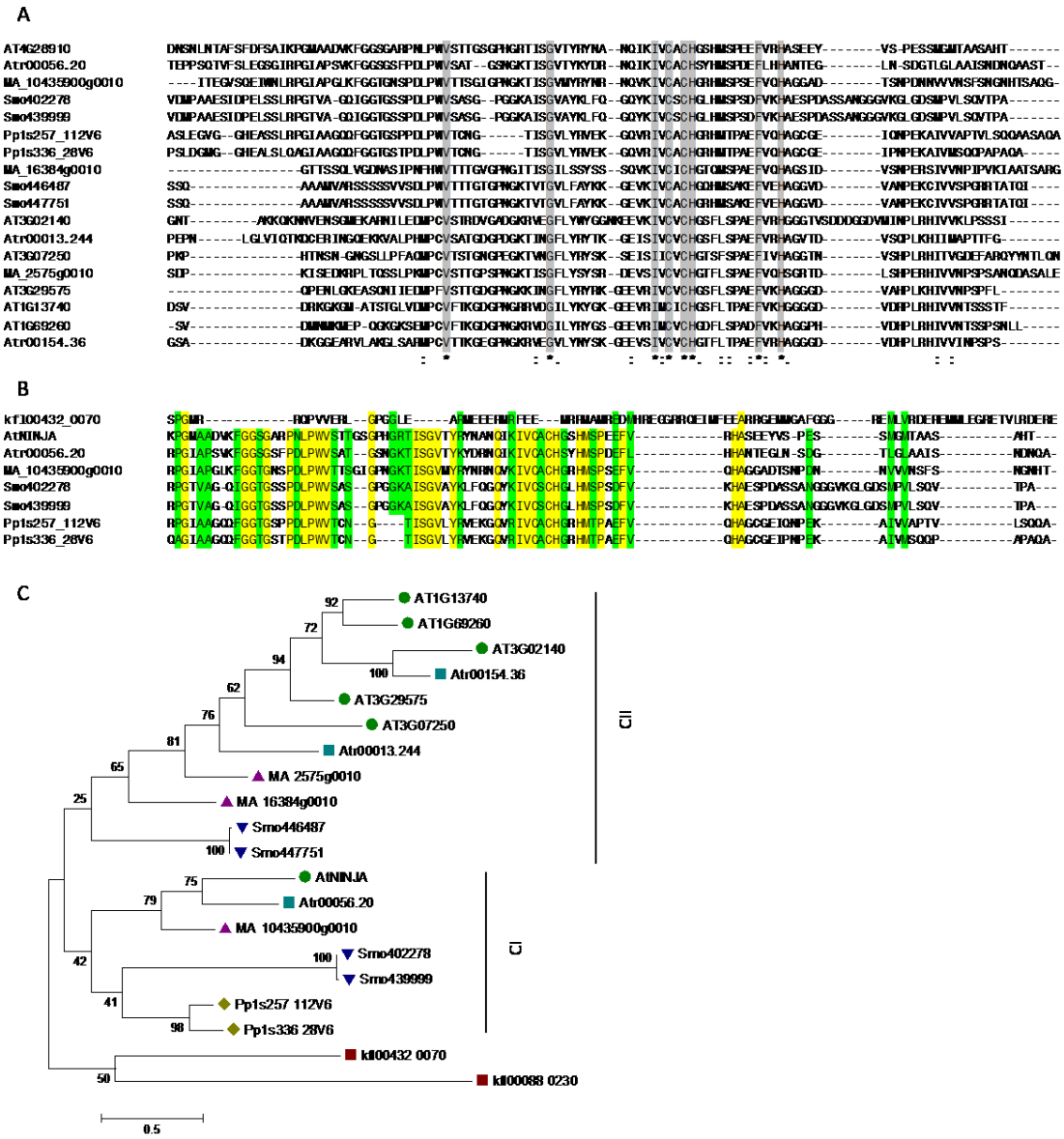


Figure 3.17. Alignment and phylogenetic analysis of the VICH domain family. A) The alignment of 18 proteins revealed conserved residues of the VICH domain family. Conserved residues are highlighted in gray. B) Kfl00332_0070 was compared to various NINJA proteins from land plants by sequence alignment. Conserved and similar residues are highlighted in yellow and green, respectively. C) The unrooted maximum likelihood tree for VICH domain proteins was constructed with MEGA5.2 using the ClustalW alignment, Equal input model, uniform rates among sites, 100 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange heuristic method with a strong branch swap filter starting from an initial neighborhood joining tree.

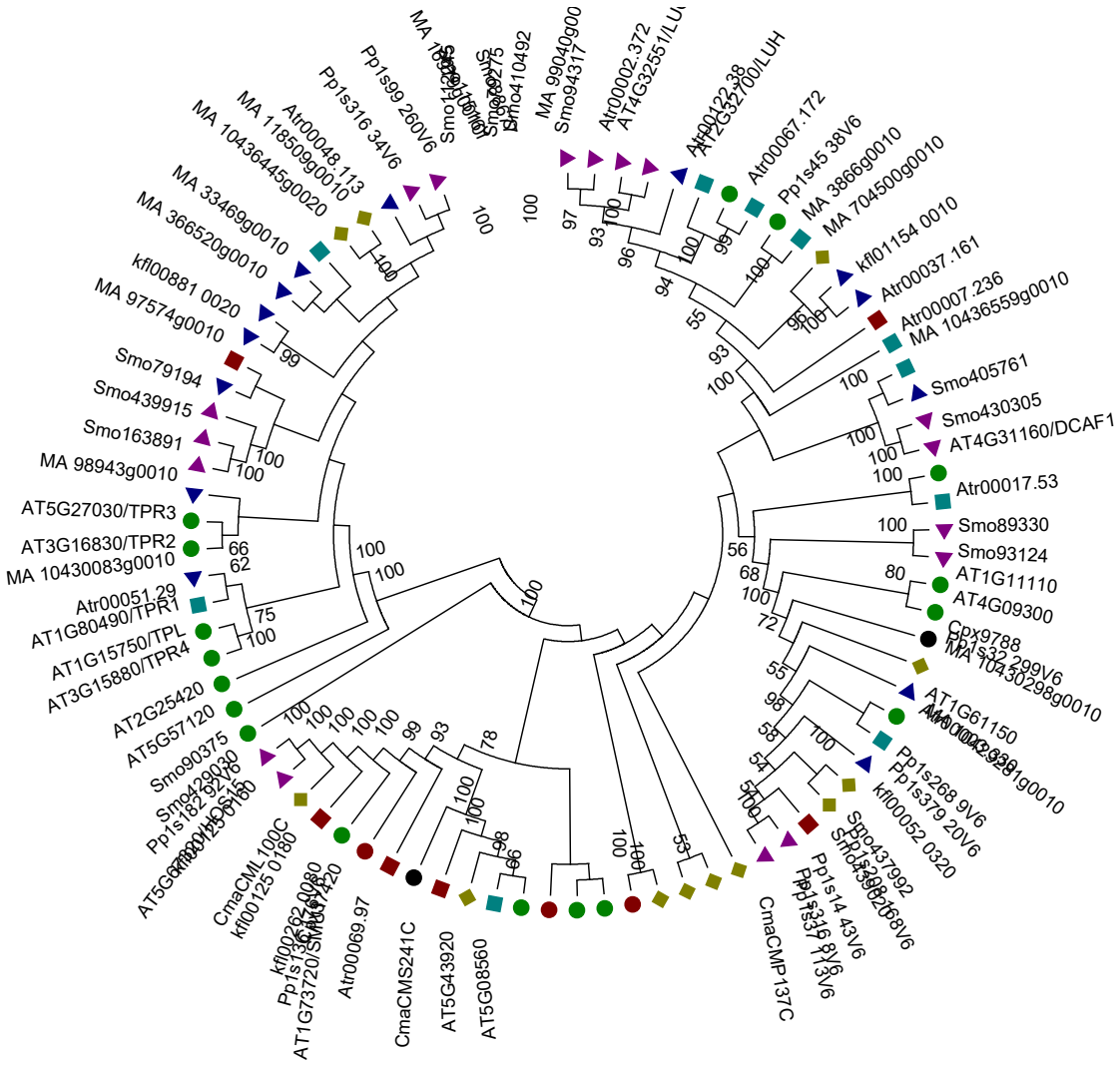


Figure 3.18. Phylogenetic analysis of the Archaeplastida LisH family. The minimum evolution phylogenetic tree of LisH domain containing proteins was constructed using 2 *C. paradoxa*, 3 *C. merolae*, 6 *K. flaccidum*, 12 moss, 17 spikemoss, 14 spruce, 10 Amborella, and 17 Arabidopsis sequences. The LisH family was constructed with MEGA5.2 using the ClustalW alignment, p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange algorithm starting from an initial neighborhood joining tree. The tree was not rooted.

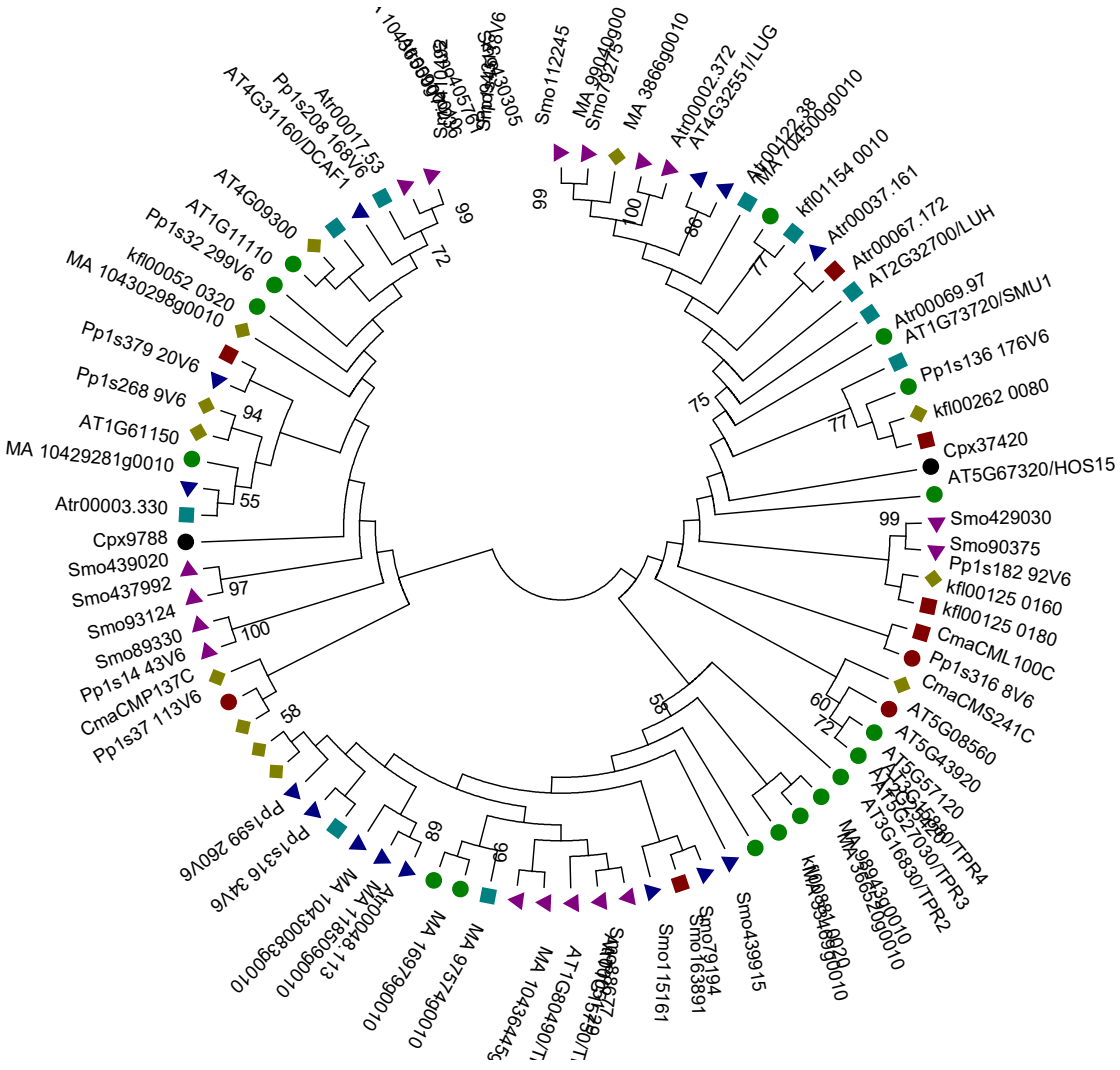


Figure 3.19. Maximum likelihood tree of LisH domain proteins. The maximum likelihood phylogenetic tree of LisH domain containing proteins was constructed using 2 *C. paradoxa*, 3 *C. merolae*, 6 *K. flaccidum*, 12 moss, 17 spikemoss, 14 spruce, 10 Amborella, and 17 sequences. The unrooted maximum likelihood tree was constructed with MEGA5.2 using the ClustalW alignment, Equal input model, uniform rates among sites, 100 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange heuristic method with a strong branch swap filter starting from an initial neighborhood joining tree.

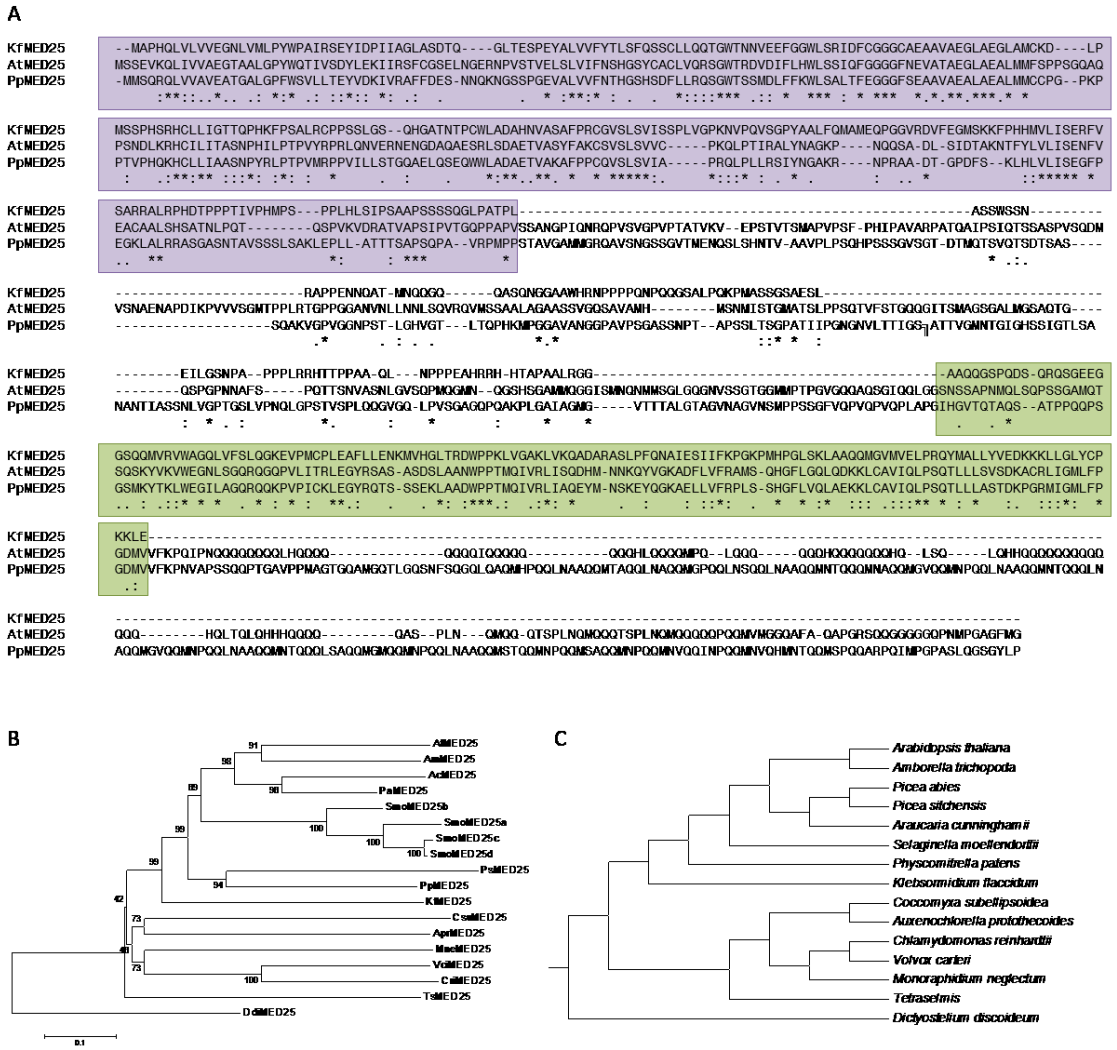


Figure 3.20. Alignment and phylogenetic analysis of MED25 in the Viridiplantae. A) The pairwise protein sequences were aligned for MED25 subunits from *K. flaccidum*, moss, and Arabidopsis. The von Willebrand and ACID domains are indicated in purple and green, respectively. B) The minimum evolution tree of the MED25 family was constructed with MEGA5.2 using the ClustalW alignment, p-distance method, 1000 bootstrap replications, pairwise deletion of gaps or missing sequences, and a nearest-neighbor interchange algorithm starting from an initial neighborhood joining tree. The tree was rooted with the outgroup *Dictyostelium discoideum* sequence. C) The phylogenetic tree of species was constructed with organisms used for assembly of the MED25 family tree.

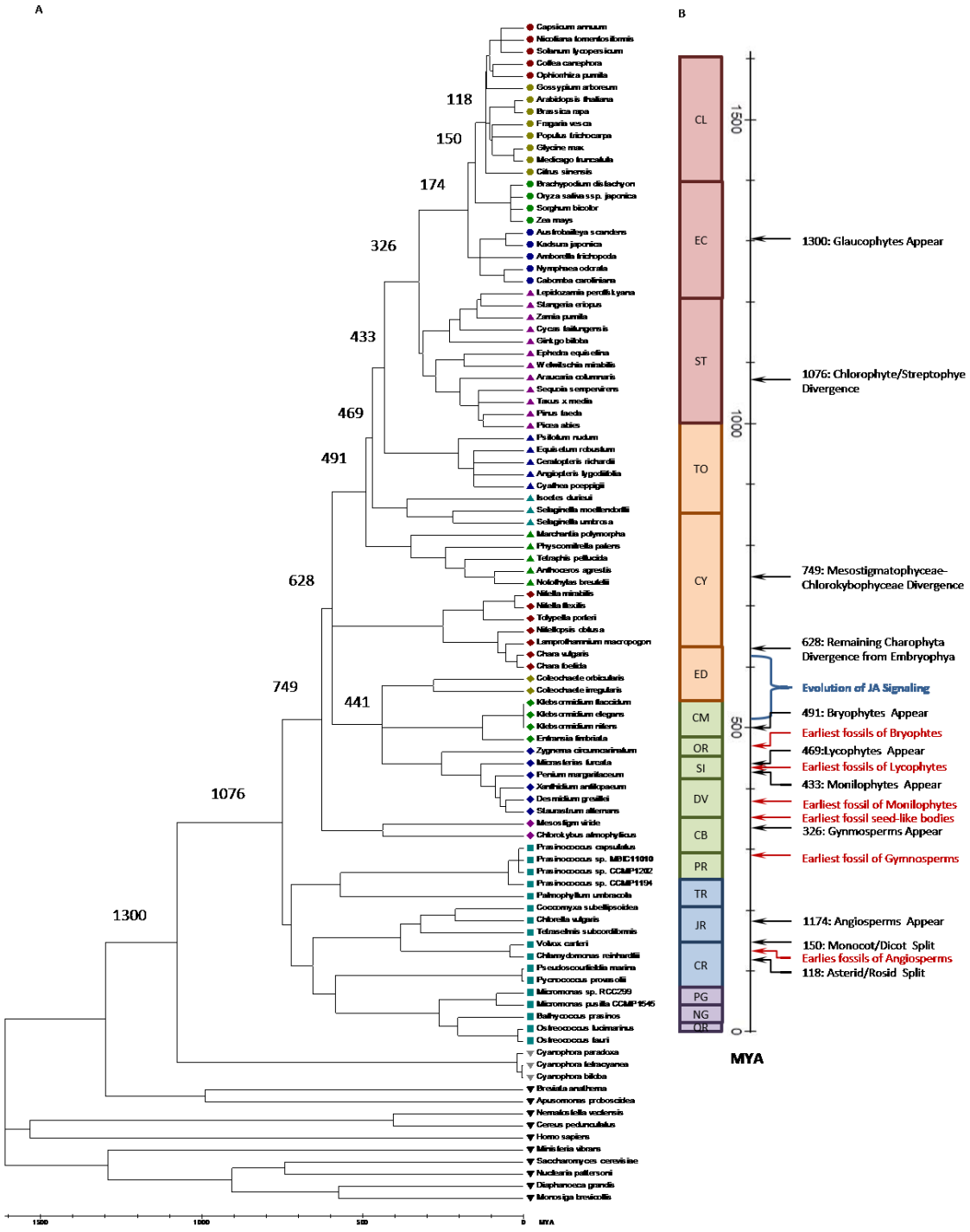
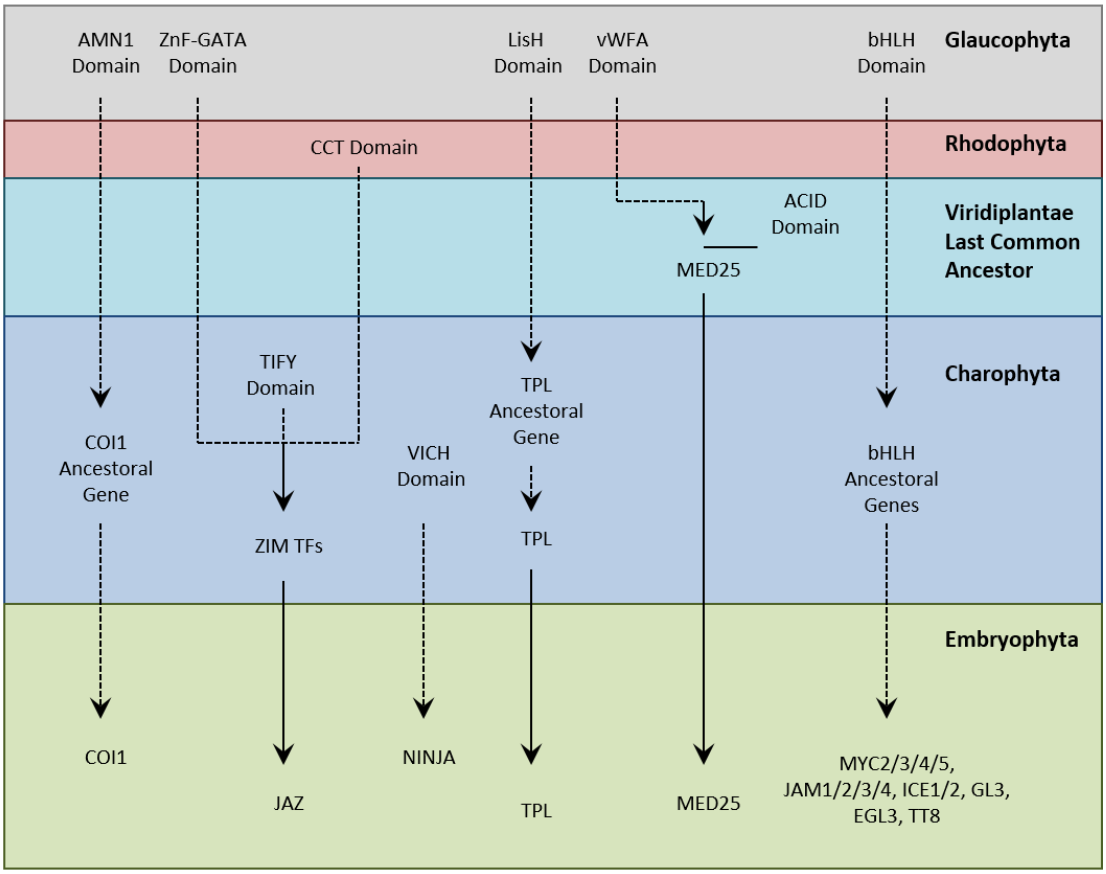


Figure 3.21. Molecular clock analysis and divergence times within the Viridiplantae. A) Sequences of rRNA from 98 species were used to construct a maximum likelihood tree for the molecular clock analysis. The tree was rooted with uni- and opisthokont sequences. The initial minimum evolution tree was generated with a maximum composite likelihood model with transition and transversion substitutions, a gamma distribution of 0.45 with homogeneous patterns among lineages, pairwise deletion of gaps or missing sequences, and a nearest-neighbor

interchange algorithm starting from an initial neighborhood joining tree. Divergence times were estimated using a maximum likelihood statistics method and general time reversible model with a gamma distribution of 5, partial deletion with a 95% site coverage filter cutoff, and a moderate branch swap filter. A divergence time of 150 million years ago (MYA) between monocots and dicots was used to date the tree. Uni- and opisthokont sequences, glaucophyte, chlorophyte, charophyte, non-angiosperm embryophytes, and angiosperms are indicated with black inverted triangles, gray inverted triangles, squares, diamonds, triangles, and circles, respectively. Within charophytes, Charophyceae, Coleochaetophyceae, Klebsormidiophyceae, Zygnematophyceae, and Mesostigmatophyceae/Chlorokybophyceae are represented by red, yellow, green, blue, and purple colors, respectively. Bryophytes, lycophytes, pterophytes, and gymnosperms are represented by green, teal, blue, and purple triangles, respectively. The scale is in MYA. Key divergence times are indicated above their respective branches. B) The Meso-Proterozoic (CL, Calymmian; EC, Ectasian; and ST, Stenian periods), Neo-Proterozoic (TO, Tonian; CY, Cryogenian; and ED, Ediacaran periods), Paleozoic (CM, Cambrian; OR, Ordovician; SI, Silurian; DV, Devonian; CB, Carboniferouis; and PR, Permian periods), Mesozoic (TR, Triassic; JR, Jurassic; and CR, Cretaceous periods), and Cenozoic Eras (PG, Palogene; NG, Neogene; and QR, Quaternary periods) are indicated in red, orange, green blue, and purple, respectively. Key events in plant evolution are denoted and dated in black text. The earliest fossils from key evolutionary events are indicated with red text. The time of jasmonate signaling evolution is bracketed and labeled in blue text. The timeline is in MYA.

A



B

Species	AMN1	bHLH	CCT	LisH	vWFA	WRKY	ZnF-GATA
Cyanophora paradoxa	7	6	0	2	30	0	1
Cyanidioschyzon merolae	5	1	3	3	8	0	7
Bathycoccus prasinos	2	1	5	3	14	4	9
Chlamydomonas reinhardtii	16	10	8	3	32	1	17
Chlorella variabilis	10	2	4	4	22	1	11
Coccomyxa subellipsoidea	13	6	5	3	17	1	7
Micromonas pusilla CCMP1545	7	2	3	2	14	2	10
Micromonas pusilla RCC299	4	2	4	6	17	2	8
Ostreococcus lucimarinus	2	1	5	1	12	2	10
Ostreococcus tauri	2	1	4	1	13	2	5
Volvox carteri	12	5	6	2	21	3	11
Klebsormidium flaccidum	51	11	7	6	40	2	15

Figure 3.22. A proposed model for JA signaling evolution. A) A proposed model for when organisms acquired JA signaling components and their progenitor domains during the Viridiplantae evolution. Solid lines indicate the presence of a JA signaling component protein. Dashed lines indicate the absence of a signaling ortholog but the necessary protein domain is present. Domains are listed under the category of organism in which they could first be identified, tracing as far back as the Glaucophyta. B) The numbers of the AMN1, bHLH, CCT, LisH, vWFA, WRKY and ZnF-GATA superfamily domains occurring in each species are presented.

CHAPTER FOUR: MED25 REGULATION OF ANTHOCYANIN BIOSYNTHESIS

Abstract

Mediator is a multi-subunit complex which interacts with transcription factors and other machinery to recruit RNA Polymerase II for transcription. MEDIATOR COMPLEX SUBUNIT 25 (MED25) recognizes a variety of transcription factors involved in plant growth and development, abiotic and biotic stress tolerance, and phytohormone signaling. However, little information exists about this complex in regulation of secondary metabolites. Here I show that MED25 affects anthocyanin accumulation in response to both jasmonate and sucrose treatment. Analysis of publicly available microarrays indicated MED25 is needed for proper expression of multiple anthocyanin biosynthetic genes. Gene expression analysis indicated *med25* mutants exhibits reduced expression of nearly all biosynthetic genes of the pathway, from *PAL* to *UF3GT*, either before or after jasmonate treatment. Conversely, only the expression of *TT8*, *PAP1*, and *PAP2* transcription factors were significantly down-regulated in *med25* plants. MED25 was found to interact with GL3 in yeast-two hybrid and protoplast assay systems. These results suggest MED25 is a key regulator of anthocyanin biosynthesis.

Introduction

Plants produce a diversity of natural products with a multitude of biological functions. One category of natural products synthesized by plants is the flavonoids, which include stilbenes, aurones, flavones, flavonols, isoflavonoids, and anthocyanins. Flavonoids function in plant development (Maloney et al., 2014), attracting pollinators or seed dispersers (Bradshaw and Schemske, 2003), nitrogen fixation (Firmin et al., 1986), photoprotection (Emiliani et al., 2013), and plant defense to pathogens (Treutter, 2005; Zhang et al., 2013), herbivores (Johnson and Dowd, 2004) and stress (Nakabayashi et al., 2014). Additionally, flavonoids provide numerous health benefits to humans as antioxidants, antimicrobial, anti-inflammatory and having anticancer activity (Butelli et al., 2008).

Synthesis of anthocyanins begins by the conversion of phenylalanine into cinnamate by PHENYLALANINE LYASE (PAL). CINNAMATE 4-HYDROXYLASE (C4H) and 4-COUMARATE:COA LIGASE (4CL) lead to the formation of *p*-coumaryl-CoA which is then extended with three malonyl CoA molecules by CHALCONE SYNTHASE (CHS), forming the polyketide naringenin chalcone. Naringenin undergoes isomerization and hydroxylation to form dihydrokaempferol, a dihydroflavanol. Further reduction by DIHYDROFLAVONOL REDUCTASE (DFR) directs the pathway towards anthocyanin biosynthesis. Conversely, reduction of dihydroflavonols by FLAVONOL SYNTHASE (FLS) directs the pathway into forming flavonols. PAL, C4H, and 4CL are referred to as the general flavonoid pathway as products from these enzymes can also flow into lignin biosynthesis. The anthocyanin pathway proper is classified into early (CHS, CHI, and F3H) and late (DFR, LDOX, and UF3GT) stages which are precisely regulated by various transcription factors (TF).

Production of anthocyanins is controlled by a variety of transcription regulators including, ARF (Liu et al., 2007), B-Box (Chang et al., 2011; Holtan et al., 2011; Gangappa et al., 2013), basic helix-loop-helix (bHLH) (Shirley et al., 1995; Nesi et al., 2000; Zhang et al., 2003; Feyissa et al., 2009), bZIP (Ang et al., 1998), DELLA and JAZ repressors (Xie et al., 2016), homeobox (Kubo et al., 1999), LDB (Rubin et al., 2009), R2-R3 MYB (Borevitz et al., 2000; Gonzalez et al., 2008) R3 MYB (Matsui et al., 2008; Zhu et al., 2009), NAC (Morishita et al., 2009), SPL (Gou et al., 2011), TCP (Li and Zachgo, 2013) WD40 (Walker et al., 1999), and WRKY (Devaiah et al., 2007) factors and microRNAs (Gou et al., 2011; Yang et al., 2013). TFs controlling anthocyanin biosynthesis are often characterized as regulators of early and late portions of the pathway as these sub-pathways are targeted by distinct sets of factors. PRODUCTION OF FLAVONOL GLYCOSIDES1, 2, and 3 (MYB12, MYB11, and MYB111 respectively), regulate early pathway genes (Stracke et al., 2007). A well characterized MYB-bHLH-WD40 (MBW) complex is the predominant regulator of the late pathway (Payne et al., 2000; Zhang et al., 2003). The MBW complex is comprised of PRODUCTION OF ANTHOCYANIN 1 (MYB75/PAP1), MYB90/PAP2, MYB113 or MYB114 MYB factors (Borevitz et al., 2000; Gonzalez et al., 2008); GLABRA3 (GL3), ENHANCER OF GLABRA3 (EGL3) or TRANSPARENT TESTA8 (TT8) bHLH factors (Zhang et al., 2003; Zimmermann et al., 2004; Feyissa et al., 2009); and the WD40 protein TRANSPARENT TESTA GLABRA1 (TTG1) cofactor (Walker et al., 1999). The R3 MYBs CAPRICE (CPC) and MYBL2 suppress anthocyanin production by competing with the R2-R3 MYBs for binding to the MBW complex.

Environmental signals including drought, high light, nitrogen and phosphate deficiency, salt, temperature, and UV-stress, as well as sugars (Christie et al., 1994; Landry et al., 1995; Teng et al., 2005; Vanderauwera et al., 2005; Solfanelli et al., 2006; Jiang et al., 2007; Oh et al., 2011; Nakabayashi et al., 2014; Nemie-Feyissa et al., 2014) and the phytohormones abscisic acid (ABA), gibberellin (GA) ,and jasmonate are known to activate anthocyanin biosynthesis (Hattori et al.,

1992; Jiang et al., 2007; Loreti et al., 2008). The phytohormone jasmonate (JA) is a key elicitor of anthocyanin biosynthesis (Feys et al., 1994; Loreti et al., 2008). JA signaling occurs through a CORONATINE INSENSITIVE 1 (COI1)-JASMONATE ZIM-DOMAIN (JAZ) coreceptor complex which bind JA (Chini et al., 2007; Thines et al., 2007). In the presence of JA, the F-box protein COI1 interacts with JAZ, leading to polyubiquitination of JAZ and degradation via the 26S proteasome (Chini et al., 2007; Thines et al., 2007). The JA coreceptor (COI1) has impaired anthocyanin accumulation (Feys et al., 1994). Furthermore, the JAZs repressors interact with GL3, EGL3, and TT8 inhibiting transcription of anthocyanin biosynthetic genes freeing GL3, EGL3, and TT8 for activation of the late anthocyanin pathway (Qi et al., 2011). Anthocyanin biosynthesis is further induced by JA upregulation of PAP1, PAP2, and GL3 expression (Shan et al., 2009).

Mediator is a 25+ subunit complex which interacts with chromatin modifiers and transcription factors to recruit RNA POLYMERASE II complex to gene promoters. Identification of the Mediator complex in plant revealed several additional subunits not present in yeast or metazoan (Bäckström et al., 2007). Studies of the plant Mediator have revealed multiple subunits function in growth, development and defense (Dhawan et al., 2009; Kidd et al., 2009; Xu and Li, 2011; Sundaravelpandian et al., 2012; Xu and Li, 2012; Zhang et al., 2013; Zheng et al., 2013; Lai et al., 2014; Zhu et al., 2014). In Arabidopsis, the Mediator complex has been implicated in regulating natural product biosynthesis and accumulation; however, much remains unknown (Stout et al., 2008; Kidd et al., 2009; Bonawitz et al., 2014). Double mutants of MEDIATOR COMPLEX SUBUNIT 5b (MED5b) and MED5a, also called REDUCED EPIDERMIS FLUORESCENCE (REF4) and REF4-RESEMBLING 1 (RFR1) respectively, were found to have increased accumulation of early phenylpropanoids (Bonawitz et al., 2012). A semidominant mutant of MED5b ($REF4^{D}$) severely reduces plant growth and phenylpropanoid accumulation (Stout et al., 2008). Overexpression of *PAP1* in the $REF4^{D}$ mutant revealed compromised accumulation of anthocyanins, suggesting that PAP1 probably functions through MED5b (Bonawitz et al., 2012). More recently, the *med5a/b* double mutant was shown to partially reduce the compromised lignin accumulation and dwarf phenotypes of the *ref8*, a mutant p-COUMAROYLSHIKIMATE 3'-HYDROXYLASE (Besseau et al., 2007; Bonawitz et al., 2014). Levels of flavonoids are also reduced in *med5a/b ref8* triple mutants, suggesting that flavonoid hyperaccumulation in the *ref8* mutant is not due to metabolic overflow (Bonawitz et al., 2014). MED25, also known as PHYTOCHROME FLOWERING TIME 1, has been reported to positively regulate jasmonate signaling and anthocyanin accumulation in Arabidopsis (Kidd et al., 2009). While expression of PHENYLALANINE AMMONIUM LYASE 1 (PAL1) expression is down regulated in a *med25* mutant (Kidd et al., 2009), the exact mechanism regulating the accumulation of anthocyanins has not been elucidated. Collectively, these studies suggest a role of the Mediator complex in regulation of anthocyanin biosynthesis and accumulation.

Previous work has implicated MED25 in regulating anthocyanin accumulation by an unknown mechanism. My work indicates MED25 is a master regulator of jasmonate induced anthocyanin biosynthesis in Arabidopsis, affecting the expression of early and late portions of the pathway. I find that MED25 interacts with the jasmonate regulated transcriptional activator GL3, a component of the MBW complex important for activation of late anthocyanin pathway genes. Furthermore, I find that MED25 can interact with JAZ1, presumably limiting Mediator activation of plant defenses, including genes for flavonoid biosynthesis.

Results

MED25 Regulates Anthocyanin Accumulation in Response to Jasmonate and Sucrose

Previous reports indicate that the *med25* mutant is impaired in jasmonate signaling and accumulates less anthocyanin (Kidd et al., 2009). To confirm these finding I first measured the jasmonate response using the previously reported *med25-2* null mutant (Figure 4.1A, Kidd et al., 2009). As previously reported the growth of *med25* seedlings roots, with or without jasmonate added to the media, was increased compared to WT (Figure 4.1B). However, root growth while exposed to jasmonate indicated only a slight insensitivity of mutant seedlings. Accumulation of anthocyanin in response to jasmonate induction was reduced in *med25* mutant seedlings after 1 day of induction (Figure 4.2A). Differences in anthocyanin accumulation were detected with 10, but not 25 µM JA. Given this information, I suspected that *med25* mutants may eventually accumulate anthocyanin levels similar to WT when grown for a longer time under induction conditions. However, results show that even after 9 days of induction the 10 µM JA concentration still showed reduced anthocyanin accumulation (Figure 4.2C).

As high sucrose levels also induce anthocyanin accumulation in Arabidopsis seedlings (Teng et al., 2005; Solfanelli et al., 2006) I sought to quantify root growth in response to sucrose in *med25* seedlings. Similar to jasmonate treatment I found anthocyanin levels were decreased in *med25* mutants after induction with 6% sucrose and constant light treatment for 24 hr (Figure 4.2C). Conversely, exposure of *med25* seedlings to high sucrose for 9 days resulted in WT anthocyanin accumulation (Figure 4.2D). These results indicate that for a short time the *med25* mutant has reduced anthocyanin accumulation in response to high sucrose.

Microarray Analysis Reveals Possible Regulation of Anthocyanin Biosynthesis Genes

To identify the role of MED25 in regulating the accumulation of the anthocyanin pathway I sought to use publicly available microarray data to investigate which genes may be miss-regulated. Two microarray datasets are available for the *med25* mutant, one in response to *Fusarium oxysporum* infection, GSE15236, and another for temperature regulation of gene expression, GSE37014, (Kidd et al., 2009; Iñigo et al., 2011). A two-way Analysis of Variance of was conducted on normalized data using the Multiexperiment Viewer (MeV) software (Saeed et al., 2003). As neither experiment was specifically designed for analyzing the anthocyanin pathway directly a significance level of 0.05, without application of a false discovery rate, was used as a cutoff. In the Fusarium and temperature array 2761 and 1338 genes were differentially expressed between the WT and med25 genotypes, respectively. Comparing the datasets revealed 208 shared genes. A hypergeometric distribution analysis was applied to determine the statistical significance between the control datasets. There was a statistically significant (p-value = 0.00069) between datasets. Unsurprisingly, differences were abundant as 2553 and 1130 genes were not shared between the arrays.

In Arabidopsis, anthocyanin accumulation occurs in response to pathogen infection (Veronese et al., 2003; Kidd et al., 2011). Infection with *F. oxysporum* induced the expression of *PAL1*, *PAL2*, *PAL3*, *PAL4*, *CHS*, *DFR*, *F3'H*, *F3'5'H*, and *UF3GT*, accounting for at least half of the biosynthetic steps necessary for anthocyanin synthesis. *F. oxysporum* infections differentially regulated the anthocyanin transcriptional regulators *MYB12*, *TTG1*, *PAP1*, and *PAP2*. Contrary to reports for *PAL1* (Kidd et al., 2009), analyzing the Fusarium experiment microarrays revealed expression of *PAL2* was significantly up regulated in the *med25* mutant. Conversely, there was a

significant down regulation of *DFR*, a key gene for jasmonate regulation of anthocyanin accumulation (Figure 4.3). The transcriptional regulators *TTG1*, *MYB113*, and *EGL3* were found to be significantly up regulated in the *med25* mutant.

Anthocyanin accumulation is also affected by temperature, with low and high temperature stresses promoting and suppressing pigmentation, respectively (Rowan et al., 2009; Catalá et al., 2011). Microarray analysis of gene expression at 16°C and 23°C, low and optimal temperatures respectively, revealed effects on nearly all biosynthetic enzymes involved in anthocyanin formation. Only *LDOX* was not significantly up-regulated by low temperature. Low temperature significantly up-regulated the expression of *MYB12*, *PAP1*, *TT8*, and *EGL3*. At both 16°C and 23°C, the *med25* mutation significantly down-regulated expression of *PAL1*, *PAL2*, *PAL4*, *C4H*, *4CL3*, *CHS*, *CHI*, *FLS1*, and *F3H* (Figure 4.3). Expression of the *MYB12* transcriptional regulator was down regulated in the *med25* mutant at both 16°C and 23°C. According to the microarray analysis, no biosynthetic genes or transcriptional regulators displayed increased expression in the *med25* mutant at either temperature. Collectively, these results suggest MED25 may affect the expression of most anthocyanin biosynthetic genes and several transcription factors involved in regulating the pathway.

Expression of Anthocyanin Biosynthetic Genes Are Reduced in med25 Mutants

To demonstrate the role of MED25 in governing anthocyanin biosynthesis I focused on jasmonate induction of the flavonoid pathway. Expressions of regulatory genes were investigated before and after jasmonate treatment. Treatment of Arabidopsis seedlings with JA successful induced accumulation of *JAZ1* transcripts (Figure 4.4A), a marker gene for JA response. Although *JAZ1* transcripts were induced in response to jasmonate treatment in the *med25* mutant, transcript levels were less than control plants. This finding suggests a disruption of endogenous jasmonate signaling in the *med25* mutant. To show jasmonate induced expression of the anthocyanin pathways I measured *DFR* transcript levels. *DFR* expression was induced 8 h after jasmonate treatment (Figure 4.4B).

In addition to DFR, I performed a comprehensive gene expression analysis of the flavonoid biosynthetic pathway to further elucidate the targets by which MED25 regulates anthocyanin accumulation. Surprisingly, the mutation of *med25* negatively affects the expression of almost all flavonoid pathways genes, both before and after JA induction. The general (PAL1, C4H, and 4CL), the early (CHS, CHI, F3H, F3'H, and F3'5'H) and late (DFR, LDOX, and UF3GT) phenylpropanoid pathways were all affected either before and/or after JA treatment (Figure 4.4C). Only CHI and F3'H were unchanged at a basal level in *med25* mutants. The basal levels of CHS and LDOX were down-regulated more than 2 fold. After JA induction for 8 h PAL1, C4H, 4CL, CHS, CHI, F3H, DFR, LDOX, and UF3GT are down-regulated (Figure 4.4D). Interestingly, expression of F3'H and F3'5'H significantly increased in the *med25* mutant upon JA induction. Similar to the basal expression level, JA induction of LDOX was suppressed by 2 fold in the *med25* mutant.

MED25 Affects the Expression of Transcription Factors Regulating the Anthocyanin Pathway

As multiple biosynthetic genes were miss-regulated in the *med25* mutant I sought to identify the transcription factors which were connected to these disruptions. To address this issue, I started by measure transcription factor expression at 0 and 8 h. Expression of *TTG1* and

EGL3 remained unaffected either before or after JA treatment (Figure 4.5A and B). In contrast *TT8* was not affected at a basal level but its induction by JA was impaired in the *med25* mutant (Figure 4.5B).

To further investigate regulatory effects on the anthocyanin pathway several TFs were with known JA-induced expressions were more thoroughly investigated with 0, 2, 4, and 8 h time points (Figure 4.5C-E). *GL3*, *PAP1*, and *PAP2* were significantly up-regulated by JA treatments as has been previously reported (Qi et al., 2011). GL3 expression did not differ between *med25* and WT at any of the times measured (Figure 4.5C). Conversely, expression of *PAP1* and *PAP2* displayed reduced expression beginning around 2 h after JA treatment and that was significantly different by 4 h after treatment (Figure 4.5D and E). These findings suggest that miss-regulation of anthocyanin accumulation in the *med25* mutant is due to down-regulation of select anthocyanin regulatory TFs.

MED25 Interacts with GLABRA3

As multiple regulatory and biosynthetic genes of the anthocyanin pathway were down regulated in *med25* I decided to test possible interactions with known regulators of anthocyanin biosynthesis. MED25 and potential transcription factor targets were tested with a yeast-2 hybrid system. The ACID domain, amino acids 551-681, of MED25 is important for interaction with transcription factor targets (Elfving et al., 2011; Ou et al., 2011; Çevik et al., 2012). The MED25 ACID domain (BD-MED25$^{228-681}$) and potential interacting factors were fused with the GAL4 binding and activation domains, respectively. Five mM 3-amino-1,2,4-triazole (3-AT) was used to reduce the strong autoactivation observed with the BD-MED25$^{228-681}$. No interaction with BD-MED25$^{228-681}$ was found for AD-TTG1$^{1-341}$, AD-PAP1$^{1-248}$, AD-PAP2$^{1-249}$, AD-TT8$^{1-518}$, or AD-EGL3$^{1-596}$ (Figure 4.6A). However, AD-GL3$^{1-637}$ interacted with BD-MED25$^{228-681}$ on triple selection plates in the presence of 10 mM 3-AT but not on quadruple selection plates (Figure 4.6A), indicating a weak interaction between these factors. Surprisingly, despite the high identity, AD-EGL3$^{1-596}$ and AD-GL3$^{1-637}$ interacted with BD-MED25$^{228-681}$ differently.

To further characterize the interaction of GL3 with MED25 I dissected GL3 into five parts consisting of the different TF domains. AD-GL3$^{1-209}$, AD-GL3$^{200-400}$, AD-GL3$^{400-495}$, and AD-GL3$^{552-637}$ consisted of the MIR, AD, bHLH, and ACT domains respectively. Furthermore, AD-GL3$^{400-637}$ consisting of the bHLH and ACT domains was also constructed. The middle and ACID domain section of MED25 was used as the binding domain, BD-MED25$^{228-681}$. Only AD-GL3$^{400-637}$ and BD-MED25$^{228-681}$ were found to interact (Figure 4.6B). Notably, the bHLH, AD-GL3$^{400-495}$, or ACT, AD-GL3$^{552-637}$, domains alone were not sufficient to allow detectable MED25 binding.

A protoplast-based transactivation assay was performed to validate the interaction of GL3 with MED25 *in vivo*. The middle and ACID domain region of MED25 was fused to the GAL4 domain. GL3 was expressed under the 35S promoter. MED25-GAL4, significantly increased protoplast activity, indicating this domain is functional *in vivo* (Figure 4.6C). Activity was further increased by the addition of GL3. In yeast, EGL3 did not interact with MED25 (Figure 4.6C). Due to the sequence similarity between EGL3 and GL3 I sought to determine whether a key factor is missing from yeast or if no interaction occurs *in vivo*. The addition of EGL3 or PAP1 did not increase activity beyond the background level induced by MED25 alone (Figure 4.6D). These results suggest that MED25 specifically interacts with GL3 *in vivo*.

JAZ1 Interaction with the ACID Domain of MED25

I speculated at the possibility of MED25 interacting with JAZ proteins. As MED25 interacts with MYC2, MYC3, MYC4, (Çevik et al., 2012; Chen et al., 2012) along with GL3, and these proteins are known target of JAZ proteins, I sought to investigate the possibility of MED25 directly interacting with JAZ proteins. I tested the interaction of the same MED25 middle and ACID domains, BD-MED25$^{228-681}$, with full length JAZ1, JAZ2, JAZ8, and JAZ10 as a representatives of the JAZ protein family which can interact with GL3 (Qi et al., 2011). Full length JAZ1, JAZ2, JAZ8, and JAZ10 were cloned into the activation domain vector forming AD-JAZ1$^{1-253}$, AD-JAZ2$^{1-249}$, AD-JAZ8$^{1-131}$, and AD-JAZ10$^{1-197}$ respectively. Interestingly, only AD-JAZ1$^{1-253}$ consistently interacted with BD-MED25$^{228-681}$ (Figure 4.7A). Colonies only grown on the HLT plate, suggesting a weak interaction between MED25 and JAZ1.

To validate the interaction between MED25 and JAZ1 in vivo I conducted a protoplast-based transactivation. As with GL3 (Figure 4.7B), I used only the middle and ACID domain of MED25 fused to GAL4 binding domain for transactivation. To insure repressor function independent of additional cofactors (NINJA and TPL), I fused JAZ1 with the SRDX repressor domain. To demonstrate functional activity of the SRDX domain I fused it to GL3. Both JAZ1-SRDX and GL3-SRDX were expressed using the 35S promoter. As expected GL3-SRDX did not transactivate the luciferase promoter in the presence of MED25 (Figure 4.7B), indicating a functional SDRX domain. Expression of JAZ1-SRDX with MED25 resulted in a slight decrease in luciferase activity (Figure 4.7B). These results further substantiate the *in vivo* interaction of MED25 with JAZ1 and GL3.

Localization of MED25

As a necessary component of transcription Mediator is expected to localize to the nucleus. However, no studies have validated this assumption for MED25. GFP was fused to the C-terminal end of a full length MED25. The GFP-MED25 plasmids was transfected into tobacco protoplast and observed under the microscope. As expected, GFP-MED25 primarily localized to the nucleus (Figure 4.7C). Possibly due to slow nuclear transport, minor fluorescence was observed in the cytosol and around the cell wall.

Microarray Analysis Reveals Overlap between GL3 and MED25 Regulated Genes

To support the interaction of MED25 with GL3 I compared genes overlapping between publicly available microarrays of *med25* and *gl3* mutant plants. Two array are available for the *med25* (Kidd et al., 2009; Iñigo et al., 2011) mutant and three for *gl3* (Jakoby et al., 2008; Morohashi and Grotewold, 2009). As the microarray experiments were conducted independently, I normalized and analyzed each array in context of the initial study to identify genes differentially regulated. Then, only data relating to genotypic differences was used to determine the overlap between MED25 and GL3 target genes. A hypergeometric distribution analysis was used to determine if a statistically significant overlap of target genes occurred. Significant overlap between MED25 and GL3 genotypes occurred in five of the six array combinations (Figure 4.8). Gene ontology (GO) analysis was conducted to identify shared target gene functions. Genes shared between the MED25 GSE12536 and GL E-GEOD-12522 revealed an enrichment for cellular amino acid derived metabolic process. Flavonoid biosynthetic processes,

intracellular transport, response to organic substances and defense response were GO categories enriched for genes shared between the MED25 GSE37014 and GL3 E-GEOD-12551 arrays. Other combinations did not reveal an enrichment for any GO category. Genes occurring in more than one array combination included ABCI18, CER1, CYP79B3, JAV1, LOX5, PMEPCRF, UMAMIT28, AT2G45520, AT3G06450, AT3G10930, AT5G04310, and AT5G16990. The significant overlap between MED25 and GL3 target genes further supports their physical interaction.

MED25 Regulates Trichome Development

To further validate that MED25 interacts with GL3 I reasoned that *med25* and *gl3* mutant plants should display similar phenotypes. GL3 is known to function in the regulation of trichome formation (Payne et al., 2000; Bernhardt et al., 2003), therefore I first sought to quantify the number of trichomes on the *med25* mutant stems. The basal 2.5 cm of inflorescence stem were section and found to have less trichomes (Figure 4.9A and B). Furthermore, there were significantly less branched trichomes, as has also been reported for *gl3* plants (Figure 4.9C). As further validation, I quantified the number of trichomes on the 3rd and 4th leaves of *med25* compared to WT. The total number of trichomes was not significantly different between *med25* and WT (Figure 4.9D and G). As expected, there were no difference in single or bifurcated trichomes, but significantly less triradiate trichomes in the *med25* mutant (Figure 4.9E-G).

To further investigate the cause of reduced trichome number and branching I analyzed expression from trichome regulatory and developmental genes *TRIPTYCHON* (*TRY*), *CAPRICE* (*CPC*), *SIAMESE* (*SIM*), and *ENHANCER OF TRY AND CPC 1* (*ETC1*). The TF MYC2 is known to interact with MED25 (Çevik et al., 2012; Chen et al., 2012). As with GL3, there is a significant overlap between MYC2 and MED25 regulated genes (Figure 4.10A). To determine if trichome development occurs specifically through GL3, I compared the *med25* mutant with *gl3* and *myc2* lines. Expression of *TRY*, *CPC*, *SIM*, and *ETC1* was slightly elevated in the *med25* mutant, but did not differ significantly from WT seedlings (Figure 4.9H-K). However, expression of *TRY*, *CPC*, and *ETC1* were significantly increased in both the *myc2* and *gl3* mutants (Figure 4.9H-K). *SIM* expression was also significantly higher in the *myc2* mutant (Figure 4.9J). Thus, MED25 regulation of trichome genes likely requires both GL3 and MYC2.

MED25 Affects the Accumulation of Proanthocyanins in Arabidopsis Seeds

The MBW complex regulates the production of seed coat mucilage and proanthocyanidins (Baudry et al., 2004). To test the impact of genetic mutations on seed mucilage I stained seed with ruthenium red. No differences were readily distinguishable between WT, *med25*, *gl3*, and *myc2* lines (Figure 4.11A).

Expression of TT8 was decreased in the *med25* mutant (Figure 4.5B). I investigated the possibility of differential accumulation of proanthocyanidins in the seed coat. As MYC2 has also been implicated in regulating flavonoid production (Dombrecht et al., 2007) I included the *myc2* mutant to differentiate any potential proanthocyanidin phenotypes. Visual inspection revealed slightly more pigmented seed coats in the MED25 overexpressor (MED25 OX) line and less in the mutant (Figure 4.11B). Both *gl3* and *myc2* has slightly lighter brown seeds coats similar to the *med25* line. To further quantify the difference in seed coat color, I extracted proanthocyanidins from the seeds of *med25*, *gl3*, *myc2*, and WT lines. Proanthocyanidins were divided into soluble

and insoluble fractions. After extraction, soluble and insoluble proanthocyanidins were acid decomposed into anthocyanidins and the absorbance at 530 nm was quantified. Soluble proanthocyanidins were significantly higher in the MED25 OX line (Figure 4.11C). In contrast, seeds of *med25* and *myc2* lines accumulated significantly higher levels insoluble proanthocyanidins (Figure 4.11C). These results indicate MED25 impacts accumulation of proanthocyanidins in Arabidopsis seeds, potentially through the MYC2 TF.

To further investigate the proanthocyanidin phenotype of *med25* mutant I analyzed the expression of TT8. Expression of TT8 in *med25* seedlings was down both before and after induction with JA (Figure 4.11D). Both *myc2* and *gl3* also displayed reductions in TT8 expression under both conditions (Figure 4.11D). Thus, visual pigmentation reduced in *med25*, *gl3*, and *myc2* mutants may be due decreased expression of the MBW complex subunit TT8.

Interactions between MEDIATOR Subunits and the MBW Complex

As both the MEDIATOR and MBW complex contain multiple subunits, I suspected additional protein-protein interactions may occur between these two complexes. A Y2H assay was conducted to determine interactions of the MBW complex with MED5, also called MED33, MED16, and CDK8. Due to endogenous restriction sites the MED5a, MED5b, and MED16 sequences, these genes were divided into four, three, and three pieces, respectively. MED16 was divided into BD-MED16$^{1\text{-}420}$, BD-MED16$^{421\text{-}800}$, and BD-MED16$^{801\text{-}1280}$. Protein fragments for MED5a were BD-MED5a$^{1\text{-}227}$, BD-MED5a$^{228\text{-}560}$, BD-MED5a$^{561\text{-}978}$, and BDMED5a$^{979\text{-}1309}$. The three protein MED5b fragments included BD-MED5b$^{1\text{-}289}$, BD-MED5b$^{290\text{-}718}$, and BD-MED5b$^{719\text{-}1276}$. BD-CDK8$^{1\text{-}470}$ spanned the full length CDK8 protein.

MED16 has been shown to interact with MED25 (Yang et al., 2014). Thus, due to proximity, I investigated the possibility that MED16 interacts with subunits of the MBW complex. However, neither BD-MED16$^{1\text{-}420}$, BD-MED16$^{421\text{-}800}$, and BD-MED16$^{801\text{-}1280}$ interacted with AD-TTG1$^{1\text{-}341}$, AD-PAP1$^{1\text{-}248}$, AD-PAP2$^{1\text{-}249}$, AD-TT8$^{1\text{-}518}$, AD-GL3$^{1\text{-}637}$ or AD-EGL3$^{1\text{-}596}$.

MED25 also interacts with CDK8 (Zhu et al., 2014). I hypothesized association of CDK8 with MED25 would also place it in close proximity to the MBW complex. Thus, I also investigated the possible interactions of CDK8 with MBW complex components. Surprisingly, similar to MED25, BD-CDK8$^{1\text{-}470}$ also interacted with AD-GL3$^{1\text{-}637}$. As MED25 also interacts with JAZ1 I tested the interaction between CDK8 and JAZ1. The presence of yeast colonies on the HLT plate indicated the interaction of BD-CDK8$^{1\text{-}470}$ and AD-JAZ1$^{1\text{-}253}$ (Figure 4.12A).

The relationship between MED5 subunits and MED25 remains poorly defined. As the REF4^{D} mutant represses anthocyanin accumulation of a PAP1^{D} mutant (Bonawitz et al., 2012) I investigated the possibility of MED5 interactions with MBW complex components. No interactions between BD-MED5a$^{1\text{-}227}$, BD-MED5a$^{228\text{-}560}$, BD-MED5a$^{561\text{-}978}$, or BDMED5a$^{979\text{-}1309}$ and AD-TTG1$^{1\text{-}341}$, AD-PAP1$^{1\text{-}248}$, AD-PAP2$^{1\text{-}249}$, AD-TT8$^{1\text{-}518}$, AD-GL3$^{1\text{-}637}$ or AD-EGL3$^{1\text{-}596}$ were observed. Similarly, no interactions between, BD-MED5b$^{290\text{-}718}$ or BD-MED5b$^{719\text{-}1276}$ with AD-TTG1$^{1\text{-}341}$, AD-PAP1$^{1\text{-}248}$, AD-PAP2$^{1\text{-}249}$, AD-TT8$^{1\text{-}518}$, AD-GL3$^{1\text{-}637}$ or AD-EGL3$^{1\text{-}596}$ occurred. However, a weak interaction between BD-MED5b$^{1\text{-}289}$ and AD-PAP2$^{1\text{-}249}$ occurred on the HLT plate (Figure 4.12B). Collectively, these interactions suggest multiple Mediator subunits interact with the MBW complex.

Discussion

Prior studies have indicated a potential role of MED25 in regulating anthocyanin production(Kidd et al., 2009). However, the mechanism by which this process occurs has not been elucidated. Here I characterize the role of the Mediator subunit MED25 in regulating biosynthesis of anthocyanins and proanthocyanidins. The *med25* mutant plants display a reduction in anthocyanin accumulation in response to sucrose and JA. Importantly, I found MED25 is a master regulator of anthocyanin biosynthesis. In Arabidopsis, all biosynthetic genes from *PAL* to *UF3GT* were down-regulated in the *med25* mutant. Analysis of regulatory genes revealed reduced expression of *PAP1*, *PAP2*, and *TT8* in the *med25* mutant. MED25 was found to directly interact with GL3 and JAZ1 in yeast. JAZ1 interaction with MED25 revealed a new mechanism by which JA controls gene expression.

Anthocyanin Accumulation Is Regulated by the MEDIATOR Complex Through MED25

A prior report indicated anthocyanin accumulation is affected by MED25 (Kidd et al., 2009). These authors induced anthocyanins by culturing detached seedlings in water for three weeks. To validated the role of MED25 in regulation of anthocyanins, I used a more traditional approach by using JA and sucrose treatments. Minor, but significant and repeatable differences, were observed in JA or sucrose treated *med25* and WT seedlings (Figure 4.2). Kidd et al. (2009) reveal a more pronounced difference between MED25 genotypes and WT plants. The contrast may be due to a combination of larger seedlings, treatments, and environmental conditions. Conditions used by Kidd et al. (2009) would elicit nutrient starvation, including nitrogen. The MBW complex needs GL3 to induce anthocyanins in response to nitrogen deficiency (Nemie-Feyissa et al., 2014). In contrast, GL3 displayed a more limited role for inducting anthocyanin accumulation in response to JA. Thus, the interaction between MED25 and GL3 could explain stronger phenotype in the report by Kidd et al. (2009).

MED25 Is a Potential Master Regulator of Flavonoid Biosynthesis

The flavonoid pathway is divided into early, middle, and late stages of anthocyanin biosynthesis. Each sector of the flavonoid pathway is mediated by a unique set of TFs. The early portion of the pathway is primarily controlled by factors involved in lignin biosynthesis (Didi et al., 2015), although the MBW component PAP1 also contributes (Bhargava et al., 2010). Although functioning in distinct tissues, MYB11, MYB12, and MYB111 regulate the middle section of the pathway and synthesis of flavonols (Mehrtens et al., 2005; Stracke et al., 2010). The MBW complex is a key regulator of late stages of the flavonoid pathway (Patra et al., 2013). The late steps of the flavonoid pathway lead to production of anthocyanins and proanthocyanidins (Patra et al., 2013).

Analysis of publicly available microarrays suggested MED25 may target multiple sections of the flavonoid pathway (Figure 4.3). Gene expression analysis supported this hypothesis (Figures 4). Indeed, all core early, middle, and late flavonoid biosynthetic genes are down-regulated in the *med25*, either before or after JA induction. Basal level expression of *PAL1*, *C4H*, *4CL*, *CHS*, *F3H*, *F3'H*, *LDOX*, and *UF3GT* was significantly reduced in *med25* seedlings. Basal expression of *CHI* and *F3'5'H* remained unchanged. Upon JA induction PAL1, *C4H*, *4CL*, *CHS*, *CHI*,

F3H, *DFR*, *LDOX*, and *UF3GT* expression was reduced relative to WT. In contrast, expression of *F3'H* and *F3'5'H* was induced in response to JA.

To determine the cause for anthocyanin biosynthetic gene down-regulation I measured expression of MBW complex transcription factors. Basal level expression of TTG1, EGL3, GL3, TT8, PAP1, and PAP2 remained unchanged compared to WT (Figure 4.5). Upon induction with JA, the expression of TT8, PAP1, and PAP2 was significantly reduced versus the control (Figure 5B, D and E). Thus, the strong reduction in PAP1 and PAP2 gene expression likely explains the decreased levels of late stage anthocyanin biosynthetic genes. Reduction of PAP1 and PAP2 transcript, and presumably the corresponding TF proteins, would limit formation of a fully functional MBW complex. With impaired MBW complex formation, maximum transcriptional activation of late stage anthocyanin biosynthetic genes may not be achieved (Gonzalez et al., 2008).

In addition to anthocyanins, MED25 affects accumulation of proanthocyanidins. TT8 is the bHLH component of the MBW complex for primary responsible for proanthocyanidin biosynthesis in the Arabidopsis seed coat (Nesi et al., 2000). The reduction of pigmentation in the *med25*, *gl3*, and *myc2* mutant testa may be explained by decreased expression of TT8. Indeed, expression of TT8 was decreased in these lines both before and after induction with JA (Figure 4.11). I found that MED25 OX and *med25* mutant lines had increased levels of soluble and insoluble proanthocyanidins, respectively (Figure 4.11C). The opposite role of overexpressor and mutant MED25 suggests it affects a proanthocyanidins polymerization. It is possible soluble and insoluble proanthocyanidins produce a darker and lighter testa color, respectively. Alternatively, as *gl3* did not display significantly altered proanthocyanidin levels, the visible testa color may be determined by one or more specific proanthocyanidin compounds. *Medicago truncatula* hairy roots overexpressing MtMYB5 or MtMYB14 accumulated more soluble and insoluble proanthocyanidins, leading to darker roots (Liu et al., 2014). However, analysis of Mt*myb5* or Mt*myb14* seeds had lower levels of soluble proanthocyanidins but displayed a darker seed coat (Liu et al., 2014). In soybean (*Glycine max*), different combinations of anthocyanins and proanthocyanidins result in yellow, brown, or black seed coats (Todd and Vodkin, 1993; Kovinich et al., 2011). These studies suggest testa color is likely a complex trait resulting from multiple metabolites. More work is needed to understand the mechanisms by which MED25 governs this process.

Evidence for the in vivo Interaction between GL3 and MED25

As a component of the MBW complex, GL3 regulates anthocyanin biosynthesis as well as root hair and trichome development (Payne et al., 2000; Bernhardt et al., 2003; Feyissa et al., 2009). In Arabidopsis, root hair development is also regulated by MED25 (Chandrika et al., 2013; Raya-González et al., 2014), suggesting a possible connection with GL3. Using a yeast-two hybrid assay, I identified MED25 interacting with GL3 (Figure 4.6A). Dissecting the MED25 and GL3 genes revealed the ACID domain and C-terminal end were responsible for the interaction (Figure 4.6B). In tobacco cells, functional activity of the truncated MED25 ACID domain was illustrated through increased basal expression of the luciferase reporter. When combined with GL3, the MED25 ACID domain significantly induced luciferase expression (Figure 4.6C). When fused with the SRDX repressor domain and coexpressed with MED25, GL3 could no longer full induce transactivation of the reporter construct (Figure 4.7B). Combined with MED25, enhanced and repressed transactivation by GL3 and GL3-SRDX in tobacco protoplast, respectively, provide evidence for their interaction *in planta*.

Furthermore, support for the GL3 and MED25 interaction occur through their shared phenotypes. Both *med25* and *gl3* mutants displayed similar visual reductions in seed coat coloration. Although *med25* exhibited increased insoluble proanthocyanidins whereas *gl3* did not, this effect may occur through the MYC2 TF (Figure 4.11C). Additionally, mutation of *gl3* disrupts leaf trichome development (Payne et al., 2000; Zhang et al., 2003). Leaves of *gl3* seedlings possess similar number of trichomes a WT plants, but display a reduced number triradiate branched trichomes (Payne et al., 2000). Mutant *med25* plants also displayed disrupted trichome development. Stems of *med25* plants possessed fewer trichomes which had less branching (Figure 4.9). In contrast, young *med25* leaves did not have a significant reduction of total trichomes. Further study revealed the number of triradiate trichomes was significantly reduced (Figure 4.9F and G). Thus, the phenotypes of *med25* and *gl3* seedlings agree, suggesting GL3 recruits MED25 for transcription of its target genes. GL3 activates the expression of CPC, ETC1, TRY, and SIM (Morohashi et al., 2007; Morohashi and Grotewold, 2009). The R3 MYBs (CPC, ETC1, ETC2, ETC3, MYBL2, and TRY) interact with bHLH factors of the MBW complex to suppress trichome development and anthocyanin biosynthesis while enhancing root hair formation (Kirik et al., 2004; Kirik et al., 2004; Dubos et al., 2008; Matsui et al., 2008; Zhu et al., 2009). Increased TRY, CPC, and ETC1 expression occurred in the *gl3* mutant (Figure 4.9H, I, and K). However, expression of TRY, CPC, ETC1, and SIM were not significantly different in *med25*. GL3 may regulate the expression of these genes independently of MED25. As GL3 also interacts with Mediator component CDK8 (Figure 4.12A), the regulation of TRY, CPC, and ETC1 could occur via this subunit. Using publicly available microarray data I compared genes miss-regulated between *med25* and *gl3*. This identified 13 to74 genes common between both microarray datasets. This overlap was significant despite the cultural and environmental conditions differences between array datasets. Overall, these experiments provide strong evidence for the interaction of GL3 with MED25 *in vivo*.

Integration of MED25 into the Anthocyanin Regulatory Network

Here, I showed GL3 interacts with MED25. MED25 also interacts with the JA master regulator MYC2 (Çevik et al., 2012; Chen et al., 2012). I used the *gl3* and *myc2* mutants to help dissect MED25 regulation of anthocyanin accumulation. When a glucocorticoid receptor fused to GL3 was induced with dexamethasone PAP2, but not PAP1, expression increased (Gonzalez et al., 2008). Addition of cycloheximide revealed GL3 induction of PAP2 occurred indirectly. MYC2 has been implicated in regulation of PAP1 gene transcription (Dombrecht et al., 2007). Thus, I sought to determine if either GL3 or MYC2 were responsible for the reduced JA induced PAP1 transcript accumulation in the *med25* mutant. Although statistically not significant, induction of PAP1 expression was impaired in the *gl3* mutant (Figure 4.10B). In contrast, mutation of *myc2* did not affect the induction of PAP1 transcripts in response to JA. GL3 transcripts accumulate in response to JA induction (Figure 4.5C) (Shan et al., 2009). Under JA elicited conditions, the accumulation of GL3 proteins presumable increases interaction with MED25. Thus, disruption of GL3 interaction with MED25, may partially explain the reduced JA induction of PAP1 transcripts. However, GL3 is also targeted at the protein level by UBIQUITIN PROTEIN LIGASE 3 (UPL3) (Patra et al., 2013). Thus, an increase in GL3 transcript levels does not necessarily equate to more protein available to interact with MED25. The role of JA on UPL3 regulation of GL3 warrants further study.

Only a few anthocyanin factors regulate multiple sectors of the flavonoid pathway. MYBL2 function to bind bHLH factors of MBW complex, repressing its activity (Dubos et al., 2008). However, overexpression of MYBL2 represses expression of both middle and late steps of the

flavonoid pathway (Dubos et al., 2008; Matsui et al., 2008). ELONGATED HYPOCOTYL5 (HY5) regulates the expression of middle and late steps of anthocyanin biosynthesis by targeting MYB12 and PAP1, respectively (Stracke et al., 2010; Shin et al., 2013). HY5 also function to induce anthocyanin accumulation through repression of MYBL2 (Wang et al., 2016). MED25 genetically interacts with HY5 to regulate light signaling and sulfate metabolism (Klose et al., 2012; Koprivova et al., 2014). Therefore, it is possible MED25 functions through the action of HY5 to regulate the entire flavonoid pathway.

In addition to MYBL2 and HY5, other regulators more specifically target components of the MBW complex. Overexpression of LBD37 or LBD38 results in suppression of anthocyanin accumulation and down-regulation of PAP1 (Rubin et al., 2009). TCP3 binds to MBW complex MYBs in addition to MYB12 and MYB111 involved in flavonol regulation (Li and Zachgo, 2013). SPL9 binds PAP1 to suppress activity of the MBW complex (Gou et al., 2011). Accumulation of the SPL9 TF is negatively regulated by the microRNA miR156. While SPL9's role appears limited to binding the MBW complex MYBs, miR156 overexpression affects both middle and late genes of flavonoid biosynthesis (Gou et al., 2011). Notably, SPL9 and miR156 function to regulate anthocyanin accumulation to nitrogen deficiency, a condition which requires GL3 (Feyissa et al., 2009; Rubin et al., 2009). The Mediator complex subunits MED17, MED18, and MED20a are required for microRNA accumulation (Kim et al., 2011). It is possible MED25 is also involved in this process. Future research will address the MED25 regulation of PAP1 and PAP2

JAZ1 Interaction with MED25 Provides a New Mechanism for Controlling JA Response

JAZ factor elicit key mechanistic controls over the JA signaling pathway. JAZ repressors bind regulators of JA signaling (Sasaki-Sekimoto et al., 2013). Anthocyanins are induced by jasmonate treatment (Loreti et al., 2008; Shan et al., 2009). Furthermore, activators of anthocyanin biosynthetic pathway, EGL3, GL3, PAP1, and TT8, are targets of JAZ (Qi et al., 2011). Presumably upon JA induction, JAZ releases these factors which can then form a functional MBW complex to activate transcription. I noted that JAZ repressors binding GL3, MYC2, MYC3, and MYC4, which are all TF targets of MED25. This led us to test for a possible interaction between JAZ factors and MED25 (Figure 4.7A). Indeed, JAZ1 was found to directly interact with MED25 in yeast and *in planta* (Figure 4.7B). This lead us to develop a new model for regulation of anthocyanin biosynthesis (Figure 4.13).

The interaction of JAZ1 with MED25 indicates a new mechanism for JA regulation of gene expression. Zhang et al. (2015) reported that JAZ9 competes with MED25 for binding MYC3. This study also demonstrated the structural change that occurs when JAZ9 binds the MYC3 TF. Thus, structural conformation changes and direct competition for target TFs are two mechanisms by which JAZ repressors exert their function. The interaction of JAZ1 with MED25 expands the repertoire of JAZ repressive mechanisms. TFs are known to interact with MED25 occurs through the ACID domain (Figure 4.6 and Figure 4.7) (Ou et al., 2011; Çevik et al., 2012; Chen et al., 2012). By binding directly to the ACID domain, JAZ1 likely physically obstructs TFs interaction with MED25 when no JA signal is present. Upon JA perception, JAZ is degraded by the 26S proteasome (Chini et al., 2007; Thines et al., 2007). JAZ degradation, in combination with releasing TFs to recognize and bind DNA promoter *cis*-elements, permits access to the ACID domain of MED25. Interaction of TFs with MED25 and the remaining Mediator complex can then recruit RNA Polymerase II to initiation target gene transcription. Expression of *JAZ1* by MYC2 (Chini et al., 2007; Grunewald et al., 2009) leads to JAZ1 protein accumulation to fine tune gene expression. In advent of aberrant

accumulation of JA response regulators, such *GL3* or *MYC2*, JAZ1 bound MED25 blocks TF recognition and in turn prevents downstream gene activation. As JA signaling diverts plant resources from growth to defense, a tight control of signaling is needed to optimize plant survival (Yang et al., 2012). Interestingly, expression of JAZ1 is also induced by auxin (Grunewald et al., 2009). In addition to JA signaling, MED25 also functions to regulate auxin signaling (Raya-González et al., 2014). The role of MED25 in auxin signaling remains ambiguous and my results implicate a connection with JA response. Overall, my result further emphasizes the fine control of plant hormone signaling networks and continues to expand the diverse roles JAZ factors contribute to this process. Better understanding mechanisms regulating JA signaling are necessary for genetically engineering crops with improved defenses or increased production of specialized metabolites.

Materials and Methods

Arabidopsis Growth Conditions

Arabidopsis seedlings were sterilized using 70% ethanol for 1 minute followed by 10 minutes in 30% bleach with several drops of Tween20. Seeds were rinsed then chilled at 4°C for 3 days. Seedlings root growth and anthocyanin assays were grown on ½ MS media for 4 days in 16:8 day:night growth chamber then transferred to treatment conditions. Seedlings for root growth assays were transferred to JA plates then grown an additional 7 days under 16 h long days prior to measuring final root length. Trichome number and branching were determined using a dissecting microscope. Trichomes were counted on third and fourth leaves of seedlings and stems of mature inflorescences.

Anthocyanin and Proanthocyanidin Accumulation

For anthocyanin assays, seedlings were transferred to 10 or 25 µM JA or 6% sucrose plates and then grown for a specified time under constant light. Five seedling per sample were combined and weighted. Anthocyanin was extracted overnight using 1% HCl in methanol. Absorbance at 530 and 657 nm was recorded. Absorbance of anthocyanins was calculated according to the equation $A=A_{530} - ¼(A_{657})$ as previously described (Teng et al., 2005).

Seeds for the proanthocyanidin accumulation assay were collected from mature soil grown plants. Extraction of proanthocyanidins was modified from Routaboul et al. (2006) and Pourcel et al. (2005). To extract proanthocyanidins, I weighted 7.5 mg of seed for each sample. Each sample was ground in 500 µL of 75 acetonitrile:25 H_2O. Once ground another 500 µL of 75% acetonitrile:25% water was added followed by a 1 h extraction on a plate shaker. Samples were centrifuged for 5 min. at 13,200 rpm. Supernatant was collected and stored at 4°C. The pellet was resuspended in 1 ml of 75% acetonitrile:25% water and placed on the plate shaker overnight to extract any remaining soluble proanthocyanidins. After centrifugation, samples of soluble proanthocyanidins were pooled and dried using a rotary evaporator. Each sample was dissolved in 100 µL of 75% acetonitrile:25% water. To decompose soluble proanthocyanidins into anthocyanidins, 95 µL of 95% methanol: 5% concentrate HCl was added to 30 µL of each sample. Samples were incubated at 99°C for 20 min. in a thermocycler. Absorbance was measured at A_{530} using a Nanodrop photospectrometer. For insoluble proanthocyanidins, the pellet was dried with a rotary evaporator. To each sample was added 105 uL of a 95% methanol: 5% concentrate HCl

solution. Samples were transferred to PCR tubes then incubated at 99°C for 20 min. in a thermocycler. Samples were spun down for 5 minute at 13,200 rpm in a microcentrifuge. Another 105 µL of 95% methanol: 5% concentrate HCl solution was added and samples were incubated again for another 20 min. at 99°C. Samples were pool and the absorbance at 530 nm was used to quantify insoluble proanthocyanidins.

Yeast-2 Hybrid Assay

Mediator subunits were cloned into the pBD-GAL4 plasmid. As with prior studies, the MED25 ACID domain was sufficient for interaction with TFs (Elfving et al., 2011; Ou et al., 2011; Çevik et al., 2012). Further analysis only used the middle and ACID domains fragment (BD-MED25$^{228-681}$). Due to size and number of endogenous restriction sites, MED16, MED5a, and MED5b were broken into 3 (BD-MED16$^{1-420}$, BD-MED16$^{421-800}$, and BD-MED16$^{801-1280}$), 4 (BD-MED5a$^{1-227}$, BD-MED5a$^{228-560}$, BD-MED5a$^{561-978}$, and BDMED5a$^{979-1309}$), and 3 (BD-MED5b$^{1-289}$, BD-MED5b$^{290-718}$, and BD-MED5b$^{719-1276}$) fragments, respectively. The full length CDK8 gene was cloned into pBD-GAL4 plasmid (BD-CDK8$^{1-470}$). Components of the MBW complex (AD-TTG1$^{1-341}$, AD-PAP1$^{1-248}$, AD-PAP2$^{1-249}$, AD-TT8$^{1-518}$, AD-GL3$^{1-637}$ or AD-EGL3$^{1-596}$) and JAZ proteins (AD-JAZ1$^{1-253}$, AD-JAZ2$^{1-249}$, AD-JAZ8$^{1-131}$, and AD-JAZ10$^{1-197}$) were cloned into the pAD-GAL4 plasmid. All MBW and JAZ clones were full length genes.

Yeast (*Saccharomyces cerevisiae*) Y2H Gold strain (Clontech, http://www.clontech.com) cultures were grown overnight in YPDA media. YPDA media was prepared with 4 g peptone, 2 g yeast extract, 3 mL 0.2% adenine hemisulfate (prepared in 0.1 M NaOH), and 187 mL water, then autoclaved. After autoclaving 10 mL 40% filter sterilized glucose was added. Cultures were diluted to 0.2 optical density at 600 nm (OD_{600}) in YPDA. Yeast cultures were then grown for 3 to 3.5 h until 0.4 OD_{600}. Samples containing 500 ng AD plasmids and 500 ng BD plasmid were brought up to 50 µL volume with sterile water. Twenty mg of salmon sperm DNA was dissolved in 10 ml TE buffer [10 mM tris(hydroxymethyl)aminomethane (TRIS) and 1 mM ethylenediaminetetraacetic acid (EDTA)] then denatured for 5 min at 95°C. Cell cultures were spun down at 5000 rpm for 6 min to remove YPDA. Cells were resuspended in 20 mL sterile water then spun down at 5000 rpm for 6 min to remove water. Yeast cells were then resuspended in 1 mL 0.1 M lithium acetate (LiAc) and transferred to 1.5 mL microfuge tubes. Samples were washed by spinning down 30 s at 13,000 rpm in a microcentrifuge, removing the LiAC, then resuspending in 1.5 mL of 0.1 M LiAc. In preparation for transformation, 100 µL cell culture in LiAc, 240 µL of filter sterilized 50% polyethylene glycol 3350, 36 µL of 1 M LiAc, 25 µL denature salmon sperm DNA (2 mg/mL), and 50 µL plasmid DNA was sequentially added together in a 1.5 mL microfuge tube. Samples were vortexed for 15 s using medium speed then incubated 30 min at 30°C. Samples were then heat shocked for 20 min at 42°C. Then samples were spun down for 30 s at 7000 rpm to remove supernatant. The pellet was resuspended in 300 µL of sterile water. Next, 150 µL was used to spread yeast on synthetic dropout (SD) media lacking leucine and threonine (SD-LT). Plates were grown for 3 d at 30C, then individual colonies restreaked on SD plates deficient in amino acids -HLT or -AHLT. Colony grew for 3 to 5 d. SD media consisted of 3.35 g SD yeast nitrogen base without amino acids, 320 mg supplement amino acids, 475 mL water and 7.5 g agar which was adjusted to pH 5.8. After autoclaving, 25 mL filter sterilized glucose was added to the media. When needed, 3-Amino-1,2,4-triazole (3AT) was used to suppress autoactivation stimulated colony growth on -HLT plates.

Transient Protoplast Assay

Gene interaction *in vivo* was tested using tobacco (cultivar Xanthi) cell suspension cultures. Cells grew in Murashige and Skoog's media with minimal organics supplemented with KH_2PO_4 (24 mg/L), pyridoxine HCl (0.5 mg/L), 2,4-dichlorophenoxyacetic acid (2,4-D; 0.2 mg/L), kinetin (0.1 mg/L) at pH adjusted to 5.8. Cells were collected by spinning down cultures for at 1100 rpm 4 min. Supernatant was removed and replaced with 30 mL enzyme solution (0.75% cellulose and 0.075% pectinase dissolved in MMC solution). On a shaker in the dark, cells were incubated at 26°C for 2 h. Cells were centrifuged at 1100 rpm for 3 min and then the supernatant removed. Protoplasts were washed using 20 mL of a MMC solution [mannitol (91.1 g/L), 2-morpholinoethane sulfonic acid (MES; 1.95 g/L), and $CaCl_2 \cdot 2H2O$ (1.47 g/L), pH 5.6]. Protoplasts were then resuspended in 10 mL MMC solution then layered on top of a sucrose solution [25% sucrose, MES (1.95 g/L) and $CaCl_2 \cdot 2H2O$ (1.47 g/L), pH 5.6] then centrifuged at 1100 rpm for 4 min. The ring of protoplast formed was removed then transferred to a new tube for resuspension in electroporation buffer [mannitol (91.1 g/L), KCl (5.21 g/L), and MES (975 mg/L), pH 5.6]. Samples were prepared by transferring 750 µL along with 5 µg of reporter (a modified pKYLX80 plasmid containing 5 tandem repeats of the GAL responsive element with minimal CaMV 35S promoter controlling a firefly luciferase gene) and effector plasmids (a modified pBlueScript plasmid where the TF is fused to GAL4 DNA-binding domain and controlled by a mirabilis mosaic virus promoter) into an electroporation cuvette (0.4 cm gap, 200 V, 950 µF; Bio-Rad, Hercules, CA). Samples were transferred to a new tube, centrifuged for 3 min at 1100 rpm, and the supernatant removed. Each sample was resuspended in 1 mL of protoplast culture medium [mannitol (91.1 g/L), MES (1.95 g/L), $CaCl_2 \cdot 2H2O$ (1.47 g/L), KH_2PO_4 (27 mg/L), KNO_3 (101 mg/L), $MgSO_4$ (120 mg/L), KI (2 mg/L) and sucrose (30 g/L) at pH 5.6] then moved to an agarose-coated 25 mm petri dish. Plates were coated by using 1 mL of an agarose solution (9 g mannitol, 1 g low-melting agarose, in 100 mL water). Samples were then cultured for 20-22 h at 26°C.

Protoplasts were collected and placed in 1.5 mL tubes. Samples were centrifuged at 3000 rpm for 3 min and the supernatant removed. Activity was measured using a luciferase assay system according to directions (Promega, Madison, WI). One hundred µL of 1X lysis buffer was added to each tube. Samples were vortexed for 30s to lyses cells. Supernatant was transferred to a new tube after centrifugation for 2 min at 13,200 rpm. Finally, 10 µL cell lysate was added to 50 µL luciferase assay reagent and fluorescence measured using a luminometer (model no. TD2020; Turner Designs, San Jose, CA).

Microarray analysis

Microarray data was downloaded from NCBI and EMBL-EBI Array Express. Arrays were normalized using RMAExpress (Bolstad et al., 2003). Using the MeV software data were analyzed with a Two-way ANOVA with a p-value of 0.05 (Saeed et al., 2003). Probes corresponding to known anthocyanin regulatory and biosynthetic pathway genes were identified. Overlap between microarrays was calculated using a hypergeometric distribution analysis.

Gene Expression

Gene expression from three independent experiments were collected and analyzed. Samples were grown on ½ MS plates for 2 week under a 16:8 day:night photoperiod growth

chamber set at 22C. RNA was extracted using a RNeasy ® Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufactures instructions. cDNA was synthesized as previously described by Schluttenhofer et al. (2014). Each sample was quantified using three technical replicates. At least three biological replicates were used for each experiment.

Statistical Analyses

Statistical analyses of gene expression consisted of Student's T-test calculated in Microsoft Excel. For mutli-time point gene expression experiments, data was analyzed using ANOVA according to the GLM procedure of SAS version 9.3 (SAS Institute, Cary, NC). Hypergeometric distribution analyses determining significant overlap between microarray datasets were calculated using Microsoft Excel.

Accessions

The *med25-3*, myc2 (*jin1-9*), and *gl3- 4*, lines have been previously characterized (Anderson et al., 2004; Kidd et al., 2009; Patra et al., 2013). Overexpression and complement lines were previously reported (Cerdan and Chory, 2003). Identification number for genes used in this study are *PAL1* (AT2G37040), *C4H* (AT2G30490), *4CL1* (AT1G51680), *CHS* (AT5G13930), *CHI* (AT3G55120), *F3H* (AT3G51240), *F3'H/CYP75B1* (AT5G07990), *F3'5'H* (AT4G12330), *DFR* (AT5G42800), *LDOX* (AT4G22880), *UF3GT/UGT79B1* (AT5G54060), *PAP1* (AT1G56650), *PAP2* (AT1G66390), *GL3* (AT5G41315), *EGL3* (AT1G63650), *TT8* (AT4G09820), *TTG1* (AT5G24520), *MED25* (AT1G25540), *MED5a* (AT3G23590), *MED5b* (AT2G48110), *MED16* (AT4G04920), *CDK8* (AT5G63610), *JAZ1* (AT1G19180), *JAZ2* (AT1G74950), *JAZ8* (AT1G30135), and *JAZ10* (AT5G13220).

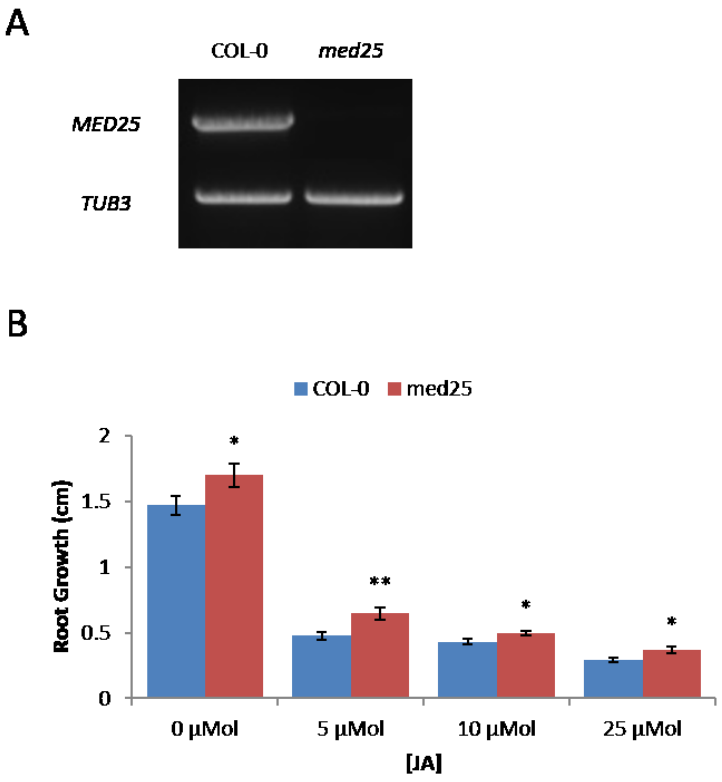


Figure 4.1. Validation of the *med25* mutant. (A) Real-time polymerase chain reaction confirms the med-3 line is a knock-out mutant. (B) Confirming previous results, *med25* mutants are more resistance to JA.

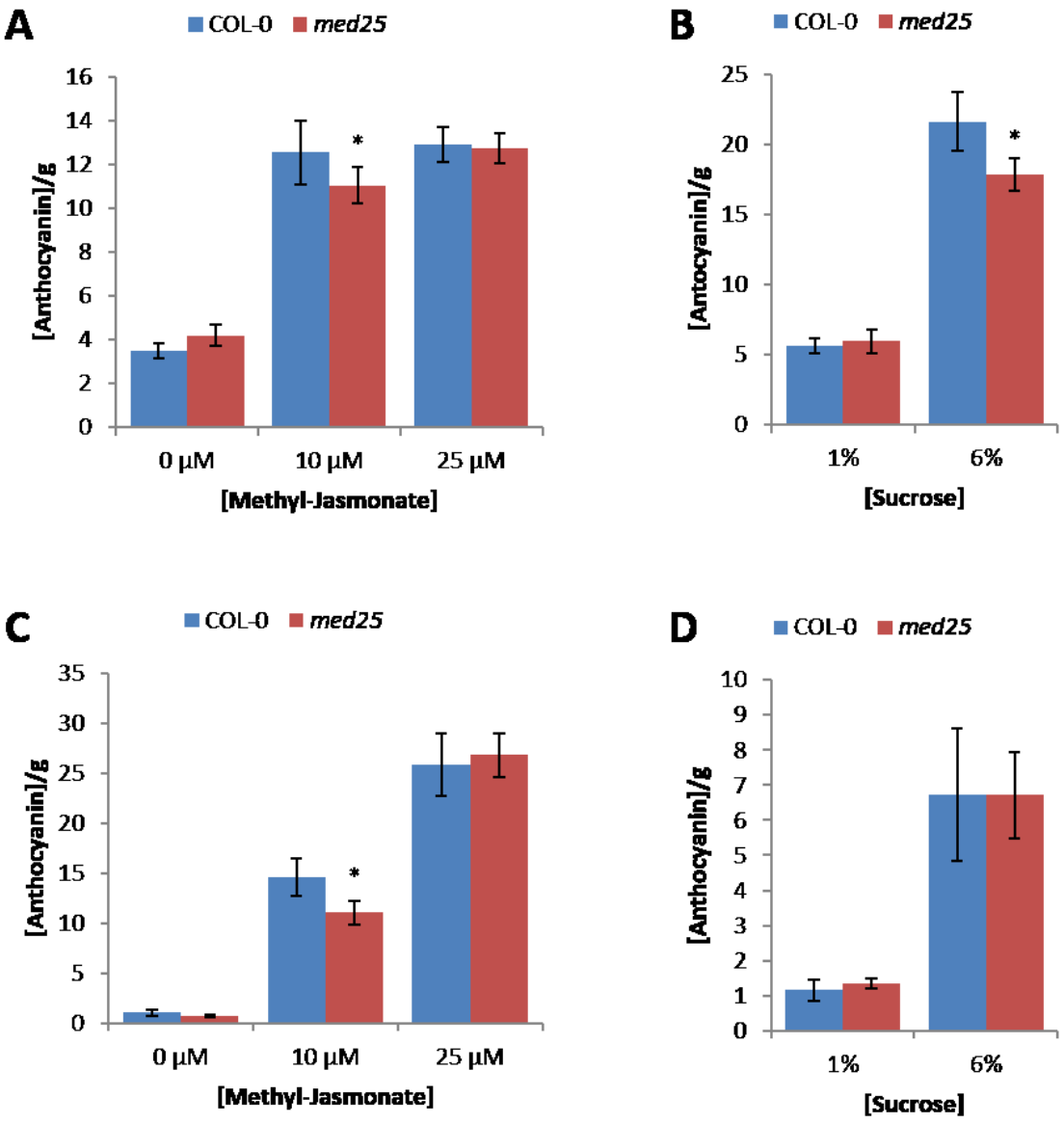


Figure 4.2. MED25 impacts accumulation of anthocyanins. Anthocyanin production in wild-type COL-0 and *med25* was induced with 10 or 25 µM methyl jasmonate (JA), or 6% sucrose. One week old seedlings were transferred to treatment media for 3 (A and B) or 9 days (C and D). * indicates p-value < 0.05.

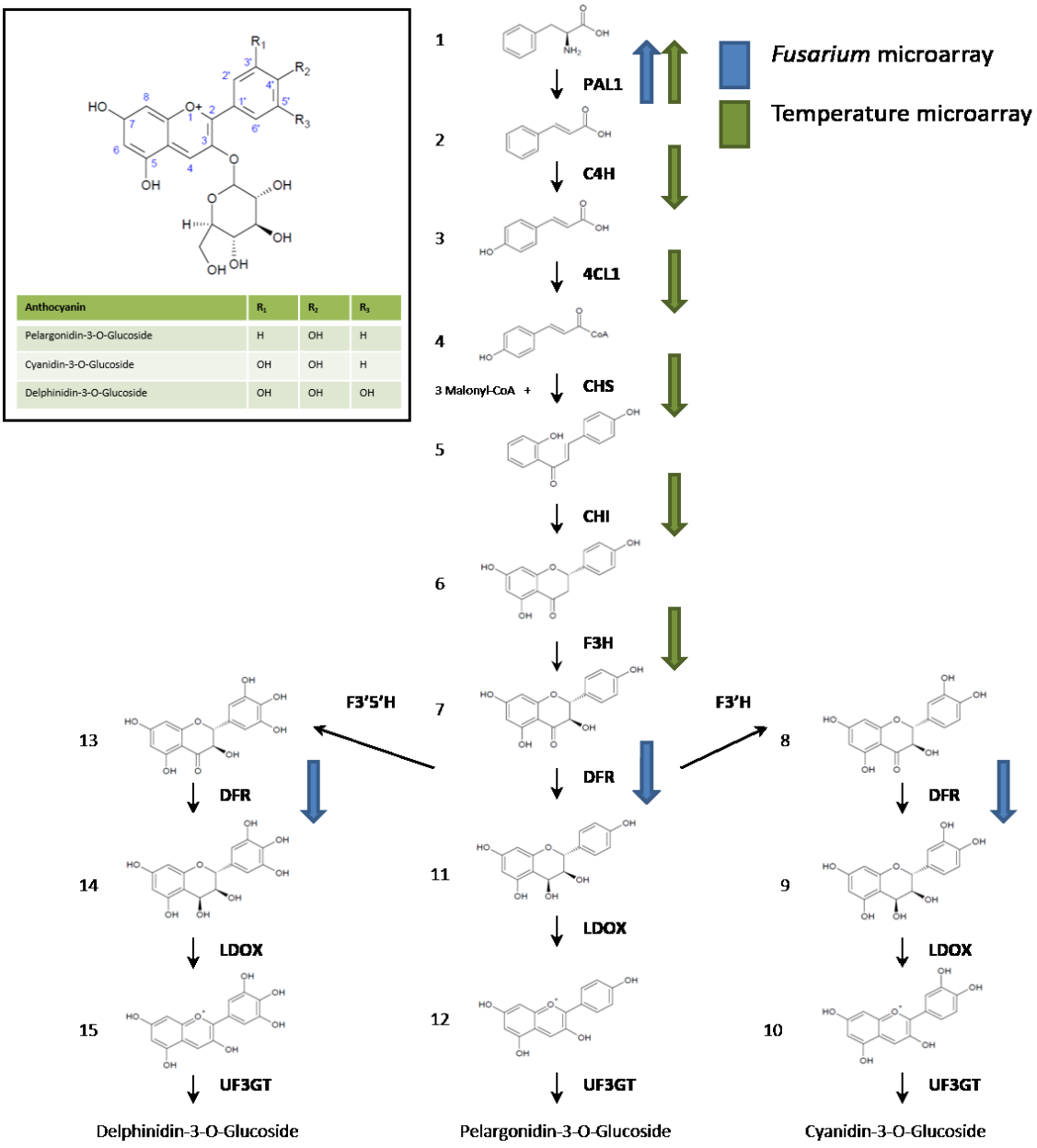


Figure 4.3. The metabolic pathway for the synthesis of anthocyanins. Arrow indicate direction genes changed in either of two publicly available microarray datasets. Metabolites are 1) phenylalanine, 2) cinnamate, 3) *p*-coumarate, 4) *p*-coumaroyl-CoA, 5) naringenin, 6) naringenin, 7) dihydrokaempferol, 8) dihydroquercetin, 9) leucocyanidin, 10) cyanidin, 11) leucopelargonidin, 12) pelargonidin, 13) dihydromyricetin, 14) leucodelphinidin, and 15) delphinidin. Enzymes, listed in bold are: PHENYLALANINE AMMONIUM-LYASE (PAL), (C4H), (4CL), CHALCONE SYNTHASE (CHS), CHALCONE ISOMERASE (CHI), FLAVANONE 3-HYDROXYLASE (F3H), FLAVANONE 3'-HYDROXYLASE/CYP75B1 (F3'H), FLAVANONE 3'5'-HYDROXYLASE/CYP706A7 (F3'5'H), DIHYDROFLAVONOL REDUCTASE (DFR), LEUCOANTHOCYANIN DIOXYGENASE (LDOX), UDP-GLUCOSE:FLAVONOID 3-O-GLUCOSYLTRANSFERASE/UGT79B1 (UF3GT).

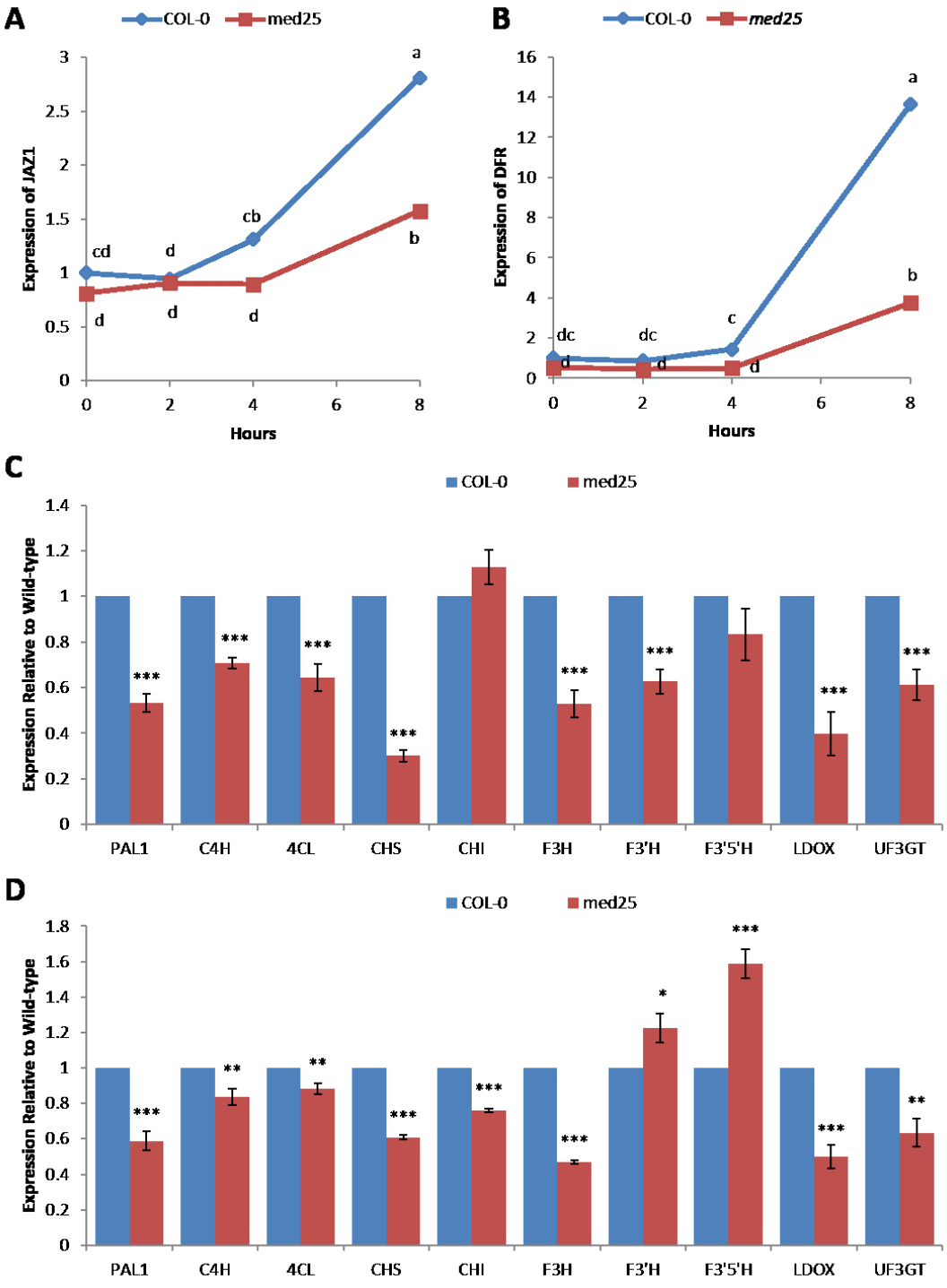


Figure 4.4. Expression of anthocyanin biosynthetic genes are regulated by MED25. Gene expression (A) of a jasmonate inducible control gene JASMONATE ZIM-DOMAIN PROTEIN1 (JAZ1), DFR (B), and anthocyanin biosynthetic genes were measured in the wild-type and *med25* mutant (C) before and (D) after 8 hours of 100 µM JA treatment. Bars indicate standard errors of samples performed with three biological replicates.

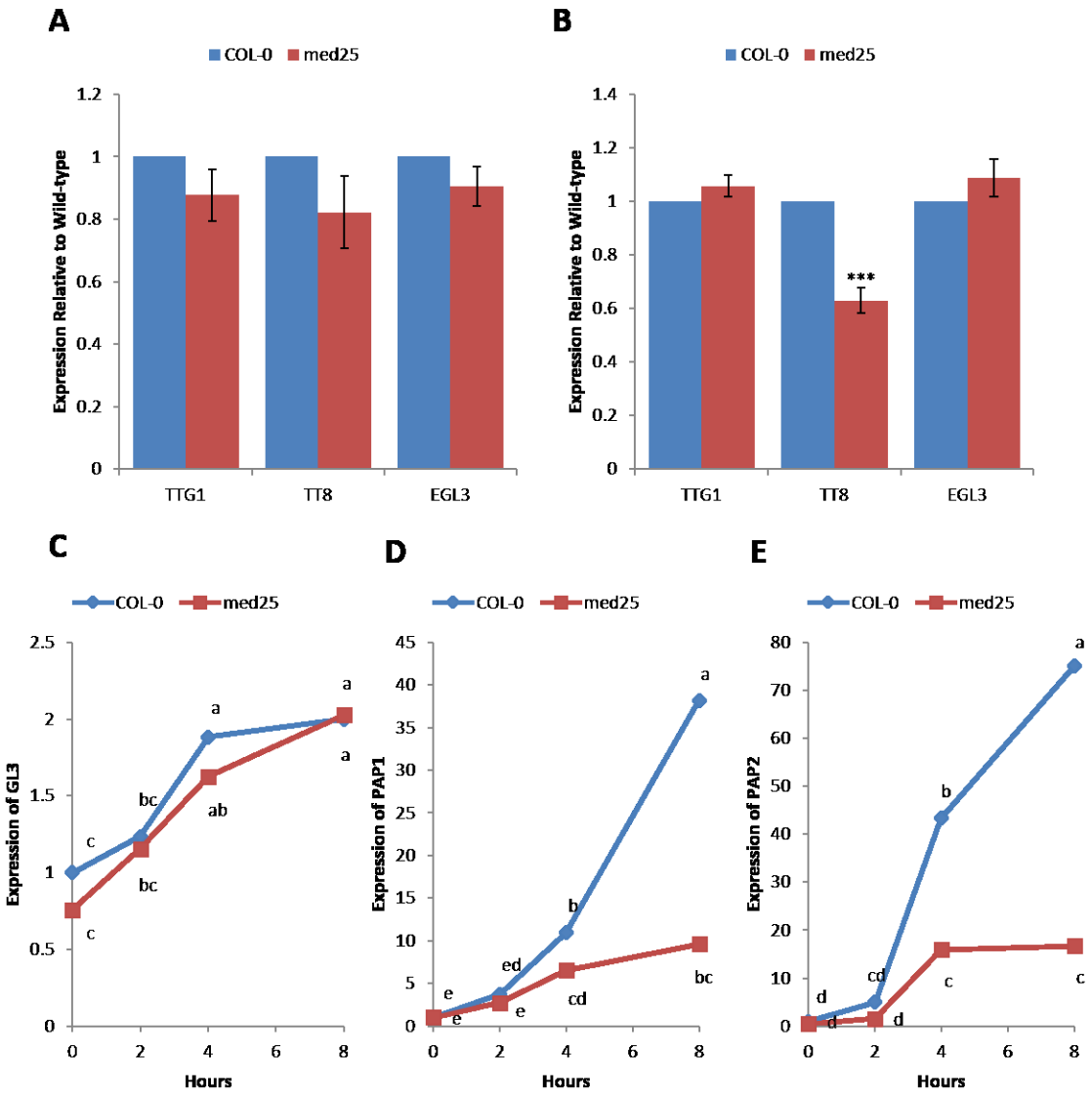


Figure 4.5. Expression of anthocyanin regulatory genes is dependent upon MED25. Expression of MYB-bHLH-WD40 complex members TTG1, TT8, and EGL3 in wild-type and the *med25* mutant were measured before (A) and after 8 hours (B) of JA induction. A JA induced time course analysis of the jasmonate regulated genes GL3 (C), PAP1/MYB75 (D), and PAP2/MYB90 (E) was performed in the wild-type and *med25* mutant seedlings. Bars indicate standard errors of samples performed with three biological replicates.

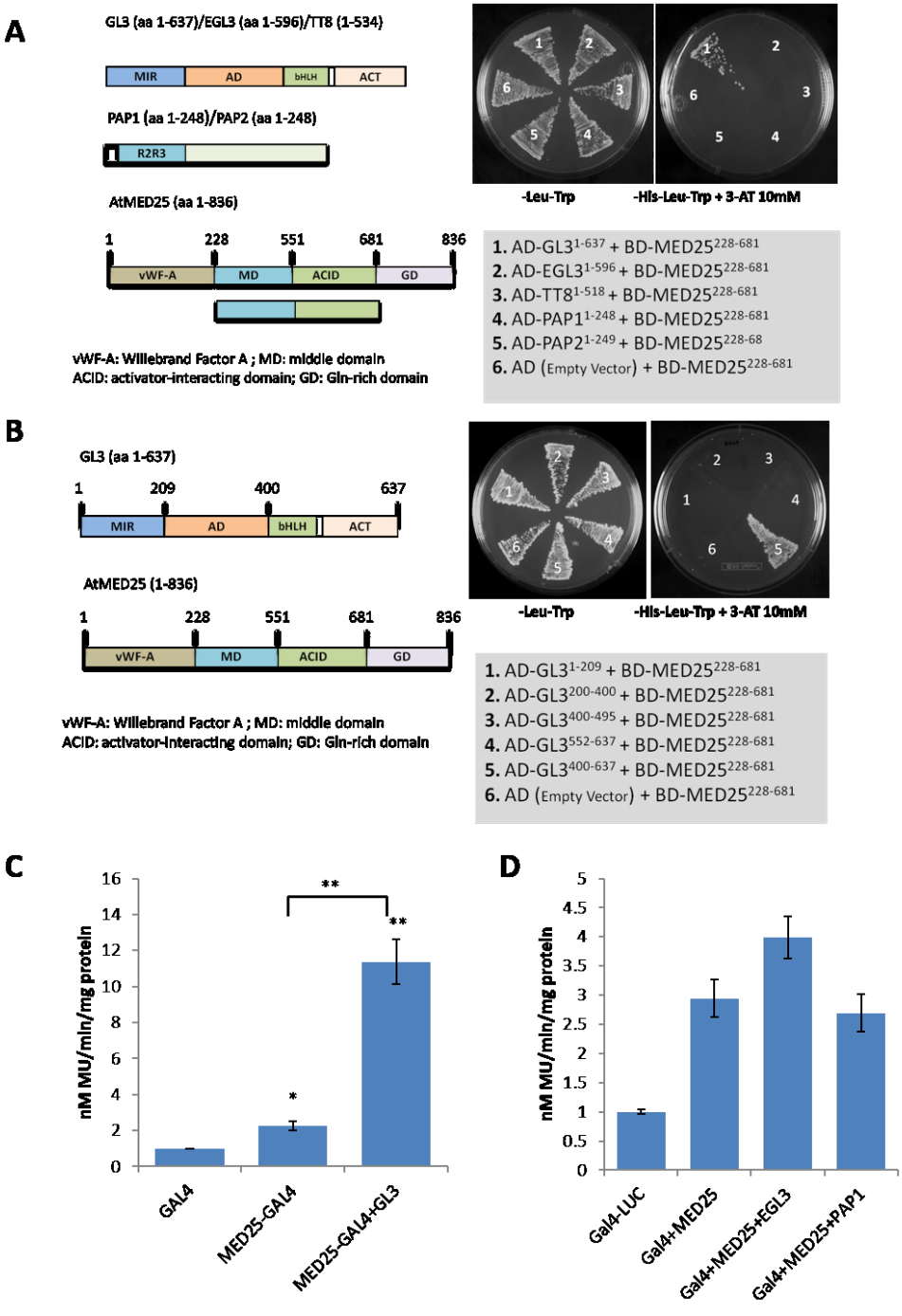


Figure 4.6. MED25 interacts with GL3, but not other members of the MBW complex. The interaction of MED25 with MBW complex subunits was tested using a yeast-2 hybrid system. (A) MED25 was fused with the GAL4 binding domain (BD) and MBW transcription factors were fused with the GAL4 activation domain (AD). (B) The region of GL3 which interacts with the MED25 ACID domain was identified using a Y2H system. (A and B) 10 mM 3-AT was included to suppress autoactivation of colony growth caused by MED25. (C and D) MED25 was fused with the GAL4 BD and was used to activate a luciferase reporter in tobacco protoplast. MBW complex subunits were used without a GAL4 AD fusion. Bars indicate standard errors of samples.

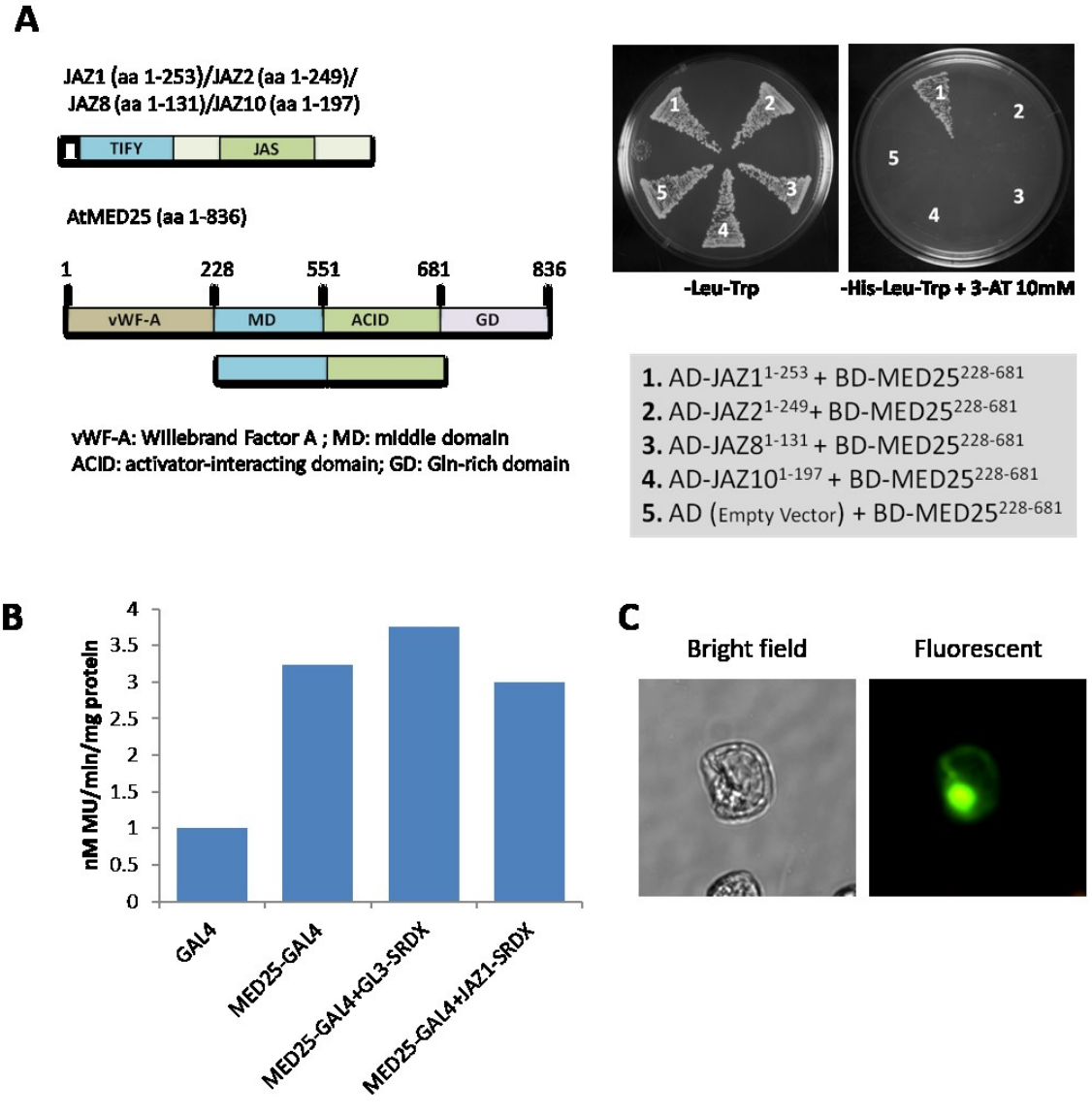


Figure 4.7. MED25 interacts with JAZ1. The interaction of MED25 with JAZ proteins was tested using a yeast-2 hybrid system. (A) MED25 was fused with the GAL4 BD and JAZ factors were fused with the GAL4 activation domain (AD). 10 mM 3-AT was included to suppress autoactivation of colony growth caused by MED25. (B) MED25 was fused with the GAL4 BD and was used to activate a luciferase reporter in tobacco protoplast. Both GL3 and JAZ1 were used without a GAL4 AD fusion. The SRDX repression domain was fused to GL3 and JAZ1. (C) MED25 fused with GFP revealed protein localization.

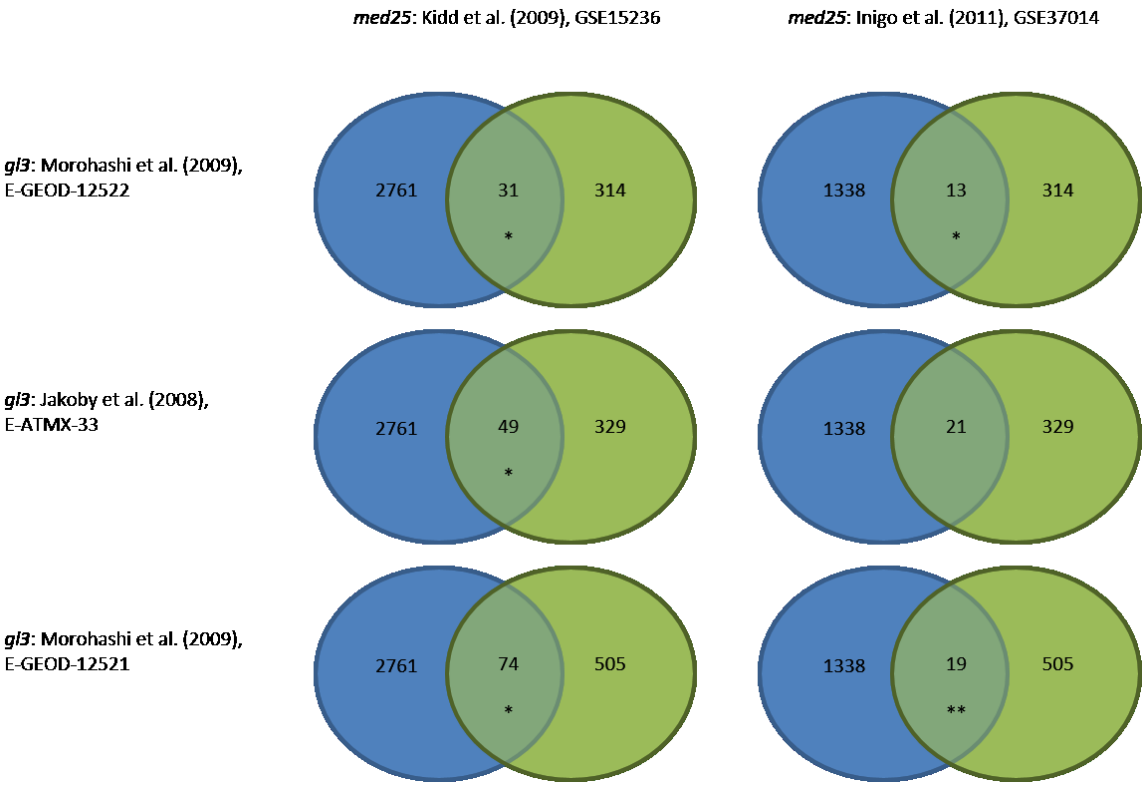


Figure 4.8. MED25 and GL3 regulate common target genes. A hypergeometric distribution analysis of *med25* and *gl3* regulated genes was performed using publicly available microarray datasets. Identifiers of arrays used to identify MED25 and GL3 differentially expressed genes are located at the top and left portions of the figure, respectively. Genes significantly different in *med25* and *gl3* microarrays are indicated in blue and green, respectively. The number of gene overlapping is depicted in the overlapping regions of circles. A population size of 21197 genes was used. Microarray control probes were excluded from analysis. * and ** indicate significance at p-value = 0.05 or 0.01, respectively.

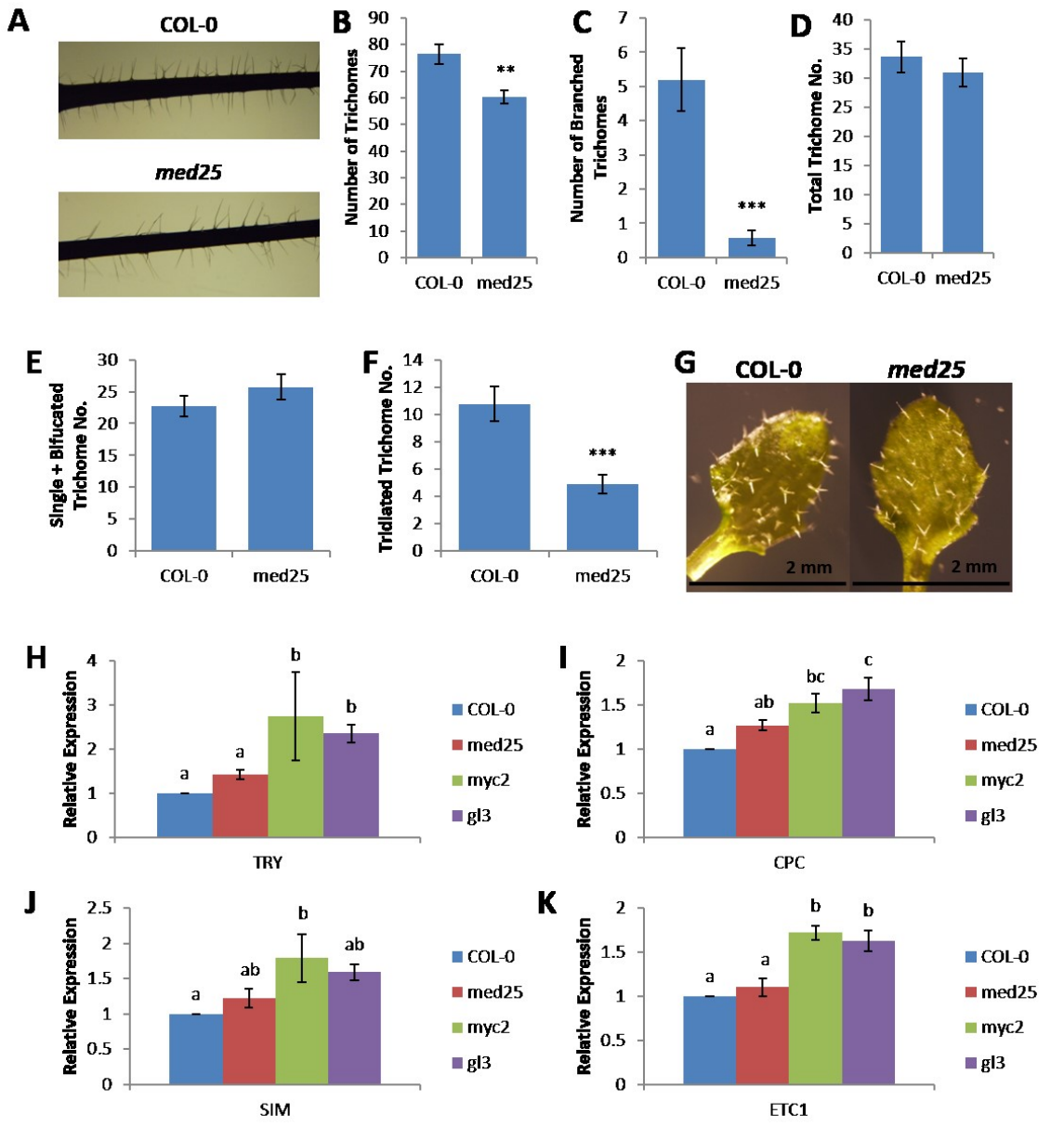


Figure 4.9. Regulation of trichome development by MED25. Stem trichomes of WT and *med25* mutant plants (A). The number of total trichomes (B) and branched trichomes (C) on stems of WT and *med25* mutant plants. The number of total (D), single and bifurcated (E), and triradiate (F) trichomes on third and fourth leaves of WT and *med25* seedlings (G). Expression of trichome development genes *TRIPTYCHON* (*TRY*), *CAPRICE* (*CPC*), *SIAMESE* (*SIM*), and *ENHANCER OF TRY AND CPC 1* (*ETC1*) in wild-type, *med25*, *myc2*, and *gl3* mutants were measured (H-K). Bars indicate standard errors of samples.

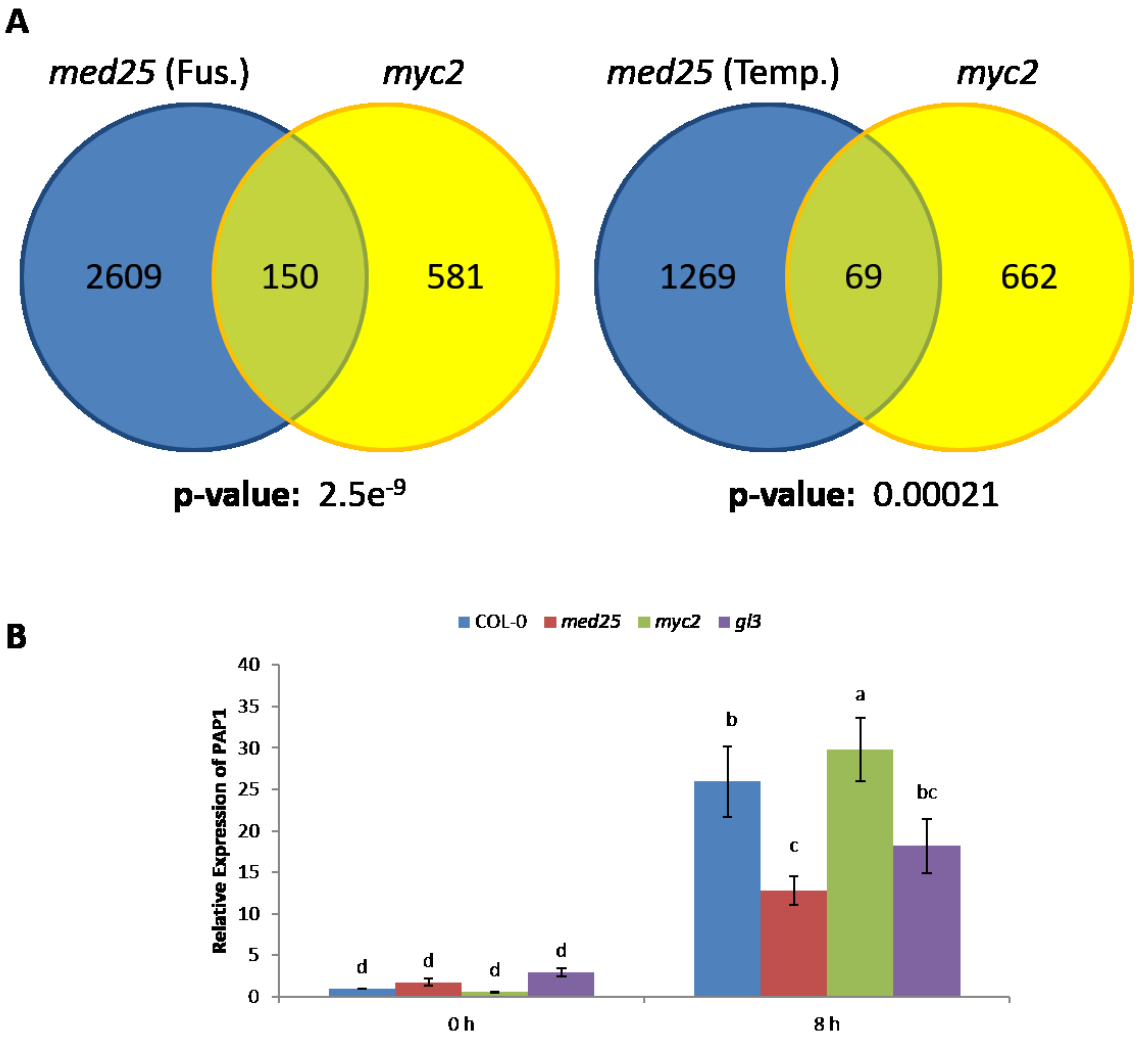


Figure 4.10. MED25 and MYC2 regulate common target genes. (A) A hypergeometric distribution analysis of *med25* and *gl3* regulated genes was performed using publicly available microarray datasets. Genes significantly different in *med25* and *gl3* microarrays are indicated in blue and green, respectively. The number of gene overlapping is depicted in the overlapping regions of circles. (B) Gene expression of PAP1 was measured in WT, *med25*, *myc2*, and *gl3* seedlings. Different letters indicate statistically significant differences at p-value = 0.05. Bars indicate standard errors of samples performed with three biological replicates.

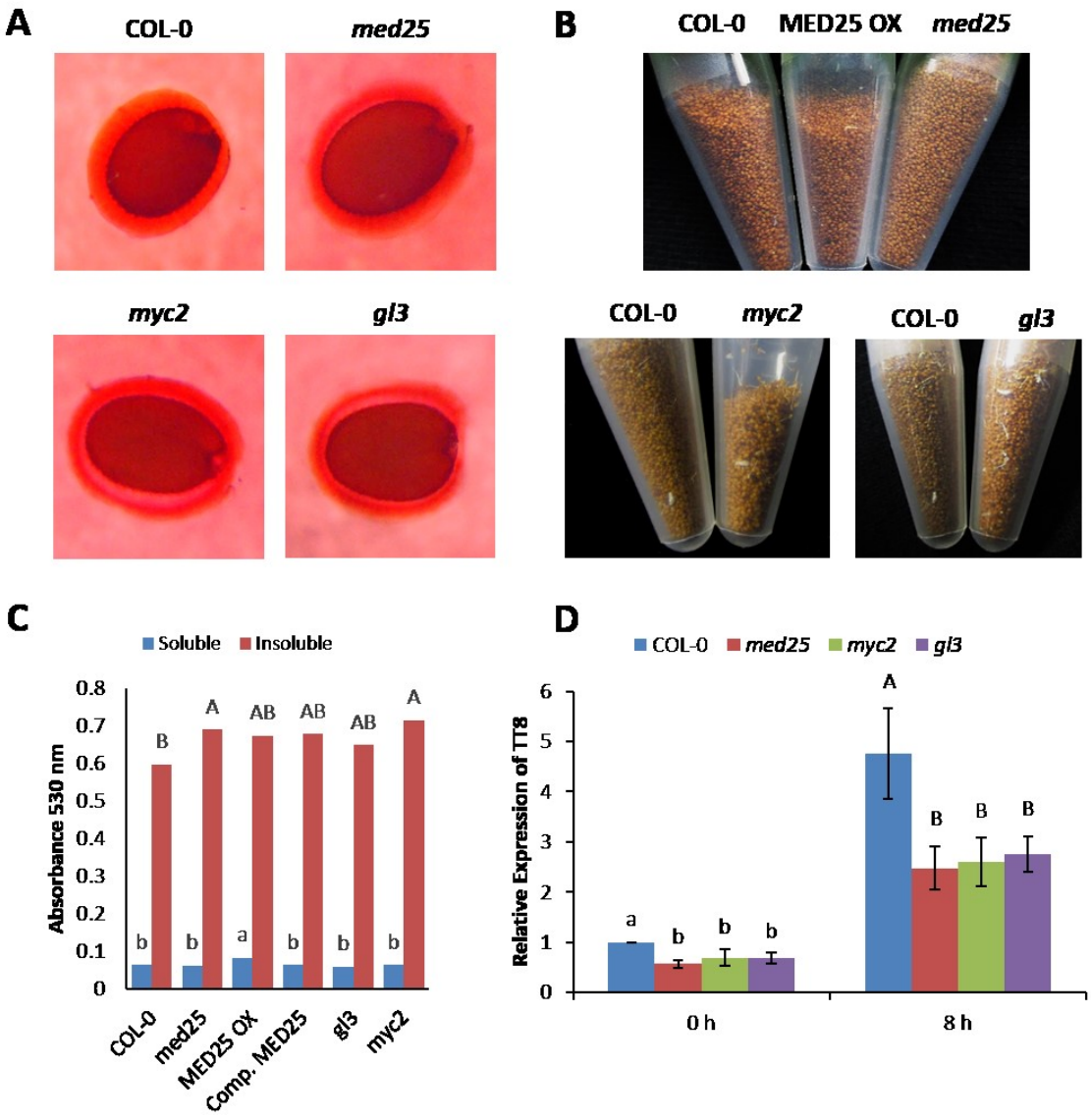


Figure 4.11. Accumulation of mucilage and proanthocyanidins. (A) WT *med25*, *myc2*, and *gl3* mutant seeds were imbibed for 1 h then stained with ruthenium red. (B) Visible differences in testa color were observed between WT, MED25 OX, *med25*, *myc2*, and *gl3* seeds. (C) Soluble and insoluble proanthocyanidins were extracted and quantified from MED25 lines and related mutant lines. (D) Gene expression of *TT8* was measured before and 8 h after induction with JA. Bars indicate standard errors of samples performed with three biological replicates.

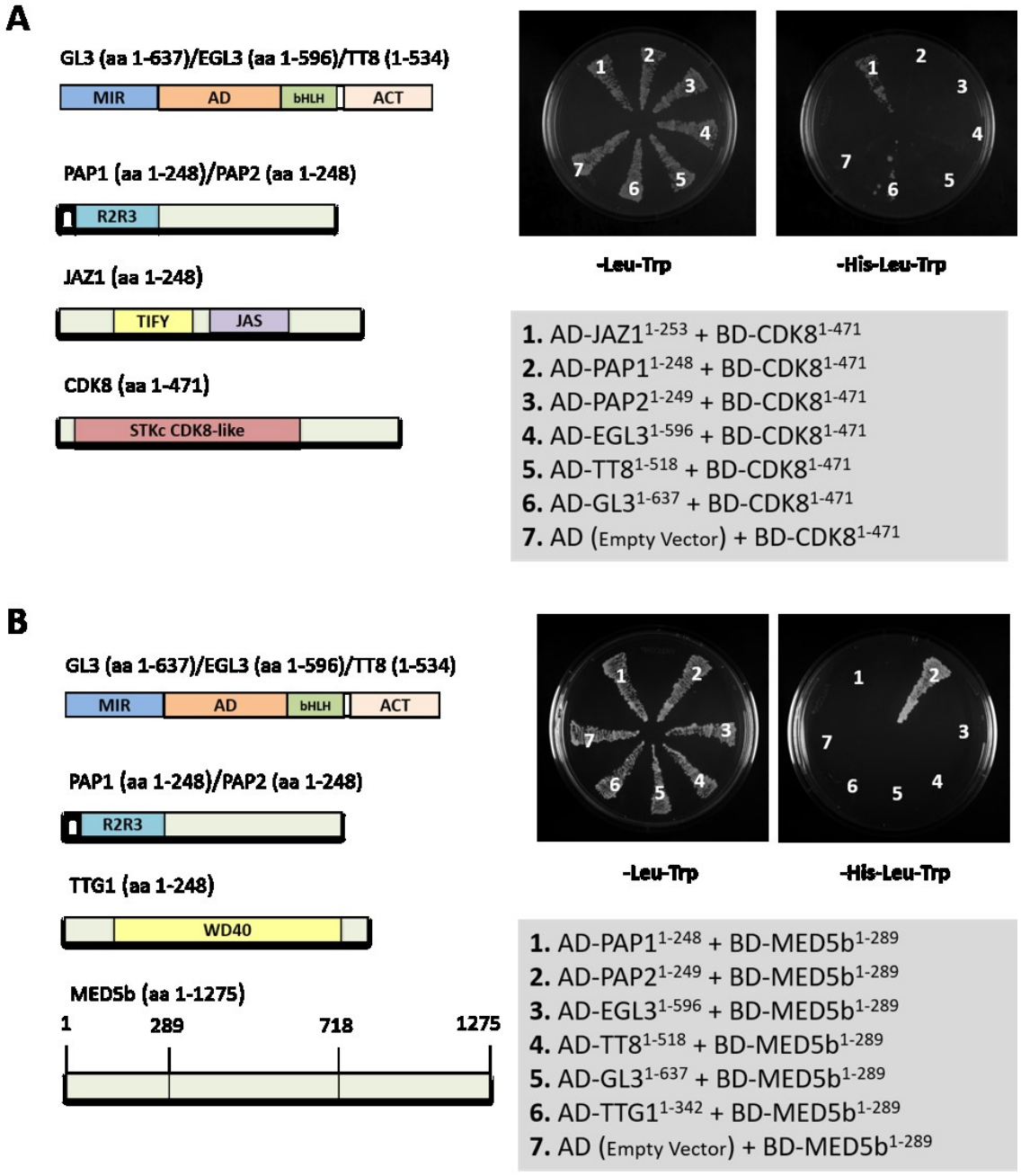


Figure 4.12. Interaction of the MBW and Mediator complex subunits. (A) Interaction of CDK8 and MBW complex subunits was test using a yeast-2 hybrid system. CDK8 was fused with the GAL4 binding domain (BD), while MBW transcription factors and JAZ1 were fused with the GAL4 activation domain (AD). (B) Interaction of MED5b and MBW complex subunits was test using a yeast-2 hybrid system. MED5b was fused with the GAL4 binding domain (BD) and MBW transcription factors were fused with the GAL4 activation domain (AD).

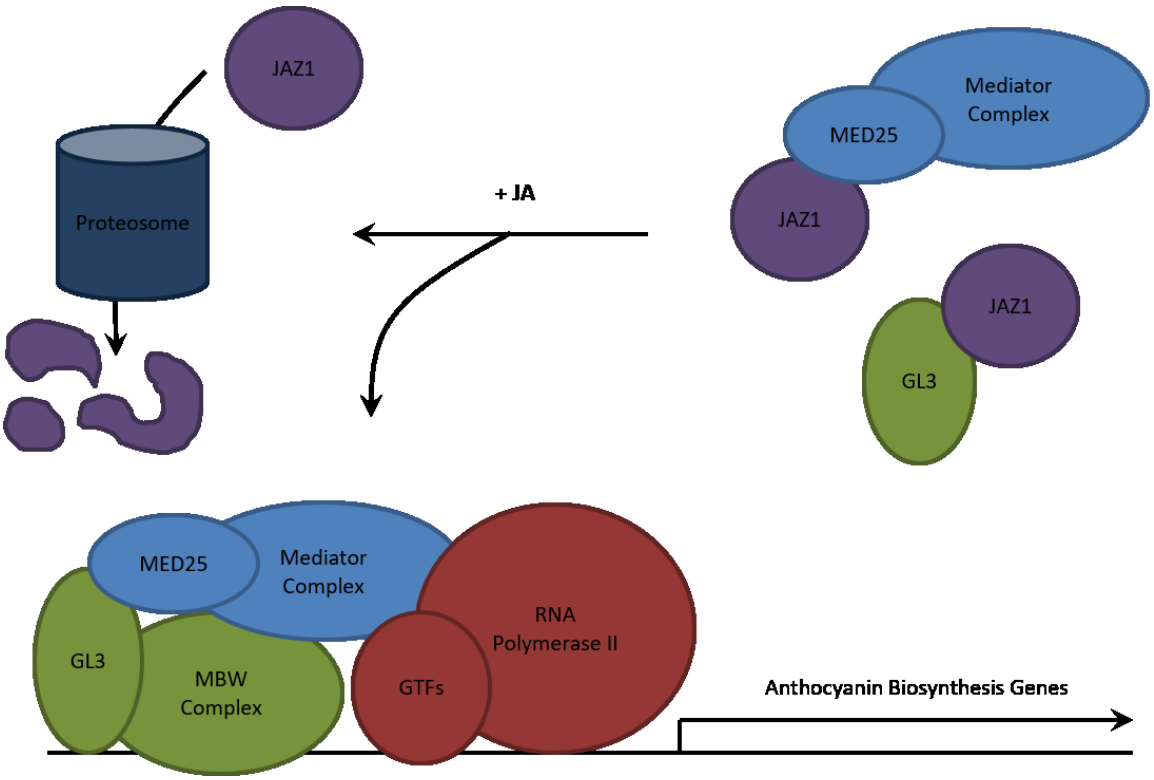


Figure 4.13. A model for the regulation of anthocyanins by MED25. I propose a mechanism where GL3 and JAZ1 compete for the interaction with MED25 to regulate expression of anthocyanin biosynthetic genes. In the absence of JA JAZ1 binds both MED25 and GL3, preventing accidental induction of anthocyanin production. Upon JA induction, JAZ1 gets degraded by the proteasome subsequently freeing both MED25 and GL3 which permits their physical interaction. The MBW complex binds promoters of late stage anthocyanin biosynthetic genes. The Mediator complex then recruits RNA Polymerase II and associated general transcription factors (GTF) for gene transcription.

CHAPTER FIVE: SUMMARY AND FUTURE DIRECTIONS

Summary

Chapters two of this work has been published (Schluttenhofer et al., 2014). Additionally, portions of the Introduction have been published in a prior review (Patra et al., 2013). Chapter three is ready for submission. Several experiments remain to finish chapter four. Additionally, future studies need to be conducted to help address further question raised by my research.

MED25 Regulation of Anthocyanins

To complete the MED25 work, several experiments need completed. Foremost, further evidence is needed to validate the interaction of MED25 with JAZ1. I plan to accomplish this by performing a GST pull-down assay *in vitro* and a pull-down assay *in vivo*. A pull-down assay will also be conducted to further support the MED25 interaction with GL3. For the *in vitro* pull-down assay using *Escherichia coli*, the GST will be fused to the N-terminal end of the MED25 ACID domain. The HIS domain will be fused to the C-terminal region of the JAZ1 protein. The ACID-GST protein will be bound to the column of glutathione containing beads and then washed to remove excess protein. The prey protein (JAZ1-HIS) will be eluted through the column and washed to remove unbound protein. The bound proteins will be eluted and run on an SDS-PAGE gel to reveal protein interaction. For the *in vivo* pull-down assay eGFP and FLAG will be fused to the C-terminal end of full length MED25 and JAZ1, respectively. Constructs will be infiltrated into *Nicotiana benthamiana* leaves. M2 agarose gel will be used to bind the JAZ1-FLAG protein. Then, plant protein extract containing MED25-eGFP will we run through the column. Bound protein will be eluted and run on an SDS-PAGE gel to show the presence of both proteins. As a Y2H assay indicated the interaction of CDK8 with GL3 and JAZ1, I plan to validate this interaction *in vivo* using a protoplast-based transactivation assay. This will be performed by fusing the CDK8 to the GAL4 domain. An increase in background activation will be used to indicate a functional CDK8 protein *in vivo*. As JAZ1 in combination with MED25 led to increased trans-activation, I will use the construct where JAZ1 is fused with the SRDX repression domain. After cotransfection into protoplast, a decrease in transactivation will indicate the interaction of CDK8-GAL4 and JAZ1-SRDX *in vivo*. Similarly, the interaction between MED5b and PAP2 will be validated *in vivo* using the protoplast-based transactivation assay system. As with MED25, MED5b will be fused with the GAL4 domain. Cotransfection of MED5b with PAP2 should result in activation of the luciferase reporter construct. For all protoplast-based transactivation assays, expression of the GUS reporter will be used to normalize for differences in protein levels. Interaction of proteins *in vivo* will further support the physical association of these factors.

GL3 plays a key role in regulating induction of anthocyanins in response to nitrogen deficiency (Feyissa et al., 2009). As MED25 function to regulate anthocyanin by its interaction with GL3, but not EGL3, I suspect the anthocyanin phenotype will be more prominent under nitrogen deficiency than induction by JA. Thus, I plan to compare the response of *gl3* and *med25* mutants, MED25 complemented, and MED25 OX lines to WT plants stressed by nitrogen deficiency. I expect *med25* and *gl3* mutants will display an impairment of anthocyanin accumulation under this condition. To further investigate this phenotype, expression of anthocyanin and proanthocyanidin genes will be tested in both med25 and MED OX lines. This

data will add to the list of phenotypes shared by *med25* and *gl3* mutants and provide further support that interaction of these factors is needed for activation of flavonoid biosynthesis.

The expression of *TT8* and *PAP1* was reduced in *med25* mutants. To determine if either GL3 or MYC2 were involved in regulating these factors I tested *TT8* expression in the respective mutant lines. PAP1 expression was reduced in the *myc2* mutant, although this was not statistically significant. Additionally, *TT8* expression was significantly reduced in both the *gl3* and *myc2* mutant. To validate the findings in mutant lines, I will clone the PAP1 and TT8 promoters as well as the GL3 and MYC2 TFs. Using a transient transactivation assay I will test whether GL3 and MYC2 can activate the TT8 promoter. Similarly, I will test if MYC2 activates the PAP1 promoters. PAP1 and TT8 promoters will be cloned in front of the luciferase reporter. GL3 and MYC2 will be expressed using the pBlueScript plasmid containing the GUS reporter which permits normalization for variation in protein levels. Promoter and TF containing plasmids will cotransfected into tobacco protoplasts. The activity of the GUS and luciferase reporters will be used to quantify gene activation.

Future Projects

Despite the progress I have made here, many question regarding JA signaling and regulation of specialized metabolism remain. For example, does MED25 function to regulate other classes of specialized metabolites? How does JA signaling interact with other phytohormone pathways to regulate TIA production? What other families of TF are involved in regulating TIA production? Do orthologs retain a conserved function in regulating specialized metabolism between species? If not, what causes orthologs to function differently? Ongoing and future project will address these questions.

Here I demonstrated MED25 interacts with the TF GL3. The role for the MED25 interaction with GL3 may be to regulate anthocyanin accumulation and trichome production. Future projects will include analyzing how MED25 affects the regulation, biosynthesis, and accumulation of other specialized metabolites. Gene expression analysis of the phenylpropanoid pathway suggests MED25 regulates flavonol and lignin biosynthesis. Flavonols contribute to regulation of auxin signaling by suppressing polar auxin transport (Yin et al., 2014). Indeed, the *med25* mutant is altered in polar auxin transport (Raya-González et al., 2014), but the causal mechanism remains to be identified. Disruption of flavonol signaling in *med25* could result from decreased levels of endogenous flavonols, thereby increasing polar auxin transport. Enhanced polar auxin transport would result in increased auxin response phenotypes, including longer primary and lateral roots, as well as initiation of more adventitious and lateral roots. These phenotypes occur in the *med25* mutant (Raya-González et al., 2014). How MED25 regulates alkaloids and terpene also remains to be studied. As a factor important for JA signaling response, the interaction of MED25 with Catharanthus TFs should reveal important insights into the TIA regulatory network.

To date, most studies in Catharanthus have focused on JA signaling. However, other phytohormones also impact TIA production (Pan et al., 2010). How does JA signaling interact with other phytohormone pathways to regulate TIA production? Auxin has long been known to repress TIA production (Goddijn et al., 1992; Gantet et al., 1997; Whitmer et al., 1998), yet it remains poorly understood how this occurs. Particularly, how different phytohormones cooperate to regulate TIA production remains unknown. Furthermore, it is possible inhibition of

phytohormones which suppress the TIA pathway could boost metabolite yield. Several auxin inhibitors are known, including the flavonol naringenin and the synthetic compound 2,3,5-triiodobenzoic acid (TIBA). If either of these can induce TIA production remains untested. Combining JA and TIBA could possibly enhance TIA biosynthesis and accumulation or one hormone may override the other. In either circumstance, our knowledge of TIA regulation will be enhanced.

With the recently sequence Catharanthus genome published (Kellner et al., 2015), promoters for TIA biosynthetic and regulatory genes can now be easily acquired. Additionally, this allows for the easy identification of whole TF families. This information can be used to address the question, what other families of TF are involved in regulating TIA production? Analysis of promoters will provide candidate gene families important for TIA pathway regulation. Phylogenetic analysis of these important TF families can be used to identify orthologs of known regulators of specialized metabolism occurring in other species. Recently, correlation of TFs with pathway genes has proven a successful approach to identifying candidate regulators (Van Moerkercke et al., 2015; Moerkercke et al., 2016). Correlation of an entire TF family with TIA biosynthetic and regulatory genes can reveal uncharacterized candidate factors. Combining both phylogenetic and correlative methods can provide promising candidate regulators for further study. Characterization of candidate factors will establish the role of additional TF families in regulating TIA production. Overall, identification of additional regulators will help clarify the regulatory network governing TIA biosynthesis. Eventually, pathway specific master regulators will be discovered that can be used for genetically engineering improved TIA production in Catharanthus plants, hairy roots, or cell cultures.

Orthologs are genes evolutionarily descended from a common ancestral sequence but occur in distinct species due to one or more speciation events. While not a formal part of the definition, orthologs are believed to retain similar functions across species. Indeed, the process of genome annotation inherently involves using orthologous relationships to predict gene function. However, the role of orthologs in specialized metabolism remains unclear. Many plants do not produce the same specialized metabolites (e.g. Arabidopsis does not produce TIA found in Catharanthus). Yet, some components of phytohormone (e.g. MYC2) or other signaling pathways are conserved across species. Do these orthologs retain a conserved function in regulating specialized metabolism between species? If not, what causes orthologs to function differently? Or do both conserved and non-conserved functions occur between orthologs? With the falling cost of transcriptome and genome sequencing, the past several years have observed an increase in genes of medicinal plants. This has been accompanied by increased identification of biosynthetic genes and regulators in specialized metabolism. For some factors, there are now enough orthologs to compare functions across species. And differences between orthologs are already easily apparent. For example, consider the orthologous Arabidopsis and Catharanthus MYC2 sequences. In Arabidopsis, MYC2 regulates the production of glucosinolate, flavonoids, and sequiterpenes (Dombrecht et al., 2007; Hong et al., 2012), but in Catharanthus it regulates expression of TIAs by targeting ORCA3 (Zhang et al., 2011). While both factors regulated specialized metabolism, the regulators activating or repressing the pathways and biosynthetic enzymes are vastly different. Furthermore, Arabidopsis does not have a clear ORCA3 ortholog. Thus, while the MYC2 genes are orthologs and they even retain a conserved function in JA response to elicit specialized metabolism, there must be considerable differences in the transcriptional regulatory networks controlled by these factors. Comparison of orthologs between medicinal species can thus provide information on how TFs retained or diverged functions during plant evolution. From a metabolic engineering perspective, evolutionary

differences between species is important. If the transferred gene does not have a conserved role between species or the target pathway is not regulated by a similar network as the host factor, then an unanticipated phenotype could result. On the other hand, use of functionally divergent factors could potentially lead to ectopic biosynthetic gene expression and result in novel chemical products.

Since their identification as valuable anti-cancer drugs in the 1960's, researchers have sought to improve vinblastine and vincristine yield. Five decades later and this goal remains unaccomplished. However, the last two decades has seen rapid developments in understanding JA signaling. Promisingly, during the last half decade important strides have been made in understanding TIA biosynthesis and regulation. The rapid development of new research technologies now allows studies which were once impossible. All the aforementioned future projects are underway and should further improve our understanding of JA signaling and TIA regulation. Armed with knowledge from these studies, perhaps a half century old goal will finally be attainable, or at least one step closer to accomplishment.

REFERENCES

Aerts RJ, De Luca V (1992) Phytochrome Is Involved in the Light-Regulation of Vindoline Biosynthesis in Catharanthus. Plant Physiology **100:** 1029-1032

Aerts RJ, Gisi D, De Carolis E, De Luca V, Baumann TW (1994) Methyl jasmonate vapor increases the developmentally controlled synthesis of alkaloids in Catharanthus and Cinchona seedlings. The Plant Journal **5:** 635-643

Alam P, Abdin MZ (2011) Over-expression of HMG-CoA reductase and amorpha-4,11-diene synthase genes in Artemisia annua L. and its influence on artemisinin content. Plant Cell Reports **30:** 1919-1928

Aldridge D, Galt S, Giles D, Turner W (1971) Metabolites of Lasiodiplodia theobromae. Journal of the Chemical Society C: Organic**:** 1623-1627

Anderson JP, Badruzsaufari E, Schenk PM, Manners JM, Desmond OJ, Ehlert C, Maclean DJ, Ebert PR, Kazan K (2004) Antagonistic Interaction between Abscisic Acid and Jasmonate-Ethylene Signaling Pathways Modulates Defense Gene Expression and Disease Resistance in Arabidopsis. The Plant Cell **16:** 3460-3479

Ang L-H, Chattopadhyay S, Wei N, Oyama T, Okada K, Batschauer A, Deng X-W (1998) Molecular Interaction between COP1 and HY5 Defines a Regulatory Switch for Light Control of Arabidopsis Development. Molecular Cell **1:** 213-222

Aquil S, Husaini AM, Abdin MZ, Rather GM (2009) Overexpression of the HMG-CoA Reductase Gene Leads to Enhanced Artemisinin Biosynthesis in Transgenic Artemisia annua Plants. Planta Med **75:** 1453-1458

Asada K, Salim V, Masada-Atsumi S, Edmunds E, Nagatoshi M, Terasaka K, Mizukami H, De Luca V (2013) A 7-Deoxyloganetic Acid Glucosyltransferase Contributes a Key Step in Secologanin Biosynthesis in Madagascar Periwinkle. The Plant Cell Online **25:** 4123-4134

Bäckström S, Elfving N, Nilsson R, Wingsle G, Björklund S (2007) Purification of a Plant Mediator from Arabidopsis thaliana Identifies PFT1 as the Med25 Subunit. Molecular Cell **26:** 717-729

Bai Y, Meng Y, Huang D, Qi Y, Chen M (2011) Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. Genomics **98:** 128-136

Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. The Plant Journal **39:** 366-380

Bernhardt C, Lee MM, Gonzalez A, Zhang F, Lloyd A, Schiefelbein J (2003) The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the Arabidopsis root. Development **130:** 6431-6439

Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M (2007) Flavonoid Accumulation in Arabidopsis Repressed in Lignin Synthesis Affects Auxin Transport and Plant Growth. The Plant Cell **19:** 148-162

Besseau S, Kellner F, Lanoue A, Thamm AMK, Salim V, Schneider B, Geu-Flores F, Höfer R, Guirimand G, Guihur A, Oudin A, Glevarec G, Foureau E, Papon N, Clastre M, Giglioli-Guivarc'h N, St-Pierre B, Werck-Reichhart D, Burlat V, De Luca V, O'Connor SE, Courdavault V (2013) A Pair of Tabersonine 16-Hydroxylases Initiates the Synthesis of Vindoline in an Organ-Dependent Manner in Catharanthus roseus. Plant Physiology **163:** 1792-1803

Bhargava A, Mansfield SD, Hall HC, Douglas CJ, Ellis BE (2010) MYB75 Functions in Regulation of Secondary Cell Wall Formation in the Arabidopsis Inflorescence Stem. Plant Physiology **154:** 1428-1438

Bolstad BM, Irizarry RA, Åstrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics **19:** 185-193

Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, Chapple C (2014) Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. Nature **509:** 376–380

Bonawitz ND, Soltau WL, Blatchley MR, Powers BL, Hurlock AK, Seals LA, Weng J-K, Stout J, Chapple C (2012) REF4 and RFR1, Subunits of the Transcriptional Coregulatory Complex Mediator, Are Required for Phenylpropanoid Homeostasis in Arabidopsis. Journal of Biological Chemistry **287:** 5434-5445

Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation Tagging Identifies a Conserved MYB Regulator of Phenylpropanoid Biosynthesis. The Plant Cell **12:** 2383-2393

Bradshaw HD, Schemske DW (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. Nature **426:** 176-178

Briggs DEG (2015) The Cambrian explosion. Current Biology **25:** R864-R868

Burlat V, Oudin A, Courtois M, Rideau M, St-Pierre B (2004) Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of Catharanthus roseus implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-derived primary metabolites. The Plant Journal **38:** 131-141

Butelli E, Titta L, Giorgio M, Mock H-P, Matros A, Peterek S, Schijlen EGWM, Hall RD, Bovy AG, Luo J, Martin C (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. Nat Biotech **26:** 1301-1308

Caron J-B, Scheltema A, Schander C, Rudkin D (2006) A soft-bodied mollusc with radula from the Middle Cambrian Burgess Shale. Nature **442:** 159-163

Catalá R, Medina J, Salinas J (2011) Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. Proceedings of the National Academy of Sciences **108:** 16475-16480

Causier B, Ashworth M, Guo W, Davies B (2012) The TOPLESS Interactome: A Framework for Gene Repression in Arabidopsis. Plant Physiology **158:** 423-438

Cerdan PD, Chory J (2003) Regulation of flowering time by light quality. Nature **423:** 881-885

Çevik V, Kidd BN, Zhang P, Hill C, Kiddle S, Denby KJ, Holub EB, Cahill DM, Manners JM, Schenk PM, Beynon J, Kazan K (2012) MEDIATOR25 Acts as an Integrative Hub for the Regulation of Jasmonate-Responsive Gene Expression in Arabidopsis. Plant Physiology **160:** 541-555

Chabner BA (2011) Drug Shortages — A Critical Challenge for the Generic-Drug Market. New England Journal of Medicine **365:** 2147-2149

Chandrika N, Tsai Y-H, Schmidt W (2013) PFT1-controlled ROS balance is critical for multiple stages of root hair development in Arabidopsis. Plant Signaling & Behavior **8:** e24066

Chang C-SJ, Maloof JN, Wu S-H (2011) COP1-Mediated Degradation of BBX22/LZF1 Optimizes Seedling Development in Arabidopsis. Plant Physiology **156:** 228-239

Chebbi M, Ginis O, Courdavault V, Glévarec G, Lanoue A, Clastre M, Papon N, Gaillard C, Atanassova R, St-Pierre B, Giglioli-Guivarc'h N, Courtois M, Oudin A (2014) ZCT1 and ZCT2 transcription factors repress the activity of a gene promoter from the methyl erythritol phosphate pathway in Madagascar periwinkle cells. Journal of Plant Physiology **171:** 1510-1513

Chen G-H, Sun J-Y, Liu M, Liu J, Yang W-C (2014) SPOROCYTELESS Is a Novel Embryophyte-Specific Transcription Repressor that Interacts with TPL and TCP Proteins in Arabidopsis. Journal of Genetics and Genomics **41:** 617-625

Chen R, Jiang H, Li L, Zhai Q, Qi L, Zhou W, Liu X, Li H, Zheng W, Sun J, Li C (2012) The Arabidopsis Mediator Subunit MED25 Differentially Regulates Jasmonate and Abscisic Acid Signaling through Interacting with the MYC2 and ABI5 Transcription Factors. The Plant Cell Online

Chini A, Fonseca S, Ferna´ndez G, Adie B, Chico JM, Lorenzo O, Garcı´a-Casado G, Lo´pez-Vidriero I, Lozano FM, Ponce MR, Micol JL, R. Solano1 (2007) The JAZ Family of Repressors Is the Missing Link in Jasmonate Signalling. Nature **448:** 666-671

Christie P, Alfenito M, Walbot V (1994) Impact of low-temperature stress on general phenylpropanoid and anthocyanin pathways: Enhancement of transcript abundance and anthocyanin pigmentation in maize seedlings. Planta **194:** 541-549

Christov C, Pouneva I, Bozhkova M, Toncheva T, Fournadzieva S, Zafirova T (2001) Influence of Temperature and Methyl Jasmonate on Scenedesmus Incrassulatus. Biologia Plantarum **44:** 367-371

Collén J, Hervé C, Guisle-Marsollier I, Léger JJ, Boyen C (2006) Expression profiling of Chondrus crispus (Rhodophyta) after exposure to methyl jasmonate. Journal of Experimental Botany **57:** 3869-3881

Collu G, Unver N, Peltenburg-Looman AMG, van der Heijden R, Verpoorte R, Memelink J (2001) Geraniol 10-hydroxylase, a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. FEBS Letters **508:** 215-220

Costa MMR, Hilliou F, Duarte P, Pereira LG, Almeida I, Leech M, Memelink J, Barceló AR, Sottomayor M (2008) Molecular Cloning and Characterization of a Vacuolar Class III Peroxidase Involved in the Metabolism of Anticancer Alkaloids in Catharanthus roseus. Plant Physiology **146:** 403-417

Courdavault V, Papon N, Clastre M, Giglioli-Guivarc'h N, St-Pierre B, Burlat V (2014) A look inside an alkaloid multisite plant: the Catharanthus logistics. Current Opinion in Plant Biology **19:** 43-50

Czerpak R, Piotrowska A, Szulecka K (2006) Jasmonic acid affects changes in the growth and some components content in alga Chlorella vulgaris. Acta Physiologiae Plantarum **28:** 195-203

Dai X, Sinharoy S, Udvardi M, Zhao PX (2013) PlantTFcat: an online plant transcription factor and transcriptional regulator categorization and analysis tool. BMC Bioinformatics **14:** 321

De León IP, Hamberg M, Castresana C (2015) Oxylipins in moss development and defense. Frontiers in Plant Science **6:** 483

De Leon IP, Schmelz EA, Gaggero C, Castro A, ÁLvarez A, Montesano M (2012) Physcomitrella patens activates reinforcement of the cell wall, programmed cell death and accumulation of evolutionary conserved defence signals, such as salicylic acid and 12-oxo-phytodienoic acid, but not jasmonic acid, upon Botrytis cinerea infection. Molecular Plant Pathology **13:** 960-974

De Luca V, Marineau C, Brisson N (1989) Molecular cloning and analysis of cDNA encoding a plant tryptophan decarboxylase: comparison with animal dopa decarboxylases. Proceedings of the National Academy of Sciences **86:** 2582-2586

De Luca V, St Pierre B (2000) The cell and developmental biology of alkaloid biosynthesis. Trends Plant Sci **5:** 168-173

Devaiah BN, Karthikeyan AS, Raghothama KG (2007) WRKY75 Transcription Factor Is a Modulator of Phosphate Acquisition and Root Development in Arabidopsis. Plant Physiology **143:** 1789-1801

Dhawan R, Luo H, Foerster AM, Abuqamar S, Du HN, Briggs SD, Mittelsten Scheid O, Mengiste T (2009) HISTONE MONOUBIQUITINATION1 interacts with a subunit of the mediator complex and regulates defense against necrotrophic fungal pathogens in Arabidopsis. Plant Cell **21:** 1000-1019

Didi V, Jackson P, Hejátko J (2015) Hormonal regulation of secondary cell wall formation. Journal of Experimental Botany **66:** 5015-5027

Dombrecht B, Xue GP, Sprague SJ, Kirkegaard JA, Ross JJ, Reid JB, Fitt GP, Sewelam N, Schenk PM, Manners JM, Kazan K (2007) MYC2 Differentially Modulates Diverse Jasmonate-Dependent Functions in Arabidopsis. Plant Cell **19:** 2225-2245

Dubos C, Le Gourrierec J, Baudry A, Huep G, Lanet E, Debeaujon I, Routaboul J-M, Alboresi A, Weisshaar B, Lepiniec L (2008) MYBL2 is a new regulator of flavonoid biosynthesis in Arabidopsis thaliana. The Plant Journal **55:** 940-953

Dugé de Bernonville T, Foureau E, Parage C, Lanoue A, Clastre M, Londono MA, Oudin A, Houillé B, Papon N, Besseau S, Glévarec G, Atehortùa L, Giglioli-Guivarc'h N, St-Pierre B, De Luca V, O'Connor SE, Courdavault V (2015) Characterization of a second secologanin synthase isoform producing both secologanin and secoxyloganin allows enhanced de novo assembly of a Catharanthus roseus transcriptome. BMC Genomics **16:** 619

Edwards D, Davies KL, Axe L (1992) A vascular conducting strand in the early land plant Cooksonia. Nature **357:** 683-685

Elfving N, Davoine C, Benlloch R, Blomberg J, Brännström K, Müller D, Nilsson A, Ulfstedt M, Ronne H, Wingsle G, Nilsson O, Björklund S (2011) The Arabidopsis thaliana Med25 mediator subunit integrates environmental cues to control plant development. Proceedings of the National Academy of Sciences

Emiliani J, Grotewold E, Falcone Ferreyra ML, Casati P (2013) Flavonols Protect Arabidopsis Plants against UV-B Deleterious Effects. Molecular Plant **6:** 1376-1379

Facchini PJ, De Luca V (2008) Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. The Plant Journal **54:** 763-784

Feyissa D, Løvdal T, Olsen K, Slimestad R, Lillo C (2009) The endogenous GL3, but not EGL3, gene is necessary for anthocyanin accumulation as induced by nitrogen depletion in Arabidopsis rosette stage leaves. Planta **230:** 747-754

Feys B, Benedetti CE, Penfold CN, Turner JG (1994) Arabidopsis Mutants Selected for Resistance to the Phytotoxin Coronatine Are Male Sterile, Insensitive to Methyl Jasmonate, and Resistant to a Bacterial Pathogen. The Plant Cell **6:** 751-759

Finet C, Timme RE, Delwiche CF, Marlétaz F (2010) Multigene Phylogeny of the Green Lineage Reveals the Origin and Diversification of Land Plants. Current Biology **20:** 2217-2222

Firmin JL, Wilson KE, Rossen L, Johnston AWB (1986) Flavonoid activation of nodulation genes in Rhizobium reversed by other compounds present in plants. Nature **324:** 90-92

Flanagan PM, Kelleher RJ, Sayre MH, Tschochner H, Kornberg RD (1991) A mediator required for activation of RNA polymerase II transcription in vitro. Nature **350:** 436-438

Friis EM, Pedersen KR, Crane PR (2001) Fossil evidence of water lilies (Nymphaeales) in the Early Cretaceous. Nature **410:** 357-360

Fujii S, Yamamoto R, Miyamoto K, Ueda J (1997) Occurrence of jasmonic acid in Dunaliella (Dunaliellales, Chlorophyta). Phycological Research **45:** 223-226

Fukazawa J, Teramura H, Murakoshi S, Nasuno K, Nishida N, Ito T, Yoshida M, Kamiya Y, Yamaguchi S, Takahashi Y (2014) DELLAs Function as Coactivators of GAI-ASSOCIATED FACTOR1 in Regulation of Gibberellin Homeostasis and Signaling in Arabidopsis. The Plant Cell **26:** 2920-2938

Gangappa SN, Crocco CD, Johansson H, Datta S, Hettiarachchi C, Holm M, Botto JF (2013) The Arabidopsis B-BOX Protein BBX25 Interacts with HY5, Negatively Regulating BBX22 Expression to Suppress Seedling Photomorphogenesis. The Plant Cell **25:** 1243-1257

Gantet P, Imbault N, Thiersault M, Doireau P (1997) Inhibition of alkaloid accumulation by 2,4-D in Catharanthus roseus cell suspension is overcome by methyl jasmonate. Acta Botanica Gallica **144:** 501-508

Geerlings A, Ibañez MM-L, Memelink J, van der Heijden R, Verpoorte R (2000) Molecular Cloning and Analysis of Strictosidine β-d-Glucosidase, an Enzyme in Terpenoid Indole Alkaloid Biosynthesis in Catharanthus roseus. Journal of Biological Chemistry **275:** 3051-3056

Gerrienne P, Meyer-Berthaud B, Fairon-Demaret M, Streel M, Steemans P (2004) Runcaria, a Middle Devonian Seed Plant Precursor. Science **306:** 856-858

Geu-Flores F, Sherden NH, Courdavault V, Burlat V, Glenn WS, Wu C, Nims E, Cui Y, O'Connor SE (2012) An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. Nature **492:** 138–142

Goddijn OJM, Kam RJ, Zanetti A, Schilperoort RA, Hoge JHC (1992) Auxin rapidly down-regulates transcription of the tryptophan decarboxylase gene from Catharanthus roseus. Plant Molecular Biology **18:** 1113-1120

Goklany S, Rizvi NF, Loring RH, Cram EJ, Lee-Parsons CWT (2013) Jasmonate-dependent alkaloid biosynthesis in Catharanthus Roseus hairy root cultures is correlated with the relative expression of Orca and Zct transcription factors. Biotechnology Progress **29:** 1367-1376

Gomez B, Daviero-Gomez V, Coiffard C, Martín-Closas C, Dilcher DL (2015) Montsechia, an ancient aquatic angiosperm. Proceedings of the National Academy of Sciences **112:** 10985-10988

Góngora-Castillo E, Childs KL, Fedewa G, Hamilton JP, Liscombe DK, Magallanes-Lundback M, Mandadi KK, Nims E, Runguphan W, Vaillancourt B, Varbanova-Herde M, DellaPenna D, McKnight TD, O'Connor S, Buell CR (2012) Development of Transcriptomic Resources for Interrogating the Biosynthesis of Monoterpene Indole Alkaloids in Medicinal Plant Species. PLoS ONE **7:** e52506

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. The Plant Journal **53:** 814-827

Gou J-Y, Felippes FF, Liu C-J, Weigel D, Wang J-W (2011) Negative Regulation of Anthocyanin Biosynthesis in Arabidopsis by a miR156-Targeted SPL Transcription Factor. The Plant Cell **23:** 1512-1522

Grunewald W, Vanholme B, Pauwels L, Plovie E, Inzé D, Gheysen G, Goossens A (2009) Expression of the Arabidopsis jasmonate signalling repressor JAZ1/TIFY10A is stimulated by auxin. EMBO reports **10:** 923-928

Guirimand G, Guihur A, Poutrain P, Héricourt F, Mahroug S, St-Pierre B, Burlat V, Courdavault V (2011) Spatial organization of the vindoline biosynthetic pathway in Catharanthus roseus. Journal of Plant Physiology **168:** 549-557

Han J, Wang H, Lundgren A, Brodelius PE (2014) Effects of overexpression of AaWRKY1 on artemisinin biosynthesis in transgenic *Artemisia annua* plants. Phytochemistry **102:** 89-96

Harborne JB (2001) Secondary Metabolites: Attracting Pollinators. *In* eLS. John Wiley & Sons, Ltd

Hattori T, Vasil V, Rosenkrans L, Hannah LC, McCarty DR, Vasil IK (1992) The Viviparous-1 gene and abscisic acid activate the C1 regulatory gene for anthocyanin biosynthesis during seed maturation in maize. Genes & Development **6:** 609-618

Heijden Rvd, Jacobs DI, Snoeijer W, Hallard D, Verpoorte R (2004) The Catharanthus Alkaloids: Pharmacognosy and Biotechnology. Current Medicinal Chemistry **11:** 607-628

Hill SA, Scheckler SE, Basinger JF (1997) Ellesmeris sphenopteroides, gen. et sp. nov., a New Zygopterid Fern from the Upper Devonian (Frasnian) of Ellesmere, N.W.T., Arctic Canada. American Journal of Botany **84:** 85-103

Holland JF, Scharlau C, Gailani S, Krant MJ, Olson KB, Horton J, Shnider BI, Lynch JJ, Owens A, Carbone PP, Colsky J, Grob D, Miller SP, Hall TC (1973) Vincristine Treatment of Advanced Cancer: A Cooperative Study of 392 Cases. Cancer Research **33:** 1258-1264

Holtan HE, Bandong S, Marion CM, Adam L, Tiwari S, Shen Y, Maloof JN, Maszle DR, Ohto M-a, Preuss S, Meister R, Petracek M, Repetti PP, Reuber TL, Ratcliffe OJ, Khanna R (2011) BBX32, An Arabidopsis B-box Protein, Functions in Light Signaling by Suppressing HY5-Regulated Gene Expression And Interacting with STH2. Plant Physiology

Hong G-J, Xue X-Y, Mao Y-B, Wang L-J, Chen X-Y (2012) Arabidopsis MYC2 Interacts with DELLA Proteins in Regulating Sesquiterpene Synthase Gene Expression. The Plant Cell **24:** 2635-2648

Hori K, Maruyama F, Fujisawa T, Togashi T, Yamamoto N, Seo M, Sato S, Yamada T, Mori H, Tajima N, Moriyama T, Ikeuchi M, Watanabe M, Wada H, Kobayashi K, Saito M, Masuda T, Sasaki-Sekimoto Y, Mashiguchi K, Awai K, Shimojima M, Masuda S, Iwai M, Nobusawa T, Narise T, Kondo S, Saito H, Sato R, Murakawa M, Ihara Y, Oshima-Yamada Y, Ohtaka K, Satoh M, Sonobe K, Ishii M, Ohtani R, Kanamori-Sato M, Honoki R, Miyazaki D, Mochizuki H, Umetsu J, Higashi K, Shibata D, Kamiya Y, Sato N, Nakamura Y, Tabata S, Ida S, Kurokawa K, Ohta H (2014) Klebsormidium flaccidum genome reveals primary factors for plant terrestrial adaptation. Nat Commun **5**

Hou X, Lee LYC, Xia K, Yan Y, Yu H (2010) DELLAs Modulate Jasmonate Signaling via Competitive Binding to JAZs. Developmental cell **19:** 884-894

Hounsome N, Hounsome B, Tomos D, Edwards-Jones G (2008) Plant Metabolites and Nutritional Quality of Vegetables. Journal of Food Science **73:** R48-R65

Hu Y, Jiang L, Wang F, Yu D (2013) Jasmonate Regulates the INDUCER OF CBF EXPRESSION–C-REPEAT BINDING FACTOR/DRE BINDING FACTOR1 Cascade and Freezing Tolerance in Arabidopsis. The Plant Cell **25:** 2907-2924

Iñigo S, Alvarez MJ, Strasser B, Califano A, Cerdán PD (2011) PFT1, the MED25 subunit of the plant Mediator complex, promotes flowering through CONSTANS dependent and independent mechanisms in Arabidopsis. The Plant Journal**:** no-no

Irmler S, Schröder G, St-Pierre B, Crouch NP, Hotze M, Schmidt J, Strack D, Matern U, Schröder J (2000) Indole alkaloid biosynthesis in Catharanthus roseus: new enzyme activities and identification of cytochrome P450 CYP72A1 as secologanin synthase. The Plant Journal **24:** 797-804

Ito J, Fukaki H, Onoda M, Li L, Li C, Tasaka M, Furutani M (2016) Auxin-dependent compositional change in Mediator in ARF7- and ARF19-mediated transcription. Proceedings of the National Academy of Sciences **113:** 6562–6567

Jakoby MJ, Falkenhan D, Mader MT, Brininstool G, Wischnitzki E, Platz N, Hudson A, Hülskamp M, Larkin J, Schnittger A (2008) Transcriptional Profiling of Mature Arabidopsis Trichomes Reveals That NOECK Encodes the MIXTA-Like Transcriptional Regulator MYB106. Plant Physiology **148:** 1583-1602

Jiang C, Gao X, Liao L, Harberd NP, Fu X (2007) Phosphate Starvation Root Architecture and Anthocyanin Accumulation Responses Are Modulated by the Gibberellin-DELLA Signaling Pathway in Arabidopsis. Plant Physiology **145:** 1460-1470

Jiang L, Liu X, Xiong G, Liu H, Chen F, Wang L, Meng X, Liu G, Yu H, Yuan Y, Yi W, Zhao L, Ma H, He Y, Wu Z, Melcher K, Qian Q, Xu HE, Wang Y, Li J (2013) DWARF 53 acts as a repressor of strigolactone signalling in rice. Nature **504:** 401–405

Johnson ET, Dowd PF (2004) Differentially Enhanced Insect Resistance, at a Cost, in Arabidopsis thaliana Constitutively Expressing a Transcription Factor of Defensive Metabolites. Journal of Agricultural and Food Chemistry **52:** 5135-5138

Jose S, Gillespie AR (1998) Allelopathy in black walnut (Juglans nigra L.) alley cropping. I. Spatio-temporal variation in soil juglone in a black walnut–corn (Zea maysL.) alley cropping system in the midwestern USA. Plant and Soil **203:** 191-197

Kanaoka MM, Pillitteri LJ, Fujii H, Yoshida Y, Bogenschutz NL, Takabayashi J, Zhu J-K, Torii KU (2008) SCREAM/ICE1 and SCREAM2 Specify Three Cell-State Transitional Steps Leading to Arabidopsis Stomatal Differentiation. The Plant Cell **20:** 1775-1785

Karol KG, McCourt RM, Cimino MT, Delwiche CF (2001) The Closest Living Relatives of Land Plants. Science **294:** 2351-2353

Kelleher RJ, Flanagan PM, Kornberg RD (1990) A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell **61:** 1209-1215

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE (2015) The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protocols **10:** 845-858

Kellner F, Kim J, Clavijo BJ, Hamilton JP, Childs KL, Vaillancourt B, Cepela J, Habermann M, Steuernagel B, Clissold L, McLay K, Buell CR, O'Connor SE (2015) Genome-guided investigation of plant natural product biosynthesis. The Plant Journal **82:** 680-692

Kidd BN, Edgar CI, Kumar KK, Aitken EA, Schenk PM, Manners JM, Kazan K (2009) The Mediator Complex Subunit PFT1 Is a Key Regulator of Jasmonate-Dependent Defense in Arabidopsis. Plant Cell **21:** 2237-2252

Kidd BN, Kadoo NY, Dombrecht B, Tekeoglu M, Gardiner DM, Thatcher LF, Aitken EAB, Schenk PM, Manners JM, Kazan K (2011) Auxin Signaling and Transport Promote Susceptibility to the Root-Infecting Fungal Pathogen Fusarium oxysporum in Arabidopsis. Molecular Plant-Microbe Interactions **24:** 733-748

Kim YJ, Zheng B, Yu Y, Won SY, Mo B, Chen X (2011) The role of Mediator in small and long noncoding RNA production in Arabidopsis thaliana. EMBO J **30:** 814-822

Kirik V, Simon M, Huelskamp M, Schiefelbein J (2004) The ENHANCER OF TRY AND CPC1 gene acts redundantly with TRIPTYCHON and CAPRICE in trichome and root hair cell patterning in Arabidopsis. Developmental Biology **268:** 506-513

Kirik V, Simon M, Wester K, Schiefelbein J, Hulskamp M (2004) ENHANCER of TRYand CPC 2(ETC2) reveals redundancy in the region-specific control of trichome development of Arabidopsis. Plant Molecular Biology **55:** 389-398

Klose C, Büche C, Fernandez AP, Schäfer E, Zwick E, Kretsch T (2012) The Mediator Complex Subunit PFT1 Interferes with COP1 and HY5 in the Regulation of Arabidopsis Light Signaling. Plant Physiology **160:** 289-307

Koeduka T, Ishizaki K, Mwenda C, Hori K, Sasaki-Sekimoto Y, Ohta H, Kohchi T, Matsui K (2015) Biochemical characterization of allene oxide synthases from the liverwort Marchantia polymorpha and green microalgae Klebsormidium flaccidum provides insight into the evolutionary divergence of the plant CYP74 family. Planta **242:** 1175-1186

Koprivova A, Calderwood A, Lee B-R, Kopriva S (2014) Do PFT1 and HY5 interact in regulation of sulfate assimilation by light in Arabidopsis? FEBS Letters **588:** 1116-1121

Kováčik J, Klejdus B, Štork F, Hedbavny J, Bačkor M (2011) Comparison of Methyl Jasmonate and Cadmium Effect on Selected Physiological Parameters i Scenedesmus quadricauda (Chlorophyta, Chlorophyceae). Journal of Phycology **47:** 1044-1049

Kovinich N, Saleem A, Arnason JT, Miki B (2011) Combined analysis of transcriptome and metabolite data reveals extensive differences between black and brown nearly-isogenic soybean (Glycine max) seed coats enabling the identification of pigment isogenes. BMC Genomics **12:** 381

Krogan NT, Hogan K, Long JA (2012) APETALA2 negatively regulates multiple floral organ identity genes in Arabidopsis by recruiting the co-repressor TOPLESS and the histone deacetylase HDA19. Development **139:** 4180-4190

Krupina MV, Dathe W (1991) Occurrence of jasmonic acid in the red alga Gelidium latifolium. Zeitschrift für Naturforschung C **46:** 1127-1129

Kubo H, Peeters AJ, Aarts MG, Pereira A, Koornneef M (1999) ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. The Plant Cell **11:** 1217-1226

Kumar K, Kumar SR, Dwivedi V, Rai A, Shukla AK, Shanker K, Nagegowda DA (2015) Precursor feeding studies and molecular characterization of geraniol synthase establish the limiting role of geraniol in monoterpene indole alkaloid biosynthesis in Catharanthus roseus leaves. Plant Science **239:** 56-66

Kwon Y, Yu S-i, Park J-h, Li Y, Han J-H, Alavilli H, Cho J-I, Kim T-H, Jeon J-S, Lee B-h (2012) OsREL2, a rice TOPLESS homolog functions in axillary meristem development in rice inflorescence. Plant Biotechnology Reports **6:** 213-224

Lai Z, Schluttenhofer CM, Bhide K, Shreve J, Thimmapuram J, Lee SY, Yun D-J, Mengiste T (2014) MED18 interaction with distinct transcription factors regulates multiple plant functions. Nat Commun **5:** 3064

Landry LG, Chapple C, Last RL (1995) Arabidopsis Mutants Lacking Phenolic Sunscreens Exhibit Enhanced Ultraviolet-B Injury and Oxidative Damage. Plant Physiology **109:** 1159-1166

Lee D-S, Nioche P, Hamberg M, Raman CS (2008) Structural insights into the evolutionary paths of oxylipin biosynthetic enzymes. Nature **455:** 363-368

Lenka S, Boutaoui N, Paulose B, Vongpaseuth K, Normanly J, Roberts S, Walker E (2012) Identification and expression analysis of methyl jasmonate responsive ESTs in paclitaxel producing *Taxus cuspidata* suspension culture cells. BMC Genomics **13:** 148

Levac D, Murata J, Kim WS, De Luca V (2008) Application of carborundum abrasion for investigating the leaf epidermis: molecular cloning of Catharanthus roseus 16-hydroxytabersonine-16-O-methyltransferase. The Plant Journal **53:** 225-236

Li C, Leopold A, Sander G, Shanks J, Zhao L, Gibson S (2013) The ORCA2 transcription factor plays a key role in regulation of the terpenoid indole alkaloid pathway. BMC Plant Biology **13:** 155

Li CY, Leopold AL, Sander GW, Shanks JV, Zhao L, Gibson SI (2015) CrBPF1 overexpression alters transcript levels of terpenoid indole alkaloid biosynthetic and regulatory genes. Frontiers in Plant Science **6**

Li J, Ou-Lee TM, Raba R, Amundson RG, Last RL (1993) Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. The Plant Cell **5:** 171-179

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. Genome Research **13:** 2178-2189

Li L, Zhao Y, McCaig BC, Wingerd BA, Wang J, Whalon ME, Pichersky E, Howe GA (2004) The Tomato Homolog of CORONATINE-INSENSITIVE1 Is Required for the Maternal Control of

Seed Maturation, Jasmonate-Signaled Defense Responses, and Glandular Trichome Development. The Plant Cell **16:** 126-143

Li S, Zachgo S (2013) TCP3 interacts with R2R3-MYB proteins, promotes flavonoid biosynthesis and negatively regulates the auxin response in Arabidopsis thaliana. The Plant Journal **76:** 901-913

Linden H, Macino G (1997) White collar 2, a partner in blue-light signal transduction, controlling expression of light–regulated genes in Neurospora crassa. The EMBO Journal **16:** 98-109

Liscombe DK, Usera AR, O'Connor SE (2010) Homolog of tocopherol C methyltransferases catalyzes N methylation in anticancer alkaloid biosynthesis. Proceedings of the National Academy of Sciences **107:** 18793-18798

Liu C, Jun JH, Dixon RA (2014) MYB5 and MYB14 Play Pivotal Roles in Seed Coat Polymer Biosynthesis in Medicago truncatula. Plant Physiology **165:** 1424-1439

Liu P-P, Montgomery TA, Fahlgren N, Kasschau KD, Nonogaki H, Carrington JC (2007) Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. The Plant Journal **52:** 133-146

Liu ZC, Karmarkar V (2008) Groucho/Tup1 family co-repressors in plant development. Trends in Plant Science **13:** 137-144

Long JA, Ohno C, Smith ZR, Meyerowitz EM (2006) TOPLESS Regulates Apical Embryonic Fate in Arabidopsis. Science **312:** 1520-1523

Lorenzo O, Chico JM, Sanchez-Serrano JJ, Solano R (2004) JASMONATE-INSENSITIVE1 Encodes a MYC Transcription Factor Essential to Discriminate between Different Jasmonate-Regulated Defense Responses in Arabidopsis. Plant Cell **16:** 1938-1950

Loreti E, Povero G, Novi G, Solfanelli C, Alpi A, Perata P (2008) Gibberellins, jasmonate and abscisic acid modulate the sucrose-induced expression of anthocyanin biosynthetic genes in Arabidopsis. New Phytologist **179:** 1004-1016

Lurling M, Beekman W (2006) Palmelloids formation in Chlamydomonas reinhardtii : defence against rotifer predators? Ann. Limnol. - Int. J. Lim. **42:** 65-72

Lv Z, Wang S, Zhang F, Chen L, Hao X, Pan Q, Fu X, Li L, Sun X, Tang K (2016) Overexpression of a Novel NAC Domain-Containing Transcription Factor Gene (AaNAC1) Enhances the Content of Artemisinin and Increases Tolerance to Drought and Botrytis cinerea in Artemisia annua. Plant and Cell Physiology **57:** 1961-1971

Makhzoum A, Petit-Paly G, St. Pierre B, Bernards M (2011) Functional analysis of the DAT gene promoter using transient *Catharanthus roseus* and stable *Nicotiana tabacum* transformation systems. Plant Cell Reports **30:** 1173-1182

Maloney GS, DiNapoli KT, Muday GK (2014) The anthocyanin reduced Tomato Mutant Demonstrates the Role of Flavonols in Tomato Lateral Root and Root Hair Development. Plant Physiology **166:** 614-631

Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. Nucleic Acids Research **43:** D222-D226

Mathur S, Vyas S, Kapoor S, Tyagi AK (2011) The Mediator Complex in Plants: Structure, Phylogeny, and Expression Profiling of Representative Genes in a Dicot (Arabidopsis) and a Monocot (Rice) during Reproduction and Abiotic Stress. Plant Physiology **157:** 1609-1627

Matsui K, Umemura Y, Ohme-Takagi M (2008) AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in Arabidopsis. The Plant Journal **55:** 954-967

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends in Ecology & Evolution **19:** 661-666

Mehrtens F, Kranz H, Bednarek P, Weisshaar B (2005) The Arabidopsis Transcription Factor MYB12 Is a Flavonol-Specific Regulator of Phenylpropanoid Biosynthesis. Plant Physiology **138:** 1083-1096

Memelink J, Gantet P (2007) Transcription factors involved in terpenoid indole alkaloid biosynthesis in Catharanthus roseus. Phytochemistry Reviews **6:** 353-362

Menke FL, Champion A, Kijne JW, Memelink J (1999) A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. EMBO J **18:** 4455-4463

Miao Y, Zentgraf U (2007) The Antagonist Function of Arabidopsis WRKY53 and ESR/ESP in Leaf Senescence Is Modulated by the Jasmonic and Salicylic Acid Equilibrium. The Plant Cell Online **19:** 819-830

Miettinen K, Dong L, Navrot N, Schneider T, Burlat V, Pollier J, Woittiez L, van der Krol S, Lugan R, Ilc T, Verpoorte R, Oksman-Caldentey K-M, Martinoia E, Bouwmeester H, Goossens A, Memelink J, Werck-Reichhart D (2014) The seco-iridoid pathway from Catharanthus roseus. Nat Commun **5**

Mizukami H, Nordlov H, Lee S-L, Scott AI (1979) Purification and properties of strictosidine synthetase (an enzyme condensing tryptamine and secologanin) from Catharanthus roseus cultured cells. Biochemistry **18:** 3760-3763

Moerkercke AV, Steensma P, Gariboldi I, Espoz J, Purnama PC, Schweizer F, Miettinen K, Bossche RV, Clercq RD, Memelink J, Goossens A (2016) The basic helix-loop-helix transcription factor BIS2 is essential for monoterpenoid indole alkaloid production in the medicinal plant Catharanthus roseus. The Plant Journal **88:** 3-12

Morishita T, Kojima Y, Maruta T, Nishizawa-Yokoi A, Yabuta Y, Shigeoka S (2009) Arabidopsis NAC Transcription Factor, ANAC078, Regulates Flavonoid Biosynthesis under High-light. Plant and Cell Physiology **50:** 2210-2222

Morohashi K, Grotewold E (2009) A Systems Approach Reveals Regulatory Circuitry for *Arabidopsis* Trichome Initiation by the GL3 and GL1 Selectors. PLoS Genet **5:** e1000396

Morohashi K, Zhao M, Yang M, Read B, Lloyd A, Lamb R, Grotewold E (2007) Participation of the Arabidopsis bHLH Factor GL3 in Trichome Initiation Regulatory Events. Plant Physiology **145:** 736-746

Morris JA, Khettry A, Seitz EW (1979) Antimicrobial activity of aroma chemicals and essential oils. Journal of the American Oil Chemists' Society **56:** 595-603

Morris SC (1993) The fossil record and the early evolution of the Metazoa. Nature **361:** 219-225

Munkert J, Pollier J, Miettinen K, Van Moerkercke A, Payne R, Müller-Uri F, Burlat V, O'Connor Sarah E, Memelink J, Kreis W, Goossens A (2015) Iridoid Synthase Activity Is Common among the Plant Progesterone 5β-Reductase Family. Molecular Plant **8:** 136-152

Murata J, Roepke J, Gordon H, De Luca V (2008) The Leaf Epidermome of Catharanthus roseus Reveals Its Biochemical Specialization. The Plant Cell Online **20:** 524-542

Nakabayashi R, Yonekura-Sakakibara K, Urano K, Suzuki M, Yamada Y, Nishizawa T, Matsuda F, Kojima M, Sakakibara H, Shinozaki K, Michael AJ, Tohge T, Yamazaki M, Saito K (2014) Enhancement of oxidative and drought tolerance in Arabidopsis by overaccumulation of antioxidant flavonoids. The Plant Journal **77:** 367-379

Nemie-Feyissa D, Olafsdottir SM, Heidari B, Lillo C (2014) Nitrogen depletion and small R3-MYB transcription factors affecting anthocyanin accumulation in Arabidopsis leaves. Phytochemistry **98:** 34-40

Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L (2000) The TT8 Gene Encodes a Basic Helix-Loop-Helix Domain Protein Required for Expression of DFR and BAN Genes in Arabidopsis Siliques. The Plant Cell **12:** 1863-1878

Newman DJ, Cragg GM (2012) Natural Products As Sources of New Drugs over the 30 Years from 1981 to 2010. Journal of Natural Products **75:** 311-335

Nishii A, Takemura M, Fujita H, Shikata M, Yokota A, Kohchi T (2000) Characterization of a Novel Gene Encoding a Putative Single Zinc-finger Protein, ZIM, Expressed during the Reproductive Phase in Arabidopsis thaliana. Bioscience, Biotechnology, and Biochemistry **64:** 1402-1409

Niu Y, Figueroa P, Browse J (2011) Characterization of JAZ-interacting bHLH transcription factors that regulate jasmonate responses in Arabidopsis. Journal of Experimental Botany

O'Connor SE, Maresh JJ (2006) Chemistry and biology of monoterpene indole alkaloid biosynthesis. Natural Product Reports **23:** 532-547

Oh E, Zhu J-Y, Ryu H, Hwang I, Wang Z-Y (2014) TOPLESS mediates brassinosteroid-induced transcriptional repression through interaction with BZR1. Nat Commun **5**

Oh J, Kim Y, Kim J, Kwon Y, Lee H (2011) Enhanced level of anthocyanin leads to increased salt tolerance in arabidopsis PAP1-D plants upon sucrose treatment. Journal of the Korean Society for Applied Biological Chemistry **54:** 79-88

Oliver J, Castro A, Gaggero C, Cascón T, Schmelz E, Castresana C, Ponce de León I (2009) Pythium infection activates conserved plant defense responses in mosses. Planta **230:** 569-579

Ou B, Yin K-Q, Liu S-N, Yang Y, Gu T, Wing Hui JM, Zhang L, Miao J, Kondou Y, Matsui M, Gu H-Y, Qu L-J (2011) A High-Throughput Screening System for Arabidopsis Transcription Factors and Its Application to Med25-Dependent Transcriptional Regulation. Molecular Plant

Ouwerkerk PBF, Trimborn TO, Hilliou F, Memelink J (1999) Nuclear factors GT-1 and 3AF1 interact with multiple sequences within the promoter of the Tdc gene from Madagascar periwinkle: GT-1 is involved in UV light-induced expression. Molecular and General Genetics MGG **261:** 610-622

Pan Q, Chen Y, Wang Q, Yuan F, Xing S, Tian Y, Zhao J, Sun X, Tang K (2010) Effect of plant growth regulators on the biosynthesis of vinblastine, vindoline and catharanthine in *Catharanthus roseus*. Plant Growth Regulation **60:** 133-141

Pan Q, Wang Q, Yuan F, Xing S, Zhao J, Choi YH, Verpoorte R, Tian Y, Wang G, Tang K (2012) Overexpression of ORCA3 and G10H in *Catharanthus roseus* Plants Regulated Alkaloid Biosynthesis and Metabolism Revealed by NMR-Metabolomics. PLoS ONE **7:** e43038

Parage C, Foureau E, Kellner F, Burlat V, Mahroug S, Lanoue A, Dugé de Bernonville T, Arias Londono M, Carqueijeiro IT, Oudin A, Besseau S, Papon N, Glévarec G, Atehortua L, Giglioli-Guivarc'h N, St-Pierre B, Clastre M, O'Connor S, Courdavault V (2016) Class II Cytochrome P450 reductase governs the biosynthesis of alkaloids. Plant Physiology

Patra B, Pattanaik S, Yuan L (2013) Ubiquitin protein ligase 3 mediates the proteasomal degradation of GLABROUS 3 and ENHANCER OF GLABROUS 3, regulators of trichome development and flavonoid biosynthesis in Arabidopsis. The Plant Journal **74:** 435-447

Patra B, Schluttenhofer C, Wu Y, Pattanaik S, Yuan L (2013) Transcriptional regulation of secondary metabolite biosynthesis in plants. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms **1829:** 1236-1247

Pauw B, Hilliou FA, Martin VS, Chatel G, de Wolf CJ, Champion A, Pre M, van Duijn B, Kijne JW, van der Fits L, Memelink J (2004) Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in Catharanthus roseus. J Biol Chem **279:** 52940-52948

Pauwels L, Barbero GF, Geerinck J, Tilleman S, Grunewald W, Perez AC, Chico JM, Bossche RV, Sewell J, Gil E, Garcia-Casado G, Witters E, Inze D, Long JA, De Jaeger G, Solano R, Goossens A (2010) NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature **464:** 788-791

Payne CT, Zhang F, Lloyd AM (2000) GL3 Encodes a bHLH Protein That Regulates Trichome Development in Arabidopsis Through Interaction With GL1 and TTG1. Genetics **156:** 1349-1362

Pourcel L, Routaboul J-M, Kerhoas L, Caboche M, Lepiniec L, Debeaujon I (2005) TRANSPARENT TESTA10 Encodes a Laccase-Like Enzyme Involved in Oxidative Polymerization of Flavonoids in Arabidopsis Seed Coat. The Plant Cell **17:** 2966-2980

Prost I, Dhondt S, Rothe G, Vicente J, Rodriguez MJ, Kift N, Carbonne F, Griffiths G, Esquerré-Tugayé M-T, Rosahl S, Castresana C, Hamberg M, Fournier J (2005) Evaluation of the Antimicrobial Activities of Plant Oxylipins Supports Their Involvement in Defense against Pathogens. Plant Physiology **139:** 1902-1913

Qi T, Song S, Ren Q, Wu D, Huang H, Chen Y, Fan M, Peng W, Ren C, Xie D (2011) The Jasmonate-ZIM-Domain Proteins Interact with the WD-Repeat/bHLH/MYB Complexes to Regulate Jasmonate-Mediated Anthocyanin Accumulation and Trichome Initiation in *Arabidopsis thaliana*. The Plant Cell **23:** 1795-1814

Qi T, Wang J, Huang H, Liu B, Gao H, Liu Y, Song S, Xie D (2015) Regulation of Jasmonate-Induced Leaf Senescence by Antagonism between bHLH Subgroup IIIe and IIId Factors in Arabidopsis. The Plant Cell

Qu Y, Easson MLAE, Froese J, Simionescu R, Hudlicky T, De Luca V (2015) Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. Proceedings of the National Academy of Sciences

Rai A, Smita SS, Singh AK, Shanker K, Nagegowda DA (2013) Heteromeric and Homomeric Geranyl Diphosphate Synthases from Catharanthus roseus and Their Role in Monoterpene Indole Alkaloid Biosynthesis. Molecular Plant **6:** 1531-1549

Raman V, Ravi S (2011) Effect of salicylic acid and methyl jasmonate on antioxidant systems of Haematococcus pluvialis. Acta Physiologiae Plantarum **33:** 1043-1049

Raya-González J, Ortiz-Castro R, Ruíz-Herrera LF, Kazan K, López-Bucio J (2014) PHYTOCHROME AND FLOWERING TIME1/MEDIATOR25 Regulates Lateral Root Formation via Auxin Signaling in Arabidopsis. Plant Physiology **165:** 880-894

Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP (2006) GenePattern 2.0. Nat Genet **38:** 500-501

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud P-F, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin-I T, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S-i, Yamaguchi K, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu S-H, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The Physcomitrella Genome Reveals Evolutionary Insights into the Conquest of Land by Plants. Science **319:** 64-69

Rietveld WJ (1983) Allelopathic effects of juglone on germination and growth of several herbaceous and woody species. Journal of Chemical Ecology **9:** 295-308

Robson F, Costa MMR, Hepworth SR, Vizir I, Pin˜eiro M, Reeves PH, Putterill J, Coupland G (2001) Functional importance of conserved domains in the flowering-time gene CONSTANS demonstrated by analysis of mutant alleles and transgenic plants. The Plant Journal **28:** 619-631

Roepke J, Salim V, Wu M, Thamm AMK, Murata J, Ploss K, Boland W, De Luca V (2010) Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. Proceedings of the National Academy of Sciences **107:** 15287-15292

Routaboul J-M, Kerhoas L, Debeaujon I, Pourcel L, Caboche M, Einhorn J, Lepiniec L (2006) Flavonoid diversity and biosynthesis in seed of Arabidopsis thaliana. Planta **224:** 96-107

Rowan DD, Cao M, Lin-Wang K, Cooney JM, Jensen DJ, Austin PT, Hunt MB, Norling C, Hellens RP, Schaffer RJ, Allan AC (2009) Environmental regulation of leaf colour in red 35S:PAP1 Arabidopsis thaliana. New Phytologist **182:** 102-115

Rubin G, Tohge T, Matsuda F, Saito K, Scheible W-R (2009) Members of the LBD family of transcription factors repress anthocyanin synthesis and affect additional nitrogen responses in Arabidopsis. The Plant Cell **21:** 3567-3584

Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M (2003) TM4: a free, open-source system for microarray data management and analysis. Biotechniques **34:** 374

Salim V, Wiens B, Masada-Atsumi S, Yu F, De Luca V (2014) 7-Deoxyloganetic acid synthase catalyzes a key 3 step oxidation to form 7-deoxyloganetic acid in Catharanthus roseus iridoid biosynthesis. Phytochemistry **101:** 23-31

Salim V, Yu F, Altarejos J, De Luca V (2013) Virus-induced gene silencing identifies Catharanthus roseus 7-deoxyloganic acid-7-hydroxylase, a step in iridoid and monoterpene indole alkaloid biosynthesis. The Plant Journal **76:** 754-765

Sasaki-Sekimoto Y, Jikumaru Y, Obayashi T, Saito H, Masuda S, Kamiya Y, Ohta H, Shirasu K (2013) Basic Helix-Loop-Helix Transcription Factors JASMONATE-ASSOCIATED MYC2-LIKE1 (JAM1), JAM2, and JAM3 Are Negative Regulators of Jasmonate Responses in Arabidopsis. Plant Physiology **163:** 291-304

Schilling B, Kaiser R, Natsch A, Gautschi M (2010) Investigation of odors in the fragrance industry. Chemoecology **20:** 135-147

Schluttenhofer C, Pattanaik S, Patra B, Yuan L (2014) Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. BMC Genomics **15:** 502-522

Schluttenhofer C, Yuan L (2015) Regulation of Specialized Metabolism by WRKY Transcription Factors. Plant Physiology **167:** 295-306

Schröder G, Unterbusch E, Kaltenbach M, Schmidt J, Strack D, De Luca V, Schröder J (1999) Light-induced cytochrome P450-dependent enzyme in indole alkaloid biosynthesis: tabersonine 16-hydroxylase. FEBS Letters **458:** 97-102

Schulz E, Tohge T, Zuther E, Fernie AR, Hincha DK (2016) Flavonoids are determinants of freezing tolerance and cold acclimation in Arabidopsis thaliana. Scientific Reports **6:** 34027

Schwab W, Davidovich-Rikanati R, Lewinsohn E (2008) Biosynthesis of plant-derived flavor compounds. The Plant Journal **54:** 712-732

Schweizer F, Fernández-Calvo P, Zander M, Diez-Diaz M, Fonseca S, Glauser G, Lewsey MG, Ecker JR, Solano R, Reymond P (2013) Arabidopsis Basic Helix-Loop-Helix Transcription Factors MYC2, MYC3, and MYC4 Regulate Glucosinolate Biosynthesis, Insect Performance, and Feeding Behavior. The Plant Cell **25:** 3117-3132

Scott AC, Chaloner WG (1983) The Earliest Fossil Conifer from the Westphalian B of Yorkshire. Proceedings of the Royal Society of London B: Biological Sciences **220:** 163-182

Shaikhali J, de Dios Barajas-Lopéz J, Ötvös K, Kremnev D, Garcia AS, Srivastava V, Wingsle G, Bako L, Strand Å (2012) The CRYPTOCHROME1-Dependent Response to Excess Light Is Mediated through the Transcriptional Activators ZINC FINGER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM LIKE1 and ZML2 in Arabidopsis. The Plant Cell **24:** 3009-3025

Shan X, Zhang Y, Peng W, Wang Z, Xie D (2009) Molecular mechanism for jasmonate-induction of anthocyanin accumulation in Arabidopsis. Journal of Experimental Botany **60:** 3849-3860

Sheard LB, Tan X, Mao H, Withers J, Ben-Nissan G, Hinds TR, Kobayashi Y, Hsu F-F, Sharon M, Browse J, He SY, Rizo J, Howe GA, Zheng N (2010) Jasmonate perception by inositol-phosphate-potentiated COI1-JAZ co-receptor. Nature **468:** 400-405

Sheehan H, Moser M, Klahre U, Esfeld K, Dell'Olivo A, Mandel T, Metzger S, Vandenbussche M, Freitas L, Kuhlemeier C (2016) MYB-FL controls gain and loss of floral UV absorbance, a key trait affecting pollinator preference and reproductive isolation. Nat Genet **48:** 159-166

Shikata M, Matsuda Y, Ando K, Nishii A, Takemura M, Yokota A, Kohchi T (2004) Characterization of Arabidopsis ZIM, a member of a novel plant-specific GATA factor gene family. Journal of Experimental Botany **55:** 631-639

Shin DH, Choi M, Kim K, Bang G, Cho M, Choi S-B, Choi G, Park Y-I (2013) HY5 regulates anthocyanin biosynthesis by inducing the transcriptional activation of the MYB75/PAP1 transcription factor in Arabidopsis. FEBS Letters **587:** 1543-1547

Shirley BW, Kubasek WL, Storz G, Bruggemann E, Koornneef M, Ausubel FM, Goodman HM (1995) Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. The Plant Journal **8:** 659-671

Shyu C, Figueroa P, DePew CL, Cooke TF, Sheard LB, Moreno JE, Katsir L, Zheng N, Browse J, Howe GA (2012) JAZ8 Lacks a Canonical Degron and Has an EAR Motif That Mediates Transcriptional Repression of Jasmonate Responses in Arabidopsis. The Plant Cell Online **24:** 536-550

Siberil Y, Benhamron S, Memelink J, Giglioli-Guivarc'h N, Thiersault M, Boisson B, Doireau P, Gantet P (2001) Catharanthus roseus G-box binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. Plant Mol Biol **45:** 477-488

Singh G, Gavrieli J, Oakey JS, Curtis WR (1998) Interaction of methyl jasmonate, wounding and fungal elicitation during sesquiterpene induction in Hyoscyamus muticus in root cultures. Plant Cell Reports **17:** 391-395

Smith MR (2012) Mouthparts of the Burgess Shale fossils Odontogriphus and Wiwaxia: implications for the ancestral molluscan radula. Proceedings of the Royal Society B: Biological Sciences **279:** 4287-4295

Solfanelli C, Poggi A, Loreti E, Alpi A, Perata P (2006) Sucrose-Specific Induction of the Anthocyanin Biosynthetic Pathway in Arabidopsis. Plant Physiology **140:** 637-646

Song S, Qi T, Fan M, Zhang X, Gao H, Huang H, Wu D, Guo H, Xie D (2013) The bHLH Subgroup IIId Factors Negatively Regulate Jasmonate-Mediated Plant Defense and Development. PLoS Genet **9:** e1003653

Song S, Qi T, Huang H, Ren Q, Wu D, Chang C, Peng W, Liu Y, Peng J, Xie D (2011) The Jasmonate-ZIM Domain Proteins Interact with the R2R3-MYB Transcription Factors MYB21 and MYB24 to Affect Jasmonate-Regulated Stamen Development in Arabidopsis. The Plant Cell Online **23:** 1000-1013

Sottomayor M, López-Serrano M, DiCosmo F, Ros Barceló A (1998) Purification and characterization of α-3′,4′-anhydrovinblastine synthase (peroxidase-like) from Catharanthus roseus (L.) G. Don. FEBS Letters **428:** 299-303

St-Pierre B, Laflamme P, Alarco A-M, D V, Luca E (1998) The terminal O-acetyltransferase involved in vindoline biosynthesis defines a new class of proteins responsible for coenzyme A-dependent acyl transfer. The Plant Journal **14:** 703-713

St-Pierre B, Vazquez-Flota FA, De Luca V (1999) Multicellular Compartmentation of Catharanthus roseus Alkaloid Biosynthesis Predicts Intercellular Translocation of a Pathway Intermediate. The Plant Cell
11: 887-900

Staswick PE, Su W, Howell SH (1992) Methyl jasmonate inhibition of root growth and induction of a leaf protein are decreased in an Arabidopsis thaliana mutant. Proceedings of the National Academy of Sciences of the United States of America **89:** 6837-6840

Stintzi A, Weber H, Reymond P, Browse J, Farmer EE (2001) Plant defense in the absence of jasmonic acid: The role of cyclopentenones. Proceedings of the National Academy of Sciences **98:** 12837-12842

Stout J, Romero-Severson E, Ruegger MO, Chapple C (2008) Semidominant Mutations in Reduced Epidermal Fluorescence 4 Reduce Phenylpropanoid Content in Arabidopsis. Genetics **178:** 2237-2251

Stracke R, Favory J-J, Gruber H, Bartelniewoehner L, Bartels S, Binkert M, Funk M, Weisshaar B, Ulm R (2010) The Arabidopsis bZIP transcription factor HY5 regulates expression of the PFG1/MYB12 gene in response to light and ultraviolet-B radiation. Plant, Cell & Environment **33:** 88-103

Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B (2007) Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. The Plant Journal **50:** 660-677

Stracke R, Jahns O, Keck M, Tohge T, Niehaus K, Fernie AR, Weisshaar B (2010) Analysis of PRODUCTION OF FLAVONOL GLYCOSIDES-dependent flavonol glycoside accumulation in Arabidopsis thaliana plants reveals MYB11-, MYB12- and MYB111-independent flavonol glycoside accumulation. New Phytologist **188:** 985-1000

Stumpe M, Göbel C, Faltin B, Beike AK, Hause B, Himmelsbach K, Bode J, Kramell R, Wasternack C, Frank W (2010) The moss Physcomitrella patens contains cyclopentenones but no jasmonates: mutations in allene oxide cyclase lead to reduced fertility and altered sporophyte morphology. New Phytologist **188:** 740-749

Sun G, Dilcher DL, Zheng S, Zhou Z (1998) In Search of the First Flower: A Jurassic Angiosperm, Archaefructus, from Northeast China. Science **282:** 1692-1695

Sun J, Xu Y, Ye S, Jiang H, Chen Q, Liu F, Zhou W, Chen R, Li X, Tietz O, Wu X, Cohen JD, Palme K, Li C (2009) Arabidopsis ASA1 Is Important for Jasmonate-Mediated Regulation of Auxin Biosynthesis and Transport during Lateral Root Formation. The Plant Cell **21:** 1495-1511

Sundaravelpandian K, Chandrika NNP, Schmidt W (2012) PFT1, a transcriptional Mediator complex subunit, controls root hair differentiation through reactive oxygen species (ROS) distribution in Arabidopsis. New Phytologist **197:** 151-161

Suttipanta N, Pattanaik S, Gunjan S, Xie CH, Littleton J, Yuan L (2007) Promoter analysis of the *Catharanthus roseus* geraniol 10-hydroxylase gene involved in terpenoid indole alkaloid biosynthesis. Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression **1769:** 139-148

Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L (2011) The Transcription Factor CrWRKY1 Positively Regulates the Terpenoid Indole Alkaloid Biosynthesis in *Catharanthus roseus*. Plant Physiology **157:** 2081-2093

Suttipanta N, Pattanaik S, Kulshrestha M, Patra B, Singh SK, Yuan L (2011) The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in Catharanthus roseus. Plant Physiol **157:** 2081-2093

Swisher CC, Wang Y-q, Wang X-l, Xu X, Wang Y (1999) Cretaceous age for the feathered dinosaurs of Liaoning, China. Nature **400:** 58-61

Szemenyei H, Hannon M, Long JA (2008) TOPLESS Mediates Auxin-Dependent Transcriptional Repression During Arabidopsis Embryogenesis. Science **319:** 1384-1386

Talora C, Franchi L, Linden H, Ballario P, Macino G (1999) Role of a white collar-1–white collar-2 complex in blue-light signal transduction. The EMBO Journal **18:** 4961-4968

Tao Q, Guo D, Wei B, Zhang F, Pang C, Jiang H, Zhang J, Wei T, Gu H, Qu L-J, Qin G (2013) The TIE1 Transcriptional Repressor Links TCP Transcription Factors with TOPLESS/TOPLESS-RELATED Corepressors and Modulates Leaf Development in Arabidopsis. The Plant Cell **25:** 421-437

Teakle GR, Gilmartin PM (1998) Two forms of type IV zinc-finger motif and their kingdom-specific distribution between the flora, fauna and fungi. Trends in Biochemical Sciences **23:** 100-102

Teng S, Keurentjes J, Bentsink L, Koornneef M, Smeekens S (2005) Sucrose-Specific Induction of Anthocyanin Biosynthesis in Arabidopsis Requires the MYB75/PAP1 Gene. Plant Physiology **139:** 1840-1852

Teplitski M, Chen H, Rajamani S, Gao M, Merighi M, Sayre RT, Robinson JB, Rolfe BG, Bauer WD (2004) Chlamydomonas reinhardtii Secretes Compounds That Mimic Bacterial Signals and Interfere with Quorum Sensing Regulation in Bacteria. Plant Physiology **134:** 137-146

Thaler JS, Humphrey PT, Whiteman NK (2012) Evolution of jasmonate and salicylate signal crosstalk. Trends in Plant Science **17:** 260-270

Thines B, Katsir L, Melotto M, Niu Y, Mandaokar A, Liu G, Nomura K, He SY, Howe GA, Browse J (2007) JAZ Repressor Proteins Are Targets of the SCF-COI1 Complex during Jasmonate Signalling. Nature **448:** 661-665

Thissen D, Steinberg L, Kuang D (2002) Quick and Easy Implementation of the Benjamini-Hochberg Procedure for Controlling the False Positive Rate in Multiple Comparisons. Journal of Educational and Behavioral Statistics **27:** 77-83

Thomma BPHJ, Eggermont K, Penninckx IAMA, Mauch-Mani B, Vogelsang R, Cammue BPA, Broekaert WF (1998) Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. Proceedings of the National Academy of Sciences **95:** 15107-15111

Todd JJ, Vodkin LO (1993) Pigmented Soybean (Glycine max) Seed Coats Accumulate Proanthocyanidins during Development. Plant Physiology **102:** 663-670

Treimer JF, Zenk MH (1979) Purification and Properties of Strictosidine Synthase, the Key Enzyme in Indole Alkaloid Formation. European Journal of Biochemistry **101:** 225-233

Treutter D (2005) Significance of Flavonoids in Plant Resistance and Enhancement of Their Biosynthesis. Plant Biology **7:** 581-591

Tsahar E, Friedman J, Izhaki I (2002) Impact on fruit removal and seed predation of a secondary metabolite, emodin, in Rhamnus alaternus fruit pulp. Oikos **99:** 290-299

Ueda J, Miyamoto K, Aoki M, Hirata T, Sato T, Momotani Y (1991) Identification of jasmonic acid in Chlorella and Spirulina. Bulletin of the University of Osaka Prefecture **43:** 103-108

Ueda J, Miyamoto K, Sato T, Momotani Y (1991) Identification of jasmonic acid from Euglena gracilis Z as a plant growth regulator. Agricultural and biological chemistry **55:** 275-276

Vakil RJ (1949) A clinical trail of Rauwolfia Serpentina in essential hypertension. British heart journal **11:** 350-355

van der Fits L, Memelink J (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science **289:** 295-297

van der Fits L, Zhang H, Menke FLH, Deneka M, Memelink J (2000) A Catharanthus roseus BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene Str and is induced by elicitor via a JA-independent signal transduction pathway. Plant Molecular Biology **44:** 675-685

Van Moerkercke A, Fabris M, Pollier J, Baart GJE, Rombauts S, Hasnain G, Rischer H, Memelink J, Oksman-Caldentey K-M, Goossens A (2013) CathaCyc, a Metabolic Pathway Database Built from Catharanthus roseus RNA-Seq Data. Plant and Cell Physiology

Van Moerkercke A, Steensma P, Schweizer F, Pollier J, Gariboldi I, Payne R, Vanden Bossche R, Miettinen K, Espoz J, Purnama PC, Kellner F, Seppänen-Laakso T, O'Connor SE, Rischer H, Memelink J, Goossens A (2015) The bHLH transcription factor BIS1 controls the iridoid branch of the monoterpenoid indole alkaloid pathway in Catharanthus roseus. Proceedings of the National Academy of Sciences **112:** 8130-8135

Vanderauwera S, Zimmermann P, Rombauts S, Vandenabeele S, Langebartels C, Gruissem W, Inzé D, Van Breusegem F (2005) Genome-Wide Analysis of Hydrogen Peroxide-Regulated Gene Expression in Arabidopsis Reveals a High Light-Induced Transcriptional Cluster Involved in Anthocyanin Biosynthesis. Plant Physiology **139:** 806-821

Vazquez-Flota F, De Carolis E, Alarco A-M, De Luca V (1997) Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in Catharanthus roseus (L.) G. Don. Plant Molecular Biology **34:** 935-948

Ventola CL (2011) The Drug Shortage Crisis in the United States. Pharmacy and Therapeutics **36:** 740-742

Veronese P, Narasimhan ML, Stevenson RA, Zhu JK, Weller SC, Subbarao KV, Bressan RA (2003) Identification of a locus controlling Verticillium disease symptom response in Arabidopsis thaliana. The Plant Journal **35:** 574-587

Vom Endt D, Soares e Silva M, Kijne JW, Pasquali G, Memelink J (2007) Identification of a Bipartite Jasmonate-Responsive Promoter Element in the Catharanthus roseus ORCA3 Transcription Factor Gene That Interacts Specifically with AT-Hook DNA-Binding Proteins. Plant Physiology **144:** 1680-1689

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC (1999) The TRANSPARENT TESTA GLABRA1 Locus, Which Regulates Trichome Differentiation and Anthocyanin Biosynthesis in Arabidopsis, Encodes a WD40 Repeat Protein. The Plant Cell **11:** 1337-1349

Wang C, Liu Y, Li S-S, Han G-Z (2015) Insights into the Origin and Evolution of the Plant Hormone Signaling Machinery. Plant Physiology **167:** 872-886

Wang L, Kim J, Somers DE (2013) Transcriptional corepressor TOPLESS complexes with pseudoresponse regulator proteins and histone deacetylases to regulate circadian transcription. Proceedings of the National Academy of Sciences **110:** 761-766

Wang Q, Yuan F, Pan Q, Li M, Wang G, Zhao J, Tang K (2010) Isolation and functional analysis of the *Catharanthus roseus* deacetylvindoline-4-*O*-acetyltransferase gene promoter. Plant Cell Reports **29:** 185-192

Wang Y, Wang Y, Song Z, Zhang H (2016) Repression of MYBL2 by both microRNA858a and HY5 Leads to the Activation of Anthocyanin Biosynthetic Pathway in Arabidopsis. Molecular Plant

Wei S (2010) Methyl jasmonic acid induced expression pattern of terpenoid indole alkaloid pathway genes in *Catharanthus roseus* seedlings. Plant Growth Regulation **61:** 243-251

Wellman CH, Osterloff PL, Mohiuddin U (2003) Fragments of the earliest land plants. Nature **425:** 282-285

Whitmer S, Verpoorte R, Canel C (1998) Influence of auxins on alkaloid accumulation by a transgenic cell line of Catharanthus roseus. Plant Cell, Tissue and Organ Culture **53:** 135-141

Wickett NJ, Mirarab S, Nguyen N, Warnow T, Carpenter E, Matasci N, Ayyampalayam S, Barker MS, Burleigh JG, Gitzendanner MA, Ruhfel BR, Wafula E, Der JP, Graham SW, Mathews S, Melkonian M, Soltis DE, Soltis PS, Miles NW, Rothfels CJ, Pokorny L, Shaw AJ, DeGironimo L, Stevenson DW, Surek B, Villarreal JC, Roure B, Philippe H, dePamphilis CW, Chen T, Deyholos MK, Baucom RS, Kutchan TM, Augustin MM, Wang J, Zhang Y, Tian Z, Yan Z, Wu X, Sun X, Wong GK-S, Leebens-Mack J (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. Proceedings of the National Academy of Sciences **111:** E4859-E4868

Winter D, Vinegar B, Nahal H, Ammar R, Wilson GV, Provart NJ (2007) An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets. PLoS ONE **2:** 1-12

Xie D-X, Feys BF, James S, Nieto-Rostro M, Turner JG (1998) COI1: An Arabidopsis Gene Required for Jasmonate-Regulated Defense and Fertility. Science **280:** 1091-1094

Xie Y, Tan H, Ma Z, Huang J (2016) DELLA Proteins Promote Anthocyanin Biosynthesis via Sequestering MYBL2 and JAZ Suppressors of the MYB/bHLH/WD40 Complex in Arabidopsis thaliana. Molecular Plant **9:** 711-721

Xu J, Zhang Y (2010) How significant is a protein structure similarity with TM-score = 0.5? Bioinformatics **26:** 889-895

Xu R, Li Y (2011) Control of final organ size by Mediator complex subunit 25 in Arabidopsis thaliana. Development **138:** 4545-4554

Xu R, Li Y (2012) The Mediator complex subunit 8 regulates organ size in Arabidopsis thaliana. Plant Signaling & Behavior **7:** 182-183

Yamamoto K, Takahashi K, Mizuno H, Anegawa A, Ishizaki K, Fukaki H, Ohnishi M, Yamazaki M, Masujima T, Mimura T (2016) Cell-specific localization of alkaloids in Catharanthus roseus stem tissue measured with Imaging MS and Single-cell MS. Proceedings of the National Academy of Sciences

Yamamoto Y, Ohshika J, Takahashi T, Ishizaki K, Kohchi T, Matusuura H, Takahashi K (2015) Functional analysis of allene oxide cyclase, MpAOC, in the liverwort Marchantia polymorpha. Phytochemistry **116:** 48-56

Yang D-L, Yao J, Mei C-S, Tong X-H, Zeng L-J, Li Q, Xiao L-T, Sun T-p, Li J, Deng X-W, Lee CM, Thomashow MF, Yang Y, He Z, He SY (2012) Plant hormone jasmonate prioritizes defense over growth by interfering with gibberellin signaling cascade. Proceedings of the National Academy of Sciences **109:** E1192–E1200

Yang F, Cai J, Yang Y, Liu Z (2013) Overexpression of microRNA828 reduces anthocyanin accumulation in Arabidopsis. Plant Cell, Tissue and Organ Culture (PCTOC) **115:** 159-167

Yang Y, Ou B, Zhang J, Si W, Gu H, Qin G, Qu L-J (2014) The Arabidopsis Mediator subunit MED16 regulates iron homeostasis by associating with EIN3/EIL1 through subunit MED25. The Plant Journal **77:** 838-851

Yin R, Han K, Heller W, Albert A, Dobrev PI, Zažímalová E, Schäffner AR (2014) Kaempferol 3-O-rhamnoside-7-O-rhamnoside is an endogenous flavonol inhibitor of polar auxin transport in Arabidopsis shoots. New Phytologist **201:** 466-475

Yu F, De Luca V (2013) ATP-binding cassette transporter controls leaf surface secretion of anticancer drug components in Catharanthus roseus. Proceedings of the National Academy of Sciences

Yu Z-X, Li J-X, Yang C-Q, Hu W-L, Wang L-J, Chen X-Y (2012) The Jasmonate-Responsive AP2/ERF Transcription Factors AaERF1 and AaERF2 Positively Regulate Artemisinin Biosynthesis in Artemisia annua L. Molecular Plant **5:** 353-365

Zafra-Stone S, Yasmin T, Bagchi M, Chatterjee A, Vinson JA, Bagchi D (2007) Berry anthocyanins as novel antioxidants in human health and disease prevention. Molecular Nutrition & Food Research **51:** 675-683

Zhai Q, Zhang X, Wu F, Feng H, Deng L, Xu L, Zhang M, Wang Q, Li C (2015) Transcriptional Mechanism of Jasmonate Receptor COI1-Mediated Delay of Flowering Time in Arabidopsis. The Plant Cell **27:** 2814-2828

Zhang F, Gonzalez A, Zhao M, Payne CT, Lloyd A (2003) A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development **130:** 4859-4869

Zhang F, Wang Y, Li G, Tang Y, Kramer EM, Tadege M (2014) STENOFOLIA Recruits TOPLESS to Repress ASYMMETRIC LEAVES2 at the Leaf Margin and Promote Leaf Blade Outgrowth in Medicago truncatula. The Plant Cell **26:** 650-664

Zhang F, Yao J, Ke J, Zhang L, Lam VQ, Xin X-F, Zhou XE, Chen J, Brunzelle J, Griffin PR, Zhou M, Xu HE, Melcher K, He SY (2015) Structural basis of JAZ repression of MYC transcription factors in jasmonate signalling. Nature **525:** 269–273

Zhang H, Hedhili S, Montiel G, Zhang Y, Chatel G, Pré M, Gantet P, Memelink J (2011) The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. The Plant Journal **67:** 61-71

Zhang J, Subramanian S, Zhang Y, Yu O (2007) Flavone Synthases from Medicago truncatula Are Flavanone-2-Hydroxylases and Are Important for Nodulation. Plant Physiology **144:** 741-751

Zhang X, Yao J, Zhang Y, Sun Y, Mou Z (2013) The Arabidopsis Mediator Complex Subunits MED14/SWP and MED16/SFR6/IEN1 Differentially Regulate Defense Gene Expression in Plant Immune Responses. The Plant Journal **75:** 484-497

Zhang Y, Butelli E, De Stefano R, Schoonbeek H-j, Magusin A, Pagliarani C, Wellner N, Hill L, Orzaez D, Granell A, Jones Jonathan DG, Martin C (2013) Anthocyanins Double the Shelf Life of Tomatoes by Delaying Overripening and Reducing Susceptibility to Gray Mold. Current Biology **23:** 1094-1100

Zheng Z, Guan H, Leal F, Grey PH, Oppenheimer DG (2013) *Mediator Subunit18* Controls Flowering Time and Floral Organ Identity in *Arabidopsis*. PLoS ONE **8:** e53924

Zhu H-F, Fitzsimmons K, Khandelwal A, Kranz RG (2009) CPC, a Single-Repeat R3 MYB, Is a Negative Regulator of Anthocyanin Biosynthesis in Arabidopsis. Molecular Plant **2:** 790-802

Zhu Y, Schluttenhofer CM, Wang P, Fu F, Thimmapuram J, Zhu J-K, Lee SY, Yun D-J, Mengiste T (2014) CYCLIN-DEPENDENT KINASE8 Differentially Regulates Plant Immunity to Fungal Pathogens through Kinase-Dependent and -Independent Functions in Arabidopsis. The Plant Cell **26:** 4149-4170

Zhu Z, An F, Feng Y, Li P, Xue L, A M, Jiang Z, Kim J-M, To TK, Li W, Zhang X, Yu Q, Dong Z, Chen W-Q, Seki M, Zhou J-M, Guo H (2011) Derepression of ethylene-stabilized transcription

factors (EIN3/EIL1) mediates jasmonate and ethylene signaling synergy in Arabidopsis. Proceedings of the National Academy of Sciences **108:** 12539-12544

Zhu Z, Xu F, Zhang Y, Cheng YT, Wiermer M, Li X, Zhang Y (2010) Arabidopsis resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. Proceedings of the National Academy of Sciences **107:** 13960-13965

Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF (2004) Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. The Plant Journal **40:** 22-34

Zwenger S (2008) Plant terpenoids: applications and future potentials. Biotechnology and Molecular Biology Reviews **3:** 1-7

VITA

Birthplace:

Shelbyville, Kentucky

Education:

Purdue University, West Lafayette, IN — May 2011
Masters in Plant Pathology

Purdue University, West Lafayette, IN — May 2009
Bachelor of Science in Horticulture Science
Bachelor of Science in Plant Genetics and Breeding
Associates in Agronomy

Professional Positions:

Graduate Research Assistant & Graduate Teaching Assistant — 8/09-8/11
Purdue University
Supervisor: Dr. Tesfaye Mengiste

Undergraduate Research Assistant — 8/05-5/09
Purdue University
Supervisor: Dr. Cary Mitchell

Summer Scholar — 5/09-8/09
Samuel Roberts Nobel Foundation
Supervisor: Dr. Maria Monteros

Intern — 5/08-8/08
Monsanto
Supervisor: Dr. Martin Medina

Scholastic and Professional Honors

American Society for Plant Biologist member
Kentucky Academy of Science member
Gamma Sigma Delta member
American Society for Horticulture Scientist Collegiate Scholar

Publications:

- **Craig Schluttenhofer** and Ling Yuan. 2015. Regulation of Specialized Metabolism by WRKY Transcription Factors. Plant Physiology. 167(2): 295-306
- **Craig Schluttenhofer**, Sitakanta Pattanaik, Barunava Patra, and Ling Yuan. Analyses of *Catharanthus roseus* and *Arabidopsis thaliana* WRKY transcription factors reveal involvement in jasmonate signaling. BCM Genomics. 15(1): 502-522
- Barunava Patra*, **Craig Schluttenhofer***, Yongmei Wu, Sitakanta Pattanaik, and Ling Yuan. 2013. Transcriptional regulation of secondary metabolite biosynthesis in plants. Biochimica et Biophysica Acta. 1829(11) 1236-1247 (* co-first authors)
- Yingfang Zhu, **Craig M. Schluttenhofer**, Pengcheng Wang, Fuyou Fu, Jyothi Thimmapuram, Jian-Kang Zhu, Sang Yeol Lee, Dae-Jin Yun, Tesfaye Mengiste. 2014.

CYCLIN-DEPENDENT KINASE8 Differentially Regulates Plant Immunity to Fungal Pathogens through Kinase-Dependent and -Independent Functions in Arabidopsis. The Plant Cell. 26(10): 4149-4170

- Zhibing Lai, **Craig Schluttenhofer**, Ketaki Bhide, Jacob Shreve, Jyothi Thimmapuram, Sang Yeol Lee, Dae-Jin Yun, and Tesfaye Mengiste. 2014. MED18 interaction with distinct transcription factors regulates multiple plant functions. Nature Communications. 5: 3064 (doi: 10.1038/ncomms4064)
- **Craig Schluttenhofer**, Gioia Massa, and Cary Mitchell. 2011. Use of uniconazole to control plant height for an industrial / pharmaceutical maize platform. Industrial Crops and Products. 33(3): 720-726

Name on Final Copy:

Craig Michael Schluttenhofer