RE-ROUTING THE PHENYLPROPANOID PATHWAY AND ITS IMPLICATIONS FOR PLANT GROWTH

by

Fabiola Muro Villanueva

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Biochemistry

West Lafayette, Indiana

December 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Clint Chapple, Chair
Department of Biochemistry

Dr. Natalia Dudareva
Department of Biochemistry

Dr. Jeremy R. Lohman
Department of Biochemistry

Dr. Maureen McCann
Department of Biological Sciences

Approved by:
Dr. Andrew Mesecar

Dedicated to my parents and siblings

ACKNOWLEDGMENTS

I would like to express my deepest appreciation to my advisor Dr. Clint Chapple for helping me become the scientist I am today. In my experience, his focus relies on properly training young scientists to test hypotheses and not in "proving" hypotheses. I must thank him for reading and carefully revising every single piece of scientific writing I have written during my time in his lab, because it made me improve my writing significantly. I also appreciate the freedom he gave me to explore multiple projects without demanding positive outcomes and the excitement he shared even with the smallest results. He created an environment where making questions was welcomed and not as evidence of a lack in knowledge or a sign of weakness. In contrast to a constantly guided training style, freedom leads to a less linear research path, but I would not have had any other way, because it made me grow.

I would like to offer my special thanks to my thesis committee members: Dr. Natalia Dudareva, Dr. Jeremy Lohman and Dr. Maureen McCann for their useful advice throughout these years. I also appreciate the effort they made to keep track of the many projects I carried throughout my Ph.D. and I am thankful for the constant encouragement they gave me, especially during the last couple of years. I am particularly grateful with Dr. Brian Dilkes, whom I consider part of my thesis committee, for all the bright ideas and enlightening discussions about my research projects despite not having an official role on my thesis committee.

The work presented in this thesis would not had been possible without the contributions from Dr. John Ralph and Dr. Hoon Kim from the University of Wisconsin-Madison. They performed NMR analysis of our samples and kindly donated valuable compounds used in the experiments hereby presented. Additionally, I thank Dr. Pete Pascuzzi for providing me very valuable assistance on RNA sequencing analysis. I wish to acknowledge the help provided by the many undergrads and rotation students that worked with me throughout the years, particularly, to Tyler Bouse, Catherine Egley and Zhiwei Luo whose work made part of this thesis.

I thank my PULSe mentors Dr. Rachel M. McCoy and Dr. Brittany F. Peterson for taking their role as mentors very seriously and guiding me throughout my Ph.D., in addition to serving as incredible role models. I am grateful to have had Dr. Peng Wang and Dr. Jeong Im Kim as lab mates, because they not only taught me many technical aspects of research, but they became encouraging friends that continue to mentor me even after their departure from Purdue.

Additionally, I appreciate all the help my lab mate Dr. Jeffrey Simpson provided me and for all the useful discussions. I must also thank Dr. Laura K. Henry for her friendship and encouragement and for helping me develop hobbies which helped me to maintain a healthier life at Purdue.

I am forever grateful with my parents and siblings for always believing in me and teaching me by example the value of hard work, responsibility, and perseverance. I thank John and Fran Foster who adopted me as their grandchild and all the Foster family which became my family away from home. Lastly, but not least to my boyfriend René for his patient support throughout these years.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

STATEMENT OF PUBLISHED AND COLLABORATIVE WORK

Chapter 1 has been published as a review in Current Opinion in Biotechnology on April 2019 on Issue 56 pages 202-205 under the DOI: 10.1016/j.copbio.2018.12.008. Xiangying Mao contributed to writing this manuscript.

Chapters 2-4 will be published in the near future.

ABSTRACT

The phenylpropanoid pathway gives rise to a wide variety of specialized metabolites, but the majority of carbon flux going through this pathway is directed towards the synthesis of the lignin monomers: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. Lignin is a major impediment in biomass saccharification, which negatively affects animal feed and biofuel production. In an effort to improve biomass for the latter purposes, researchers have altered the polymer through genetic manipulations and generated biomass with lower recalcitrance to saccharification; however, in many cases these efforts have resulted in plant dwarfism. To date, we do not have a full understanding of the extent of lignin modifications a plant is able to tolerate without affecting its growth. More importantly, the mechanism that links dwarfism and modifications in lignin content and composition remains unknown. To contribute to answering these questions, we designed a strategy to incorporate a novel monomer into the lignin of *Arabidopsis thaliana*. We used mutants in genes that code for enzymes and regulators of the phenylpropanoid pathway to redirect the pathway's flux towards the synthesis of *p*-coumaraldehyde and prevent the incorporation of *p*-coumaryl alcohol. Despite being mutated for the genes typically considered to be required for monolignol biosynthesis, the plants we generated continue to incorporate *p*-coumaryl alcohol into their lignin. This result suggests that the pathway's architecture has not been completely elucidated and that there are more enzymes involved in lignification than previously thought. Additionally, we explored the connection between perturbations in phenylpropanoid metabolism and plant growth, by using an inducible system to track the changes in gene expression and metabolism that occur when phenylpropanoid metabolism is restored in a lignin biosynthetic mutant. The use of an inducible system allowed us to not only determine the metabolic processes affected in this mutant, but the proximal sequence of events that lead to restored growth when a functional copy of the mutant gene is induced. Finally, we redirected the flux through the pathway to assess the effects of simultaneously modulating lignin content and composition. Through this project we discovered that redirecting phenylpropanoid flux towards the synthesis of sinapyl alcohol in lignin-deficient mutant backgrounds, results in plant dwarfism. The growth impairment of these mutants can be overcome by providing exogenous coniferyl alcohol, suggesting that dwarfism in these mutants is caused by deficiency in coniferyl alcohol and/or derivatives thereof and not lignin alone. Altogether these

projects allowed us to define the cellular processes affected by perturbations in phenylpropanoid homeostasis and the role of other phenylpropanoids besides lignin in this process.

CHAPTER 1. LINKING PHENYLPROPANOID METABOLISM, LIGNIN DEPOSITION, AND PLANT GROWTH INHIBITION

1.1 Abstract

Lignin, a polymer found in the plant secondary cell wall, is a major contributor to biomass' recalcitrance towards saccharification. Due to this negative impact toward the value of lignocellulosic crops, there is a special interest in modifying the content and composition of this important plant biopolymer. For many years this endeavor has been hindered by the plant growth inhibition that is often associated with manipulations to phenylpropanoid metabolism. Although the actual mechanism by which dwarfism arises remains unknown, recent advances in tissue-specific lignin complementation and better understanding of phenylpropanoid transcriptional regulation has made it possible to disentangle lignin modification from perturbations in plant development.

1.2 Introduction

Lignin, a polymer found in the plant secondary cell wall, contributes substantially to the recalcitrance of biomass towards saccharification [1]. Monolignols, the building blocks of this naturally heterogeneous polymer, are the quantitatively most important products of phenylpropanoid biosynthesis. The pathway begins with the deamination of phenylalanine, and in some plants tyrosine, followed by a series of hydroxylations, acyl-transfer reactions, and methylations which give rise to numerous compounds including flavonoids and hydroxycinnamic acid conjugates as well as the three canonical monolignols: *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol [2•]. Polymerization of these compounds produces hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin units, respectively.

More than three decades ago it was observed that manipulation of phenylpropanoid pathway enzymes has an effect on growth and development when mung bean seedlings were grown in the presence of l-α-aminooxy-β-phenylpropanoic acid, an inhibitor of phenylalanine ammonia lyase [3,4]. This treatment resulted in decreased lignin content, collapsed xylem vessels and reduction in growth [3,4]. More recently, in an effort to modify biomass so that it is better suited for biofuel purposes, researchers have altered the polymer through mutation of phenylpropanoid pathway

genes or through other stable genetic manipulations such as RNA interference (RNAi) [2•,5]. In many cases this again led to growth inhibition [6]. In the well-studied model system *Arabidopsis thaliana,* these growth defects ranged from decreased rosette size to failure to bolt and even stalling at the cotyledon stage [7–10]and in poplar and alfalfa, reduced stem height, internode length, diameter, and volume [11–15]. For the most part, mutation or down-regulation of genes that code for the core enzymes, such as phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate: CoA ligase (4CL), hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), *p*-coumaroyl shikimate 3′-hydroxylase (C3′H), caffeoyl shikimate esterase (CSE), and cinnamoyl-CoA reductase (CCR), exhibit at least one of the phenotypes described earlier [8,10–13,16–22]. In contrast, perturbation of steps later in the pathway, such as ferulate 5-hydroxylase (F5H), caffeic acid *O*-methyltransferase (COMT) and cinnamyl alcohol dehydrogenase (CAD), lead to little in the way of growth defects [13,23–27].

Despite advances in understanding and manipulating lignin biosynthesis, there is a recurring question that remains unanswered: why do some lignin mutants exhibit the impaired growth that interferes with the broader deployment of lignin engineered plants? This phenomenon has been referred to as lignin modification-induced dwarfism (LMID) [28] implying that it can be attributed directly to the lack of lignin as a structural component required to support water transport, but other models for LMID have been proposed that associate dwarfism with the lack or hyperaccumulation of pathway end products or the activation of cell wall integrity surveillance mechanisms. At this time, the actual cause or causes of the LMID phenotype remain unknown but are under active investigation. This review will survey some of the most recent advances in this developing area.

1.3 Vasculature defects may lead to dwarfism

Lignin is deposited in multiple cell types including the fiber cells of stems where it provides mechanical strength to the plant body [29]. Perhaps even more critically, lignin provides the rigidity and hydrophobicity to the secondary cell walls of tracheary elements that allows them to withstand the negative pressure of water transport [30]. This important function of lignin was demonstrated clearly by the characterization of the *irregular xylem 4* (*irx4*) mutant [31] and other mutants defective in lignin biosynthetic genes [10,32] and their regulators [33–36], as well as plants in which phenylpropanoid genes were downregulated [16, 8]. In all cases, the plants'

tracheary elements exhibited deformed walls that indicated that they had collapsed due to the tensions generated during transpiration. Vascular collapse is not limited to plants with low lignification: the characterization of some of the cellulose- and hemicellulose-deficient *irx* mutants revealed that deficiencies in these polymers also result in xylem collapse, highlighting the interdependency of multiple cell wall components for vascular integrity [30,37,38]. Because reductions in lignin and other cell wall polymers result in weak secondary cell walls and vascular collapse, it has been proposed that the ensuing drop in vascular conductivity leads to growth impairment (Figure 1.1) [16,30,37]. The first evidence for this came from the analysis of mutants perturbed in interfascicular fiber development [29,39]. Although these mutants exhibit pendant phenotypes, they neither have collapsed xylem vessels nor are they dwarfed [29,39]. Complementation of mutants using genes driven by their cognate promoters ameliorates vascular collapse but does not demonstrate whether all cell types require lignification or only a subset of them do. To address this issue and building on earlier studies [40,41], a synthetic promoter containing repeats of the *SECONDARY WALL NAC BINDING ELEMENT* (*SNBE*) was used to drive the expression of *CCR1* in the *ccr1* mutant [28]. This approach resulted in vessel-specific expression, restoration of vascular integrity, and restored growth [28]. In a complementary approach, Smith *et al.* demonstrated that downregulation of *CCR1* via a microRNA driven by the fiber-specific promoter *AtPRX64* does not lead to vasculature collapse nor impaired growth [42]. These studies support the vasculature collapse model described earlier by demonstrating that dwarfism can be avoided provided that proper lignification is maintained in xylem vessels. They also suggest that it may be possible to reduce total lignin content by targeting lignin-modification strategies to fiber cells. These findings not only demonstrate the importance of xylem vessel lignification on plant growth, but also open opportunities for biomass tailoring as these plants have little to no losses in yield and their biomass is more amenable to saccharification due to their lower lignin content [28,40-42].

1.4 Altered soluble phenylpropanoid metabolite levels may lead to dwarfing

Other lines of research have suggested that manipulation of flux through the phenylpropanoid pathway may lead to dwarfism because of hypo- or hyper-accumulation of compounds that are not involved in lignification *per se*, but are pathway intermediates or derivatives thereof (Figure 1.1). For example, the *ref3* mutant accumulates high levels of

cinnamate-derived compounds due to its reduced C4H activity [10]. Cinnamic acid affects auxin transport and signaling [43–45] and Steenackers *et al.* recently demonstrated that *cis*-cinnamic acid, formed by the photoisomerization of *trans*-cinnamic acid [45], is the specific isomer capable of inhibiting auxin efflux in *Arabidopsis* roots [46]. Its hyperaccumulation and the subsequent impairment in auxin signaling may be responsible for the dwarf phenotype of *ref3-2* [10] or may augment the LMID in the mutant.

Transgenics of alfalfa and *Arabidopsis* with downregulated *HCT*, the enzyme that catalyzes the conversion of *p*-coumaroyl-CoA to *p*-coumaroyl shikimate, are severely dwarf and accumulate high levels of flavonoids [8,20]. Although flavonoids play important roles in a plant's life cycle, their synthesis is dispensable for plant growth and development as mutants devoid of flavonoids are able to grow and reproduce [47,48]. It was suggested that inhibition of auxin transport by flavonoids leads to dwarfing in *HCT-RNAi* plants because down regulation of *CHALCONE SYNTHASE* (*CHS*) by RNAi appeared to lead to growth restoration [49]. These findings were later refuted when *CHS* null alleles were used to test the same hypothesis [50]. In addition to flavonoids, *HCT-RNAi* plants (alfalfa and *Arabidopsis*) also accumulate high levels of the stress hormone salicylic acid (SA), whose biosynthesis shares enzymatic steps with lignin biosynthesis [15,51]. The content of SA negatively correlates with stem height in HCT-deficient plants, leading to the hypothesis that the hyperaccumulation of SA may be the cause of dwarfism in these plants [51]. Preventing the biosynthesis of SA by knocking out isochorismate synthase in *Arabidopsis* restores the growth of the *HCT-RNAi* plants without increasing total lignin content [51]. This mechanism of dwarfing may be unique to HCT-deficient lines given that other phenylpropanoid mutants exhibit dwarf phenotypes and decreased SA levels [18] or increased SA levels and normal growth [52].

Xue *et al.* attributed the dwarf phenotype of *ccr1* to the hyperaccumulation of ferulic acid which acts to antagonize the reactive oxygen species burst necessary for cell proliferation exit, leading to improper leaf development in *ccr1* [53]. These results have been challenged recently by De Meester *et al.* who restored expression of *CCR1* in xylem vessels and found that it was sufficient to restore the growth of *ccr1,* despite continued high levels of ferulic acid [28]. Although this finding casts doubt on the original model, the sites for the putative ferulic acid effects remain unknown. If its accumulation and subsequent reactive oxygen species burst inhibition leads to

dwarfism through xylem-localized events, the original model may still explain how the deleterious phenotype arises.

Hypo-accumulation of phenylpropanoids with intrinsic growth-promoting activity, or of compounds that indirectly facilitate growth could also explain the growth inhibition seen in plants with altered phenylpropanoid metabolism. Substances that meet the first criterion include isomers of dehydrodiconiferyl glucoside (DCG) that behave like cytokinins in that they promote growth and cell division [54,55]. Genetic evidence for the relevance of these compounds in normal growth and development is still missing, possibly because their synthesis is intertwined with that of monolignols. For example, the phenotype of the *caffeic acid O-methyltransferase* (*comt*) *caffeoyl-CoA 3-O-methyl transferase* (*ccoaomt*) double mutant could be attributed to the lack of DCGs; however, it could also be due to LMID [9]. Another possible player is fraxetin, a compound derived from coniferyl alcohol that allows plants to thrive under low-iron environments [56,57]. Although this compound only becomes relevant under suboptimal growth conditions [56,57], the discovery of its biosynthetic pathway serves as a reminder that there is still much ground to be explored in terms of phenylpropanoids that are relevant for growth and development beyond structural components such as lignin.

1.5 Dwarfism could be triggered by a cell wall integrity system

Plant cell walls are critical to cell morphogenesis and participate in direct interactions with abiotic and biotic stressors. During both processes, cell wall integrity must be maintained. It has been proposed that plants possess a sensing mechanism (Figure 1.1) devoted to monitoring and maintaining cell wall architecture and biochemistry as well as activating defense responses to fend off pathogens [58]. This hypothesis is based on a well characterized system in yeast that monitors the integrity of the cell wall [58,59]. Once activated, this system triggers a cascade of transcriptional changes that aid in overcoming damage [58,59]. The reallocation of resources towards the production of defense compounds often comes at the expense of plant growth [60]. Although the function of this putative monitoring system may be to maintain cell wall integrity, the constitutive activation of defense responses by cell wall modifications may lead to dwarfing. Presumably, receptors in the plasma membrane are the initial sensors that detect cell wall damage. For example, members of the plasma membrane localized receptor-like kinase (RLK) family, WALL ASSOCIATED KINASES (WAKs) can bind oligogalacturonides that are recognized as

Damage-Associated Molecular Patterns (DAMPs) and result in transcriptional changes of genes associated with defense and cell wall expansion [61,62]. Furthermore, the RLK THESEUS1 was identified as a suppressor for the impaired growth of a *CELLULOSE SYNTHASE A6* mutant [63]. In the double mutant, cellulose content was not restored, further supporting the idea that dwarfism is not the direct outcome of structural deficits [63].

The biochemical changes commonly found in the cell walls of phenylpropanoid mutants have also been suggested to activate the cell wall integrity system, because several of these mutants exhibit constitutive activation of defense responses. For instance, HCT-downregulated alfalfa and *Arabidopsis* deposit low levels of lignin that is primarily composed of H-subunits and exhibit increased expression of pathogenesis response (PR) genes [15,51]. Similarly, microarray analysis revealed that defense response genes are mis-regulated in plants lacking or overexpressing F5H [64•], which as a result deposit primarily G or S lignin, respectively, even though lignin content is unaltered [65]. Further, the addition of aqueous extracts of cell walls originating from these two backgrounds to wild-type *Arabidopsis* cell cultures induced expression of some of the defense genes measured in the genetically modified plants [64]. These results support the existence of a cell wall integrity system that senses perturbations in the cell wall and activates immune responses [64•]. The degree and way in which a particular mutant's cell wall chemistry is modified may dictate the scope of the defense response activated by the cell wall integrity system, which could in turn impact the severity of growth impairment.

The *C3′H*-deficient *ref8-1* mutant synthesizes primarily H-lignin and is severely dwarfed, a phenotype previously ascribed to its very low lignin content [32]. Surprisingly, mutation of the genes encoding two subunits of the transcriptional coregulator Mediator (*MED5*A and *MED5B*) in *ref8-1* resulted in restored growth and enhanced deposition of the novel H-lignin found in the mutant [52]. In related research, a point mutation in *MED5B* in the *Arabidopsis ref4-3* mutant was found to lead to a dwarf phenotype and decreased soluble phenylpropanoids and lignin [67]. A suppressor screen for *ref4-3* identified several intragenic suppressors as well as extragenic suppressors in the genes encoding Mediator subunits MED2, MED16, and MED23 [67]. All the suppressors restored growth to varying degrees, but unlike the rest of the suppressors identified, one of the intragenic suppressors and the *med2* suppressor did not restore normal phenylpropanoid metabolism [67]. Together, the restoration of growth in *med5a/5b ref8* and in several of the *ref4-3* suppressor mutants demonstrates that, at least in these mutant contexts, inhibition of growth and

phenylpropanoid metabolism can be genetically disentangled. Interestingly, *med16* was also identified as a suppressor of *cob-6,* a mutant with decreased cellulose content and altered pectin methylesterification [68]. This finding serves as an indication that Mediator may play a key role in a pathway that senses cell wall perturbations as a whole and is not limited to sensing perturbations in the phenylpropanoid pathway [68].

1.6 Conclusion

The plasticity of lignin biosynthesis is an attractive feature of this polymer that has been exploited with the end goal of improving biomass for biofuel production. Due to the economic importance of this biopolymer, research efforts have been channeled into decoupling lignin content manipulation from negative impacts on plant growth. The discovery of vessel-specific complementation and a better understanding of phenylpropanoid transcriptional regulatory mechanisms may make this possible. Despite that, we do not have an answer for the question raised earlier: why do some lignin deficient mutants exhibit impaired growth? If the answer is the drop in vascular conductivity caused by low lignification, then vessel complementation of any given lignin deficient mutant should be sufficient to rescue their growth. On the other hand, if it is due to the hypo- or hyper-accumulation of a phenylpropanoid metabolite, the mode of action may vary from mutant to mutant. If dwarfism arises from the activation of a cell wall integrity system, in principle the mechanism may be common to all mutants with defective lignin deposition. The currently available evidence cannot discriminate between the models presented here (Figure 1.1), nor address the possibility that more than one of them is correct. If phenylpropanoids are important for growth in multiple contexts, then it is possible that different mechanisms explain the growth perturbations seen in different plants or that multiple mechanisms contribute to dwarfing in a given lignin-modified plant.

1.7 References

1. Zeng Y, Zhao S, Yang S, Ding S-Y: Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr Opin Biotechnol* 2014, 27:38–45.
2. Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD: Designer lignins: Harnessing the plasticity of lignification. *Curr Opin Biotechnol* 2016, 37:190–200.

3. Amrhein N, Frank G, Lemm G, Luhmann H-B: Inhibition of lignin formation by L-α-amino-β-phenylpropanoic acid, an inhibitor of phenylalanine ammonia-lyase. *Eur J Cell Biol* 1983, 29:139–144.
4. Smart CC, Amrhein N: The influence of lignification on the development of vascular tissue in *Vigna radiata* L. *Protoplasma* 1985, 124:87–95.
5. Wang P, Dudareva N, Morgan JA, Chapple C: Genetic manipulation of lignocellulosic biomass for bioenergy. *Curr Opin Chem Biol* 2015, 29:32–39.
6. Bonawitz ND, Chapple C: Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr Opin Biotechnol* 2013, 24:336–343.
7. Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C: The Arabidopsis *REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J* 2002, 30:33–45.
8. Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M: Silencing of *Hydroxycinnamoyl-Coenzyme A Shikimate/Quinate Hydroxycinnamoyltransferase* affects phenylpropanoid biosynthesis. *Plant Cell* 2004, 16:1446–1465.
9. Do C-T, Pollet B, Thévenin J, Sibout R, Denoue D, Barrière Y, Lapierre C, Jouanin L: Both caffeoyl-Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. *Planta* 2007, 226:1117–1129.
10. Schilmiller AL, Stout J, Weng J-K, Humphreys J, Ruegger MO, Chapple C: Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in Arabidopsis. *Plant J* 2009, 60:771–82.
11. Leplé J-C, Dauwe R, Morreel K, Storme V, Lapierre C, Pollet B, Naumann A, Kang K-Y, Kim H, Ruel K, et al.: Downregulation of *Cinnamoyl-Coenzyme A Reductase* in poplar: Multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* 2007, 19:3669–3691.
12. Voelker SL, Lachenbruch B, Meinzer FC, Strauss SH: Reduced wood stiffness and strength, and altered stem form, in young antisense *4CL* transgenic poplars with reduced lignin contents. *New Phytol* 2011, 189:1096–1109.

13. Wang JP, Matthews ML, Williams CM, Shi R, Yang C, Tunlaya-Anukit S, Chen H-C, Li Q, Liu J, Lin C-Y, *et al*.: Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat Commun* 2018, 9:1579.
14. Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA: Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa* L.). *Proc Natl Acad Sci U S A* 2005, 102:16573–16578.
15. Gallego-Giraldo L, Jikumaru Y, Kamiya Y, Tang Y, Dixon RA: Selective lignin downregulation leads to constitutive defense response expression in alfalfa (*Medicago sativa* L.). *New Phytol* 2011, 190: 627-639.
16. Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet A-M: Down-regulation of *Cinnamoyl-CoA Reductase* induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 1998, 13:71–83.
17. Goujon T, Ferret V, Mila I, Pollet B, Ruel K, Burlat V, Joseleau J-P, Barrière Y, Lapierre C, Jouanin L: Down-regulation of the *AtCCR1* gene in *Arabidopsis thaliana*: Effects on phenotype, lignins and cell wall degradability. *Planta* 2003, 217:218-228.
18. Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y-H, Yu J-Q, Chen Z: Functional analysis of the Arabidopsis *PAL* gene family in plant growth, development, and response to environmental stress. *Plant Physiol* 2010, 153:1526–1538.
19. Li Y, Kim JI, Pysh L, Chapple C: Four isoforms of Arabidopsis 4-coumarate:CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol* 2015, 169:2409–2421.
20. Shadle G, Chen F, Reddy MSS, Jackson L, Nakashima J, Dixon RA: Down-regulation of *Hydroxycinnamoyl CoA:Shikimate Hydroxycinnamoyl Transferase* in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* 2007, 68:1521–1529.
21. Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P, et al.: Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* 2013, 341:1103–1106.
22. Ha CM, Escamilla-Trevino L, Serrani Yarce JC, Kim H, Ralph J, Chen F, Dixon RA: An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J* 2016, 86:363-375.

23. Chapple CCS, Vogt T, Ellis BE, Somerville CR: An Arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell* 1992, 4:1413–1424.
24. Goujon T, Sibout R, Pollet B, Maba B, Nussaume L, Bechtold N, Lu F, Ralph J, Mila I, Barrière Y, et al.: A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. *Plant Mol Biol* 2003, 51:973–989.
25. Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Chen F, Foston M, Ragauskas A, Bouton J, *et al*.: Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci U S A* 2011, 108:3803–3808.
26. Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier M-T, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, *et al*.: Red xylem and higher lignin extractability by down-regulating a *Cinnamyl Alcohol Dehydrogenase* in poplar. *Plant Physiol* 1996, 112:1479-1490.
27. Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L, Séguin A: *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* 2005, 17:2059–2076.
28. De Meester B, de Vries L, Özparpucu M, Gierlinger N, Corneillie S, Pallidis A, Goeminne G, Morreel K, De Bruyne M, De Rycke R, *et al*.: Vessel-specific reintroduction of CINNAMOYL-COA REDUCTASE 1 (CCR1) in dwarfed *ccr1* mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol* 2018, 176:611–633.
29. Zhong R, Taylor JJ, Ye Z-H: Disruption of interfascicular fiber differentiation in an Arabidopsis mutant. *Plant Cell* 1997, 9:2159–2170.
30. Turner SR, Somerville CR: Collapsed xylem phenotype of Arabidopsis Identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell Online* 1997, 9:689–701.
31. Jones L, Ennos AR, Turner SR: Cloning and characterization of *irregular xylem4* (*irx4*): A severely lignin-deficient mutant of *Arabidopsis*. *Plant J* 2001, 26:205–216.
32. Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C: Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of Arabidopsis. *Plant J* 2002, 30:47-59.

33. Stout J, Romero-Severson E, Ruegger MO, Chapple C: Semidominant mutations in *Reduced Epidermal Fluorescence 4* reduce phenylpropanoid content in Arabidopsis. *Genetics* 2008, 178:2237–2251.
34. Zhong R, Lee C, Zhou J, McCarthy RL, Ye Z-H: A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell* 2008, 20:2763–2782.
35. Zhou J, Lee C, Zhong R, Ye Z-H: MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell* 2009, 21:248–266.
36. Zhou J, Zhong R, Ye Z-H: Arabidopsis NAC domain proteins, VND1 to VND5, are transcriptional regulators of secondary wall biosynthesis in vessels. *PLoS One* 2014, 9:1–13.
37. Peña MJ, Zhong R, Zhou G-K, Richardson EA, O'Neill MA, Darvill AG, York WS, Ye Z-H: Arabidopsis *irregular xylem8* and *irregular xylem9*: Implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 2007, 19:549–563.
38. Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR: Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. *Plant J* 2007, 52:1154–1168.
39. Zhong R, Richardson EA, Ye Z-H: Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. *Planta* 2007, 225:1603–1611.
40. Yang F, Mitra P, Zhang L, Prak L, Verhertbruggen Y, Kim J-S, Sun L, Zheng K, Tang K, Auer M, *et al.*: Engineering secondary cell wall deposition in plants. *Plant Biotechnol J* 2013, 11:325–335.
41. Vargas L, Cesarino I, Vanholme R, Voorend W, de Lyra Soriano Saleme M, Morreel K, Boerjan W: Improving total saccharification yield of Arabidopsis plants by vessel-specific complementation of *caffeoyl shikimate esterase* (*cse*) mutants. *Biotechnol Biofuels* 2016, 9:1–16.
42. Smith RA, Schuetz M, Karlen SD, Bird D, Tokunaga N, Sato Y, Mansfield SD, Ralph J, Samuels AL: Defining the diverse cell populations contributing to lignification in Arabidopsis stems. *Plant Physiol* 2017, 174:1028–1036.

43. Van Overbeek J, Blondeau R, Horne V: Trans-cinnamic acid as an anti-auxin. *Am J Bot* 1951, 38:589–595.
44. Wong WS, Guo D, Wang XL, Yin ZQ, Xia B, Li N: Study of *cis*-cinnamic acid in *Arabidopsis thaliana*. *Plant Physiol Bioch* 2005, 43:929–937.
45. Liu C-m, Xu Z-h, Chua N-H: Auxin polar transport is essential for the establishment of bilateral symmetry during early plant embryogenesis. *Plant Cell* 1993, 5:621–630.
46. Steenackers W, Klíma P, Quareshy M, Cesarino I, Kumpf RP, Corneillie S, Araújo P, Viaene T, Goeminne G, Nowack MK, *et al*.: *cis*-Cinnamic acid is a novel, natural auxin efflux inhibitor that promotes lateral root formation. *Plant Physiol* 2017, 173:552–565.
47. Koes RE, Quattrocchio F, Mol JNM: The flavonoid biosynthetic pathway in plants: Function and evolution. *BioEssays* 1994, 16:123–132.
48. Burbulis IE, Iacobucci M, Shirley BW: A null mutation in the first enzyme of flavonoid biosynthesis does not affect male fertility in Arabidopsis. *Plant Cell* 1996, 8:1013–1025.
49. Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M: Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* 2007, 19:148–162.
50. Li X, Bonawitz ND, Weng J-K, Chapple C: The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *Plant Cell* 2010, 22:1620–1632.
51. Gallego-Giraldo L, Escamilla-Trevino L, Jackson LA, Dixon RA: Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc Natl Acad Sci U S A* 2011, 108:20814–20819.
52. Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, *et al*.: Disruption of Mediator rescues the stunted growth of a lignin-deficient *Arabidopsis* mutant. *Nature* 2014, 509:376–380.
53. Xue J, Luo D, Xu D, Zeng M, Cui X, Li L, Huang H: CCR1, an enzyme required for lignin biosynthesis in Arabidopsis, mediates cell proliferation exit for leaf development. *Plant J* 2015, 83:375–387.
54. Binns AN, Chen RH, Wood HN, Lynn DG: Cell division promoting activity of naturally occurring dehydrodiconiferyl glucosides: Do cell wall components control cell division? *Proc Natl Acad Sci U S A* 1987, 84:980–984.

55. Teutonico RA, Dudley MW, Orr JD, Lynn DG, Binns AN: Activity and accumulation of cell division-promoting phenolics in tobacco tissue cultures. *Plant Physiol* 1991, 97:288–297.

56. Tsai H-H, Rodriguez-Celma J, Lan P, Wu Y-C, Vélez-Bermúdez IC, Schmidt W: Scopoletin 8-hydroxylase-mediated fraxetin production is crucial for iron mobilization. *Plant Physiol* 2018, 177:194-207.

57. Rajniak J, Giehl RFH, Chang E, Murgia I, von Wirén N, Sattely ES: Biosynthesis of redox-active metabolites in response to iron deficiency in plants. *Nat Chem Biol* 2018, 14:442–450.

58. Hamann T: The plant cell wall integrity maintenance mechanism – A case study of a cell wall plasma membrane signaling network. *Phytochemistry* 2015, 112:100–109.

59. Levin DE: Regulation of cell wall biogenesis in *Saccharomyces cerevisiae*: The cell wall integrity signaling pathway. *Genetics* 2011, 189:1145–1175.

60. Stamp N: Out of the quagmire of plant defense hypotheses. *Q Rev Biol* 2003, 78:23–56.

61. Decreux A, Messiaen J: Wall-associated kinase WAK1 interacts with cell wall pectins in a calcium-induced conformation. *Plant Cell Physiol* 2005, 46:268–278.

62. Brutus A, Sicilia F, Macone A, Cervone F, De Lorenzo G: A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. *Proc Natl Acad Sci U S A* 2010, 107:9452-9457.

63. Hématy K, Sado P-E, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou JP, Höfte H: A receptor-like kinase mediates the response of *Arabidopsis* cells to the inhibition of cellulose synthesis. *Curr Biol* 2007, 17:922–931.

64. Gallego-Giraldo L, Posé S, Pattathil S, Peralta AG, Hahn MG, Ayre BG, Sunuwar J, Hernandez J, Patel M, Shah J, *et al*.: Elicitors and defense gene induction in plants with altered lignin compositions. *New Phytol* 2018, 219:1235–1251.

65. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C: Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc Natl Acad Sci U S A* 1998, 95:6619–6623.

66. Bonawitz ND, Soltau WL, Blatchley MR, Powers BL, Hurlock AK, Seals LA, Weng J-K, Stout J, Chapple C: REF4 and RFR1, subunits of the transcriptional coregulatory complex Mediator, are required for phenylpropanoid homeostasis in *Arabidopsis*. *J Biol Chem* 2012, 287:5434–5445.
67. Dolan WL, Dilkes BP, Stout JM, Bonawitz ND, Chapple C: Mediator complex subunits MED2, MED5, MED16, and MED23 genetically interact in the regulation of phenylpropanoid biosynthesis. *Plant Cell* 2017, 29:3269-3285.
68. Sorek N, Szemenyei H, Sorek H, Landers A, Knight H, Bauer S, Wemmer DE, Somerville CR: Identification of MEDIATOR16 as the *Arabidopsis* COBRA suppressor MONGOOSE1. *Proc Natl Acad Sci U S A* 2015, 112:16048–16053.

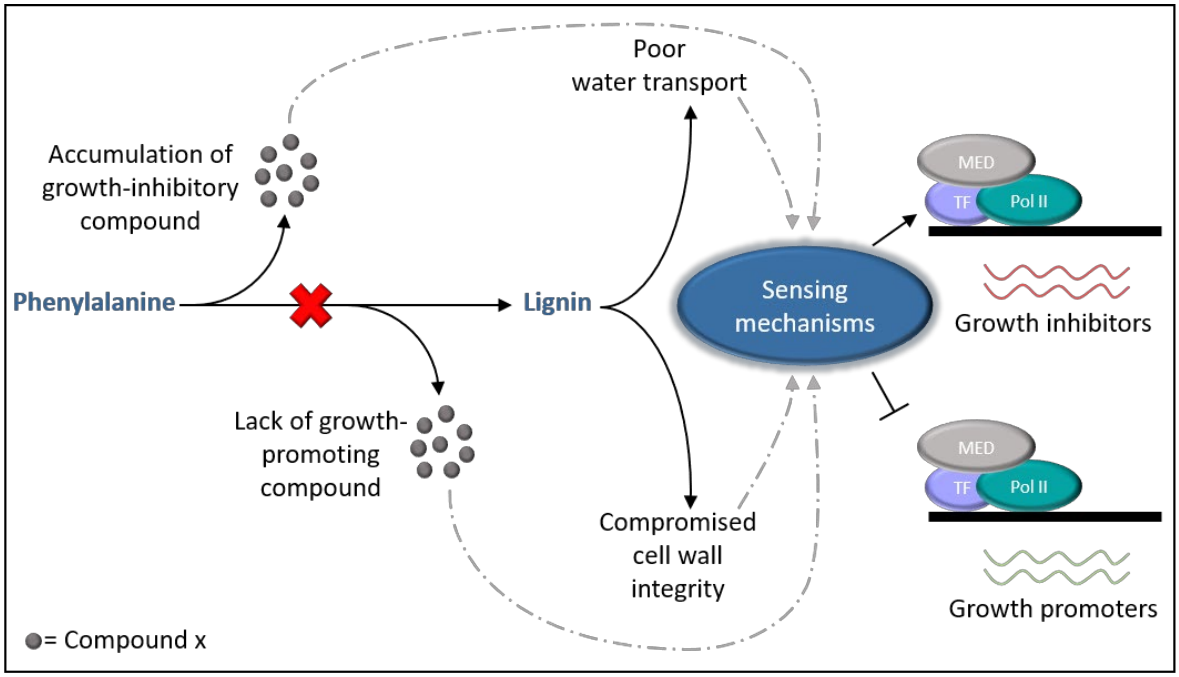


Figure 1.1. Models that link alterations in phenylpropanoid metabolism and lignin deposition to growth inhibition.

Knockout or downregulation of phenylpropanoid pathway genes in many cases leads to growth inhibition and could do so through one or more mechanisms. First, a block in the phenylpropanoid pathway could lead to the accumulation of a growth-inhibitory compound or prevent the synthesis of a growth-promoting compound (depicted as compound x) resulting in growth inhibition. Second, a reduction in lignin content could result in weak xylem vessel cell walls, vascular collapse, and reduced water transport. Finally, if alterations in phenylpropanoid metabolism result in compromised cell wall integrity, a sensing mechanism could detect these changes and activate a transcriptional response that requires Mediator.

CHAPTER 2. METABOLIC ENGINEERING OF P-COUMARALDEHYDE-DERIVED LIGNIN

2.1 Abstract

Lignin contributes substantially to the recalcitrance of biomass towards saccharification. In order to circumvent this problem, researchers have focused on altering the lignin polymer through genetic manipulations, although, in a number of cases these efforts have resulted in an undesired yield penalty. However, recent findings have shown that by knocking out two subunits (MED5A and MED5B) of the transcriptional regulatory complex Mediator the stunted growth phenotype of plants homozygous for a mutant allele of *p*-coumaroyl shikimate 3´-hydroxylase (*C3´H*), *ref8-1,* can be alleviated. Furthermore, these plants synthesize a lignin polymer almost entirely derived from *p*-coumaryl alcohol. Plants deficient in cinnamyl alcohol dehydrogenase (CAD) are notable in that they incorporate substantial levels of coniferaldehyde and sinapaldehyde into their lignin. Aldehyde-enriched lignin is a desirable trait, since it leads to greater cell wall digestibility. We tested the hypothesis that by stacking mutations in *CADC and CADD* on a *med5a/5b ref8-1* genetic background we would block the biosynthesis of *p*-coumaryl alcohol, making *p*-coumaraldehyde available for polymerization into a novel kind of lignin. We have found that these plants (*med5a/5b ref8-1 cadc cadd*) are viable and derivatization followed by reductive cleavage (DFRC) as well as 2D-NMR demonstrated that these plants continue to synthesize *p*-coumaryl alcohol despite being mutated for the CADs typically considered to be required for monolignol biosynthesis. Enzyme activity tests showed that even in the absence of CADC and CADD there is high activity in stems. This observation and the fact that the genome of *A. thaliana* encodes for 9 CAD isoforms suggest that another CAD may be catalyzing the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol. We tested the involvement of other CAD enzymes in the synthesis of *p*-coumaraldehyde in *med5a/5b ref8-1 cadc cadd* by knocking out the genes that code for these enzymes using the CRISPR/Cas9 system. Lignin analysis demonstrate that despite knocking out these genes in *med5a/5b ref8-1 cadc cadd*, plants continue to deposit *p*-coumaryl alcohol-derived lignin.

2.2 Introduction

The global economy is based upon petroleum-based industries, unfortunately such dependency on fossil fuels is associated with many problems. Biofuels represent a potential alternative to oil, which would alleviate many of the world's current environmental, economic and political problems. Bioethanol is one of the most popular biofuels available nowadays, but its production competes for the use of crops and arable land [1] that could be used to feed the increasing world population.

Lignocellulosic biomass is a highly abundant resource for the production of biofuels and commodity chemicals [2]. In evaluating their profitability, competition for farmland need not be a complicating factor because plant waste can be used for their production [3]. However, in order to extract its fermentable sugars, this biomass requires expensive pre-treatments that add significantly to the cost of biofuel production [4]. Recalcitrance of biomass towards saccharification can be primarily attributed to lignin, a polymer found in the plant secondary cell wall [5]. Lignin is made of polymerized monolignols which are themselves product of phenylpropanoid biosynthesis. This pathway gives rise to a wide variety of plant secondary metabolites, but the main products are the monolignols *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol (Figure 2.1). Polymerization of these alcohols by radical coupling occurs in the apoplast and results in hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin, respectively.

Lignin polymerization in the vasculature is neither a template-based nor an enzymatically guided process [6]. This unique feature allows for high plasticity when it comes to monomer incorporation into the growing polymer. In Nature, the incorporation of non-canonical monolignols into lignin is a widespread phenomenon. For instance, in some monocots the flavonoid tricin and the hydroxycinnamates ferulate and *p*-coumarate can be found cross-linked to the polymer [7]. In loblolly pine and mulberry tree mutants, the monolignol aldehyde precursors coniferaldehyde and sinapaldehyde (Figure 2.1) form a substantial component of lignin [8,9]. Lignification was a key feature that allowed plants to colonize terrestrial ecosystems, but why distinct lineages of plants evolved the ability to incorporate atypical monomers is still an open question [10]. Although the evolutionary reason behind this flexibility in monomer incorporation remains unknown, it has certainly served to the advantage of humans attempting to introduce novel monomers into lignin with new functionalities via metabolic engineering.

Lignin has been the target of many metabolic engineering efforts to improve biomass for biofuel production; mainly because lignin content is negatively correlated with sugar release upon biomass enzymatic treatment [5]. Lignin imparts the necessary rigidity and hydrophobicity to plant vascular tissues [11], hence complete removal of lignin biosynthesis is not a viable option. However, successful engineering examples in monocots and dicots have shown that even modest reductions in lignin content can result in an increase in saccharification [12].

Modifying monolignol ratios and incorporation of novel monomers have proven to be effective approaches for reducing biomass recalcitrance [13]. By modulating the expression of ferulate 5-hydroxylase (F5H) the native G to S monomer ratio can be dramatically altered [14–18]. Knocking out *F5H* leads to the accumulation of G-lignin, but plants overexpressing F5H are nearly devoid of G-lignin and instead deposit S-lignin [14–18]. The shift to S-lignin subunits in transgenic poplar overexpressing *F5H* reflects in an increase in pulping efficiency [16] and enzymatic digestion [19]. The diversity of natural and engineered lignin variants indicates that there is an ample range of compounds that can be accepted as lignin monomers. Only a few basic principles dictate if a compound is a substantial component of lignin, including its production level, availability at the lignification zone, and its ability to undergo radical coupling [20]. These basic requirements result in a large repertoire of possible compounds that could be engineered into lignin or may already be available in Nature, but are yet to be discovered.

H-lignin is naturally a minor component in plant's lignin, making only 2% of the total lignin in Arabidopsis thaliana. The reaction immediately downstream of the branch point between H- and G-lignin synthesis is catalyzed by the enzyme *p*-coumaroyl shikimate 3′-hydroxylase (C3′H) and a mutant in the corresponding gene, *ref8-1*, is a dwarf plant that deposits primarily H-lignin (Figure 2.1) [21]. Elimination of two paralogs of the MED5 subunit of the transcriptional co-regulator Mediator alleviates the stunted growth of *ref8-1* plants by blocking the transcriptional reprogramming that leads to dwarfing in the mutant, while retaining their H-rich composition [22]. Upon enzymatic treatment, the lignocellulosic material from *med5a/5b ref8-1* plants yields more than double the amount of glucose when compared to wild type [22].

In addition to their presence in natural mutants, hydroxycinnamaldehydes have been successfully incorporated into lignin by mutation or downregulation of Cinnamyl Alcohol Dehydrogenases (CADs) in species such as *Populus, Arabidopsis, Medicago*, and *Sorghum* [23–26]. CADs catalyze the last step in monolignol biosynthesis by reducing hydroxycinnamaldehydes

to hydroxycinnamyl alcohols (Figure 2.1). In *Arabidopsis*, polysaccharide hydrolysis of *cadc cadd* biomass resulted in ~100% increase in glucose release [27].

H-lignin and aldehyde-rich lignins are desirable traits, because they lead to greater cell wall digestibility. Combination of these two traits could potentially lead to an enhancement in cell wall digestibility beyond that of the individual traits. This hypothetical polymer would be made of p-coumaraldehyde subunits. To our knowledge, this compound has only been found weakly bound to lignin in Cucumis sativa under biotic stress [28] or as the end units of H-lignin, but not as a substantial component of lignin. To test if this polymer could be made *in muro*, we used the high H-lignin background *med5a/5b ref8-1* as a genetic background into which the *cadc*, *cadd* and two other *cad* mutations were added.

2.3 Methods

2.3.1 Plant materials

Arabidopsis thaliana, ecotype Columbia-0 plants were grown in Propagation Mix (Sun Gro Horticulture) supplemented with Osmocote® Plus (ICL Specialty Fertilizers) at 22°C under long day conditions (16 h light/ 8 h dark). The *med5a/5b ref8-1* triple mutant and its genotyping information has been described previously [22]. The *cadc* (SAIL_1265_A06) and *cadd* (SAIL_776_B06) T-DNA lines were obtained from the Arabidopsis Biological Resource Center (The Ohio State University, Columbus, OH) and crossed to generate the *cadc cadd* mutant. Homozygous double mutants were identified by PCR amplification using primers CC5583 and CC5584 for *CADC* and CC5585 and CC5586 for *CADD* (Table 2.1).

2.3.2 RNA extraction and qRT-PCR

RNA from four-week-old stems was extracted using the QIAGEN RNeasy® Plant Kit and DNAse-treated using the Invitrogen™ TURBO DNA-free™ kit following manufacturer's instructions. Complementary DNA was synthesized using 1 µg of RNA and the Invitrogen™ MultiScribe™ Reverse Transcriptase. Quantitative qRT-PCR was performed using Applied Biosystems™ SYBR™ Green PCR Master Mix, gene specific primers and *ACTIN2* (At3g18780) as the internal standard (Table 2.1). Fold change expression was determined by the ΔΔCt method.

2.3.3 Histochemical staining

Histochemical staining on hand-made cross sections of 6-week-old inflorescence stems was performed using toluidine blue-O, phloroglucinol-HCl and Maüle reagent as described by Chapple *et al*., (1992) and Bonawitz *et al*., (2014) [22,29].

2.3.4 Lignin analysis

Lignin monomers were quantified by gas chromatography after derivatization followed by reductive cleavage (DFRC) [30]. Total lignin content was measured by the thioglycolic acid (TGA) method [31].

2.3.5 Protein extraction and CAD activity assays

Four biological replicates consisting of two, five-week-old stems each, were harvested for each genotype. Stems without axillary stems, cauline leaves and siliques were flash frozen in liquid nitrogen and ground with mortar and pestle. Tissue was suspended in four volumes of 100 mM Tris-HCl (pH 7.5), 5% (v/v) glycerol and 15 µM β-mercaptoethanol. Samples were centrifuged 10 min at 16,000 x *g* at 4°C and the supernatant was transferred to a new tube and centrifuged again with the same parameters. A 250 µL fraction of the supernatant was desalted in disposable columns packed with Sephadex G-50 (medium) pre-soaked in extraction buffer. Protein was quantified using the Bradford assay [32] from Bio-Rad and bovine serum albumin was used as a standard.
CAD enzyme assays were performed using a modified version from Wyrambik and Grisebach (1975) [33]. The 200 µL reverse reaction consisted of 3 mM cinnamyl alcohol, 500 µM $NADP^+$ in 200 mM Tris-HCl (pH 8.8). The reaction was started by the addition of 20 µL of desalted protein and incubated at 25°C for 20 min. Assays were stopped by the addition of 24 µL of glacial acetic acid. Samples were centrifuged 1 h at 16,000 x *g* at 4°C prior to analysis by HPLC.
Forward assays were carried using an NADPH regenerating system that consisted of 500 µM $NADP^+$, 5 mM glucose-6-phosphate and 0.5 units of glucose-6-phosphate dehydrogenase in 200 mM potassium phosphate (pH 6.5) in a final volume of 200 µL. 180 µL aliquots containing the NADPH regenerating system and 340 µM cinnamaldehyde were pre-incubated for 5 min at 30°C. Assays were started by the addition of 20 µL of desalted protein and were incubated 30 min at 30°C. Reactions were stopped as described above and analyzed by HPLC.

2.3.6 Generation of *cada, cadg* and *cadd* using the CRISPR/Cas9

The *cada* knockout was generated by constructing two sgRNAs targeting *CADA*. The first sgRNA was generated by cloning the complementary primers CC4942 and CC4943 into the *Bbs*I-digested psgR-Cas9-At plasmid [34] to generate psgR-Cas9-At-target1. The second sgRNA was generated in a similar manner using primers CC4944 and CC4945 and resulted in psgR-Cas9-At-target2. The AtU6 promoter and the sgRNA from psgR-Cas9-At-target2 were PCR amplified with primers containing *Kpn*I and *Eco*RI overhangs as described in [34] and cloned into psgR-Cas9-At-target1 to generate psgR-Cas9-At-target1&2. The cassette containing both sgRNAs and the Cas9 gene were excised from psgR-Cas9-At-target1&2 by digesting with *Eco*RI and *Hin*dIII and cloned into the binary vector pCAMBIA1300 resulting in psgR-Cas9-At-CADA. The construct was introduced into *med5a/5b ref8-1 cadc cadd* plants via *Agrobacterium tumefaciens*-mediated transformation [35]. Transformants were selected on plates with hygromycin and transferred to soil. Genomic DNA (gDNA) was extracted from the transformants and PCR amplified with primer pair CC4949/CC5009 for target 1 and CC5012/CC5007 for target 2. Amplicons were sequenced by the Purdue Genomics Center with primers CC5010 and CC5011 for target 1 and C4950 and CC5013 for target 2.

The *cadg* knockout was generated by transforming *med5a/5b ref8-1 cadc cadd* plants with psgR-Cas9-At-CADG. The psgR-Cas9-At-CADG plasmid was constructed as described above using primers CC5494 and CC5495 for target 1 and CC5496 and CC5497 for target 2. After selection, gDNA was extracted from the transformants and both targets were PCR amplified with primers CC5418 and CC5236. The amplicon was sequenced by Eurofins Genomics LLC using primer CC5417 for target 1 and CC5236 for target 2. One of the transformants carried a single 'A' insertion in target 1 that generated a new *Bsm*AI site. A homozygous line for this edit was selected by genotyping the next generations using the primer pair CC5416 and CC5417 and digesting the amplicon with *Bsm*AI.

CADD was targeted with a single sgRNA which was generated by cloning the complementary primers CC5847 and CC5848 into psgR-Cas9-At. The plasmid was digested with *Kpn*I and *Hin*dIII to excise the sgRNA and Cas9 cassette and cloned into pCAMBIA1300 to generate psgR-Cas9-At-CADD. The sgRNA was designed to target an existing *Bsa*AI site, therefore after transformation of *med5a/5b ref8-1 cadc cadd,* genomic DNA from T1 plants was PCR amplified with primers CC5861 and CC5862 and digested with *Bsa*AI. DNA from plants that

generated amplicons that did not digest with *Bsa*AI were used for amplification of the corresponding genomic region and were sequenced with primer CC5863 by Eurofins Genomics LLC.

CADG and *CADD* targets were designed using the online tool E-CRISP (http://www.e-crisp.org/E-CRISP/) [36]. Primer sequence information can be found on Table 2.1.

2.3.7 CCR heterologous expression and enzyme assays

Arabidopsis thaliana CCR1 and CCR2 were cloned into the Gateway® pDEST^{TM}17 vector and expressed in *Escherichia coli* BL21 (DE3) cells as fusion proteins with an N-terminus 6X-His tag. As previously described in Li *et al.* (2010) [37] in order to generate the substrate *p*-coumaroyl CoA, 4-coumarate ligase 1 (4CL1) was expressed in *E. coli* BL21 (DE3) cells as a fusion protein with a 6X-His tag. For all proteins, expression was induced by 2.5 mM isopropyl β-D-thiogalactoside and incubated overnight at 18°C. Cells were harvested and lysed with lysozyme and DNAseI as described by Weng *et al.* (2011) [38]. The three fusion proteins were purified using Ni-NTA agarose beads. PD-10 (GE-Healthcare) columns were used for desalting the proteins and buffer exchange; the column was equilibrated with 100 mM Tris-HCl, 10% glycerol for 4CL and 100 mM HEPES-NaOH (pH 7.5), 10% glycerol for CCR1/2 [39]. Protein was quantified as described above.

Enzyme assays to generate *p*-coumaroyl CoA were carried by incubating 4CL1 with 200 µM *p*-coumarate, 220 µM Coenzyme A, 1 mM ATP, 1mM $MgCl_2$ in 100 mM Tris-HCl (pH 8.0) for 30 min at 30°C. Reactions were coupled with both CCR enzymes and 1mM NADPH and incubated at 30°C. The reactions were terminated by acidification with glacial acetic acid and the products analyzed by HPLC.

2.4 Results

2.4.1 *p*-coumaryl alcohol is the main constituent of lignin in *med5a/5b ref8-1 cadc cadd*

To determine whether *p*-coumaraldehyde can be incorporated into lignin, it was necessary to first generate genetic material in which *p*-coumaraldehyde would accumulate at high levels. As mentioned earlier, *med5a/5b ref8-1* synthesize lignin primarily derived from *p*-coumaryl alcohol [22]. In an attempt to block the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol in *med5a/5b*

ref8-1, we crossed it to *cadc cadd* and identified *med5a/5b ref8-1 cadc, med5a/5b ref8-1 cadd,* and *med5a/5b ref8-1 cadc cadd* plants among the segregating population. The addition of *cadc*, *cadd*, or *cadc cadd* mutations to *med5a/5b ref8-1* does not alter its growth substantially (Figure 2.2), although the quintuple mutants do have reduced stem diameter (data not shown).

We next analyzed stem cross sections to assess whether knocking out *CADC* and *CADD* in *med5a/5b ref8-1* plants affects vascular integrity, as is commonly observed in other mutants of the lignin biosynthetic pathway with collapsed xylem [21,40,41]. In particular, the *med5a/5b ref8-1* mutant exhibits partial vascular collapse in a small portion of vessels [22]. Untreated stem cross sections (Figure 2.3A) as well as those stained with toluidine blue-O (Figure 2.3B) from the quadruple and quintuple mutant do not exhibit further xylem collapse when compared to *med5a/5b ref8-1,* which indicates that the addition of *cadc cadd* mutations does not affect vascular integrity. To obtain an indication of lignin content and composition of the quadruple and quintuple mutants we performed histochemical staining. Phloroglucinol-HCl stains hydroxycinnamaldehyde end units of lignin and in wild-type plants it renders a purple color, whereas in high H mutants it results in a dark red pigmentation of the vasculature (Figure 2.3C) [22]. Quadruple and quintuple mutant stems also exhibit a dark red pigmentation when treated with phloroglucinol-HCl (Figure 2.3C). The Maüle reagent is used to differentiate between G (brown) and S (red) lignin and in wild-type *Arabidopsis* plants the vascular bundles are stained brown whereas the interfascicular fibers are stained red [29]. High H mutants have a brighter red pigmentation of the interfascicular fibers while *cadc cadd* mutants do not acquire any pigmentation at this region [22,24]. As shown in Figure 2.3D, the *med5a/5b ref8-1 cadc, med5a/5b ref8-1 cadd* and *med5a/5b ref8-1 cadc cadd* mutants stain in a similar manner as the *med5a/5b ref8-1* mutant. Altogether, the results from histochemical staining suggested the quadruple and quintuple mutants have similar lignin composition to *med5a/5b ref8-1*.

Histochemical staining provides an indication of lignin content and composition, but it does not unequivocally inform about subunit composition or provide quantitative information of lignin content. To address this issue, we performed derivatization followed by reductive cleavage (DFRC), thioglycolic acid (TGA) and NMR analysis on quadruple and quintuple mutants stem tissue. The DFRC method cleaves aryl-ether-type linkages, therefore the total amount of lignin will appear to be lower in plants with lignins with alternate linkages [30]. An example of this are *cadc cadd* plants which have cinnamaldehydes cross-coupled by 8-O-4 linkages and when

analyzed by the DFRC appear to have residual levels of lignin (Figure 2.4) [27]. However, when these plants are analyzed by alternate methods these plants only have a 30% reduction in total lignin [24,27]. The quadruple and quintuple mutants, rather than exhibiting a similar DFRC profile as that of *cadc cadd,* unexpectedly seem to have a lignin composition solely comprised of *p*-coumaryl alcohol units (Figure 2.4). Cell wall analysis by HSQC 2D-NMR provides unique signatures for the different lignin monomers and linkages. The analysis showed that the aldehyde, aromatic, and aliphatic regions of the spectrum of *med5a/5b ref8-1 cadc cadd* are indistinguishable from those of *med5a/5b ref8-1* (data not included). To further explore this observation, synthetic dimers and oligomers of *p*-coumaraldehyde were included in the sample set for NMR analysis but, none of the characteristic signatures of these synthetic compounds were observed by NMR in the quintuple mutant. To test if there was an effect on total lignin content, lignin was quantified by the thioglycolic acid (TGA) method [31]. TGA analysis revealed that the total lignin content of the quintuple is the same as that of the *med5a/5b ref8-1* mutant (Table 2.2). In conclusion, mutation of *CADC* and *CADD* in *med5a/5b ref8-1* appears to have no effect in total lignin content nor composition.

2.4.2 Despite the absence of CADC and CADD, plants continue to exhibit CAD activity

Arabidopsis mutants in *CADC* and *CADD* deposit a lignin mainly comprised of hydroxycinnamaldehyde subunits [27]. The fact that *med5a/5b ref8-1 cadc cadd* plants continue to incorporate *p*-coumaryl alcohol in its lignin suggests that there is another enzyme(s) that is able to catalyze the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol. To test this possibility, we performed CAD enzyme activity tests on *cadc cadd* and *med5a/5b ref8-1 cadc cadd* stem tissue and found that they have lower activity towards *p*-coumaraldehyde, coniferaldehyde and sinapaldehyde than wild-type and *med5a/5b ref8-1* plants. Nevertheless, despite the absence of CADC and CADD, plants continue to exhibit CAD activity (Figure 2.5), which was unexpected for *cadc cadd* because it deposits lignin primarily made from hydroxycinnamaldehydes. These observations and the fact that the genome of *A. thaliana* encodes for 9 *CAD* paralogs [42] suggests that another CAD may catalyze the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol.

2.4.3 Knocking out CAD paralogs does not have an effect on lignin composition

Working under the assumption that if there is a CAD specific for *p*-coumaryl alcohol synthesis, we anticipated that this CAD would be upregulated in genetic backgrounds that synthesize high levels of *p*-coumaryl alcohol. Accordingly, we mined an RNA-seq data set generated from *ref8-1 and med5a/5b ref8-1* rosettes [22]. We searched for CAD transcripts that were upregulated in both of these backgrounds and as seen on Table 2.3, *CADD* and *CADA* meet this criterion; however, we had already tested the effect of knocking out *CADD* and therefore selected *CADA* as a candidate to move forward. In addition, because the RNA-seq data available was from rosettes which are not highly lignifying tissues, we performed qRT-PCR using cDNA from *med5a/5b ref8-1 cadc cadd* stems and found that the *CADG* transcript is upregulated in this background (Figure 2.6). To test the hypothesis that these CADs are catalyzing the reduction of *p*-coumaraldehyde in the quintuple mutant, we generated *CADA* and *CADG* knockouts in *med5a/5b ref8-1 cadc cadd* using the CRISPR/Cas9 system [34].

To test whether knocking out *CADA* and *CADG* in the quintuple mutant can lead to the deposition of *p*-coumaraldehyde-derived lignin, independent single strand guide RNAs (sgRNAs) were designed to target two regions upstream of the CAD active site in *CADA* and *CADG*. Two sgRNAs were designed to target each gene with the expectation that two double strand breaks would result in a substantial deletion in the gene that can be readily identifiable by PCR (Mao *et al.*, 2013). For *CADA,* no edits were identified in the target 2 region, therefore selection of transformants in the subsequent generations was done by sequencing the target 1 region. A line carrying a homozygous 'T' insertion in the third exon was used for further experiments. For *CADG,* one of the transformants carried a single 'A' insertion in target 1 located in the fourth exon, that generated a new *BsmA*I restriction site. Plants with a single base pair insertion upstream of the active site were selected for both CADs. Based on the location of the CRISPR/Cas9-mediated edits in the coding sequence we protein function is almost certainly disrupted.

To determine whether *CADA* and/or *CADG* contribute to the quintuple mutant's total CAD activity, we performed CAD activity assays in stems. As seen in Figure 2.5, there is no statistical difference in CAD activity between *med5a/5b ref8-1 cadc cadd* and *med5a/5b ref8-1 cadc cadd cada* nor *med5a/5b ref8-1 cadc cadd cadg* (Figure 2.5). In addition, to determine if mutation of *CADA* and/or *CADG* had an effect on lignin quantity or composition we analyzed the sextuple mutants' lignin by DFRC and TGA analysis., Neither lignin quantity nor composition changed in

med5a/5b ref8-1 cadc cadd cada nor *med5a/5b ref8-1 cadc cadd cadg* when compared to *med5a/5b ref8-1 cadc cadd* (Figures 2.7 and 8, Table 2.3). In conclusion, *CADA* nor *CADG* significantly contribute to *p*-coumaryl alcohol synthesis in lignifying tissue.

2.4.4 Neither CCR1 nor CCR2 can catalyze the reduction of *p*-coumaraldehyde

To expand on our candidate selection, we searched for other enzymes that could be catalyzing the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol. The enzyme 3-hydroxy-3-methyl-glutaryl-coenzyme A (HMG-CoA) reductase catalyzes the reduction of HMG-CoA to mevalonate through a mevaldehyde intermediate [43]. We tested the possibility that Cinnamoyl CoA Reductase (CCR), in a similar mechanism as HMG-CoA reductase can catalyze the two-step reduction of *p*-coumaroyl CoA directly to *p*-coumaryl alcohol, obviating the need for an alcohol dehydrogenase comparable to CADC and CADD. *Arabidopsis thaliana* CCR1 and CCR2 were heterologously expressed with an N-terminus His-tag and purified, rendering active proteins *in vitro.* Both CCRs catalyzed the reduction of *p*-coumaryl CoA to *p*-coumaraldehyde, but *p*-coumaryl alcohol synthesis could not be detected under our assay conditions (Figure 2.9) even when *p*-coumaraldehyde was provided directed as a substrate.

2.4.5 *CADD* is transcribed in *med5a/5b cadd* despite having a disrupted promoter

In addition to genotyping by PCR amplification, we tested the homozygosity of our working material by generating backcrosses. Analysis of the F1 population proceeding from the backcross of *med5a/5b ref8-1 cadc cadd* to *med5a/5b ref8-1* showed that the *reduced epidermal fluorescence* phenotype as well as normal growth of *med5a/5b ref8-1* could be recapitulated (data not shown). On the contrary, the backcross of *med5a/5b ref8-1 cadc cadd* to *cadc cadd* does not exhibit the orange pigmentation in the interfascicular fibers, characteristic of the *cadc cadd* mutant (Figure 2.10). SAIL_776_B06 is a T-DNA line with an insertion in the *CADD* promoter that abolishes transcription in a *cadc* genetic background (Figure 2.11). Although genotyping data indicates that the quintuple mutant is indeed homozygous for T-DNA insertions in *CADC* and *CADD,* transcript analysis shows that *CADD* is transcribed in the quintuple mutant (Figure 2.11). Whereas in *cadc cadd* no *CADD* transcript can be detected, surprisingly, elimination of *MED5A* and *MED5B* in SAIL_776_B06 results in *CADD* transcription as seen in *med5a/5b ref8-1 cadc*

cadd and *med5a/5b cadc cadd* (Figure 2.11). These findings suggest that the activation of CADD expression in the *med5a/5b* background may be responsible for the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol.

2.4.6 *med5a/5b ref8-1 cadc/d CRISPR-CADD* knockouts deposit lignin derived from *p*-coumaryl alcohol

In order to test the hypothesis that the basal expression of CADD is catalyzing the synthesis of *p*-coumaryl alcohol in *med5a/5b ref8-1 cadc/d,* an sgRNA was designed to target a region upstream of the CAD active site. The sgRNA targeted a *Bsa*AI site located in the fourth exon of *CADD* which allowed screening for homozygous mutants by genotyping. Putative homozygous mutants were identified in the T1 by the loss of the *Bsa*AI site. A transheterozygous T2 line, segregating for an 'A' insertion and an 'AC' deletion was used to test the effect of knocking out *CADD* in *med5a/5b ref8-1 cadc/d.* Knocking out *CADD* in *med5a/5b ref8-1 cadc/d* resulted in plants with compromised growth (Figure 2.12) and decreased fertility. As seen in Figure 2.5, *med5a/5b ref8-1 cadc/d CRISPR-CADD* plants have the same activity towards coniferaldehyde and *p*-coumaraldehyde as the *med5a/5b ref8-1 cadc/d* mutant. Lignin analysis (Figure 2.13) revealed the *CRISPR-CADD* plants have the same lignin composition as the *med5a/5b ref8-1 cadc/d* mutant.

2.5 Discussion

The saccharification potential and the absence of a growth deficit of biomass containing aldehyde-derived lignins makes them an attractive substitute for biomass with alcohol-rich lignins. Aldehyde-rich lignocellulosic material not only presents a prospect for the reduction of biofuel production's cost, but it also has demonstrated to be a superior animal feedstock [44]. Hydroxycinnamaldehyde-derived lignins exist in Nature, and they have also been generated in many plant species through metabolic engineering [8,9,23–26]. The strategy necessary for generating coniferaldehyde- and/or sinapaldehyde-rich lignins has been well established in multiple species [23–26]. The recent discovery that *p*-coumaryl alcohol can be incorporated into lignin as the primary component in *med5a/5b ref8*-1 plants [22], and the availability of a strategy to generate aldehyde-derived lignins, led us to explore the possibility of generating plants that would deposit *p*-coumaraldehyde-derived lignin. In an attempt to do so, we crossed *cadc/d* to

med5a/5b ref8-1 and analyzed the lignin of the resulting *med5a/5b ref8-1 cadc/d* plants. Like the *med5a/5b ref8-1* plants, the *med5ab/5b ref8-1 cadc/d* plants deposited *p*-coumaryl alcohol-derived lignin. Therefore, we tested the involvement of other CAD isoforms searching for a CAD that could be catalyzing the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol in *med5a/5b ref8-1 cadc/d* plants. We demonstrated that CADA and CADG do not contribute substantially to lignification in *med5a/5b ref8-1* plants.

2.5.1 *cadc cadd* plants continue to exhibit CAD activity

There are nine CAD isoforms in A. thaliana, namely CADC, CADD, CAD1, CADA, CADB1, CADB2, CADG, CADE, and CADF [42]. The involvement of these CADs in stem lignification has been tested by using T-DNA lines [45], with the exception of *CADF*, because there is no detectable transcription of this gene in wild-type plants [45]. CAD1 has been shown to have an impact in young stem lignification, but *CAD1* knockouts do not exhibit differences in either lignin content or composition when compared to their wild-type counterparts at maturity [45]. CADC and CADD are the alcohol dehydrogenases that contribute most significantly to lignification in wild-type mature stems [24]. Despite having shown previously that the *cadc/d* mutant deposits lignin constituted of 99% aldehyde subunits [24,27], *med5a/5b ref8-1 cadc/d* plants deposit alcohol-derived lignin (Figure 2.4). To account for the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol in these plants, we first tested if there was detectable CAD activity in *med5a/5b ref8-1 cadc/d*. Surprisingly, *med5a/5b ref8-1 cadc/d* plants exhibited ~50% of wild-type CAD activity towards *p*-coumaraldehyde and coniferaldehyde, and *cadc/d* plants did as well (Figure 2.5). *In vitro* studies revealed that, although orders of magnitudes lower than CADC and CADD, CADB1, CADB2, CADE and CADF exhibit activity towards *p*-coumaraldehyde, coniferaldehyde and sinapaldehyde [46]. CADA and CADG could not be expressed *in vitro* [46]. Based on promoter fusion and transcript analysis, CAD1, CADB1, CADB2, CADG and/or CADA [45,47] may contribute to the remaining CAD activity measured in *cadc/d* stems. Expression reports for *CADE* show either no expression in stems or very low levels of expression, making it an unlikely candidate [45,47]. As mentioned earlier, the *cadc/d* mutant deposits lignin primarily derived of aldehyde subunits, therefore the high activity towards coniferaldehyde in *cadc/d* stems (Figure 2.5) suggests these other CAD(s) are located in non-lignifying tissues or in subcellular

compartments that do not have access to coniferaldehyde. Otherwise, we would expect to see higher coniferyl alcohol in *cadc/d*.

2.5.2 Additional CADs in lignin biosynthesis

The high level of alcohol-derived lignin in *med5a/5b ref8-1 cadc/d* prompted us to investigate what other enzyme(s) could be responsible for the reduction of *p*-coumaraldehyde. This enzyme should be specific for *p*-coumaraldehyde, but it should not be able to catalyze the reduction of coniferaldehyde or sinapaldehyde to their corresponding alcohols, otherwise *cadc/d* would incorporate alcohols instead of aldehydes into its lignin. Based on transcriptomic analysis we identified CADA (Table 2.3) and CADG (Figure 2.6) as potential candidates contributing to lignification in *med5a/5b ref8-1 cadc/d*. *CADA* and *CADG* single knockouts did not show a reduction in total CAD activity nor lignin content [45]. Nevertheless, the bulk of monolignol synthesis in wild-type plants goes towards G and S monomers. H-lignin only makes 2% of the lignin in wild-type *Arabidopsis* and measuring the contribution of CADs to this small fraction would be challenging to perform with precision. Knocking out *CADA* and/or *CADG*, in *med5a/5b ref8-1 cadc/d* did not have an effect in total CAD activity (Figure 2.5) or H-lignin content (Figure 2.7 and 2.8). It may be that other CADs were upregulated to compensate for the lack of CADC, CADD, CADA and CADG in these plants. Phylogenetic analysis clusters *CADA* with a CAD gene from aspen which has high affinity towards sinapaldehyde [48]. A similar substrate affinity would explain why CADA does not contribute significantly to the synthesis of *p*-coumaryl alcohol. CADG is highly divergent and the only member of its phylogenetic class, suggesting that it may have neofunctionalized after it diverged [45]. Taken together, these data leave the function of CADA and CADG in *Arabidopsis* stems an open question.

The remaining CADs, namely CAD1 and CADB1 and CADB2 did not seem likely candidates to be catalyzing the synthesis of *p*-coumaryl alcohol in *med5a/5b ref8-1 cadc/d*. As stated previously, a putative candidate must be able to catalyze the reduction of *p*-coumaraldehyde, but not of coniferaldehyde or sinapaldehyde, otherwise the presence of this enzyme would result in hydroxycinnamyl alcohol incorporation into the lignin of *cadc/d*. In studies attempting to complement *cadc/d,* CAD1 was the only CAD that could partially complement *cadc/d* by synthesizing both G and S subunits [45], making it a poor candidate for an H-lignin specific CAD. In these experiments, the complementation was carried out using the *CADD* promoter [45] which

reinforces the idea that other CADs are normally expressed in non-lignifying cell types. Based on the strong phenotype of *cadc/d* [24,27], the high CAD activity measured in *cadc/d* stems is not contributing significantly to the synthesis of G and S monolignols. This activity is likely the result of tissue type homogenization in our assays. As mentioned earlier, CADB1 and CADB2 have been characterized in vitro, and CADB1 has higher affinity for coniferaldehyde than for *p*-coumaraldehyde and CADB2 has nearly equal K_m values for these substrates, again making these enzymes unlikely candidates [46]. The role of these additional CAD genes in planta remains to be determined, although it has been suggested that they may be involved in the synthesis of other phenylpropanoids, particularly lignans [49].

2.5.3 Neither CCR1 nor CCR2 catalyze the reduction of hydroxycinnamaldehydes

HMG-CoA reductase catalyzes the two-step reduction of 3-hydroxy-3-methylglutaryl-CoA to mevalonate through a mevaldehyde intermediate [43]. Given this biochemical precedent, we tested the hypothesis that *A. thaliana* CCR1 and/or CCR2 can catalyze the reduction of *p*-coumaryl CoA to *p*-coumaraldehyde and subsequently to *p*-coumaryl alcohol, but only catalyze the reduction of feruloyl CoA to coniferaldehyde. Our results indicate that CCR1 and CCR2 can catalyze the reduction of the CoA esters to their corresponding aldehydes (Figure 2.9), as has been previously described in literature [39], but cannot facilitate the two-step reduction of the CoA esters to the corresponding alcohols. Neither can these enzymes reduce the aldehyde substrates to their corresponding alcohols. CCR1 has been crystallyzed in both the apo- and co-factor bound states [50]. No significant rearrangement of overall enzyme structure has been reported upon substrate and cofactor binding [50]. In contrast, HMG-CoA reductase undergoes large conformational changes and of particular note is the movement of the flap domain that protects the active site from exposure to solvent [43]. This conformational change seems to protect the mevaldehyde intermediate from dissociating before the second reduction step, which otherwise is likely to occur because the affinity of HMG-CoA reductase towards mevaldehyde is low [43]. The lack of a similar conformational change that would protect the active site in CCR1 and possibly CCR2 is consistent with the observation that these two enzymes are only capable to catalyze a single reduction per catalytic cycle.

2.5.4 Leaky expression of CADD in *med5a/5b* plants

By crossing *med5a/5b ref8-1 cadc/d* to *cadc/d*, we made the serendipitous finding that CADD is expressed in *med5a/5b* backgrounds carrying a homozygous T-DNA insertion in CADD (Figure 2.11), even though it is expressed at levels below the limit of detection in an otherwise wild-type background. The *cadd* T-DNA line has a promoter insertion that results in *cadd* being a transcriptional null (Figure 2.11). The T-DNA insertion in CADD only disrupts a portion of the promoter region leaving the coding region intact, which if transcribed and translated should lead to a functional enzyme. MED5A and MED5B negatively regulate phenylpropanoid metabolism by directly interacting with transcription factors and RNA polymerase II [51]. The absence of MED5 in the *cadd* T-DNA line may lead to expression of CADD and in this way explain the restored synthesis of monolignols in *med5a/5b cadc/d* and *med5a/5b ref8-1 cad/d* (Figure 2.13). To test this hypothesis, we used the CRISPR/Cas9 system to target an exon in CADD in *med5a/5b ref8-1 cadc/d* plants. Our results show that introducing a frameshift mutation into the open reading frame of CADD had no effect in the lignin composition of *med5a/5b ref8-1 cadc/d* (Figure 2.13), but it did have a negative effect on growth (Figure 2.12) and fertility. It has yet to be determined whether there was a reduction of total lignin content in these plants. Additionally, it remains to be tested if the effects in growth and fertility are indeed from knocking out CADD or they are the result of a non-target effect from the CRISPR/Cas9 system.

As seen in Figure 2.11, the transcript levels of *CADC* also increased in *med5a/5b* backgrounds carrying a homozygous T-DNA insertion in *CADC*. The T-DNA in *CADC* is located in the coding region and translation of this interrupted transcript would result in a truncated protein. The active site of CADC is made from residues from both monomers [52]. However, only a small portion of the amino acids that form part of the active site would be truncated from one monomer of the homodimer, leaving all the necessary residues for catalysis in the other monomer intact. It is not known whether these truncated proteins could still form a functional dimer or form a heterodimer with another CAD as has been reported in other species [53]. The formation of a functional dimer is an unlikely scenario, but if it does, its active site would have to change in such a way that it could use *p*-coumaraldehyde, but not coniferaldehyde or sinapaldehyde as substrates. Although the efficiency of these truncated enzyme would be low, the increase in expression may result in sufficient levels to catalyze the synthesis of *p*-coumaryl alcohol. Further *in silico* work would be required to assess the severity of the CADC truncation in enzyme structure and function.

Finally, we cannot exclude the possibility that a non-specific enzyme is catalyzing the reduction of *p*-coumaraldehyde to *p*-coumaryl alcohol in the mutants hereby described.

2.6 References

1. Nonhebel S: Renewable energy and food supply: Will there be enough land? *Renew Sustain Energy Rev* 2005, 9:191–201.
2. McCann MC, Carpita NC: Biomass recalcitrance: A multi-scale, multi-factor, and conversion-specific property. *J Exp Bot* 2015, 66:4109–4118.
3. U.S. Department of Energy, U.S. Department of Agriculture: Biomass as feedstock for a bioenergy and bioproducts industry: The technical feasibility of a billion-ton annual supply. 2005, https://www1.eere.energy.gov/bioenergy/pdfs/final_billionton_vision_report2.pdf
4. Himmel ME, Ding S, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD: Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* 2007, 315:804-807.
5. Zeng Y, Zhao S, Yang S, Ding S-Y: Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Curr Opin Biotechnol* 2014, 27:38–45.
6. Ralph J, Brunow G, Harris PJ, Dixon RA, Schatz PF, Boerjan W: Lignification: Are lignins biosynthesized via simple combinatorial chemistry or via proteinaceous control and template replication? In *Recent Advances in Polyphenol Research*, doi:10.1002/9781444302400.ch2
7. Ralph J, Lapierre C, Boerjan W: Lignin structure and its engineering. *Curr Opin Biotechnol* 2019, 56:240–249.
8. Ralph J, Mackay JJ, Hatfield RD, O'Malley DM, Whetten RW, Sederoff RR: Abnormal lignin in a loblolly pine mutant. *Science* 1997, 277:235–239.
9. Yamamoto M, Tomiyama H, Koyama A, Okuizumi H, Liu S, Vanholme R, Goeminne G, Hirai Y, Shi H, Takata N, et al.: A century-old mystery unveiled: Sekizaisou is a natural lignin mutant. *Plant Physiol* 2020, 182:1821–1828.
10. Renault H, Werck-Reichhart D, Weng JK: Harnessing lignin evolution for biotechnological applications. *Curr Opin Biotechnol* 2019, 56:105–111.

11. Turner SR, Somerville CR: Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell* 1997, 9:689–701.
12. Wang P, Dudareva N, Morgan JA, Chapple C: Genetic manipulation of lignocellulosic biomass for bioenergy. *Curr Opin Chem Biol* 2015, 29:32–39.
13. Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD: Designer lignins: Harnessing the plasticity of lignification. *Curr Opin Biotechnol* 2016, 37:190–200.
14. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C: Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. *Proc Natl Acad Sci U S A* 1998, 95:6619–6623.
15. Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chappie C: Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J* 2000, 22:223–234.
16. Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD: Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *J Agric Food Chem* 2003, 51:6178–6183.
17. Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA: Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa L.*). *Proc Natl Acad Sci* 2005, 102:16573–16578.
18. Stewart JJ, Akiyama T, Chapple C, Ralph J, Mansfield SD: The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol* 2009, 150:621–635.
19. Yang H, Zhang X, Luo H, Liu B, Shiga TM, Li X, Kim JI, Rubinelli P, Overton JC, Subramanyam V, et al.: Overcoming cellulose recalcitrance in woody biomass for the lignin-first biorefinery. *Biotechnol Biofuels* 2019, 12:1–18.
20. Vanholme R, De Meester B, Ralph J, Boerjan W: Lignin biosynthesis and its integration into metabolism. *Curr Opin Biotechnol* 2019, 56:230–239.
21. Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C: Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of Arabidopsis. *Plant J* 2002, 30:47-59.

22. Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, et al.: Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* 2014, 509:376–380.
23. Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier MT, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, et al.: Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol* 1996, 112:1479-1490.
24. Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L, Séguin A: *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* 2005, 17:2059–2076.
25. Zhao Q, Tobimatsu Y, Zhou R, Pattathil S, Gallego-giraldo L, Fu C: Loss of function of cinnamyl alcohol dehydrogenase 1 leads to unconventional lignin and a temperature-sensitive growth defect in *Medicago truncatula*. *Proc Natl Acad Sci U S A* 2013, 110:13660–13665.
26. Pillonel C, Mulder MM, Boon JJ, Forster B, Binder A: Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor L. Moench*. *Planta* 1991, 185:538–544.
27. Anderson NA, Tobimatsu Y, Ciesielski PN, Ximenes E, Ralph J, Donohoe BS, Ladisch M, Chapple C: Manipulation of guaiacyl and syringyl monomer biosynthesis in an Arabidopsis cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell* 2015, 27:2195–2209.
28. Varbanova M, Porter K, Lu F, Ralph J, Hammerschmidt R, Daniel Jones A, Day B: Molecular and biochemical basis for stress-induced accumulation of free and bound *p*-coumaraldehyde in cucumber. *Plant Physiol* 2011, 157:1056–1066.
29. Chapple CC, Vogt T, Ellis BE, Somerville CR: An Arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell* 1992, 4:1413–1424.
30. Lu F, Ralph J: The DFRC method for lignin analysis. 2. Monomers from isolated lignins. *J Agric Food Chem* 1998, 46:547–552.
31. Campbell MM, Ellis BE: Fungal elicitor-mediated responses in pine cell cultures - I. Induction of phenylpropanoid metabolism. *Planta* 1992, 186:409–417.

32. Bradford M: A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976, 72:248-254.
33. Wyrambik D, Grisebach H: Purification and properties of isoenzymes of cinnamyl-alcohol dehydrogenase from soybean-cell-suspension cultures. *Eur J Biochem* 1975, 59:9–15.
34. Mao Y, Zhang H, Xu N, Zhang B, Gou F, Zhu JK: Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol Plant* 2013, 6:2008–2011.
35. Clough SJ, Bent AF: Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J* 1998, 16:735–743.
36. Heigwer F, Kerr G, Boutros M: E-CRISP: Fast CRISPR target site identification. *Nat Methods* 2014, 11:122–123.
37. Li X, Bonawitz ND, Weng J-K, Chapple C: The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *Plant Cell* 2010, 22:1620–1632.
38. Weng J-K, Akiyama T, Ralph J, Chapple C: Independent recruitment of an *O*-methyltransferase for syringyl lignin biosynthesis in *Selaginella moellendorffii*. *Plant Cell* 2011, 23:2708–24.
39. Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, Chen F, Dixon RA: Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc Natl Acad Sci U S A* 2010, 107:17803–17808.
40. Jones L, Ennos AR, Turner SR: Cloning and characterization of *irregular xylem4* (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*. *Plant J* 2001, 26:205–216.
41. Schilmiller AL, Stout J, Weng J-K, Humphreys J, Ruegger MO, Chapple C: Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in Arabidopsis. *Plant J* 2009, 60:771–82.
42. Costa MA, Collins RE, Anterola AM, Cochrane FC, Davin LB, Lewis NG: An *in silico* assessment of gene function and organization of the phenylpropanoid pathway metabolic networks in *Arabidopsis thaliana* and limitations thereof. *Phytochemistry* 2003, 64:1097–1112.
43. Haines BE, Wiest O, Stauffacher C V: The increasingly complex mechanism of HMG-CoA reductase. *Acc Chem Res* 2013, 46:2416–2426.

44. Sattler SE, Funnell-Harris DL, Pedersen JF: Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci* 2010, 178:229–238.
45. Eudes A, Pollet B, Sibout R, Do CT, Séguin A, Lapierre C, Jouanin L: Evidence for a role of *AtCAD 1* in lignification of elongating stems of *Arabidopsis thaliana*. *Planta* 2006, 225:23–39.
46. Kim SJ, Kim MR, Bedgar DL, Moinuddin SGA, Cardenas CL, Davin LB, Kang CH, Lewis NG: Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis. *Proc Natl Acad Sci U S A* 2004, 101:1455–1460.
47. Kim SJ, Kim KW, Cho MH, Franceschi VR, Davin LB, Lewis NG: Expression of cinnamyl alcohol dehydrogenases and their putative homologues during *Arabidopsis thaliana* growth and development: Lessons for database annotations? *Phytochemistry* 2007, 68:1957–1974.
48. Li L, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL: The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. *Plant Cell* 2001, 13:1567–1586.
49. Suh MC, Kim MJ, Hur CG, Bae JM, Park YI, Chung CH, Kang CW, Ohlrogge JB: Comparative analysis of expressed sequence tags from *Sesamum indicum* and *Arabidopsis thaliana* developing seeds. *Plant Mol Biol* 2003, 52:1107–1123.
50. Haiyun P, Zhou R, Louie G V., Mühlemann JK, Bomati EK, Bowman ME, Dudareva N, Dixon RA, Noel JP, Wang X: Structural studies of cinnamoyl-CoA reductase and cinnamyl-alcohol dehydrogenase, key enzymes of monolignol biosynthesis. *Plant Cell* 2014, 26:3709–3727.
51. Bonawitz ND, Soltau WL, Blatchley MR, Powers BL, Hurlock AK, Seals LA, Weng JK, Stout J, Chapple C: REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in Arabidopsis. *J Biol Chem* 2012, 287:5434–5445.
52. Youn B, Camacho R, Moinuddin SGA, Lee C, Davin LB, Lewis NG, Kang C: Crystal structures and catalytic mechanism of the Arabidopsis cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org Biomol Chem* 2006, 4:1687–1697.

53. Hawkins S, Boudet AM: Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from the periderm of *Eucalyptus gunnii* hook. *Plant Physiol* 1994, 104:75–84

Table 2.1. List of primers used in this project.

Primer name	Sequence (5'-> 3')	Usage
CC4918	GAC CTT TAA CTC TCC CGC TAT G	*ACTIN2* qRT-PCR fwd
CC4819	GAG ACA CAC CAT CAC CAG AAT	*ACTIN2* qRT-PCR rv
CC4975	TCT GGA GTT ACA ACG ATG TCT	*CADC* qRT-PCR fwd
CC4983	CTT TAA GCC CAC TCG CCA TT	*CADC* qRT-PCR rv
CC4976	ACA TCA ATG GTC AAC CTA CAC	*CADD* qRT-PCR fwd
CC4984	CTC TTA GGC CTG GTT GTT TC	*CADD* qRT-PCR rv
CC4972	TGA TCA TGA TGG CTC TGT TAC T	*CADG* qRT-PCR fwd
CC4981	AGA GAT TTA CCA GGT TGA TTC ATA	*CADG* qRT-PCR rv
CC5079	ATC GTT TCG GAT ATT GAG CTC ATA AA	*CADE/F* qRT-PCR fwd [24]
CC5080	TCT CAG CTG ACG ACT CAG GGA GTA AA	*CADE/F* qRT-PCR fwd [24]
CC4977	GCA ACC TAT AAC GGA GTT CA	*CADA* qRT-PCR fwd
CC4985	GTT TAT CCG GTC CGG TCA AT	*CADA* qRT-PCR rv
CC4978	ACT TTG ACG ATA CCA TGA CCC AT	*CADB1* qRT-PCR fwd
CC4986	TGG TGC GGC ACC GTC TAA T	*CADB1* qRT-PCR rv
CC4980	TGC GAT TGG ATC CGA TGG AA	*CAD1* qRT-PCR fwd
CC4988	GCT TCC CTG CCT CAG TCA TA	*CAD1* qRT-PCR rv
CC4979	CAA TGA TGG CGA CGA GAA CT	*CADB2* qRT-PCR fwd
CC4987	TGT TAT GGT GTT GTC GTA GTA CG	*CADB2* qRT-PCR rv
CC4942	GAT TAT AAC GGA GTT CAC CAC GA	*CADA* CRIPSR/Cas9 target 1 fwd
CC4943	AAA CTC GTG GTG AAC TCC GTT AT	*CADA* CRIPSR/Cas9 target 1 rv
CC4944	GAT TTC GAC CAC TGG GAA ATC TA	*CADA* CRIPSR/Cas9 target 2 fwd
CC4945	AAA CTA GAT TTC CCA GTG GTC GA	*CADA* CRIPSR/Cas9 target 2 rv
CC5009	GCT AGT TAC ATT GCC AGA TT	*CADA* target 1 amplicon fwd

Table 2.1 continued

Primer name	Sequence (5'-> 3')	Usage
CC4949	GTA CTT CAT AGG ACT GTA CAT CG	*CADA* target 1 amplicon rv
CC5012	GTT TGT AAC GTA ACA TAT AGG CAT	*CADA* target 2 amplicon fwd
CC5007	CAT CAT GGT TTC ATC ATA CTC G	*CADA* target 2 amplicon rv
CC5010	GTG AAA TAG GGA ACA AAG TAA GT	*CADA* target 1 sequencing fwd
CC5011	GCC ATA TTT GAT TTG AAT TAC CTG	*CADA* target 1 sequencing rv
CC4950	CGA TGT ACA GTC CTA TGA AGT AC	*CADA* target 2 sequencing fwd
CC5013	CTA AAT AAA GTT GAA ACC TGA TAA CG	*CADA* target 2 sequencing rv
CC5494	GAT TGA TTC CAA GGG ATA GTC CAC	*CADG* CRIPSR/Cas9 target 1 fwd
CC5495	AAA CGT GGA CTA TCC CTT GGA ATC	*CADG* CRIPSR/Cas9 target 1 fwd
CC5496	GAT TGA ACG TAT GTT AAC TCC TGC	*CADG* CRIPSR/Cas9 target 2 fwd
CC5497	AAA CGC AGG AGT TAA CAT ACG TTC	*CADG* CRIPSR/Cas9 target 2 rv
CC5418	CTT CAT CTG GTC ATG GTC	*CADG* amplicon fwd
CC5236	ATG AGT TCT TCA GAG AGT GTG	*CADG* amplicon rv/ target 2 sequencing rv
CC5417	GTT GTT AGT GCC TAT ATG AGT T	*CADG* target 1 sequencing fwd/ *cadg* CAPs marker rv
CC5416	CTG CAT AGA TAG ATT CAC AAG TT	*cadg* CAPs marker fwd
CC5847	GAT TGG AGC TGA TGA TTA CGT GAT	*CADD* CRIPSR/Cas9 fwd
CC5848	AAA CAT CAC GTA ATC ATC AGC TCC	*CADD* CRIPSR/Cas9 rv
CC5861	CAT GAT ATC AAG CGT CGA GAT	*cadd* CAPs marker fwd
CC5862	TTG TGG TCA AGA TTC CAG AAG G	*cadd* CAPs marker rv
CC5863	CAG AAC TCA AGC ATC TCC T	*CADD* sequencing fwd

Table 2.2 Total lignin quantification by the thioglycolic acid method.
A_{280nm}= Absorbance at 280nm. S.D.= standard deviation among biological replicates (n=3).

Genotype	A_{280nm} mg cell wall^{-1} ± S.D.
med5a/5b ref8-1	1.24±0.12
med5a/5b ref8-1 cada	1.24±0.09
med5a/5b ref8-1 cadc cadd	1.19±0.05
med5a/5b ref8-1 cadc cadd cada	1.23±0.07

Table 2.3. Fold change expression of *AtCAD* genes relative to wild type determined by RNA-seq.

Gene	*ref8-1*	*med5a/b ref8-1*
AtCAD1 (CADG)	1.28	1
AtCAD2 (CADE)	1	1
AtCAD3 (CADF)	1	1
AtCAD4 (CADC)	1	1.30
AtCAD5 (CADD)	3.70	1.69
AtCAD6 (CADA)	2.86	1.64
AtCAD7 (CADB1)	1.82	1
AtCAD8 (CADB2)	1	1
AtCAD9 (CAD1)	1	1

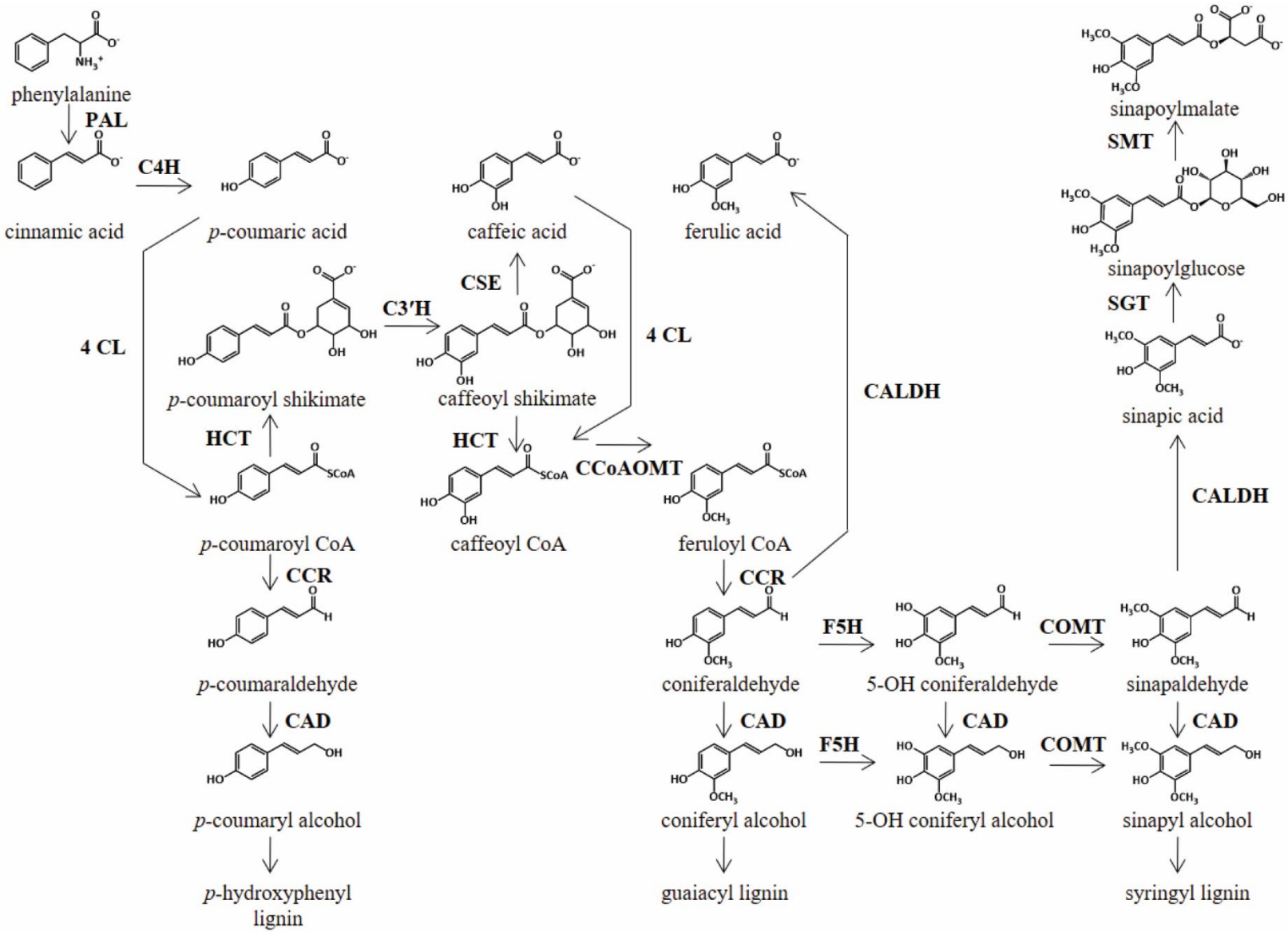


Figure 2.1. The phenylpropanoid pathway.

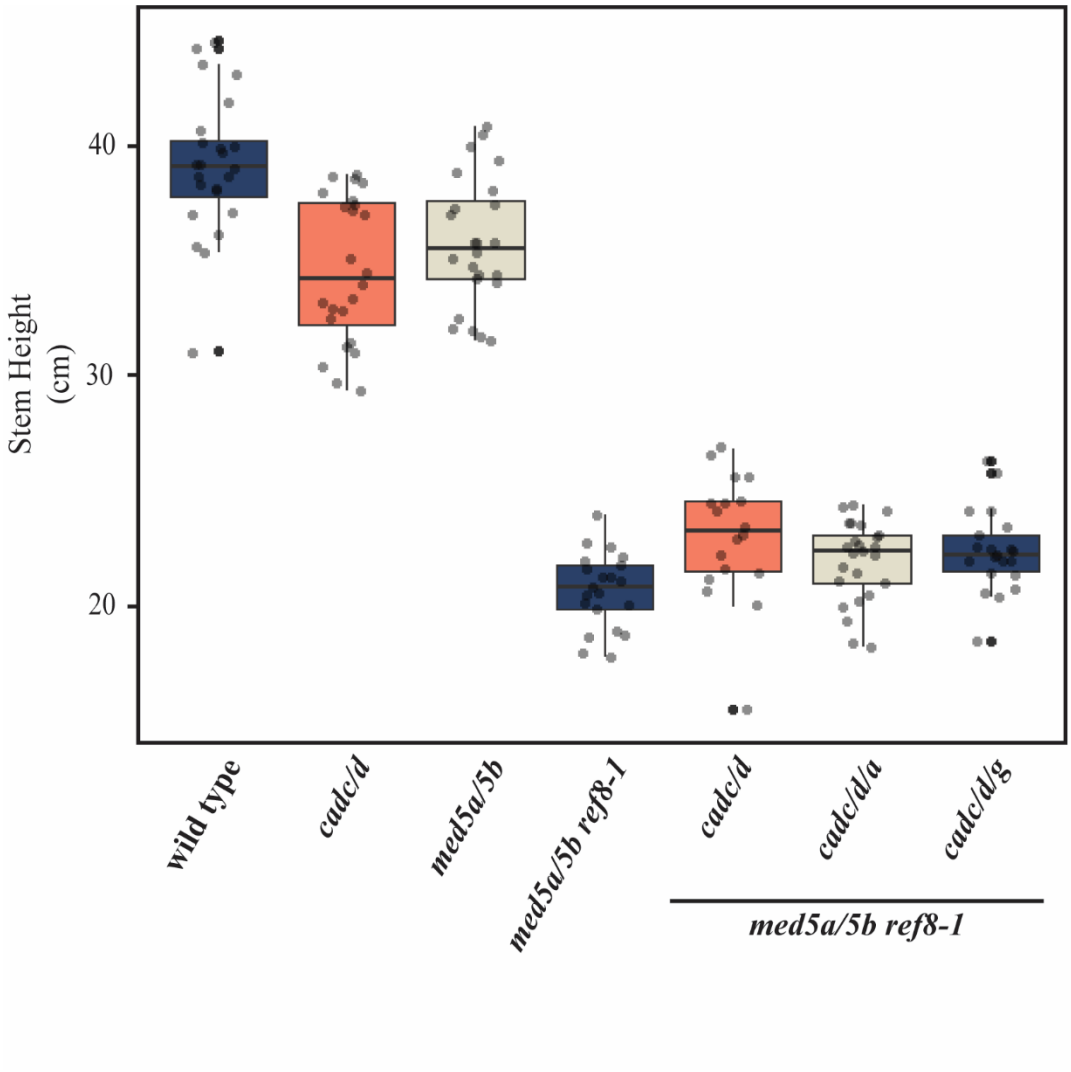


Figure 2.2. Height measurements of 6 weeks-old stems.
Box plots depict the median, first and third quartiles. Dots represent biological replicates and outliers are represented by darker dots (Tukey). Measurements are from wild type (*n*=24), *cadc/d* (*n*=24), *med5a/5b* (*n*=24), *med5a/5b ref8-1* (*n*=21), *med5a/5b ref8-1 cadc cadd* (*n*=18), *med5a/5b ref8-1 cadc/d/a* (*n*=25), *med5ab/5b ref8-1 cadc/d/g* (*n*=25).

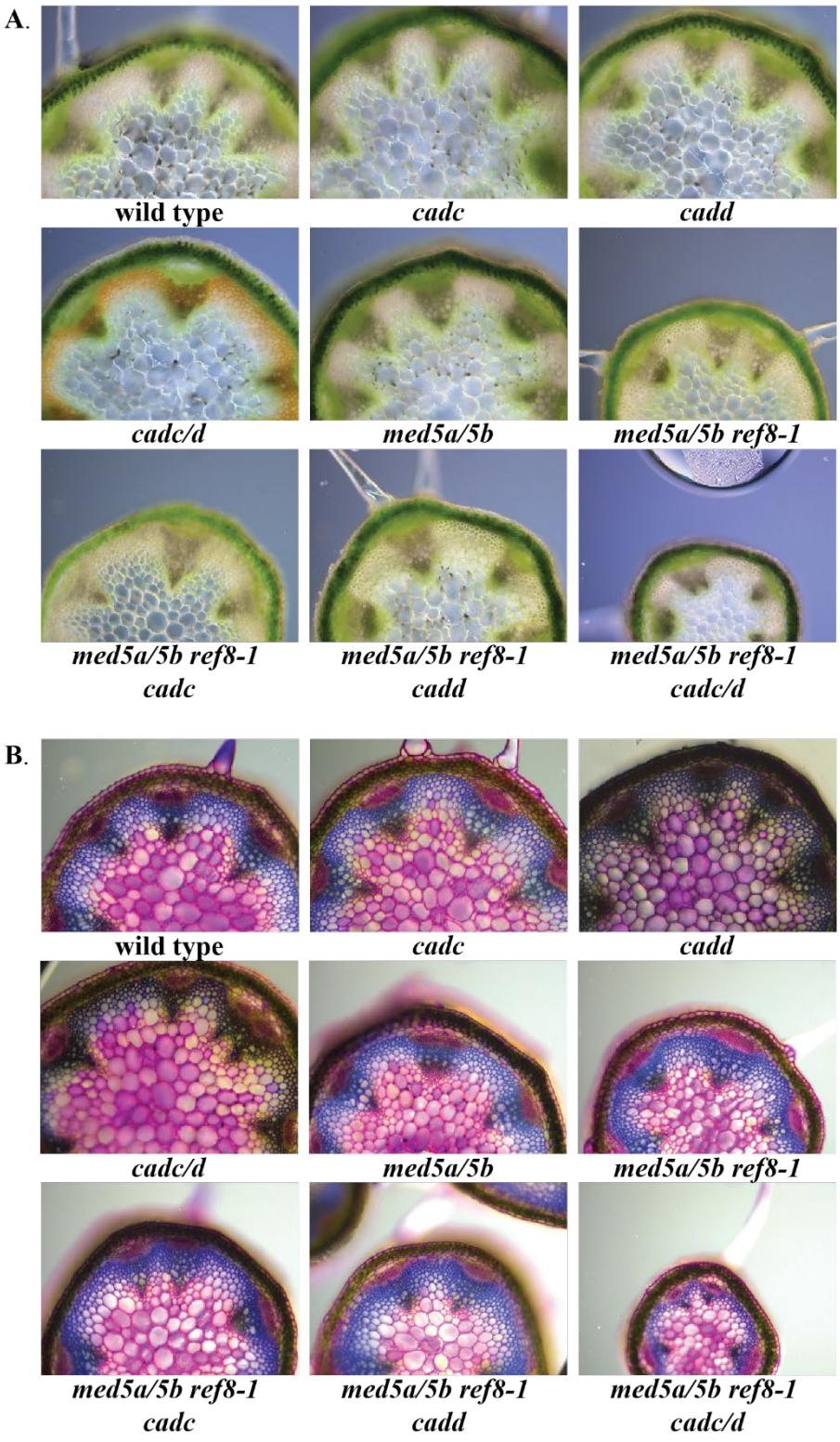
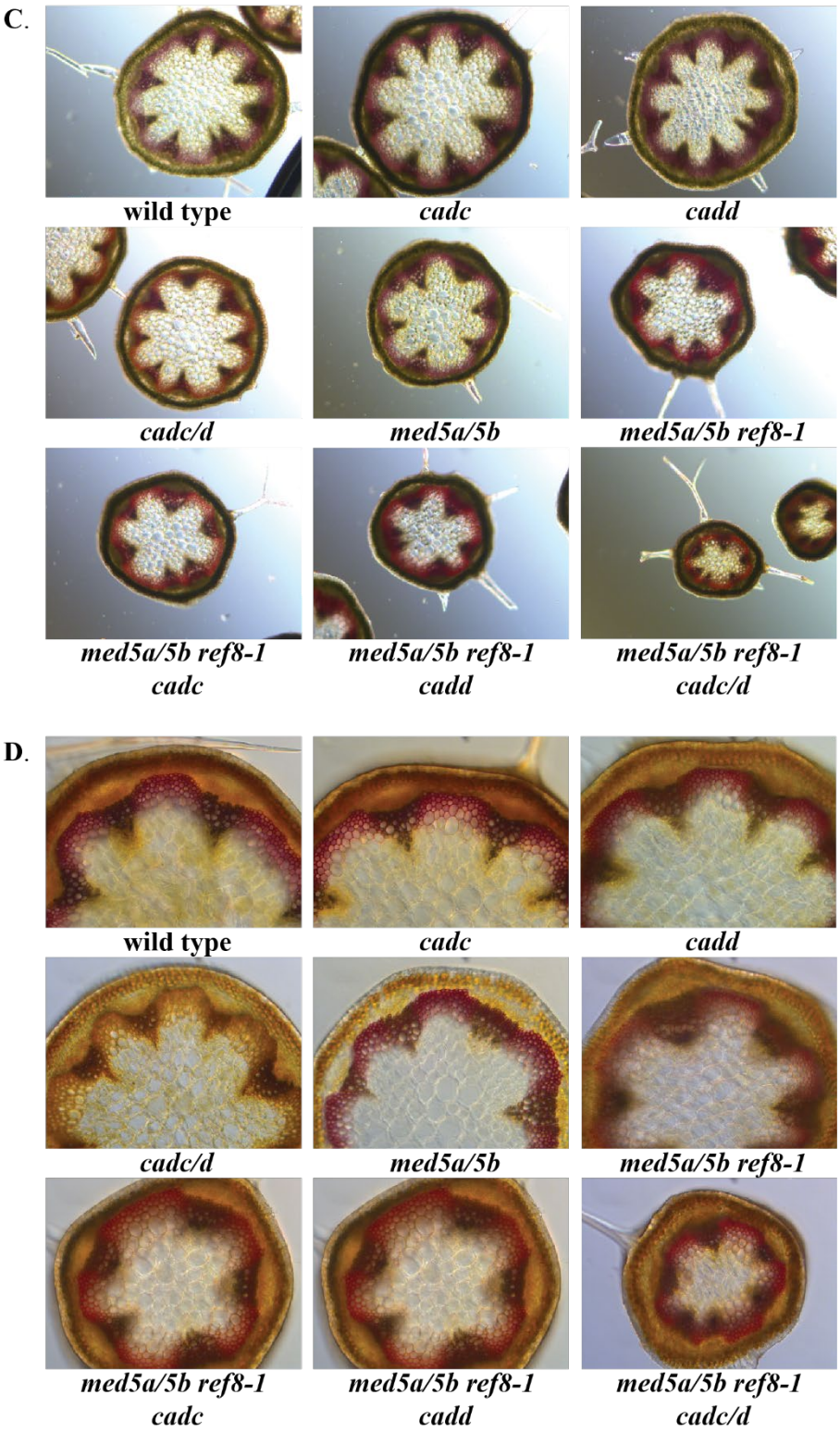


Figure 2.3. Stem cross sections of 6 weeks-old plants.
(A) Unstained, (B) stained with toluidine blue-O. Stained with (C) phloroglucinol-HCl and (D) Maüle reagent.

Figure 2.3 continued



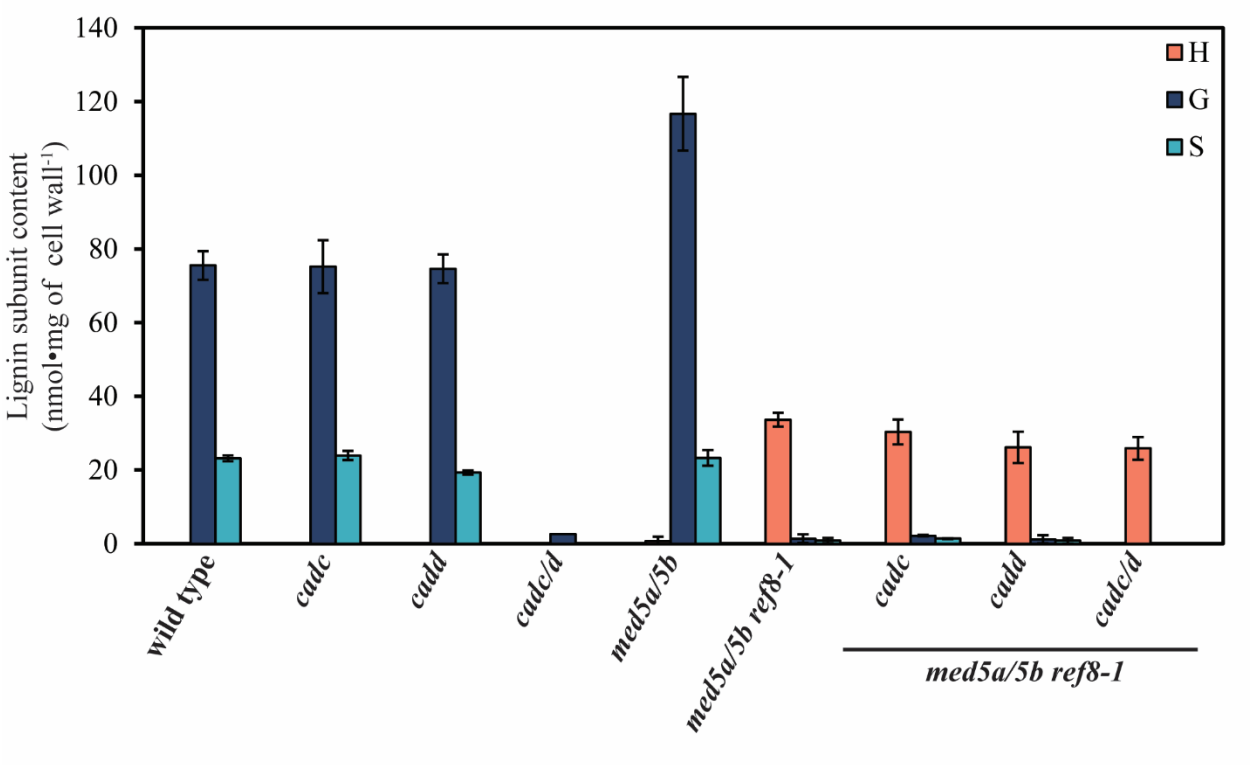


Figure 2.4. Lignin monomer composition determined by gas chromatography using the derivatization followed by reductive cleavage (DFRC) method.
Error bars indicate the standard deviation among three biological replicates.

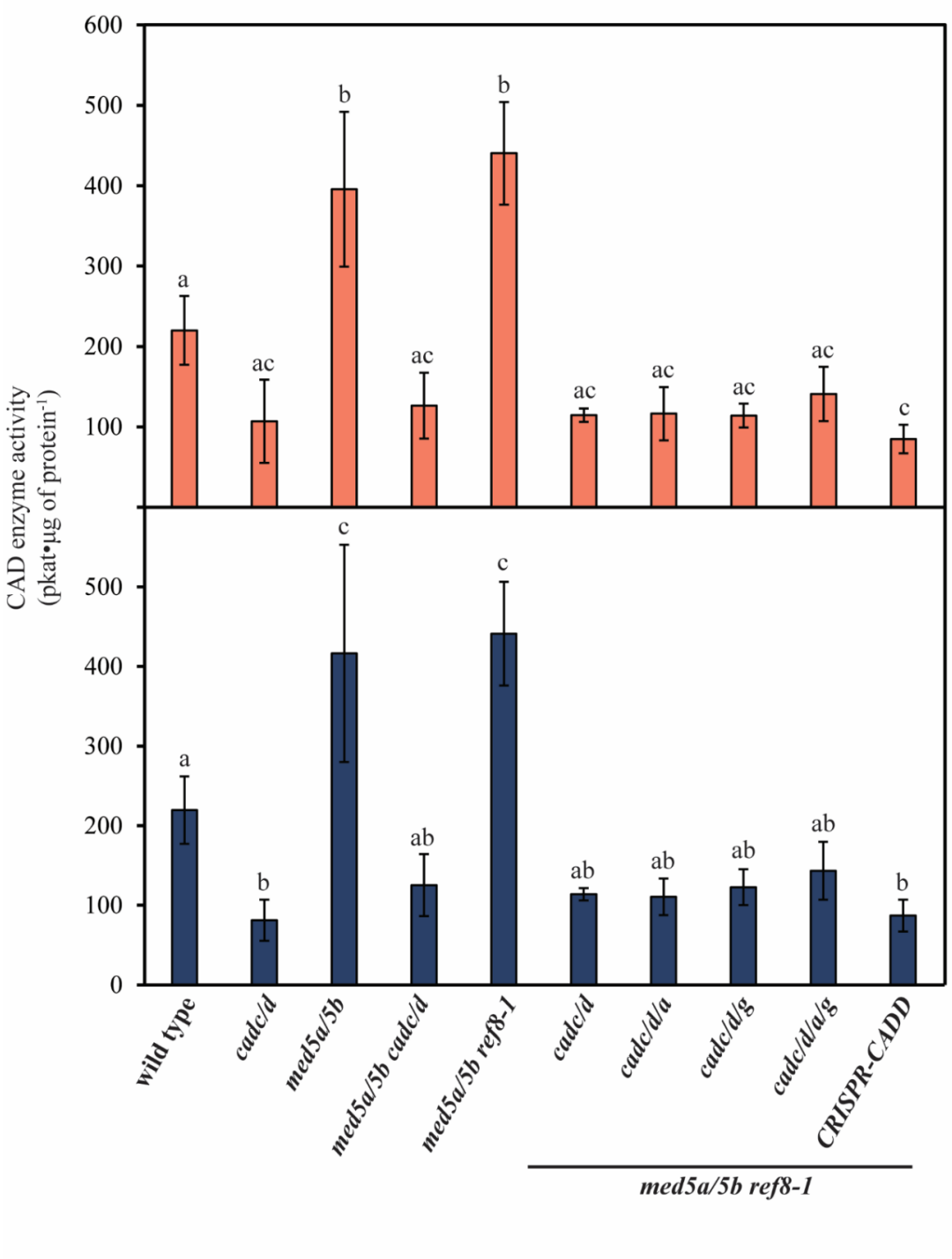


Figure 2.5. CAD enzyme activity in 4 weeks-old stems.
Quantification of *p*-coumaryl alcohol (top) and coniferyl alcohol (bottom) generated in CAD assays using *p*-coumaraldehyde and coniferaldehyde as substrates, respectively. Error bars represent standard deviation among biological replicates (n=4). Letters represent difference between genotypes determined by one-way ANOVA and Tukey's honest significant difference test ($p < 0.05$).

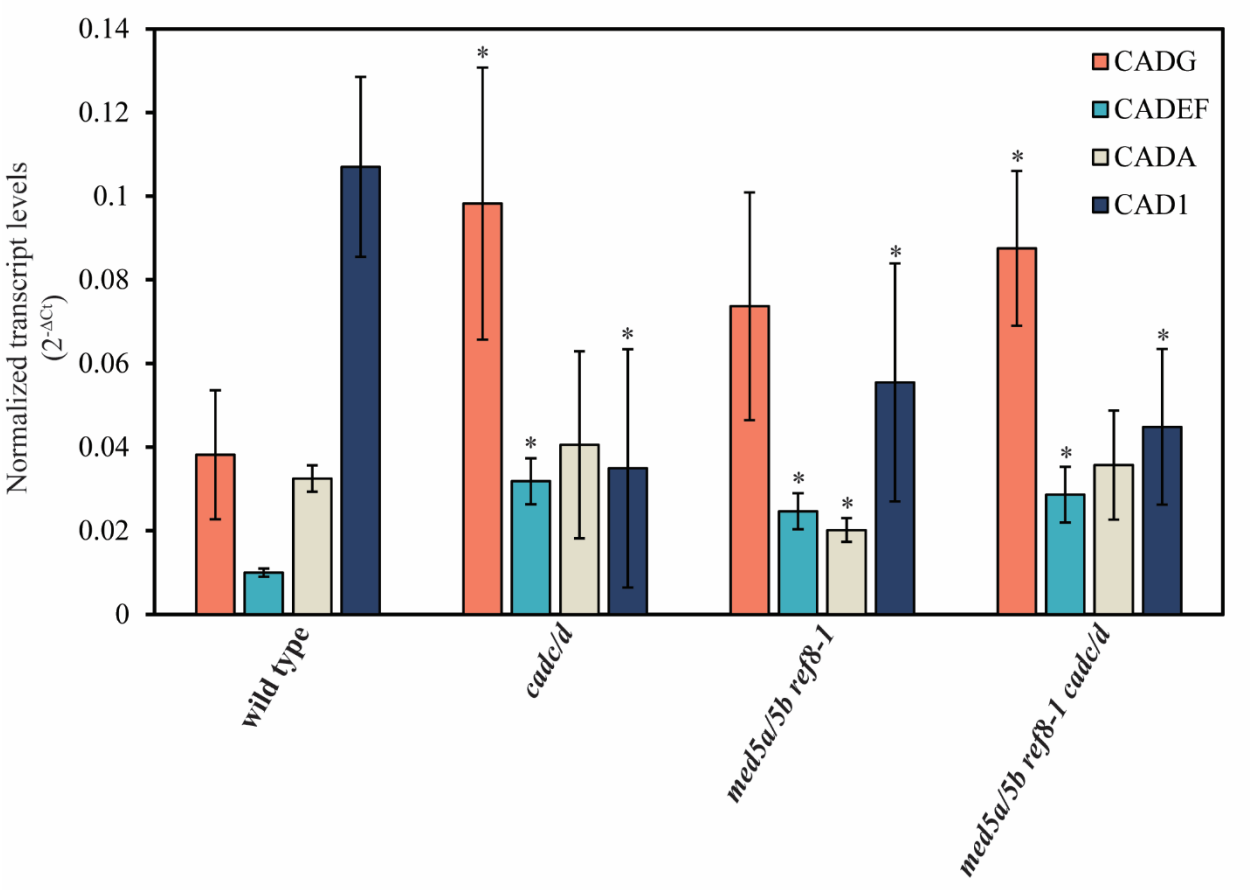


Figure 2.6. Expression of *CAD* genes in 4 weeks-old *Arabidopsis* stems.
Relative expression measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (n=3). Transcript levels significantly different from wild type ($p < 0.05$), Student's t-test are marked with (*).

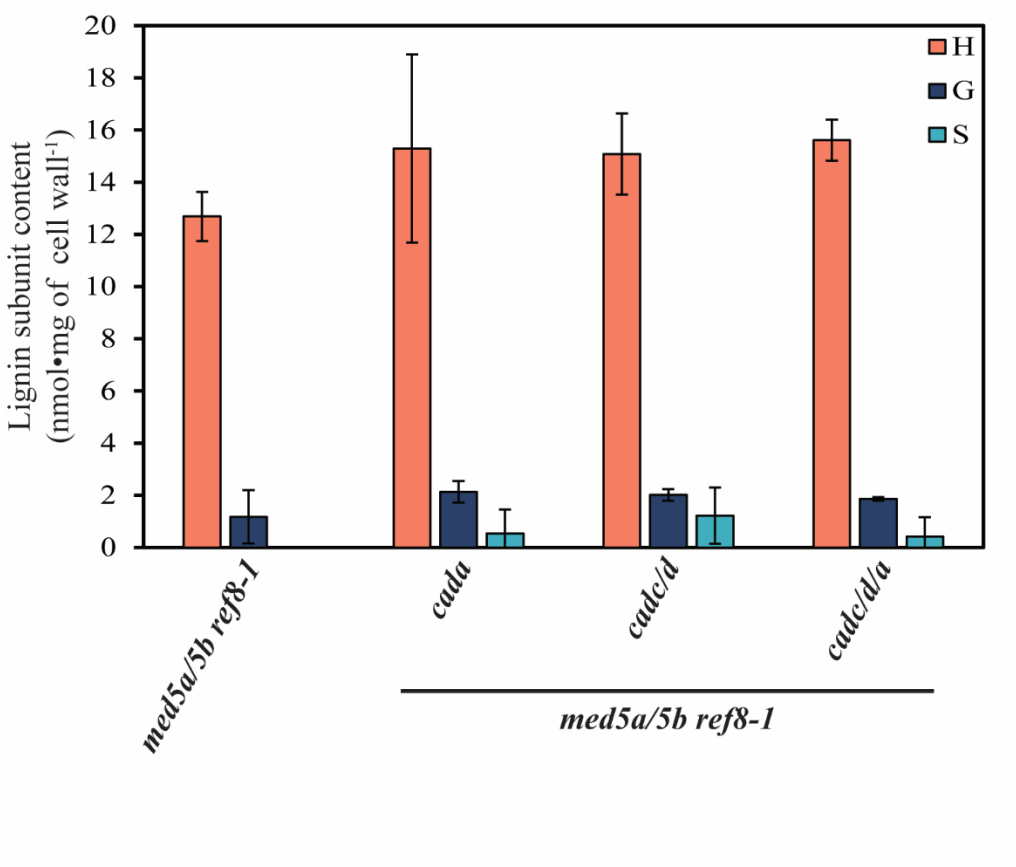


Figure 2.7. Lignin monomer composition of *CADA* knockouts determined by gas chromatography using the DFRC method.
Error bars indicate the standard deviation among three biological replicates (*n*=3).

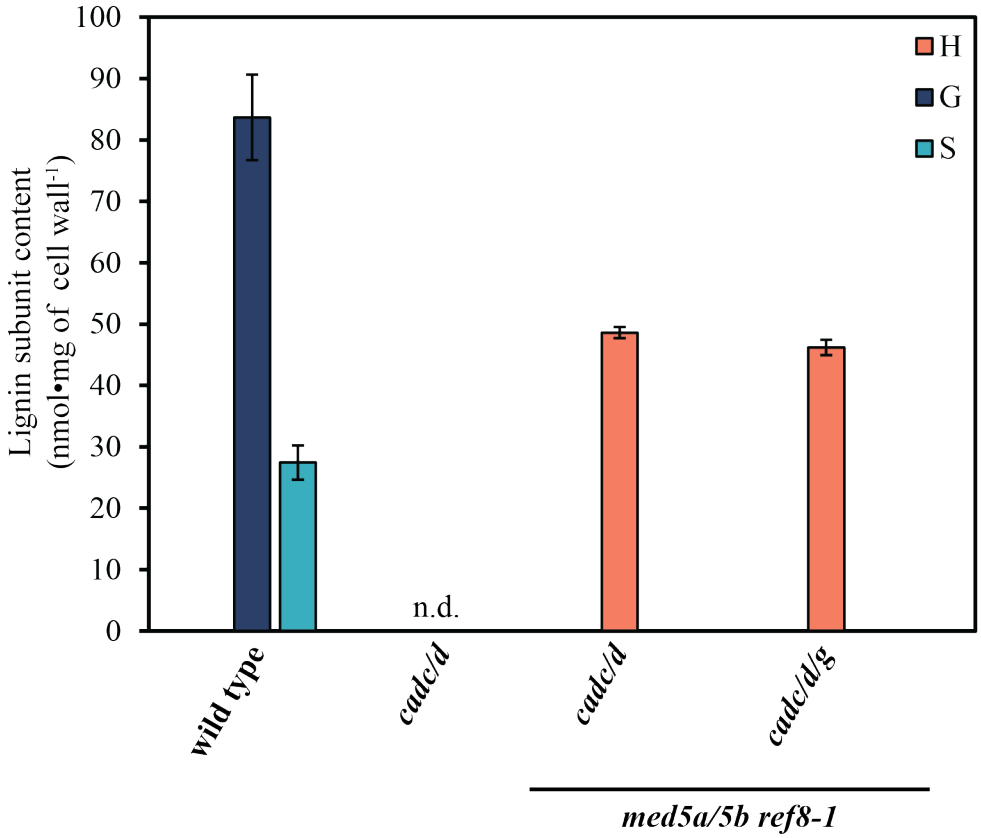


Figure 2.8. Lignin monomer composition of *CADG* knockouts determined by gas chromatography using the DFRC method.
Error bars indicate the standard deviation among three biological replicates (*n*=3). Products not detected (n.d.).

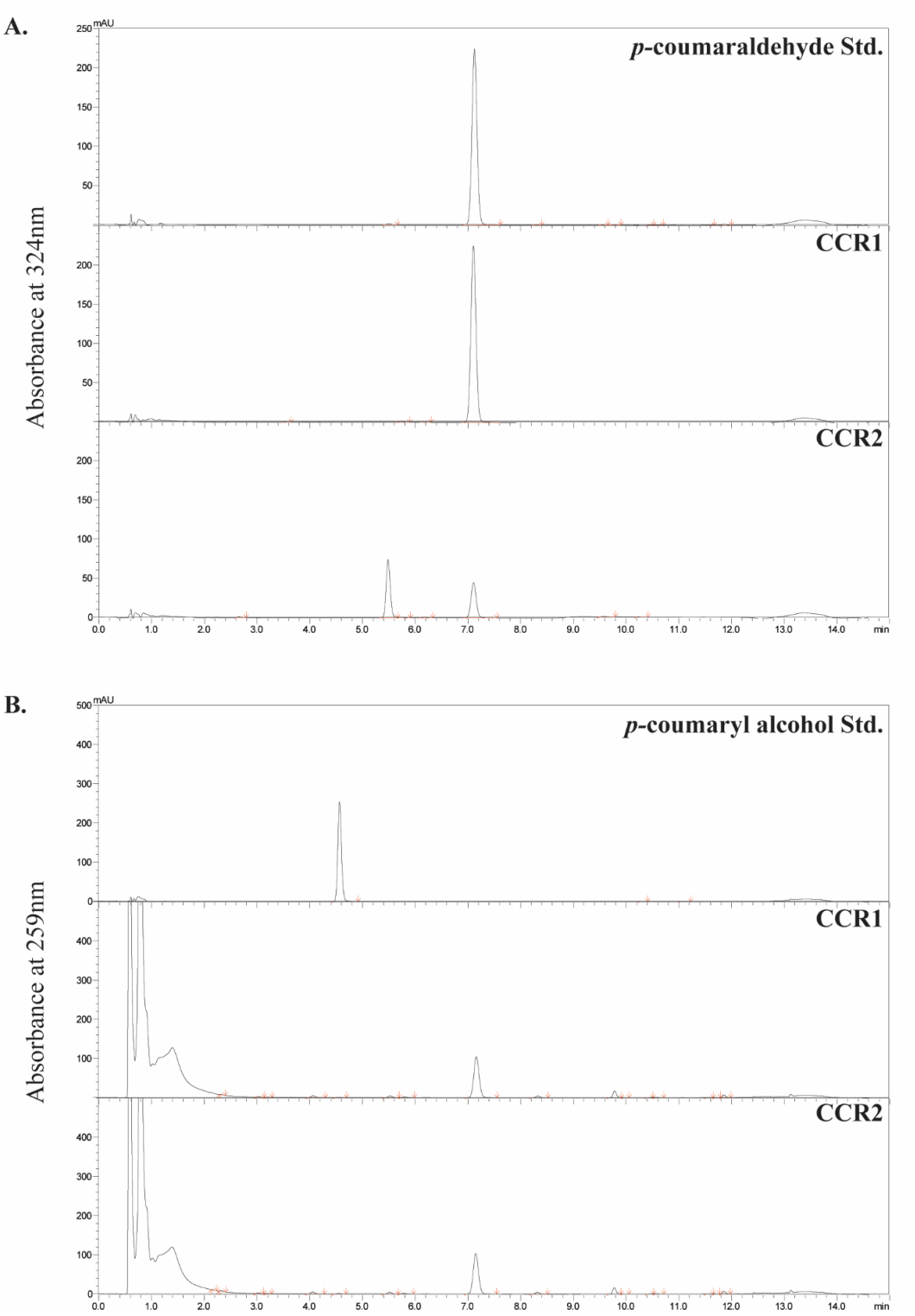


Figure 2.9. CCR assays with heterologously expressed CCR1 and CCR2 from *Arabidopsis thaliana.*

(A) Products from 4CL1 reaction coupled with CCR1 or CCR1. (B) Products of CCR reaction with *p*-coumaraldehyde as substrate.

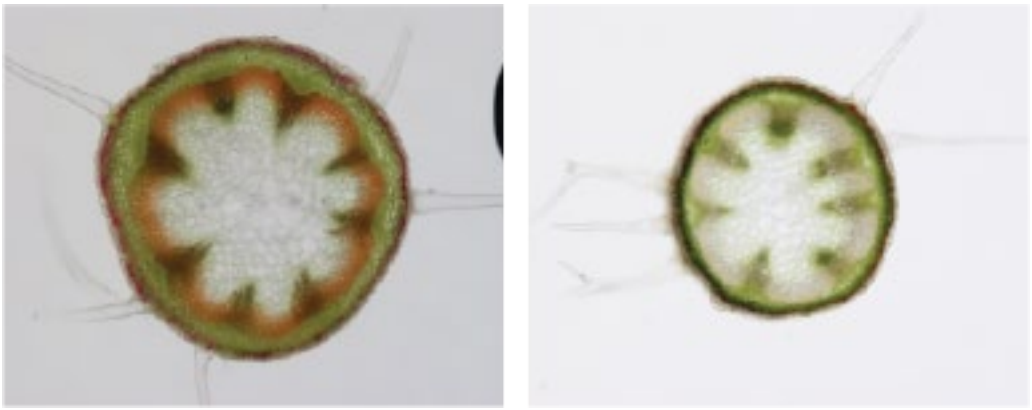


Figure 2.10. Stem cross-sections of 7 weeks-old plants.
Left, *cadc cadd*. Right, F1 of a cross between *cadc cadd* and *med5a/5b ref8-1 cadc cadd.*

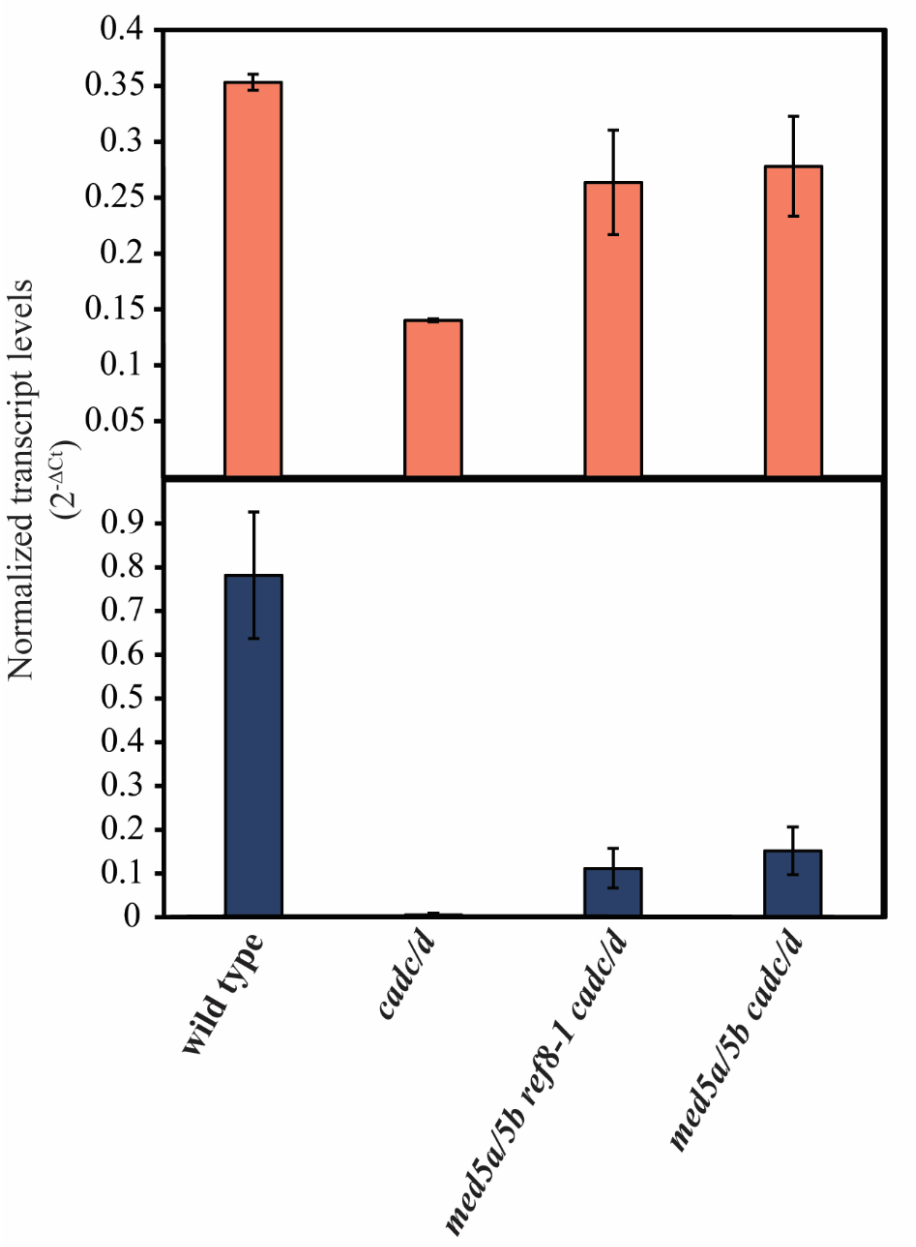


Figure 2.11. Expression of *CADC* and *CADD* genes in 4 weeks-old stems.
Relative expression of *CADC* (top) and *CADD* (bottom) measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (*n*=3).

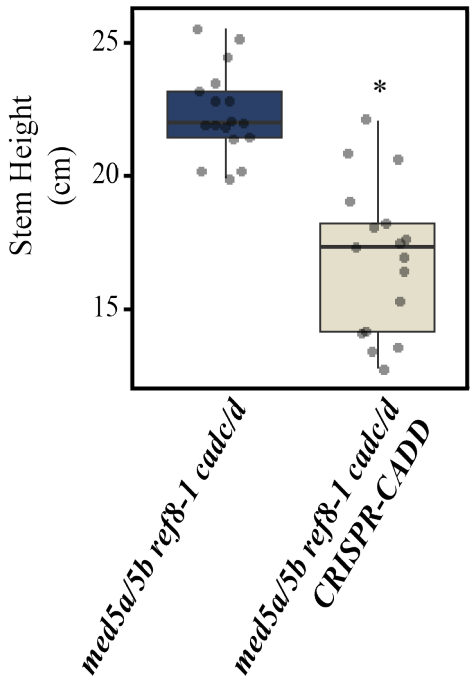


Figure 2.12. Height measurements of 5 weeks-old stems.
Box plots depicting the median, first and third quartiles. Dots represent biological replicates (n=17) and (*) marks a significant difference between genotypes ($p < 0.0001$) by Student's t-test.

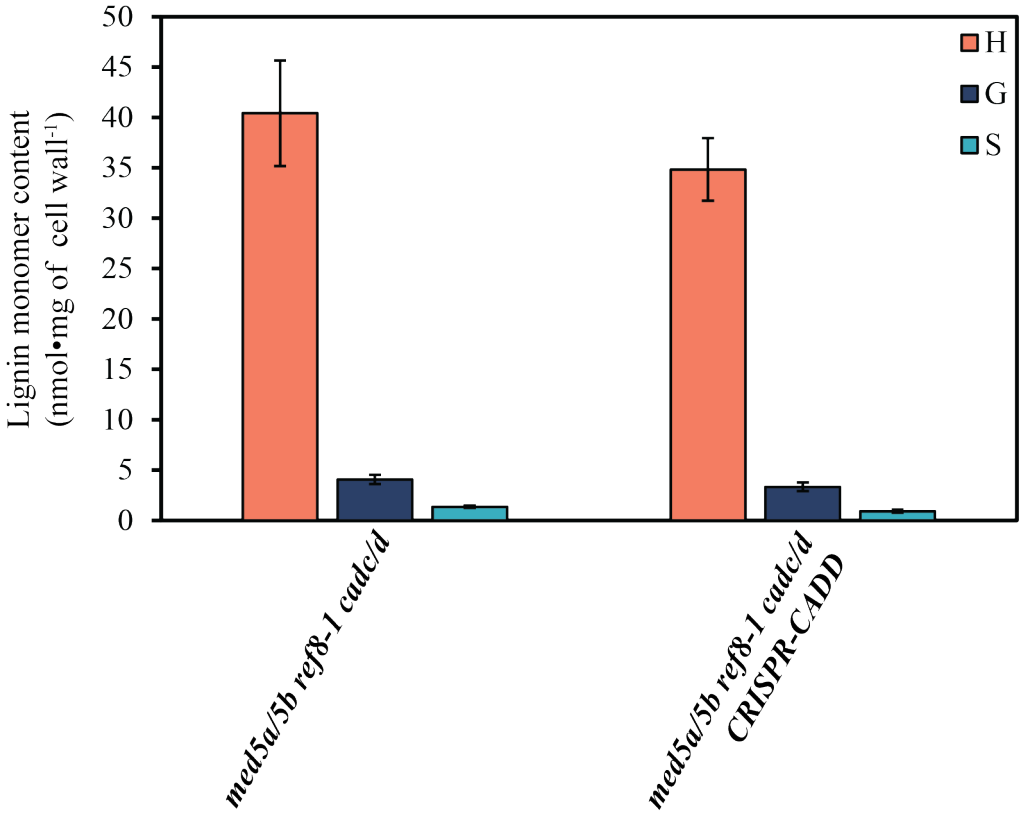


Figure 2.13. Lignin monomer composition of *CRISPR-CADD* knockouts determined by gas chromatography using the DFRC method.
Error bars indicate the standard deviation among biological replicates (*n*=3).

CHAPTER 3. THE CONNECTION BETWEEN LIGNIN ENGINEERING AND PLANT GROWTH

3.1 Abstract

Lignin contributes substantially to the recalcitrance of biomass towards saccharification. In an effort to modify biomass so that it is better suited for biofuel purposes, researchers have altered the polymer through stable genetic manipulations; however, in many cases this comes at the expense of plant growth. The enzyme *p*-coumaroyl shikimate 3'-hydroxylase (C3´H) catalyzes the conversion of *p*-coumaroyl shikimate to caffeoyl shikimate in the phenylpropanoid pathway. Previous studies have demonstrated that the *Arabidopsis thaliana C3´H*-deficient *ref8-1* mutant has a reduction in total lignin content and exhibits a dwarf phenotype. Surprisingly, elimination of two paralogs of the MED5 subunit of the transcriptional coregulator Mediator enabled *ref8-1* plants to attain a normal height despite their compromised phenylpropanoid metabolism. These findings suggest that rather than lignin being causal for the severe phenotype of *ref8-1,* the dwarfism of these plants is a consequence of downstream changes in transcription arising from perturbations in the phenylpropanoid pathway. In order to test this model, *C3´H* was introduced into *ref8-1* plants under the control of a dexamethasone inducible promoter. The induction of *C3´H* in *ref8-1* not only restores the levels of phenylpropanoids to wild-type levels, but it can also alleviate its dwarf phenotype. We performed RNA-seq and LC-MS on these plants at time points immediately following the onset of *C3´H* expression to capture the initial transcriptional and metabolic changes that occur when phenylpropanoid homeostasis is restored. Coupling these two datasets will allow us to determine the order of events that lead to dwarfing in *ref8-1* plants and provide testable hypotheses for the cause of its dwarfing. Overall, understanding the link between lignin and dwarfism will allow for better plant metabolic engineering strategies that do not compromise plant growth.

3.2 Introduction

The properties of plant secondary cell walls determine the quality of many products relevant to human daily life, including paper, livestock feed and biofuels. Among these properties is the ease by which carbohydrates can be extracted from lignocellulosic biomass [1]. This is

particularly important for the production of biofuels, because their production cost must be able to compete with that of the well-established oil industry. Unlike petroleum, lignocellulosic material is a renewable, low-cost feedstock for the production of fuel; however, the presence of lignin makes this material highly recalcitrant to enzymatic digestion which adds substantially to biofuels' production cost [2].

Lignin is a polymer derived from the phenylpropanoid pathway and is primarily found in plant vasculature where it imparts the necessary mechanical strength and hydrophobicity for water transport [3]. Genetic engineering and the discovery of natural and induced mutants has facilitated the manipulation of this polymer in planta and improved sugar extraction from plant biomass [4,5]. Nevertheless, in many cases these manipulations lead to an unexpected impairment in plant growth, known as lignin modification-induced dwarfism (LMID) [6]. Many lignin-deficient mutants have collapsed xylem vessels; therefore, dwarfism in these mutants may be the result of poor water transport [3,7]. This hypothesis is supported by a study in which a xylem-specific promoter was used to drive the expression of *Cinnamoyl CoA Reductase 1* (*CCR1*) in the dwarf *ccr1* mutant [6]. The restoration of lignification in the xylem of *ccr1* was sufficient to prevent xylem collapse and to restore plant height [6]. Alternative models for LMID propose that dwarfism can be caused by the accumulation or lack of phenylpropanoid metabolites [8–10]. For example, the plant hormone salicylic acid (SA), hyper-accumulates in some lignin-deficient mutants and this has led to the hypothesis that elevated SA levels cause dwarfism [11]. LMID can be prevented in plants deficient in *hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase* by preventing SA biosynthesis [11], but SA was found not to be causal of dwarfism in other mutants that hyper-accumulate SA [12]. Dehydrodiconiferyl glucosides (DCGs) are a group of phenylpropanoids with cytokinin-like activity whose deficiency could be implicated in LMID [13,14]. The fact that the synthesis of such phenylpropanoids is coupled with lignin synthesis has made difficult the use of traditional genetics to test this hypothesis for DCGs and other phenylpropanoid metabolites. Another model proposes growth deficit could be the outcome of a response initiated by a monitoring system which oversees the integrity of the cell wall [15]. Evidence at the mechanistic level is lacking for this monitoring system in lignin-modified plants, although there are reports that support this model at a transcriptional level [12,16]. To date, there is not sufficient evidence to support one model over the other nor to determine if two or more mechanisms contribute to LMID.

An early step in lignin biosynthesis is catalyzed by the enzyme *p*-coumaroyl shikimate 3′-hydroxylase (C3′H). A mutant carrying a defective *C3′H* gene, known as *reduced epidermal fluorescence 8* (*ref8-1*) has a reduction in total lignin content and is severely dwarf [17]. Unlike wild-type *Arabidopsis thaliana,* which deposits lignin primarily derived from the monomers coniferyl and sinapyl alcohol, the *ref8-1* mutant deposits a polymer made of *p*-coumaryl alcohol subunits [17]. Recent studies revealed that by eliminating two Mediator subunits, MED5A and MED5B, *med5a/5b ref8-1* only exhibit a mild dwarf phenotype [12]. The triple mutant also deposits a lignin derived from *p*-coumaryl alcohol, but it has wild-type lignin levels [12]. Through a suppressor screen, a mutant in an importin β-like, *GROWTH INHIBITION RELIEVED 1* (*GIR1*) was found to ameliorate the *ref8-1* growth phenotype [18]. *GIR1* is the nuclear transporter of the phenylpropanoid negative regulator MYB4 and knocking out *MYB4* also alleviates the dwarf phenotype of ref8-1 [18,19]. In contrast to *med5a/5b*, neither *gir1* nor *myb4* restore total lignin content [18]. In an attempt to understand the transcriptional changes that lead to dwarfism, Bonawitz *et al.* (2014) performed RNA sequencing (RNA-seq) on *ref8-1*. This study revealed that ~25% of the *ref8-1* transcriptome is mis-regulated [12]. This extent of transcriptional reprogramming makes it difficult to identify the differentially expressed genes that cause compromised plant growth from among those that are mis-regulated due to the severe alterations in the mutant's development [12]. In summary, the dwarf phenotype of *ref8-1* plants can be overcome by diverse manipulations at the transcriptional level, despite the mutant's compromised phenylpropanoid metabolism. These findings suggest that rather than lignin being causal for the severe phenotype of *ref8-1*, the dwarfism of these plants is a consequence of downstream changes in transcription arising from perturbations in the phenylpropanoid pathway.

In a previous study aiming to understand the link between phenylpropanoid metabolism and growth, *C3′H* was introduced into *ref8-1* plants under the control of a dexamethasone inducible promoter [20]. This study not only revealed that the induction of *C3′H* in *ref8-1* can alleviate its dwarf phenotype, but also that *C3′H* expression has distinct effects on plant growth when induced at different developmental stages [20]. In order to distinguish between the mis-regulated genes that lead to dwarfing and those that are secondarily altered in expression due to the mutant's severe growth phenotype, we have used the *C3′H* inducible system developed by Kim *et al.* (2014) to restore normal development to these plants in a controlled manner. Using this system, we performed RNA-seq and metabolite profiling at time points immediately following the

onset of *C3′H* expression to capture the initial transcriptional and metabolic changes that occur as phenylpropanoid homeostasis is being restored. Time course experiments have been instrumental in the discovery of missing links in the signal transduction of plant growth regulators, such as the small auxin up-regulated RNAs (SAURs) and jasmonate ZIM-domain (JAZ) genes [21,22]. Recently, a similar approach was taken to identify the sequence of transcriptional events that occur in response to the plant hormone abscisic acid (ABA) [23]. This study also identified the Dynamic Influencers in Gene expression (DIGs) and their homologs DIG-like proteins, a previously uncharacterized group of proteins that modulate ABA responses [23]. Our approach allowed us to determine the order of transcriptional events that lead to growth recovery in *ref8-1* plants. Furthermore, it provided a novel connection between the transcriptional changes concurrent with the fluctuations in phenylpropanoid metabolites.

3.3 Methods

3.4 Plant materials and growth conditions

The generation of the *ref8-1 C3′H* inducible (*C3′Hin*) line was described previously [20]. Columbia-0 (wild type) and *C3′Hin* seeds were surface sterilized and sown on plates containing ammonium-free Murashige and Skoog (MS) medium, 1% sucrose and 0.7% agar. After 2 days of stratification, plates were placed in a growth chamber set at 23°C and a photoperiod of 16 hours of light and 8 hours of dark. For induction experiments, seedlings were sprayed 9 days post-stratification with 20 μM dexamethasone (dex) in an aqueous solution of 0.02% dimethyl sulfoxide (DMSO) or 0.02% DMSO.

3.5 RNA extraction and qRT-PCR

Untreated seedlings were harvested in triplicate from day 6 to day 12 after stratification. RNA was extracted using the QIAGEN RNeasy® Plant Kit and DNAse-treated using the Invitrogen™ TURBO DNA-free™ kit following manufacturer's instructions. Complementary DNA was synthesized using 1 μg of RNA and the Invitrogen™ MultiScribe™ Reverse Transcriptase. Quantitative Real Time-Polymerase Chain Reaction (qRT-PCR) was performed with the primers described in primers (Table 3.1) and the Applied Biosystems™ SYBR™ Green PCR Master Mix.

3.5.1 RNA sequencing and differential expression analysis

Nine-day-old wild-type and *C3'Hin* seedlings were harvested in triplicate before treatment (time 0) and 4, 12, 24 and 48 hours after dex or DMSO treatment. Each replicate consisted of 8 seedlings. RNA was extracted as described above and DNAse-treated with the QIAGEN RNase-Free DNase in-column system. DNAse-treated RNA was submitted for quality control, library preparation and RNA sequencing to the Purdue Genomics Center. Libraries were prepared with the TECAN NuGEN Universal Plus kit, following manufacturer's instructions. Paired-end 50 bp reads were generated at a depth of 30 million using the NovaSeq 6000 Illumina® platform. Adaptors and low-quality reads (Phred < 20) were trimmed using Trimmomatic [24]. Reads were mapped to the *Arabidopsis thaliana* genome using STAR version 2.5.4b [25] and the TAIR10 genome assembly [26] and the Araport11 genome annotation [27]. Transcript counts were generated using HTSeq version 0.7.0 [28]. Differential gene expression was performed using the R Bioconductor package DESeq2 version 1.26.0 [29]. Genes with counts per million larger than 0.33 that were found in at least 3 libraries were included in the analysis. Three pairwise comparisons were done using a one factor design for each time point. One contrast evaluated the differentially expressed genes (DEGs) between wild type and *C3'Hin* treated with DMSO. The remaining contrasts evaluated the DEGs between dex and DMSO treatments for each genotype. Genes were considered differentially expressed if the adjusted p-value using the Benjamini and Hochberg method was lower than 0.05 (FDR < 0.05) and had an absolute $\log_2$ fold change equal to or higher than 1. Gene ontology (GO) analysis was performed using the online tool DAVID Bioinformatics Resource version 6.8 and using the 25,356 genes expressed in this dataset as background for the enrichment [30]. Redundant GO terms were removed using the online tool REVIGO and default parameters [31].

3.5.2 HPLC metabolite analysis

Nine-day-old wild-type and *C3'Hin* seedlings were harvested in triplicate before treatment (time 0) and 4, 12, 24, 48 and 72 hours after dex or DMSO treatment. Each replicate consisted of 10 seedlings. Metabolites were extracted in 50% methanol (10 µL mg^{-1} fresh weight) by incubating 2 h at 65°C. The supernatant was centrifuged at 14,500 rpm for 1 h and 10 µL samples were injected into a Shim-pack XR-ODS column (Shimadzu; 3.0 mm × 75 mm, 2.2 µm packing particle

size). Metabolites were separated by HPLC as described in Dolan *et al.* 2017 [16]. Sinapic acid, kaempferol and *p*-coumarate were utilized to quantify sinapoylmalate, flavonoids and *p*-coumarate esters, respectively.

3.5.3 Caffeoyl shikimate quantification

Seedlings treated for 1, 2, 4, and 8 h as described above were harvested in triplicate, with each replicate consisting of 8 seedlings. Metabolites were extracted in 75% methanol as described above. Supernatants were dried in a speed-vac and redissolved in 50% methanol. Samples were analyzed by LC/MS-MS using authentic standards for quantification essentially as described in Wang *et al.* 2018 [32].

3.5.4 Feeding and LC/MS metabolite analysis

Wild-type and *C3´Hin* seeds were sown on MS plates with 100 µM [$^{12}C_6$] or [$^{13}C_6$] phenylalanine. Nine days post-stratification, seedlings were sprayed with dex or DMSO and harvested 48 hours after treatment. Methanol-soluble metabolites were extracted as described above and 10 µL samples were injected into a Zorbax XDB-C18 column (Agilent Technologies; 2.1 mm × 150 mm, 3.5 µm packing particle size). Compounds were eluted with a 15 min gradient ranging from 5% to 35% acetonitrile in 0.1% formic acid at a rate of 0.3 mL min^{-1} in an Agilent 1290 Infinity II UPLC. The LC was coupled to an Agilent 6545 quadrupole time-of-flight tandem mass spectrometer operating in negative ion mode and an acquisition mode set to 70-1100 m/z.

3.6 Results

3.6.1 *ref8-1* exhibits a drought response that is reverted by *C3'H* induction

The *ref8-1* mutant undergoes extensive transcriptional re-programming and many of these changes have been attributed to indirect effects of the severely altered development of *ref8-1* [12]. In this study, our aim was to better understand *ref8* dwarfing by identifying the initial transcriptional and metabolic changes that lead to growth restoration when *C3'H* expression is restored in *ref8-1*. We focused on an experimental setup that would minimize the differences due to altered development in between *ref8-1* and wild type. To do so, we sought to identify a developmental stage in which the mutant's characteristic growth defects were not yet apparent by

growing wild-type and *ref8-1* seedlings on MS plates and monitoring their growth for 20 days. Under these growth conditions, the mutant's severe growth phenotype was modestly ameliorated in comparison to soil-grown plants (Figure 3.1). Mutant seedlings were visually indistinguishable from wild type until 12 days after planting (DAP) (Figure 3.1). This result narrowed down the time window in which future induction experiments would be performed.

Given that we selected a developmental stage with no visual phenotype, it was then necessary to determine if the mutant exhibited a transcriptional phenotype compared to wild-type plants at this developmental stage. A previous analysis of RNA-seq data from *ref8-1* showed that the Gene Ontology (GO) term "response to abscisic acid stimulus" was overrepresented among the upregulated genes [12]. Abscisic acid is a plant hormone involved in water stress responses and the rate-limiting step in its biosynthesis is catalyzed by the enzyme 9-cis-epoxycarotenoid dioxygenase 3 (NCED3), which was previously observed to be up-regulated in pot grown *ref8-1* plants. We monitored the expression of *NCED3* from 6 to 12 DAP in plate-grown *ref8-1* and wild-type seedlings by qRT-PCR (Figure 3.2). *NCED3* transcript levels were significantly higher in *ref8-1* when compared to wild-type plants as early as seven days after planting (Figure 3.2). Nine DAP was selected as the starting point for future experiments because at this growth stage *ref8-1* plants do not exhibit any of the mutant's characteristic growth defects (Figure 3.1), but stable changes in gene expression can already be detected. By conducting our experiments under these conditions, we expected that the secondary/pleiotropic changes in gene expression associated with impaired development would be minimized.

The *ref8-1* mutant has a characteristic soluble metabolite profile: it has low levels of sinapate esters and it hyperaccumulates flavonoids and *p*-coumarate esters [17]. Based on a previous study using the *C3'Hin* system, the altered metabolite levels of soil-grown plants can be rapidly reverted to wild-type levels by *C3'H* induction in *ref8-1* [20]. In order to identify the optimal timing for our experiments, we monitored by HPLC the changes in phenylpropanoid metabolites upon *C3'H* induction (Figure 3.3). By 12 hours after dexamethasone (dex) treatment, sinapoylmalate content had significantly increased (Figure 3.3A), whereas flavonoids and *p*-coumarate esters had decreased (Figure 3.3B and 3C). As previously reported by Kim *et al.* (2014), these data indicate that these plants turn over the hyperaccumulated compounds which originated from an impaired metabolic pathway and can rapidly return to a near wild-type state if *C3'H* expression is restored.

The timing of sinapoylmalate content increase served as an indicator of how long it takes for *C3'H* transcription and translation to restore pathway integrity. In addition, it suggested that the levels of intermediates upstream of sinapoylmalate would be affected at earlier time points. In order to determine the earliest sampling point after *C3'H* induction for our experiments, it was important to identify the earliest time point at which phenylpropanoid metabolism is affected. To do so, we assayed for the product of C3′H, caffeoyl shikimate, via LC-MS/MS from 1 to 8 hours after induction (Figure 3.3D). After only four hours of dex treatment the levels of caffeoyl shikimate were significantly higher than those of the mock-treated *ref8-1* plants (Figure 3.3D).

In order to determine the time at which the manipulation of phenylpropanoid metabolism has an impact on the transcription of other metabolic processes, we selected a marker from a different metabolic pathway. The transcript levels of *NCED3* were measured by qRT-PCR after *C3'H* induction and we found that they were restored to wild-type levels after 48 hours (Figure 3.4). Based on these findings, we selected 48 hours as the final time point for our experiment. Altogether, based on the time at which changes in phenylpropanoid metabolites and transcript levels of *NCED3* were detected, we opted to perform RNA-seq on wild-type and *ref8-1* plants 4 ,12, 24 and 48 hours after *C3'H* induction.

3.6.2 RNA-seq reveals rapid global changes in gene expression upon *C3'H* induction

To identify the changes in gene expression when *C3'H* expression is restored, we performed RNA-seq on shoot samples from nine-day old (T0) wild type and *C3'Hin* seedlings, after 4 (T4), 12 (T12), 24 (T24), and 48 (T48) hours of mock or dex treatment. Transcript data from wild-type samples was used in two ways: 1) to identify the differences between genotypes over time and 2) the changes in transcription due to dex application.

In contrast to the previous study where 1,953 genes were identified [12]as differentially expressed at least 2-fold between *ref8-1* and wild-type plants, we only detected 192 differentially expressed genes (DEGs) at T0 (Table 3.2) (adjusted $P < 0.05$). This low number may be due to the developmental stage in which we carried the experiments, supporting the notion that the majority of the DEGs detected in the previous study were due to developmental differences. An alternative or additional factor may be the sterile growth conditions used in this experiment, which eliminated the abiotic and biotic factors that may lead to changes in gene expression in pot grown plants.

Gene expression differences due to genotype were identified by comparing *C3'Hin* and wild-type mock-treated samples for each time point. These data sets revealed that the number of mis-regulated transcripts in *ref8-1* remained relatively constant in T0, T4 and T24 (Table 3.2). Unexpectedly, the number of DEGs tripled at T12 when compared to T0, which may reflect a circadian rhythm-specific effect on gene expression in *ref8-1* because at T24 there were only 260 DEGs detected (Table 3.2). This increase in DEGs at T12 could also be a single-pulse pattern of gene expression, which is an increase or decrease in transcript levels that rapidly return to their original state often seen in eukaryotic responses to environmental cues [33]. The number of differentially expressed genes increased from 192 to 591 in between T0 and T48 (Table 3.2), which based on the previous RNA-seq data is to be expected because as plant age increases the number of DEGs between wild type and *ref8-1* increases.

In order to identify the genes with altered transcription due to *C3'H* induction in *ref8-1,* we first compared dex- and mock- treated *C3'Hin* data sets and then removed the DEGs that were differentially expressed between dex- and mock- treated wild-type samples. The number of DEGs in *C3'Hin* increased over time from 38 DEGs four hours after treatment, to 741 DEGs 48 hours after treatment (Table 3.2). Restoration of gene transcription was defined as those genes that were mis-regulated in *ref8-1* compared to wild type and were differentially expressed in the opposite direction in dex- versus mock- treated *C3'Hin* samples. Transcription restoration was not detected until 12 h after treatment, with a total of 16 mis-regulated genes partially or fully restored to wild-type levels (Figure 3.5). The number of restored transcripts increased over time, ~70% of the upregulated and downregulated genes in *ref8-1* were restored or partially restored by 48 h after *C3'H* induction (Figure 3.5).

The RNA-seq data obtained by this time course experiment reflects the rapid changes in gene expression that occur when *C3'H* expression is restored. In addition, it revealed that a large proportion of changes detected after *C3'H* induction were indeed previously mis-expressed genes and not simply the response of plants to being sprayed with an exogenous chemical nor the result of *C3'H* overexpression.

3.6.3 Phenylpropanoid restoration leads to changes in responses to water deprivation and defense against bacteria

To obtain insight into the biological processes and pathways that the differentially expressed genes represent, we performed gene ontology (GO) analysis and KEGG pathway enrichment. The inputs for the analysis were the DEGs whose expression increased or decreased at least 2-fold (adjusted $P < 0.05$) due to genotype differences or *C3'H* induction. The GO terms associated with genes upregulated in *ref8-1* compared to wild type were associated with stress due to water deprivation (Figure 3.6A). In concordance with this observation, GO terms associated with the hormone abscisic acid were also among the over-represented terms (Figure 3.6A). KEGG and GO term analysis showed that the upregulated genes were also represented by terms pertaining to the hormones salicylic acid, jasmonic acid and ethylene (Figure 3.6B). In addition to abscisic acid, "ethylene-activated signaling" was consistently detected in three time points (Figure 3.6A). By 12 h after dex treatment, *C3'Hin* exhibited a gene expression profile consistent with a decrease in response to ethylene (Figure 3.6A). As early as 24 h, induction of *C3'H* led to a reduction in all water stress categories and this trend continued 48 hours after induction (Figure 3.6A).

Downregulated genes in *ref8-1* compared to wild type were represented by GO terms related to iron homeostasis and defense responses elicited by bacteria (Figure 3.6B). At T48, *C3'H* induction led to an increase in defense response-related genes in *ref8-1* (Figure 3.6B). Changes in transcriptional processes related to iron homeostasis were not captured during our induction time course.

3.6.4 Reduction in ABA biosynthesis enhances the dwarf phenotype of *ref8-1*

To determine whether there was a link between ABA gene misregulation and the *ref8-1* growth phenotype, we used mutants in the ABA biosynthetic pathway to block the synthesis of ABA in *ref8-1*. The mutants in *ABSCISIC ACID DEFICIENT 2* and *ABSCISIC ACID DEFICIENT 3, aba2* and *aba3* are impaired in the synthesis of the ABA precursor abscisic aldehyde and abscisic acid, respectively [34]. These mutants exhibit only a mild impairment in growth (Figure 3.7), in contrast, reduction of ABA levels in *ref8-1* leads to a further impairment in plant growth in *ref8-1 aba2* and *ref8-1 aba3* when compared to *ref8-1,* effect which can be mitigated by the application of ABA. These results indicate that the elevated levels of ABA in *ref8-1* may serve a protective role against the downstream effects arising from a perturbed phenylpropanoid pathway.

3.6.5 Inhibition of ethylene synthesis or signaling does not improve the growth of *ref8-1*

As mentioned earlier, genes involved in "ethylene-activated signaling" were upregulated in *C3'Hin* and by 12 h after dex treatment these transcripts were downregulated. The gene expression profile in *C3'Hin* suggests there is an increase in ethylene content in *C3'Hin.* In addition, ABA prevents ethylene hyperaccumulation to maintain shoot growth [35] which may explain why blocking ABA biosynthesis enhances the dwarf phenotype of *ref8-1.* Ethylene response can be assessed using the triple response assay which involves scoring the growth of etiolated seedlings under high ethylene concentrations [36]. The triple response consists of shorter and swelled hypocotyls and roots, and a more pronounced apical hook [36]. Mutants that overproduce ethylene or in which signaling is constitutively active, exhibit these phenotypes in the absence of ethylene [36]. We assessed the phenotype of etiolated *ref8-1* and *aba2 ref8-1* seedlings in the absence of increased ethylene. Figure 3.8A shows that there was no difference in root and shoot size or the apical hook of *ref8-1* and *aba2 ref8-1* when compared to wild-type etiolated seedlings. In order to further asses the role of ethylene in the phenotypes of *ref8-1* and *aba2 ref8-1* we grew seedlings in the presence of the ethylene synthesis inhibitor 2-aminoethoxyvinyl glycine (AVG) [36]. Figure 3.8B shows that 1 µM AVG had a negative effect in the root length of wild-type, *aba2, aba2 ref8-1* seedlings, but not on *ref8-1.* Increasing the AVG concentration to 2 µM was inconsistent with the results observed at 1 µM, because under the higher concentration *aba2* had longer roots than all other phenotypes. The effect of ethylene inhibition on root growth in *ref8-1* is inconclusive, but Lastly, we tested the ethylene signaling inhibitor silver thiosulfate [36]and observed no effect in the growth of *aba2, ref8-1* or *aba2 ref8-1* (Figure 3.8C).

3.6.6 Increased PAL activity in *ref8-1* despite the upregulation of *KFB* genes

Differential expression analysis of genes involved in phenylpropanoid regulation revealed that *KFB20*, *KFB39* and *KFB50* were upregulated in *C3'Hin* and their expression was downregulation after 12 h of dex application (Figure 3.9 and 3.10). Kelch domain F-box proteins (KFB) 1/20/39/50 regulate PAL at the posttranslational level by targeting PAL for ubiquitin mediated degradation [37,38]. Downregulation of these KFBs leads to higher PAL activity and an increase in phenylpropanoids [37–39]. In order to test if a reduction in PAL activity was responsible for the phenotypes observed in *ref8-1,* we measured PAL activity in *C3'Hin* and wild

type after mock or dex treatment application (Figure 3.11). PAL activity was higher in *C3'Hin* when compared to wild type prior to treatment (Figure 3.11). At four hours after treatment PAL activity was lower in *C3'Hin* treated with dex than in the mock treated samples, but regardless of treatment PAL activity remained higher in *C3'Hin* samples than in wild type samples until 24 hours after treatment (Figure 3.11). Similar to four hours after treatment, by 48 hours PAL activity decreased in *C3'Hin* dex-treated samples, but remained higher than wild type (Figure 3.11).

3.6.7 Early activation of WOX14 after *C3'Hin* induction

As mentioned earlier, we identified a small number of DEGs 4 h after applying dex to *C3'Hin*. One of the few transcripts that was upregulated in *C3'Hin* at this time point, and remained upregulated throughout the time course after dex application, was *WUSCHEL HOMEOBOX RELATED 14* (WOX14). WOX14 is a transcription factor involved in the regulation of cell division in the vasculature [40–42]. In order to test the role of WOX14 in the phenotype of *ref8-1,* we first tested if the increase in expression was a pleiotropic effect of *C3'H* overexpression. To test this, we crossed wild type to *C3'Hin* and measured transcript levels by qRT-PCR of both *C3'H* and *WOX14* after 4 h of dex application. As seen in Figure 3.12, the expression of *C3'H* was ~5 times lower in the wild-type background when compared to the *ref8-1* background. *C3'H* induction led to a significant increase in the expression of *WOX14* in both genetic backgrounds, although the effect was more pronounced in *ref8-1* (Figure 3.12).

3.7 Discussion

The mechanism by which dwarfism associated with lignin modification comes about has been a long-standing question in the field. The preferred hypothesis to explain the LMID phenomenon is that the reduction in lignin results in defective vessels that lead to poor water transport. This hypothesis is supported by the observation of collapsed vessels in many dwarf lignin deficient mutants [3,7]. The finding that vessel-specific complementation alleviates both the vasculature collapse and dwarfism of a lignin deficient mutant without restoring total lignin content, further supports this hypothesis [6] On the other hand, it is unknown if collapsed xylem is the direct effect of reduced lignin content or whether it occurs due transcriptional changes that arise due to perturbed phenylpropanoid metabolism. Although the vessel complementation results

imply that hypo-lignified vessels are the cause of LMID, alternative hypotheses cannot be excluded. For example, a cell wall defect sensing mechanism may be located in the xylem cells which would not be activated if lignin biosynthesis is restored in these cells. Additionally, the hyper- or hypo- accumulation of a phenylpropanoid metabolite that promotes or impairs growth may be necessary in these cell types, and again could explain how vessel-specific expression restores growth. Finally, in all these experiments, the exact cell types in which the promoters employed are active has not been fully determined. It is possible that the complementation of the dwarf phenotypes of these mutants is due to restored phenylpropanoid gene expression in non-lignified cells, and that complementation of the collapsed xylem phenotype is actually a phenomenon unrelated to the elimination of dwarfism.

In recent years, advances in sequencing technologies have enabled large scale studies that support alternative hypotheses that go beyond that of reduced water transport to explain LMID. The finding that either knocking out *MED5A* and *MED5B*, or the nuclear transporter of MYB4 relieves the dwarf phenotype of *ref8-1* [12,18] prompted us to further investigate this phenomenon with the aim of dissecting the order of transcriptional and metabolic events that lead to this recovery in growth. Using an inducible system in combination with RNA-seq and metabolomics, this study provided the opportunity to unveil the cellular processes that lead to growth recovery, and by inference, LMID, in addition to providing candidate metabolites that could be involved in these processes.

3.7.1 *ref8* exhibits response to drought despite high humidity conditions

Previous RNA-seq data indicated that soil-grown *ref8-1* plants exhibited an upregulation in transcripts involved in drought tolerance particularly those involved in response to ABA [12]. Because ABA leads to stomatal closure, an increase in ABA levels leads to reduced transpiration which in turn results in an increase in drought tolerance [43]. Despite growing *ref8-1* under high humidity conditions, we again observed upregulation of genes involved in ABA biosynthesis and response to this hormone in *ref8* plants. For example, *NCED3*, which is upregulated in *ref8-1*, is typically induced under drought conditions and this induction correlates with the levels of ABA accumulated [43]. In contrast, high humidity conditions typically lead to the degradation of ABA [44], although this does not seem to be the case in our experiment, given the high number of upregulated ABA-responsive transcripts identified in our data under high humidity conditions.

ref8-1 is not the first lignin biosynthetic mutant shown to exhibit an altered response to drought. The *pal1 pal2* mutant, which is defective in the entry step of the phenylpropanoid pathway, phenylalanine ammonia lyase, exhibits enhanced tolerance to drought [45]. Furthermore, *ref4-3* a semidominant dwarf mutant in *MED5B* that has reduced lignin content, also exhibits a transcript profile characteristic of plants undergoing water deficit [16 , Mao (unpublished)]. The activation of mechanisms to enhance drought tolerance are also not unique to lignin biosynthesis, as they have also been identified in other mutants defective in secondary cell wall deposition. Transcriptomic analysis revealed that mutants in the cellulose synthase genes *CESA4/IRREGULAR XYLEM5 (IRX5)* and *CESA8/IRX1* exhibit a strong response to drought, with 58% of their misregulated transcripts related to ABA synthesis and signaling [46]. Collapsed xylem vessels in *ref8-1* and other secondary cell wall mutants may be the direct cause of changes in ABA synthesis and response. The results are also consistent with the existence of a cell wall integrity sensing mechanism capable of detecting defects in the secondary cell wall and activating a transcriptional response that involves ABA.

3.7.2 Restoration of *C3'H* expression leads to rapid changes in gene expression beyond those related to phenylpropanoid metabolism

Bonawitz *et al.* (2014) [12] documented vast transcriptional reprogramming in *ref8-1* plants. For our experiments, we selected a developmental stage in which *ref8-1* does not exhibit strong growth defects and under conditions in which they were not subjected to abiotic and biotic stresses. The purpose of this experimental design was to identify the DEGs that were more proximal to the cause of the *ref8-1* transcriptional changes. By using this design, we detected a smaller number of DEGs. Nevertheless, our data are consistent with Bonawitz *et al.* (2014) [12] with regard to the biological processes with which the DEGs are associated. The transcriptional changes arising by induction of *C3'H* were rapid. As early as four hours after *C3'H* induction we could detect DEGs not belonging to the phenylpropanoid pathway, and by 12 hours, some of the genes that were misregulated in *ref8-1* returned to wild-type expression levels. The relevance of ABA in the response of plants to lignin deficiency was reiterated in our experiment, or at least in the context of *ref8-1*. ABA biosynthetic genes and those responsive to ABA were upregulated in *ref8-1* and induction of *C3'H* led to a restoration of their transcript levels as early as 24 hours after treatment. As mentioned earlier, in addition to *ref8-1*, there are other dwarf mutants with altered

phenylpropanoid metabolism that exhibit aberrant expression of drought responsive genes (Dolan et al., 2017; Mao *et al.,* 2019). Suppressors of these mutants' defective growth also restore a normal response to drought. For example, through a suppressor screen and a reverse genetics approach aiming to identify dwarfism suppressors of *ref4-3*, subunits of the Mediator complex, namely MED2, MED16, MED23 and CDK8 were identified [16,47]. The suppressors restored the growth phenotype in addition to restoring the expression of transcripts involved in response to water deprivation [16, Mao (unpublished)]. Altogether this data is consistent with the hypothesis that lignin-deficient mutants have impaired water transport and that may be why they exhibit a drought-like transcriptional response. Furthermore, the restoration of growth in *ref4-3 cdk8* was independent of changes in lignin content [47], consistent with the model that at least in some mutant contexts, lignification is not the direct cause of dwarfing.

The first cellular process to return to homeostasis after *C3'H* induction was the upregulation of genes associated with ethylene-activated signaling in *ref8-1*, returning to wild-type expression levels after 12 hours of treatment. Interactions between ethylene and ABA are either synergistic or antagonistic, depending on the tissue type studied [35,48,49]. In shoots, ABA inhibits ethylene synthesis to maintain growth [35]. In addition, ethylene overexpressors exhibit impaired growth [36]. Based on these observations, we used ethylene synthesis and signaling inhibitors to test the hypothesis that the upregulation of ABA biosynthesis in *ref8-1* is a mechanism to quench high levels of ethylene that may themselves be leading to dwarfing. Our results indicate that ethylene does not contribute to the dwarf phenotype of *ref8-1.*

Based on GO term analysis, a biological process that is affected in *ref8-1* is defense against bacteria. Genes associated with systemic acquired resistance (SAR) are downregulated in *ref8-1* and are restored 48 hours after *C3'H* induction. SAR is mainly mediated by the hormone salicylic acid (SA), whose biosynthesis is primarily contributed by the isochorismate pathway, with a small contribution from the phenylpropanoid pathway [50]. C3′H is downstream of SA biosynthesis in the phenylpropanoid pathway, therefore, the *ref8-1* mutation does not interfere with the phenylpropanoid-mediated synthesis of SA. Consequently, the differential expression observed in SAR genes cannot be a direct effect of changes in SA levels due to C3′H deficiency. The downregulation of genes involved in SAR could result from the increase in ABA levels, since ABA acts in an antagonistic manner upstream and downstream of SA in SAR [51]. Consistent

with this model, the restoration of expression of genes associated with SAR mediated by *C3'H* induction parallels that of some ABA responsive genes.

Genes involved in iron homeostasis were downregulated in *ref8-1* prior to treatment (T0) including *FER1, FER2, FER3* and *FRD3*, but they did not exhibit differential expression between *ref8-1* under mock treatment and wild type in later time points. There are two known links between iron and phenylpropanoids. The first one is fraxetin, a coniferyl alcohol-derived coumarin which is released by *Arabidopsis* roots under conditions of low-iron availability [52,53]. Fraxetin is released by roots into the soil in order to facilitate iron uptake [52,53]. Differential expression of iron homeostasis genes was only detected in T0, but not in later time points regardless of treatment. It is possible that the growth media and application of aqueous mock and dex treatments could have dispersed the coumarins released to the media by wild-type roots. These coumarins in turn could have complemented the deficiency of *ref8* roots, eliminating differential expression in later time points. The second association with iron homeostasis involves the Casparian strip, a structure made of lignin found in root endodermis that plays a role in controlling the diffusion of water and solutes in the root [54]. The Enhanced Suberin 1 (ESB1) mutant, which is defective in Casparian strip deposition, has altered ion homeostasis, including iron deficiency [55]. The *ref8-1* mutation may also be affecting Casparian strip deposition which like in *esb1,* would result in iron deficiency. Although a defective Casparian strip could explain the downregulation of genes associated with iron homeostasis in *ref8-1,* it is unlikely that it explains the dwarf phenotype of *ref8-1,* since *esb-1* does not exhibit impaired growth.

3.7.3 Blocking ABA biosynthesis exacerbates the dwarf phenotype of *ref8-1*

Due to the strong transcriptional response to ABA and the amount of perturbed cellular processes in *ref8-1* that could be ABA-dependent, we blocked ABA biosynthesis in ref8-1 to remove all transcriptional changes due to elevated ABA and in this way identify processes directly altered due to the loss in phenylpropanoid homeostasis. The double mutants *aba2 ref8-1* and *aba3 ref8-1* plants exhibit a more severe dwarf phenotype than *ref8-1*. This result supports the involvement of ABA in mechanisms induced to ameliorate the effects of poor water transport in *ref8-1*. Alternatively, blocking ABA biosynthesis in *ref8-1* may lead to similar responses as when ABA synthesis is reduced in *esb1,* which results in severely compromised growth [56]. Recent studies showed that ABA plays a role in reducing shoot's transpiration rates when the Casparian

strip is compromised [56]. The severe phenotype of *aba2 ref8-1* and *aba3 ref8-1* may arise by the additive effect of a defective Casparian strip together with the inability to reduce transpiration rates in the shoot due to the reduction in ABA biosynthesis.

3.7.4 Although PAL degradation is activated it does not negatively affect PAL activity

Although *ref8-1 and med5a/5b ref8-1* plants deposit primarily H-lignin, the total lignin of *ref8-1* is substantially reduced in comparison to *med5a/5b ref8-1* [12]. Both plants have the same metabolic block; therefore, a mechanism that inhibits lignin deposition in *ref8-1* must be inhibited or bypassed in *med5a/5b ref8-1*. Our RNA-seq data indicated an increase in KFB 20/39/50 expression in *ref8-1* mutants and restoration of KFB expression upon *C3'H* induction. Previous studies showed that the upregulation of *KFB1/20/39/50* is a mechanism to degrade PAL in a MED5 dependent manner [37–39]. PAL is the entry step to the phenylpropanoid pathway, and as such is a gateway point to control flux through the pathway. There are four isoforms of PAL in *Arabidopsis* plants and depending on the isoforms knocked out, mutants exhibit a varying degree of growth phenotypes [45]. This led us to test the hypothesis that the low-lignin *ref8-1* phenotype, and possibly others, could be attributed to a reduction in PAL activity. Contrary to our prediction, we measured higher levels of PAL activity in *ref8-1* when compared to wild type. RNA-seq data from Bonawitz *et al.* (2014) [12] and our data show an increase in *PAL 1* and *PAL 4* transcript levels. Although upregulation of KFB expression may lead to PAL degradation, the concurrent upregulation of *PAL* transcript levels may be able to outcompete the increase in degradation and in this way maintain high levels of PAL activity. Based on these results, it seems unlikely the increase in KFB expression is responsible for the reduced lignin content and dwarfism of *ref8-1.*

3.7.5 WOX14 may be one of the initial signals in growth restoration

WOX14 has known roles in vascular cell proliferation and lignification [40–42] and *WOX14* expression was upregulated in *ref8-1* after only two hours of *C3'H* induction and remained higher than both *ref8* and wild-type transcript levels for 48 hours. Although this could be regarded as an ectopic effect of *C3'H* expression, these data may suggest that elevated levels of this transcription factor may be associated with growth restoration. Further evidence is required in

order to differentiate between an ectopic effect of *C3'H* expression or a real role of WOX14 in the restoration of normal growth.

3.8 References

1. DeMartini JD, Pattathil S, Miller JS, Li H, Hahn MG, Wyman CE: Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. *Energy Environ Sci* 2013, 6:898–909.
2. Himmel ME, Ding S, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD: Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* 2007, 315:804-807.
3. Turner SR, Somerville CR: Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell* 1997, 9:689–701.
4. Wang P, Dudareva N, Morgan JA, Chapple C: Genetic manipulation of lignocellulosic biomass for bioenergy. *Curr Opin Chem Biol* 2015, 29:32–39.
5. Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD: Designer lignins: Harnessing the plasticity of lignification. *Curr Opin Biotechnol* 2016, 37:190–200.
6. De Meester B, de Vries L, Özparpucu M, Gierlinger N, Corneillie S, Pallidis A, Goeminne G, Morreel K, De Bruyne M, De Rycke R, et al.: Vessel-specific reintroduction of CINNAMOYL COA REDUCTASE 1 (CCR1) in dwarfed *ccr1* mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol* 2018, 176:611–633.
7. Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet A: Down-regulation of Cinnamoyl-CoA Reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 1998, 13:71–83.
8. Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M: Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* 2007, 19:148–162.
9. Schilmiller AL, Stout J, Weng J-K, Humphreys J, Ruegger MO, Chapple C: Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in Arabidopsis. *Plant J* 2009, 60:771–82.

10. Muro-Villanueva F, Mao X, Chapple C: Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr Opin Biotechnol* 2019, 56:202–208.
11. Gallego-Giraldo L, Escamilla-Trevino L, Jackson LA, Dixon RA: Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc Natl Acad Sci* 2011, 108:20814–20819.
12. Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, et al.: Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* 2014, 509:376–380.
13. Binns AN, Chen RH, Wood HN, Lynn DG: Cell division promoting activity of naturally occurring dehydrodiconiferyl glucosides: Do cell wall components control cell division? *Proc Natl Acad Sci* 1987, 84:980–984.
14. Teutonico RA, Dudley MW, Orr3 JD, Lynn DG, Binns AN: Activity and accumulation of naturally occuring dehydrodiconiferyl glucosides: do cell wall components control cell division? *Plant Physiol* 1991, 97:288–297.
15. Hamann T: The plant cell wall integrity maintenance mechanism - A case study of a cell wall plasma membrane signaling network. *Phytochemistry* 2015, 112:100–109.
16. Dolan WL, Dilkes BP, Stout JM, Bonawitz ND, Chapple C: Mediator complex subunits MED2, MED5, MED16, and MED23 genetically interact in the regulation of phenylpropanoid biosynthesis. *Plant Cell* 2017, 29:3269-3285.
17. Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C: Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of Arabidopsis. *Plant J* 2002, 30:47-59.
18. Panda C, Li X, Wager A, Chen H, Li X: An importin-beta-like protein mediates lignin-modification- induced dwarfism in Arabidopsis. *Plant J* 2020, 102:1281-129.
19. Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C: Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J* 2000, 19:6150–6161.
20. Kim JI, Ciesielski PN, Donohoe BS, Chapple C, Li X: Chemically induced conditional rescue of the *reduced epidermal fluorescence8* mutant of Arabidopsis reveals rapid restoration of growth and selective turnover of secondary metabolite pools. *Plant Physiol.* 2014, 164:584-595.

21. McClure BA, Guilfoyle T: Characterization of a class of small auxin-inducible soybean polyadenylated RNAs. *Plant Mol Biol* 1987, 9:611–623.
22. Chini A, Fonseca S, Fernández G, Adie B, Chico JM, Lorenzo O, García-Casado G, López-Vidriero I, Lozano FM, Ponce MR, et al.: The JAZ family of repressors is the missing link in jasmonate signalling. *Nature* 2007, 448:666–671.
23. Song L, Huang SSC, Wise A, Castanoz R, Nery JR, Chen H, Watanabe M, Thomas J, Bar-Joseph Z, Ecker JR: A transcription factor hierarchy defines an environmental stress response network. *Science* 2016, 354: aag1550.
24. Bolger AM, Lohse M, Usadel B: Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 2014, 30:2114–2120.
25. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013, 29:15–21.
26. Swarbreck D, Wilks C, Lamesch P, Berardini TZ, Garcia-Hernandez M, Foerster H, Li D, Meyer T, Muller R, Ploetz L, et al.: The Arabidopsis Information Resource (TAIR): Gene structure and function annotation. *Nucleic Acids Res* 2008, 36:1009–1014.
27. Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD: Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J* 2017, 89:789–804.
28. Anders S, Pyl PT, Huber W: HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015, 31:166–169.
29. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014, 15:1–21.
30. Huang DW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, 4:44–57.
31. Supek F, Bošnjak M, Škunca N, Šmuc T: Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* 2011, 6.
32. Wang P, Guo L, Jaini R, Klempien A, McCoy RM, Morgan JA, Dudareva N, Chapple C: A^{13}C isotope labeling method for the measurement of lignin metabolic flux in Arabidopsis stems. *Plant Methods* 2018, 14:1–15.

33. Yosef N, Regev A: Impulse control: Temporal dynamics in gene transcription. *Cell* 2011, 144:886–896.
34. Schwartz SH, Léon-Kloosterziel KM, Koornneef M, Zeevaart JAD: Biochemical characterization of the *aba2* and *aba3* mutants in *Arabidopsis thaliana*. *Plant Physiol* 1997, 114:161–166.
35. LeNoble ME, Spollen WG, Sharp RE: Maintenance of shoot growth by endogenous ABA: Genetic assessment of the involvement of ethylene suppression. *J Exp Bot* 2004, 55:237–245.
36. Binder BM, Schaller GE, co-editors: *Ethylene signaling: Methods and protocols.* Humana Press, Springer Science+Business Media LLC; 2007.
37. Zhang X, Gou M, Liu CJ: *Arabidopsis* Kelch repeat F-Box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell* 2013, 25:4994–5010.
38. Zhang X, Gou M, Guo C, Yang H, Liu CJ: Down-regulation of Kelch domain-containing F-box protein in Arabidopsis enhances the production of (poly)phenols and tolerance to ultraviolet radiation. *Plant Physiol* 2015, 167:337–350.
39. Kim JI, Zhang X, Pascuzzi PE, Liu CJ, Chapple C: Glucosinolate and phenylpropanoid biosynthesis are linked by proteasome-dependent degradation of PAL. *New Phytol* 2020, 225:154–168.
40. Etchells JP, Provost CM, Turner SR: Plant vascular cell division is maintained by an interaction between PXY and ethylene signalling. *PLoS Genet* 2012, 8:e1002997.
41. Etchells JP, Provost CM, Mishr L, Turner SR: WOX4 and WOX14 act downstream of the PXY receptor kinase to regulate plant vascular proliferation independently of any role in vascular organisation. *Development* 2013, 140:2224–2234.
42. Denis E, Kbiri N, Mary V, Claisse G, Conde e Silva N, Kreis M, Deveaux Y: WOX14 promotes bioactive gibberellin synthesis and vascular cell differentiation in Arabidopsis. *Plant J* 2017, 90:560–572.
43. Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K: Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in *Arabidopsis*. *Plant J* 2001, 27:325–333.

44. Okamoto M, Tanaka Y, Abrams SR, Kamiya Y, Seki M, Nambara E: High humidity induces abscisic acid 8′-hydroxylase in stomata and vasculature to regulate local and systemic abscisic acid responses in Arabidopsis. *Plant Physiol* 2009, 149:825–834.
45. Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y-H, Yu J-Q, Chen Z: Functional analysis of the Arabidopsis PAL gene family in plant growth, development, and response to environmental stress. *Plant Physiol* 2010, 153:1526–1538.
46. Hernández-Blanco C, Feng DX, Hu J, Sánchez-Vallet A, Deslandes L, Llorente F, Berrocal-Lobo M, Keller H, Barlet X, Sanchez-Rodríguez C, *et al*.: Impairment of cellulose Synthases Required for Arabidopsis Secondary Cell Wall Formation Enhances Disease Resistance. *Plant Cell* 2007, 19:890–903.
47. Mao X, Kim JI, Wheeler MT, Heintzelman AK, Weake VM, Chapple C: Mutation of Mediator subunit *CDK8* counteracts the stunted growth and salicylic acid hyperaccumulation phenotypes of an Arabidopsis *MED5* mutant. *New Phytol* 2019, 223:233–245.
48. Beaudoin N, Serizet C, Gosti F, Giraudat J: Interactions between abscisic acid and ethylene signaling cascades. *Plant Cell* 2000, 12:1103–1115.
49. Ghassemian M, Nambara E, Cutler S, Kawaide H, Kamiya Y, McCourt P: Regulation of abscisic acid signaling by the ethylene response pathway in Arabidopsis. *Plant Cell* 2000, 12:1117–1126.
50. Lefevere H, Bauters L, Gheysen G: Salicylic acid biosynthesis in plants. *Front Plant Sci* 2020, 11:1–7.
51. Yasuda M, Ishikawa A, Jikumaru Y, Seki M, Umezawa T, Asami T, Maruyama-Nakashita A, Kudo T, Shinozaki K, Yoshida S, et al.: Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in Arabidopsis. *Plant Cell* 2008, 20:1678–1692.
52. Tsai H-H, Rodriguez-Celma J, Lan P, Wu Y-C, Vélez-Bermúdez IC, Schmidt W: Scopoletin 8-Hydroxylase-Mediated Fraxetin Production is Crucial for Iron Mobilization. *Plant Physiol* 2018, 177:207.
53. Rajniak J, Giehl RFH, Chang E, Murgia I, von Wirén N, Sattely ES: Biosynthesis of redox-active metabolites in response to iron deficiency in plants. *Nat Chem Biol* 2018, 14:442–450.

54. Naseer S, Lee Y, Lapierre C, Franke R, Nawrath C, Geldner N: Casparian strip diffusion barrier in Arabidopsis is made of a lignin polymer without suberin. *Proc Natl Acad Sci U S A* 2012, 109:10101–10106.

55. Hosmani PS, Kamiya T, Danku J, Naseer S, Geldner N, Guerinot M Lou, Salt DE: Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Proc Natl Acad Sci U S A* 2013, 110:14498–14503.

56. Wang P, Calvo-Polanco M, Reyt G, Barberon M, Champeyroux C, Santoni V, Maurel C, Franke RB, Ljung K, Novak O, et al.: Surveillance of cell wall diffusion barrier integrity modulates water and solute transport in plants. *Sci Rep* 2019, 9:1–11.

57. Garapati, P., Xue, G.-P., Munné-Bosch, S., & Balazadeh, S. Transcription factor ATAF1 in Arabidopsis promotes senescence by direct regulation of key chloroplast maintenance and senescence transcriptional cascades. *Plant Physiol* 2015, 168:1122–1139.

Table 3.1. List of primers used in this project.

Primer name	Sequence (5'-> 3')	Usage
CC5535	AGT CAA CCT CGA AGC AGG GAT G	*NCED3* qRT-PCR fwd [57]
CC5536	CTC GGC TAA AGC CAA GTA AGC G	*NCED3* qRT-PCR rv [57]
CC5837	GCA AAG AAA TAG AGG AGA TGG ATA GAG A	*WOX14* qRT-PCR fwd [42]
CC5838	GAG TTG AGG AGG ACG AGT T	*WOX14* qRT-PCR rv [42]
CC4921	GCC GTG ATC TGA TGA TGA TG	*KFB01* qRT-PCR fwd [37]
CC4922	TAC CCT ACA CTT GGT TCG CC	*KFB01* qRT-PCR rv [37]
CC4923	TCC TGG TCA AAC CAA AGG AC	*KFB20* qRT-PCR fwd [37]
CC4924	AAG CGC ACA TTT CTC TTC GT	*KFB20* qRT-PCR rv [37]
CC4927	TAC GGA ACG GAA TCT CAA GG	*KFB39* qRT-PCR fwd [38]
CC4928	GTT TCC CAT CGA CGT TCA CT	*KFB39* qRT-PCR rv [38]
CC4925	AAC CGC GTA TTA CTG GAA CG	*KFB50* qRT-PCR fwd [37]
CC4926	CCG CCG ATT AAC AAC ACT TT	*KFB50* qRT-PCR rv [37]
CC2839	CAA AGG ATC AAA CGT TCA TGT GA	*C3'H* qRT-PCR fwd [20]
CC2840	CAA AGG ATC AAA CGT TCA TGT GA	*C3'H* qRT-PCR rv [20]
CC1795	GGC TAG CGA TGA CTC GCG GTA CCT AA	*aba2-1* CAPs marker fwd
CC1796	TGC TAA AAG TAA CAA TTA AAA CCA CAA	*aba2-1* CAPs marker rv

Table 3.2. Differentially expressed genes identified in the RNA-seq analysis.
Genes were considered as differentially expressed if the false discovery rate was higher than 0.05 and the absolute $\log_2$ fold change was higher than 1.

***C3'Hin* vs. wt**

Time (h)	Down	Up	Total
0	110	82	192

***C3'Hin*-mock vs. wt-mock**

Time (h)	Down	Up	Total
4	23	119	142
12	202	411	613
24	40	220	260
48	171	420	591

(*C3'Hin* -dex vs. *C3'Hin*-mock)-(wt-dex vs. wt-mock)

Time (h)	Down	Up	Total
4	23	119	142
12	202	411	613
24	40	220	260
48	171	420	591

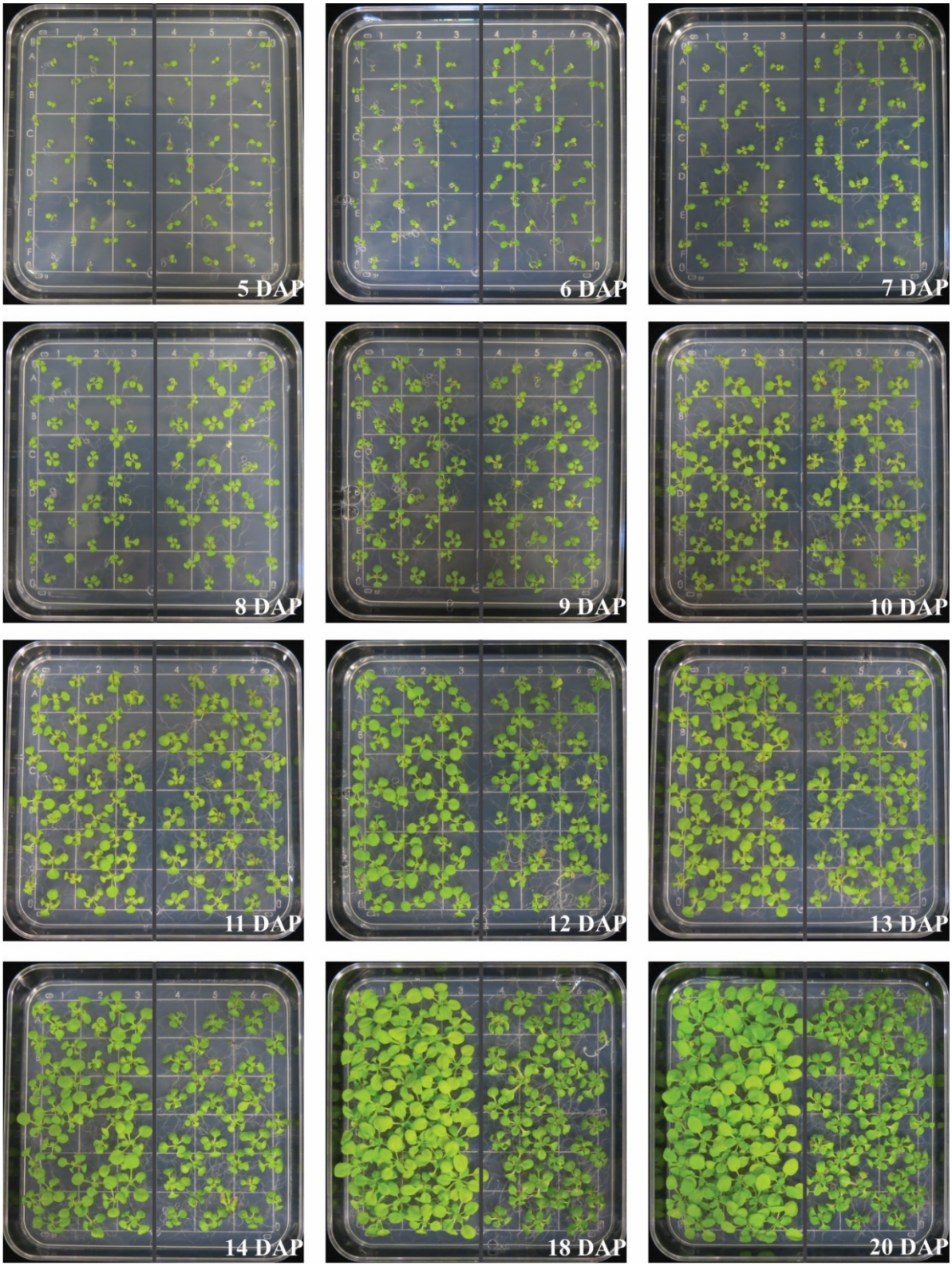


Figure 3.1. *ref8-1* and wild type growth phenotypes on MS plates over time.
Representative photographs of wild-type (left) and *ref8-1* (right) grown on MS plates for 5-14, 18 and 20 days.

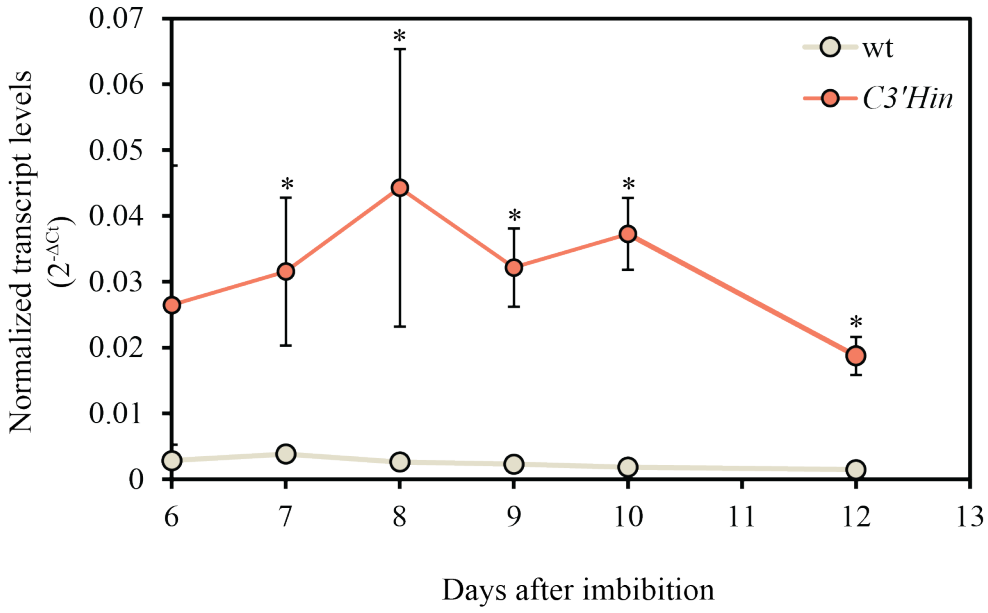


Figure 3.2. Expression of *NCED3* in plate-grown seedlings.
Relative expression measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (n=3). (*) Significantly different from wild type ($p < 0.05$), Student's *t-test.*

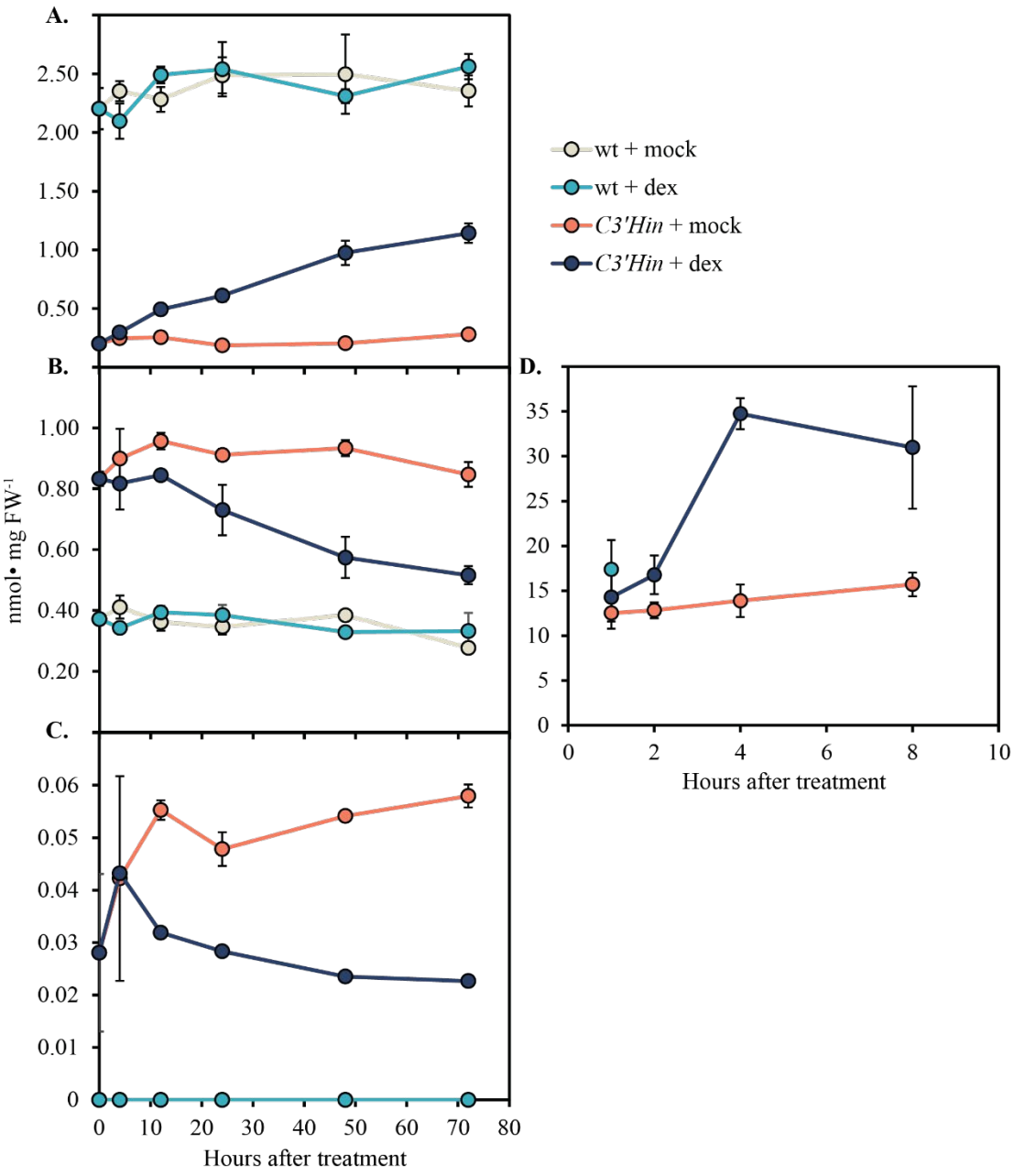


Figure 3.3. Soluble metabolites in *C3′Hin* and wild-type plants grown in the presence or absence of dexamethasone.
(A) Sinapoylmalate, (B) flavonoids (the sum of : kaempferol 3-O-[6″-O-(rhamnosyl) glucoside] 7-O-rhamnoside, kaempferol 3-O-glucoside 7-O-rhamnoside, and kaempferol 3-O-rhamnoside 7-O-rhamnoside), and (C) *p*-coumaroyl glucose content measured by HPLC. (D) Caffeoyl shikimate content quantified by LC/MS-MS. Error bars represent the standard deviation among biological replicates (*n*=3).

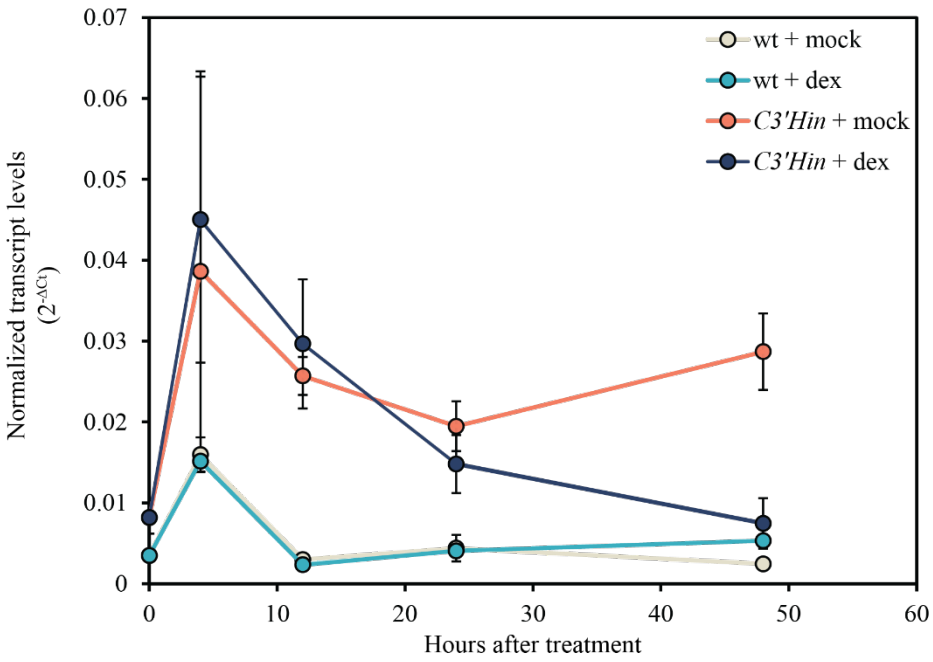


Figure 3.4. Expression of *NCED3* in *C3′Hin* and wild-type plants grown in the presence or absence of dexamethasone.
Relative expression measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (*n*=3).

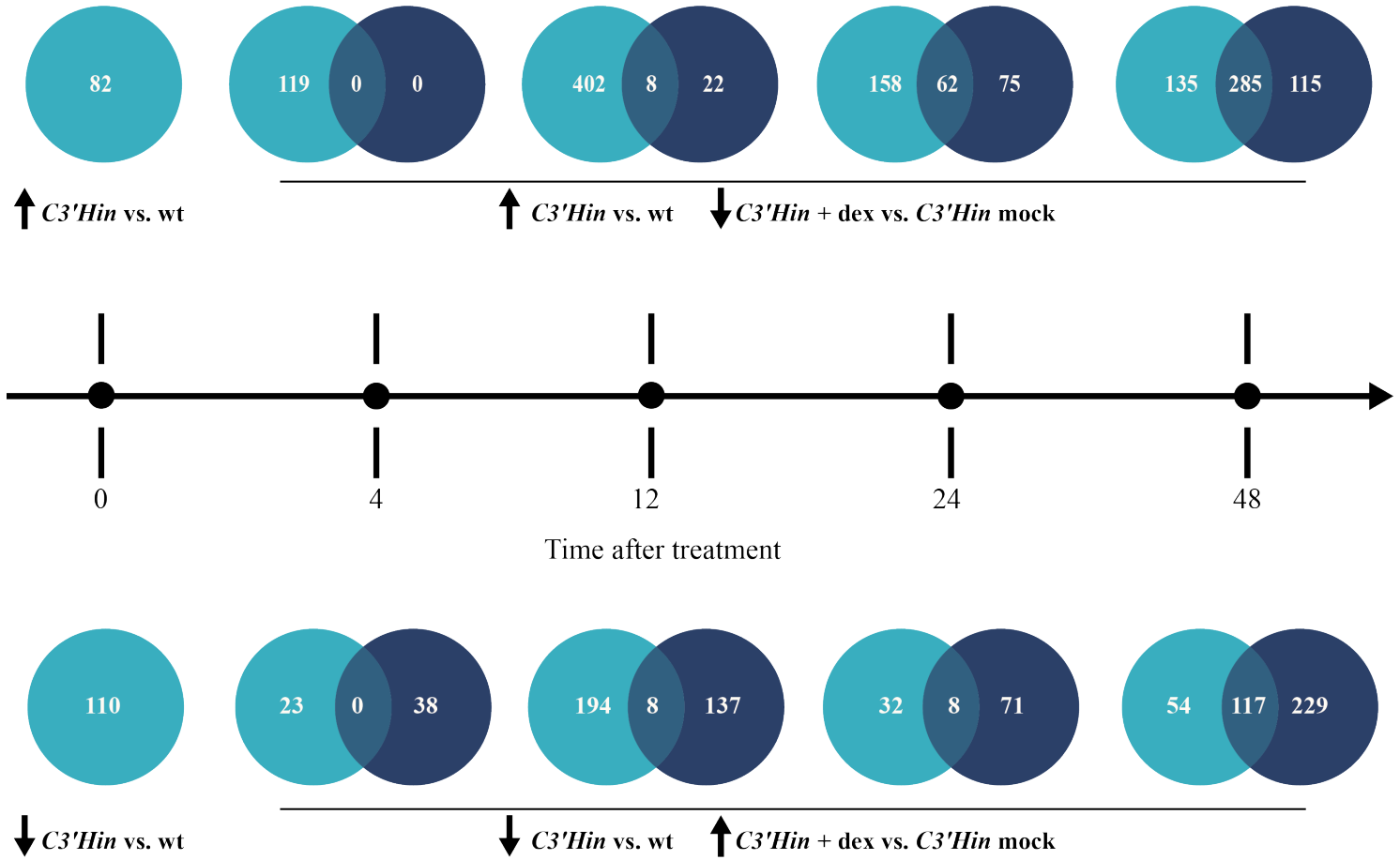


Figure 3.5. Venn diagrams depicting overlaps between differentially expressed genes over time.
FDR < 0.05, $|\log_2 FC| > 1$. Genotypic differences at each time point (light blue) were obtained by comparing *C3'Hin* mock-treated to mock-treated wild type. Treatment differences at each time point (dark blue) were obtained by comparing *C3'Hin* dex-treated to *C3'Hin* mock-treated and subtracting the transcripts that also changed in the same direction in wild type due to treatment. Arrows represent either upregulation or downregulation of gene expression.

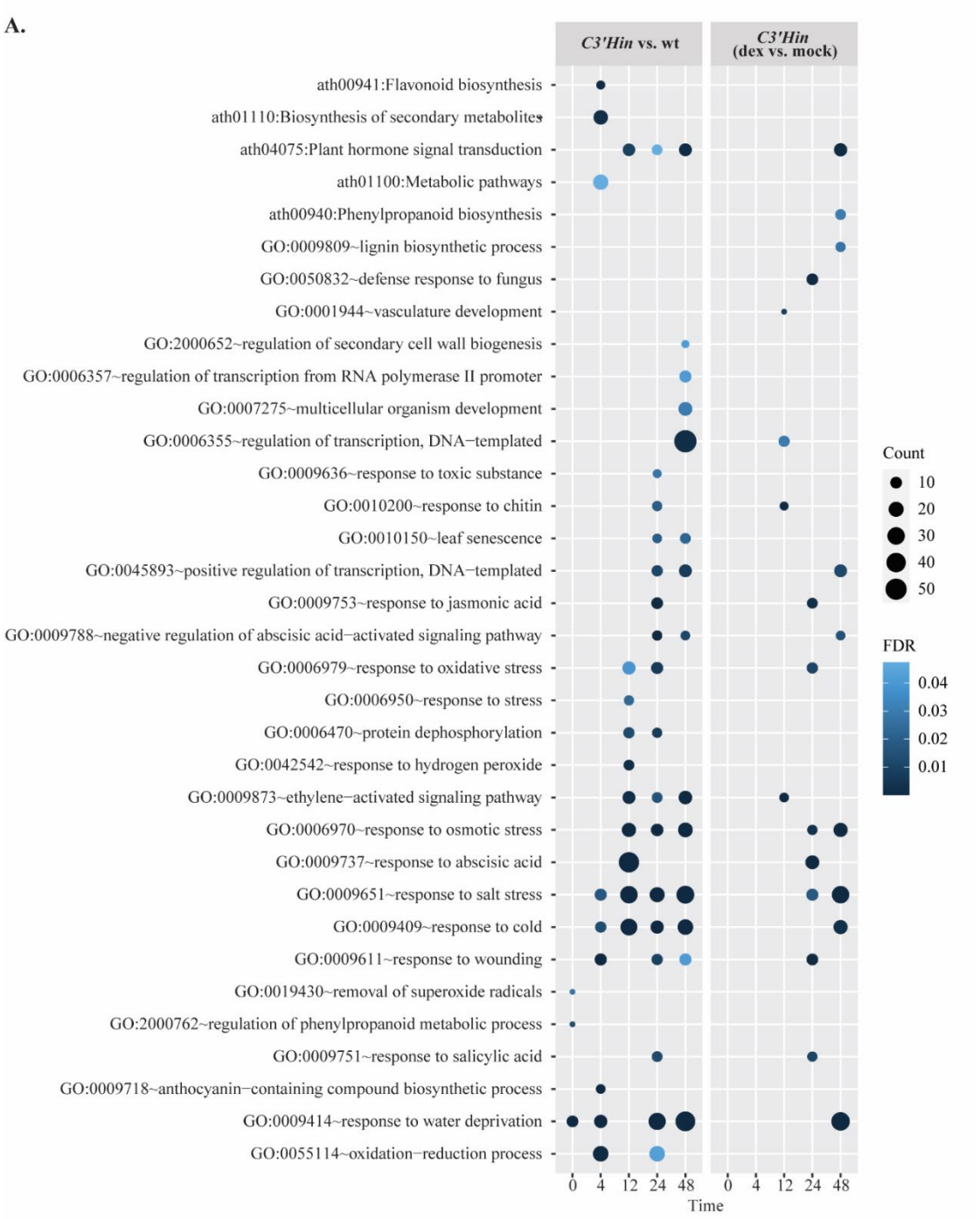
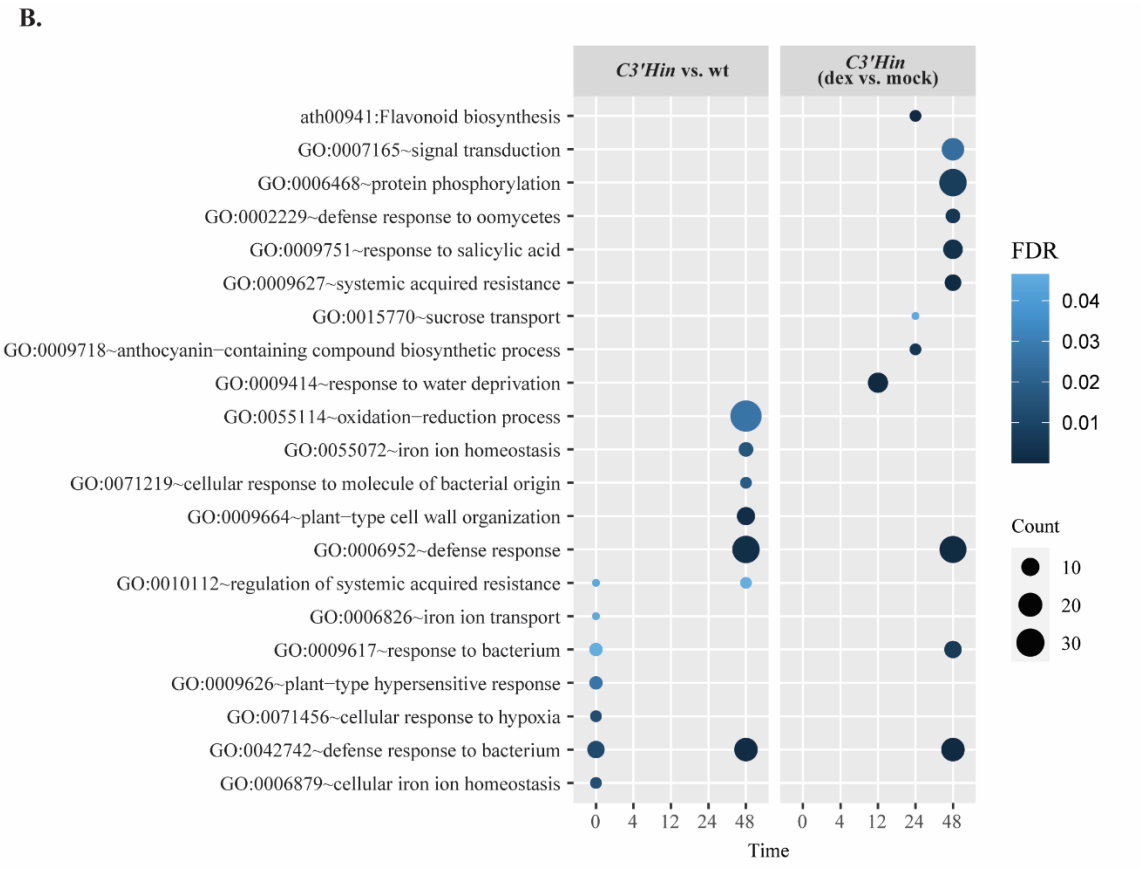


Figure 3.6. GO term and KEGG pathway enrichment analysis.
(A) Genes downregulated in *C3′Hin* compared to wild type and upregulated upon *C3′Hin* induction when compared to mock-treated plants. FDR < 0.05, $|\log_2 FC| > 1$. The size of the circle represents the number of DEGs that are associated with a particular term and the color represents the FDR value. (B) Genes upregulated in *C3′Hin* compared to wild type and downregulated upon *C3′Hin* induction when compared to mock-treated plants. FDR < 0.05, $|\log_2 FC| > 1$. The size of the circle represents the number of DEGs that are associated with a particular term and the color represents the FDR value of the GO term and KEGG enrichment analysis.

Figure 3.6 continued

B.



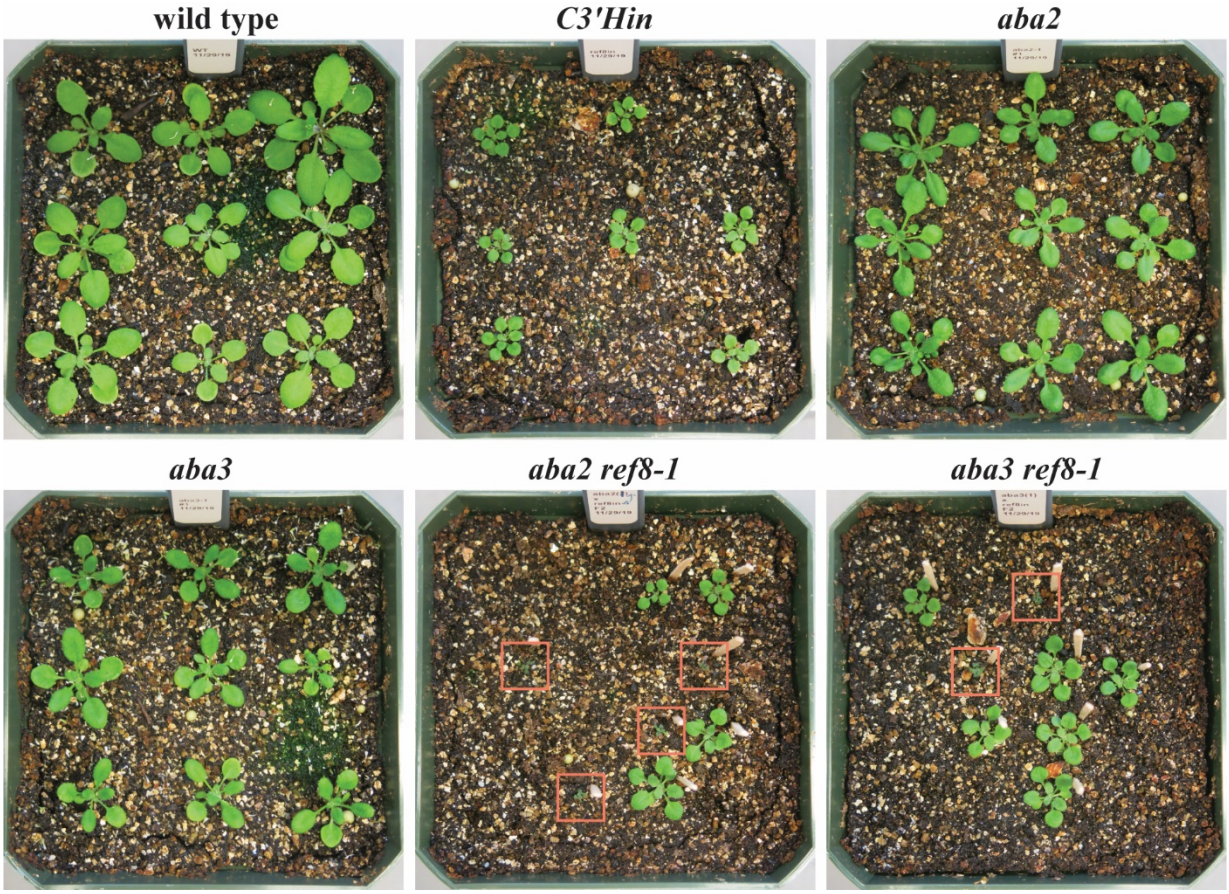


Figure 3.7. Representative photograph of 17 days-old *aba2 ref8-1* and *aba3 ref8-1* plants. Top left to bottom right: wild type, *C3'Hin*, *aba2, aba3, aba2 ref8-1, aba3 ref8-1.* Plants marked with squares are double homozygotes mutants identified in an F2 segregating population.

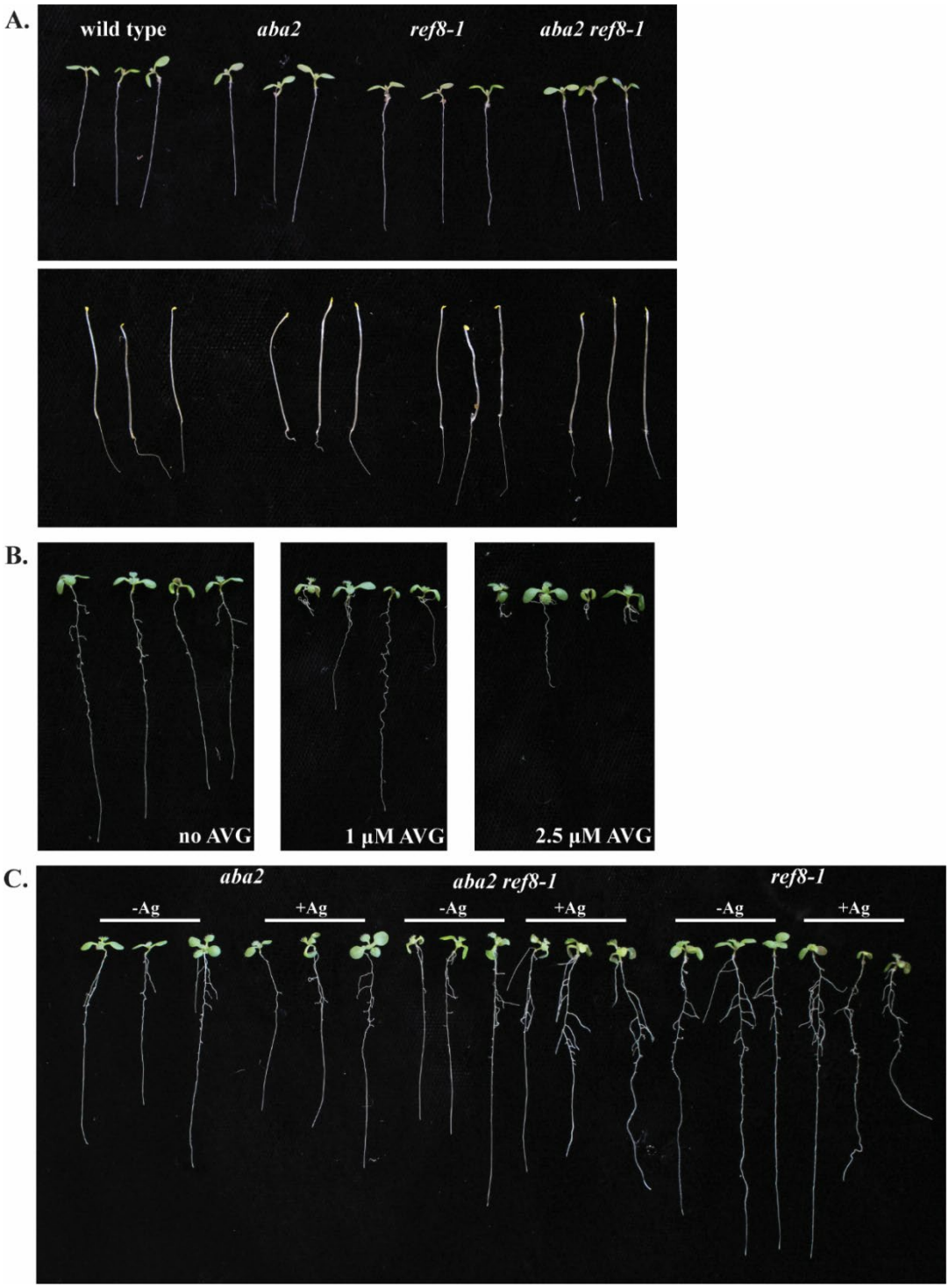


Figure 3.8. Ethylene-related phenotypes in seedlings.
(A) Five-day-old seedlings grown in 16 h light and 8 h dark (top) or grown under continuous dark (bottom). (B) Seven-days old seedlings grown in the presence of the ethylene synthesis inhibitor 2-aminoethoxyvinyl glycine (AVG). Left to right: grown without inhibitor, 1 µM of AVG and 2.5 µM of AVG. The order of genotypes is the same as in panel A. (C) Eight-day-old seedlings grown in the presence of 10 µM silver thiosulfate, an ethylene signaling inhibitor.

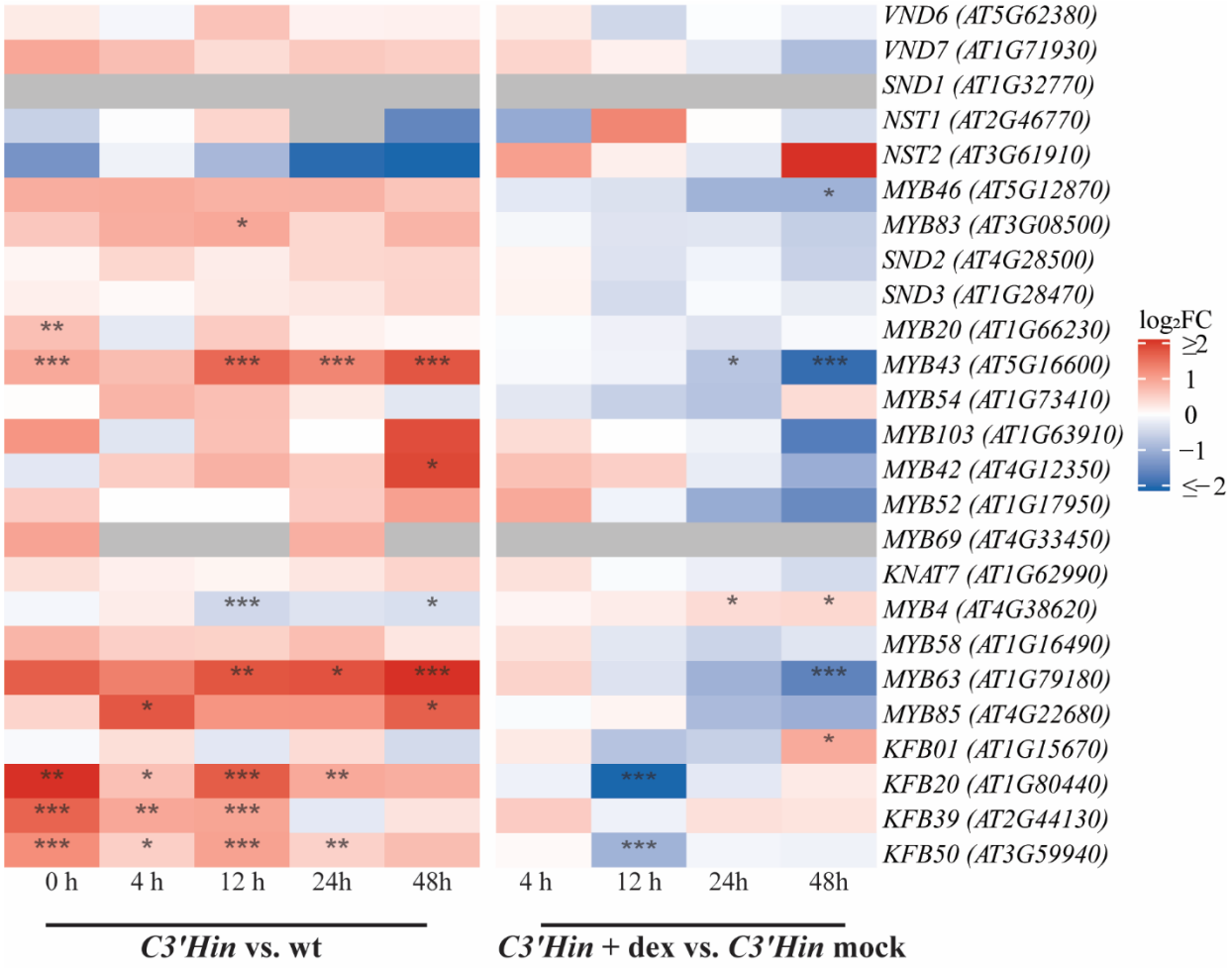
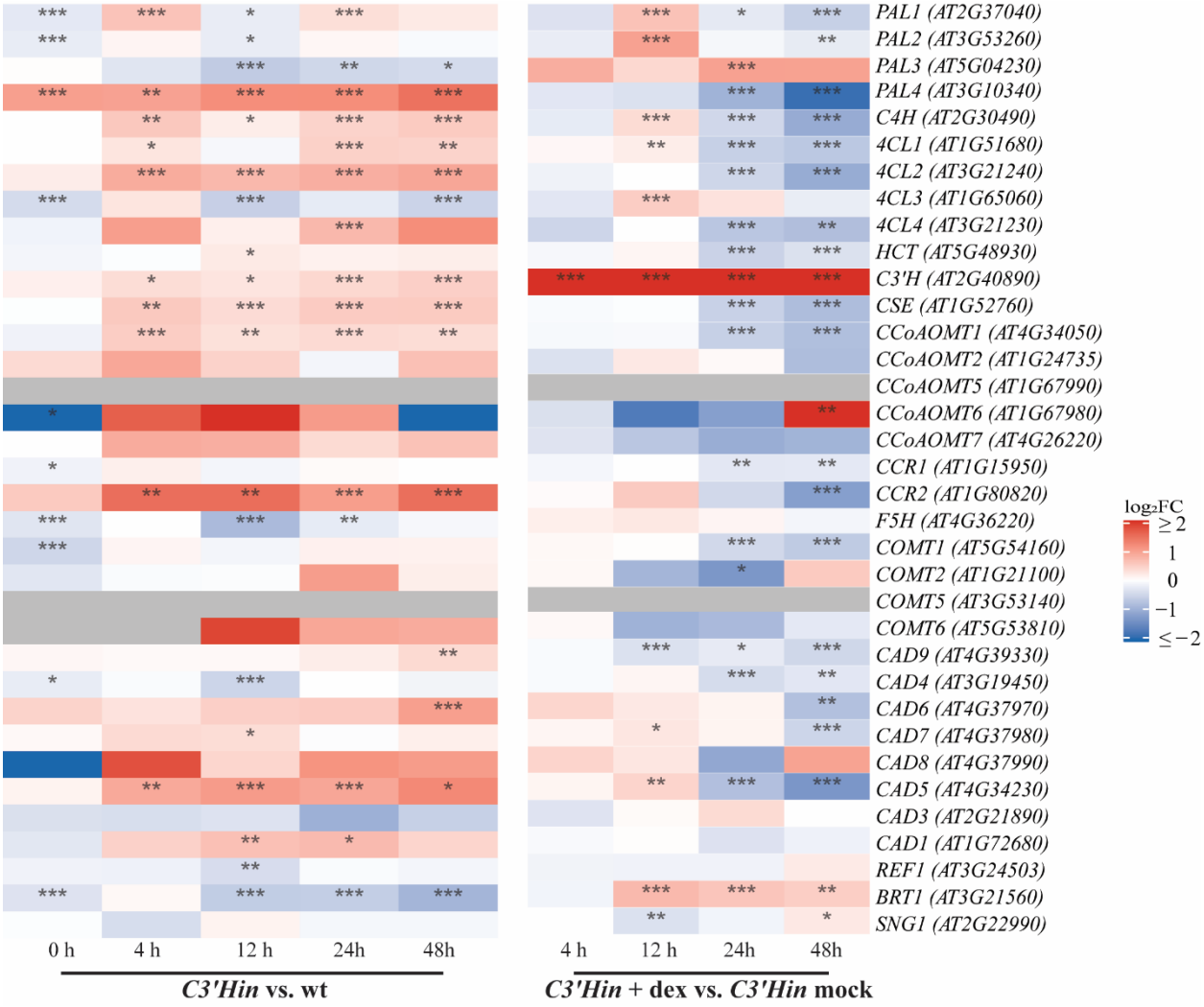


Figure 3.9. Heatmaps depicting log_2FC values for (A) phenylpropanoid regulatory genes and (B) phenylpropanoid genes.
Gray-colored boxes represent genes whose transcripts were not detected in our analysis. (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$.. Gray-colored boxes represent genes whose transcripts were not detected in our analysis. (*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$.

Figure 3.9 continued



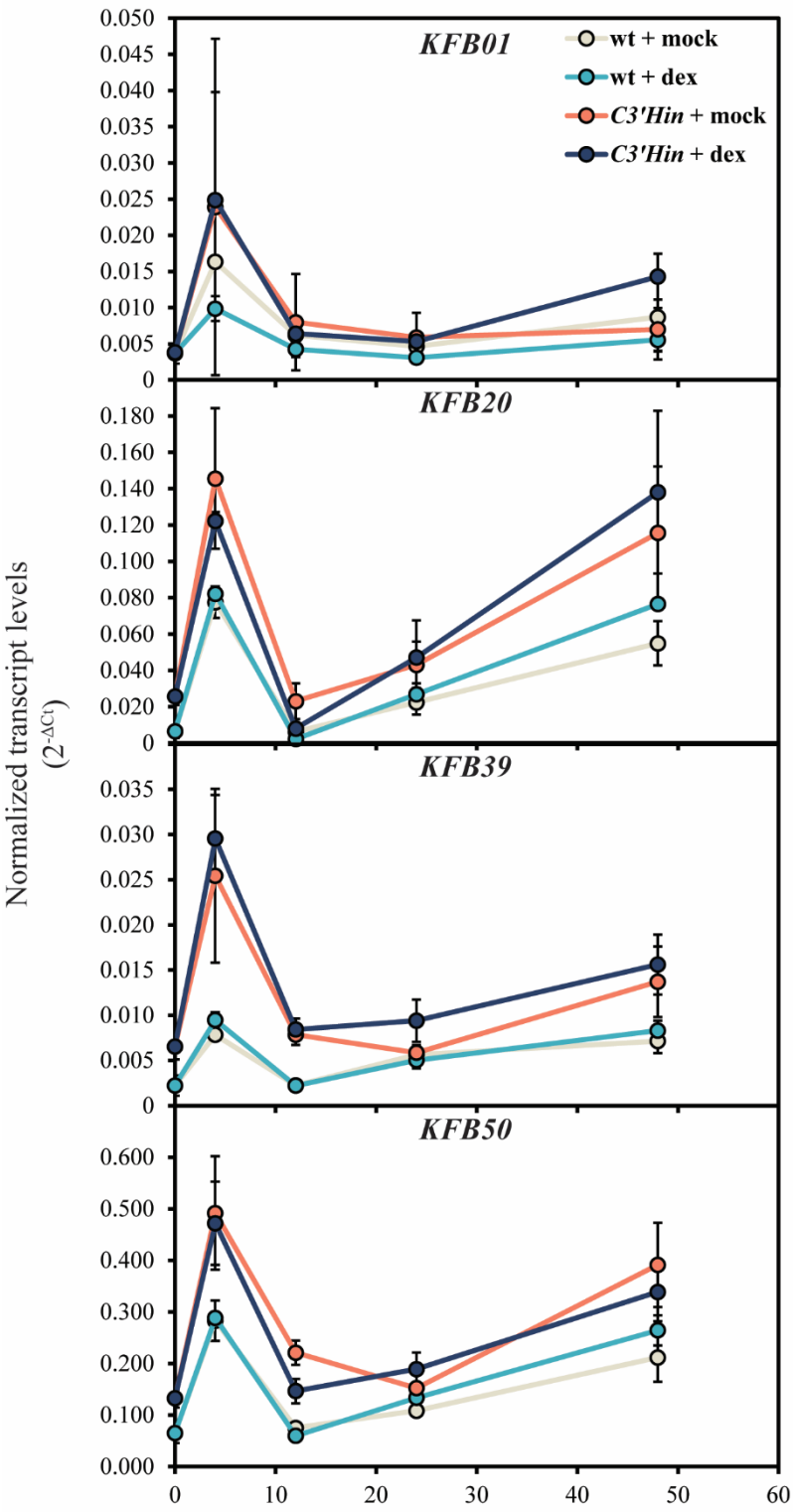


Figure 3.10. Expression of *KFB* genes in *C3′Hin* and wild-type plants grown in the presence or absence of dexamethasone.
Relative expression measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (*n*=3).

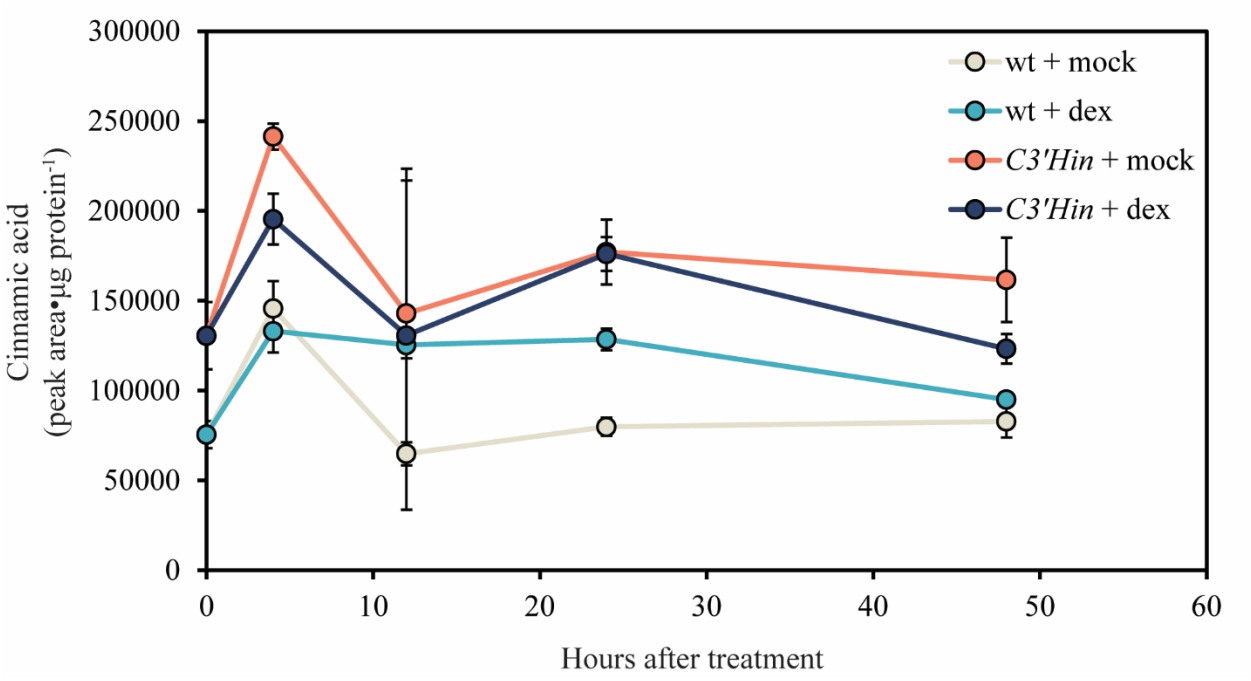


Figure 3.11. PAL activity in *C3´Hin* and wild-type plants grown in the presence or absence of dexamethasone.
Quantification of cinnamic acid generated in PAL assays using phenylalanine as substrate. Error bars represent standard deviation among biological replicates (*n*=3).

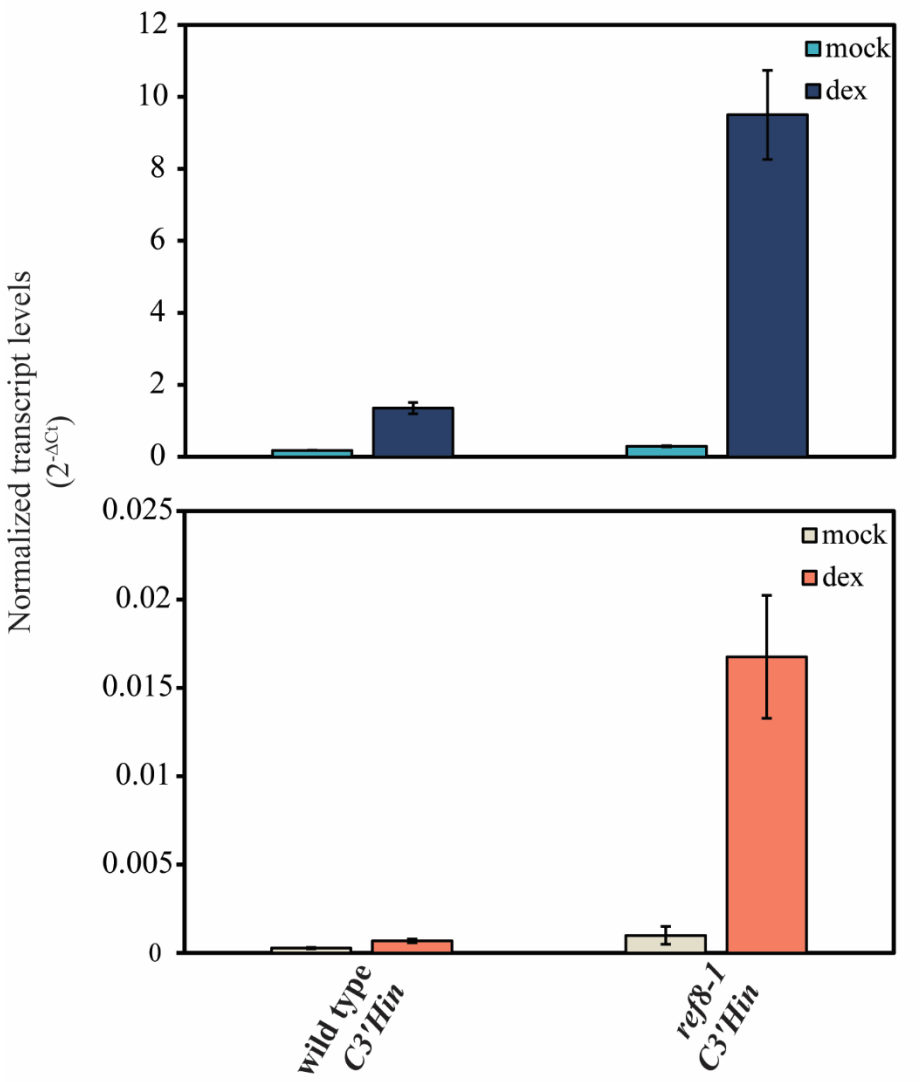


Figure 3.12. Expression of *C3'H* and *WOX14* genes in *C3´Hin* and wild-type plants grown in the presence or absence of dexamethasone 4 hours after treatment.
Relative expression of *C3'Hin* (top) and *WOX14* (bottom) measured by qRT-PCR, normalized to the reference gene At1g18780. Error bars represent standard deviation among biological replicates (*n*=3).

CHAPTER 4. A ROLE FOR CONIFERYL ALCOHOL IN PLANT GROWTH AND DEVELOPMENT

4.1 Abstract

The phenylpropanoid pathway gives rise to many specialized metabolites, but the majority of carbon flux is directed to the synthesis of lignin monomers. Lignin acts as an impediment in the extraction of carbohydrates from biomass, hence impairs the further deployment of crops for biofuel production. Significant efforts have been directed towards the manipulation of lignin content and composition as a means to decrease biomass recalcitrance to saccharification, unfortunately in many cases these manipulations come at the expense of plant growth. In this study, we tested the effect of simultaneously altering lignin composition and quantity on plant growth. To do so, we modulated the expression of ferulate 5-hydroxylase (F5H) in lignin-deficient backgrounds, to either accumulate coniferyl or sinapyl alcohol derived-lignin. Knocking out *F5H* in these mutant backgrounds had no effect on plant growth, however the overexpression of *F5H* led to a range of growth defects in all mutant backgrounds. This finding led us to the hypothesis that depletion of coniferyl alcohol by F5H or derivatives of this alcohol was responsible for the dwarf phenotype of these plants. We found that exogenous coniferyl alcohol, but not other monolignols rescued the growth of these plants. Furthermore, we identified putative coniferyl alcohol-derivatives that could be involved in this growth restoration.

4.2 Introduction

In a world with limited arable land undergoing exponential population growth it has become more important than ever to design plants with maximum productivity. The deployment of these crops should not just address the growing food demand, but rather have a multipurpose scope, meaning that their biomass could be used for the production of fibers, fuel and other chemicals. Furthermore, the development of these plants must also take into account expected rapid changes in climate. In order to accomplish these goals, it will be necessary to redirect our focus from the traditional end product approach taken in plant design, to a systems approach in which we assess how a given manipulation affects other metabolic pathways and ultimately impacts growth and development.

Regulation of plant growth and development is mediated by an intricate system that integrates internal and environmental signals. The regulatory role of the plant hormones: auxin, gibberellins, abscisic acid, cytokinins and ethylene, in this process has been studied extensively for decades [1]. The subsequent discovery of jasmonic acid, salicylic acid, brassinosteroids and strigolactones has expanded this list to include compounds that had previously been considered secondary metabolites [2–5]. Specialized or secondary metabolites are products of plant metabolism that vary widely between plant families and are often considered non-essential for plant survival [6]. This class of metabolites aids plants in resistance to biotic stressors and adapting to abiotic stress [6]. In addition, they serve as pigments and volatile compounds that provide a reproductive advantage and help ensure the next generation by attracting pollinators and seed dispersers [6]. Although specialized metabolites have been considered non-essential for plant survival, in addition to the four mentioned above, several additional secondary metabolites are now known to have roles in plant growth and development [7]. Some specialized metabolites have a direct impact on growth, while others are precursors in hormone synthesis or inhibit hormone signaling [7]. For example, phenylpyruvate is an intermediate in phenylalanine biosynthesis, which also has an effect in auxin levels because it acts as an amino acceptor in auxin biosynthesis[8]. Glucosinolates are specialized metabolites characteristic of the Brassicaceae family known for their insect deterrent properties [9]. Malinovsky *et al*. (2017) [10] demonstrated a previously unknown function of glucosinolates by showing that the aliphatic glucosinolate 3-hydroxypropyl-glucosinolate inhibits root meristem growth. Flavonoids are plant pigments derived from the phenylpropanoid pathway that can inhibit auxin transport [11] and are required for pollen development in maize and petunia [12]. Recently, they have been shown to also be involved in pollen tube growth under heat stress [13] and stomatal closure [14]. Another group of phenylpropanoids that may be involved in growth are dehydrodiconiferyl alcohol glucosides (DCGs), which are lignan glycosides that have cytokinin-like properties and promote cell-division in plant cell culture [15]. Based upon these observations, our previous definitions of hormones and secondary metabolites may have to be revisited given the importance of secondary metabolites in fine tuning growth and development.

The phenylpropanoid pathway gives rise to a variety of specialized metabolites, but the largest proportion of its flux goes towards the production of the hydroxycinnamyl alcohols, *p*-coumaryl, coniferyl and sinapyl alcohol. The latter serve as the building blocks of lignin, which is

a component of plant secondary cell walls that imparts hydrophobicity and tensile strength to vascular tissues [16,17]. Lignin content and composition have a large impact on sugar extraction from plant biomass, hence affecting the feasibility of paper and biofuel production [18]. Substantial efforts have been made to address this issue by targeting genes in the phenylpropanoid pathway to decrease lignin content [19,20]. Taking advantage of the plasticity in monolignol incorporation, researchers have also manipulated the phenylpropanoid pathway to alter the intrinsic monolignol ratios or to incorporate novel monomers that result in less recalcitrant lignocellulosic biomass [19,20]. Unfortunately, many of these efforts have resulted in dwarf plants, a phenomenon known as lignin modification induced dwarfism (LMID) [21,22]. This outcome often occurs when the genes that encode for the core enzymes in the phenylpropanoid pathway: phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate:CoA ligase (4CL), hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT), *p*-coumaroyl CoA 3′-hydroxylase (C3′H), caffeoyl shikimate esterase (CSE) and cinnamoyl-CoA reductase (CCR) are downregulated or knocked out [23–34]. On the other hand, manipulation of genes that encode enzymes further downstream in the pathway such as caffeoyl-CoA *O*-methyltransferase (CCoAOMT), ferulate 5-hydroxylase (F5H), caffeic acid *O*-methyltransferase (COMT), and cinnamyl alcohol dehydrogenase (CAD), mostly result in changes in lignin composition, but often do not affect plant growth [29,35–39].

LMID may be the result of a change in cell wall composition that leads to mechanical failure of the vascular tissues resulting in their inability to properly transport water [16,30]. This hypothesis is supported by the fact that many mutants with low lignin content exhibit collapsed xylem vessels [24,40,41]. Another hypothesis suggests that defects in the cell wall are detected by a cell wall integrity monitoring system that triggers a series of transcriptional changes which ultimately result in dwarfism [42]. Evidence for such a system exists for polysaccharide components of the cell wall [43–45]. The discovery that knocking out specific components of the Mediator complex rescues the growth of a lignin deficient mutant, may indicate this surveillance system is capable of detecting lignin-induced cell wall defects [46]. LMID has also been attributed to the hyper- or hypo- accumulation of phenylpropanoid metabolites that occurs when the pathway is blocked [22,24,47]. Elevated levels of the hormone salicylic acid (SA) have been shown to cause the dwarf phenotype of an *HCT* mutant [48], however, no such dwarfing occurs in another lignin-deficient mutant with high SA levels [46]. As mentioned earlier, DCGs, compounds derived from

monolignols, have hormone-like properties and manipulations that affect the synthesis of monolignols may also affect the levels of these compounds which in turn could affect plant growth [15]. To date, there is not enough evidence to determine which model(s) explains the LMID phenomenon.

F5H catalyzes the hydroxylation of coniferaldeyde and coniferyl alcohol to 5-hydroxy coniferaldehyde and 5-hydroxy coniferyl alcohol, respectively, which are subsequently methylated by COMT to sinapaldehyde and sinapyl alcohol, respectively. Once polymerized, coniferyl alcohol gives rise to guaiacyl (G) lignin and sinapyl alcohol forms syringyl (S) lignin. In *Arabidopsis thaliana*, knocking out *F5H* (*fah1-2*) shifts the native 3:1 G/S lignin ratio to pure G, and conversely, overexpression of *F5H* (*C4H:F5H*) redirects the flux towards nearly pure S lignin synthesis [49–53]. Neither of these modifications results in plant dwarfism in an otherwise wild-type background [49–53]. In contrast, dwarfism was observed when *F5H* was overexpressed in plants with mutant cinnamyl alcohol dehydrogenase (*CADC* and *CADD*) genes [54] The double mutant *cadc cadd* incorporates coniferaldehyde and sinapaldehyde into the lignin polymer instead of their corresponding alcohols [54]. In an effort to generate plants that incorporate either coniferaldehyde or sinapaldehyde into lignin, Anderson *et al.* (2015) [54] generated *cadc cadd fah1-2* and *cadc cadd fah1-2 C4H:F5H* plants. Knocking out *F5H* in *cadc cadd* has no effect on plant growth, but the overexpression of *F5H* in this genetic background results in severely stunted plants for reasons that remain to be determined [54].

In an attempt to assess the effects of simultaneously altering lignin quantity and composition on plant growth and development, we used the following low lignin mutants: *pal1 pal2, ref3-2* and *ccr1* which are deficient in *PAL1* and *PAL2*, *C4H* and *CCR1* respectively [24,31,32]. In order to modify the monomer ratios, we modulated the expression of F5H in all three genetic backgrounds. Knocking out *F5H* had no effect on the growth phenotype of *pal1 pal2*, *ref3-2* and *ccr1*, whereas the overexpression of *F5H* resulted in dwarf plants in all cases. These data led us to hypothesize that the overexpression of *F5H* is driving flux away from a compound derived from coniferaldehyde or coniferyl alcohol that is necessary for normal growth and the effect can only be observed in the sensitized backgrounds.

4.3 Methods

4.3.1 Plant Materials

Arabidopsis thaliana, ecotype Columbia-0 plants were grown in Propagation Mix (Sun Gro Horticulture) supplemented with Osmocote® Plus (ICL Specialty Fertilizers) at 22°C under long day conditions (16 h light/ 8 h dark). The *ref3-2, ref3-3, pal1 pal2, fah1-2 C4H:F5H* and *cadc cadd fah1-2 C4H:F5H* plants have been described previously [24,32,49,54]. The *ccr1* T-DNA line was obtained from the Arabidopsis Biological Resource Center (The Ohio State University, Columbus, OH). The *ref3-2, ref3-3, pal1 pal2,* and *ccr1* plants were crossed to *fah1-2 C4H:F5H* to generate *ref3-2 fah1-2, ref3-2 fah1-2 C4H:F5H, ref3-3 fah1-2, ref3-3 fah1-2 C4H:F5H, pal1 pal2 fah1-2, pal1 pal2 fah1-2 C4H:F5H, ccr1 fah1-2* and *ccr1 fah1-2 C4H:F5H*. Homozygous mutants were identified by PCR amplification using primers previously described.

4.3.2 Generation of inducible lines

F5H and *COMT* were amplified from cDNA using primers described on in Table 4.1 and cloned into the *pOpON* vector (dexamethasone inducible) using the Gateway® cloning system to generate *pOpOn:COMT* (*COMTin*) and *pOpOn:F5H* (*F5Hin*) [55]. *F5Hin* was introduced into *ref3-2 fah1-2, ref3-3 fah1-2 and cadc cadd fah1-2* and *COMTin* into *ccoaomt comt*, through *Agrobacterium tumefaciens*-mediated transformation [56]. Transformants were selected in the first generation (T1) on MS plates with kanamycin and/or by screening for sinapoylmalate production under UV light after application of 20 µM dexamethasone. In order to make *ref3-2 fah1-2 C4H:F5H* with inducible *C4H,* an already available *ref3-2 pOpOn:C4H* (*C4Hin*) plant was crossed to *ref3-2/+ fah1-2 C4H:F5H/-* and *ref3-2 fah1-2 C4H:F5H C4Hin* plants were identified in the F1 generation by UV screening after induction and genotyping by PCR.

4.3.3 Chemical complementation

Columbia-0 (wild type), *ref3-2/+ fah1-2 C4H:F5H, cadc/+ cadd fah1-2 C4H:F5H* seeds were surface sterilized and sown on plates containing Murashige and Skoog (MS) medium, 1% sucrose and 0.7% agar. Media was supplemented with either 200 µM coniferaldehyde, *p*-coumaryl, coniferyl, sinapyl alcohol or 0.07% dimethyl sulfoxide (DMSO).

4.3.4 HPLC metabolite analysis

Metabolites were extracted in 50% methanol (10 µL mg^{-1} fresh weight) by incubating 2 h at 65°C. The supernatant was centrifuged at 14,500 rpm for 1 hr and 10 µL samples were injected onto a Shim-pack XR-ODS column (Shimadzu; 3.0 mm × 75 mm, 2.2 µm packing particle size) and metabolites were separated using chromatographic conditions previously described [57].

4.3.5 Feeding and LC/MS metabolite analysis

Seeds from *cadc/+ cadd fah1-2 C4H:F5H* were sown on MS medium, 1% sucrose and 0.7% agar and stratified for two days. Twelve days post-stratification, homozygous *cadc cadd fah1-2 C4H:F5H* seedlings were transferred to multiwell cell culture plates with MS medium and 1% sucrose for 2 days. The medium was then replaced with MS medium, 1% sucrose supplemented with either 200 µM [$^{12}C_6$] coniferyl alcohol, [$^{13}C_6$] coniferyl alcohol or 0.07% DMSO for 24 and 48 hours. Metabolites soluble in 50% methanol were extracted as described above, dried in a speed-vac and re-dissolved in 50% methanol and analyzed by LC/MS.

4.4 Results

4.4.1 Overexpression of *F5H* in *ref3-2* leads to severe dwarfism

Lignin engineering often focuses on either decreasing lignin content or changing lignin composition. In order to assess the effects on plant growth and development of simultaneously altering both lignin properties, we modulated the expression of *F5H* in a low lignin mutant deficient in *C4H* (*ref3-2*). C4H catalyzes the second step in the phenylpropanoid pathway, the hydroxylation of cinnamic acid to *p*-coumaric acid. Plants carrying certain missense alleles in *C4H* have modest growth defects [24], (Figure 4.1) although null alleles are seedling lethal. The lignin content of two of the stronger alleles is reduced from 45% to 80% [24]. Knocking out *F5H* had no effect on the growth phenotype of *ref3-2*, whereas the overexpression of *F5H* resulted in dwarf plants (Figure 4.1). It is noteworthy that in a wild-type background neither of these manipulations of F5H have an impact on growth.

4.4.2 *ref3-2 C4H:F5H* hyper-accumulate cinnamate-derived compounds and glucosinolates

In order to test the effect of manipulating *F5H* expression in *ref3*-2 on phenylpropanoids and other metabolites, we performed untargeted LC/MS. The reduction in C4H activity in *ref3-2* leads to the accumulation of cinnamic acid derivatives not found in wild-type plants [24] namely cinnamoyl aspartate, cinnamoyl glutamate and cinnamoyl malate (Figure 4.2A). LC/MS analysis indicated that *ref3-2 fah1-2 C4H:F5H* leaves have higher levels of cinnamoylated compounds in comparison to either *ref3-2* or *ref3-2 fah1-2* (Figure 4.2A). Because of its reduced leaf expansion, the metabolite content of *ref3-2 fah1-2 C4H:F5H* may seem higher on a fresh weight basis [58], we examined the relative content of other metabolites. The increase in cinnamic acid esters in *ref3-2 fah1-2 C4H:F5H* was not accompanied with an overall increase in other hydroxycinnamic acid esters, since sinapoylmalate levels in *ref3-2 fah1-2 C4H:F5H* are the same as that of *ref3-2* (Figure 4.2A). These results suggest that the increase in cinnamoylated compounds may be due to an actual increase in the availability of cinnamic acid.

In addition to cinnamoylated compounds, LC/MS analysis revealed that among the soluble compounds identified in our dataset, glucosinolate content varied widely between genotypes. Glucosinolates are sulfur-containing metabolites that serve in plant defense [9]. The amino acid precursors of glucosinolates determine their class; aliphatic glucosinolates are derived from alanine, leucine, isoleucine, valine or methionine, aromatic glucosinolates from phenylalanine and tyrosine and indole glucosinolates are derived from tryptophan [9]. Although glucosinolates and phenylpropanoids are not synthesized by the same pathways, significant cross-talk between these two pathways has been reported [59]. LC/MS analysis of *ref3-2* leaves revealed higher levels of the indole glucosinolates indol-3-ylmethyl glucosinolate (I3M-GS) and 1-methoxy-indol-3-methyl glucosinolate and/or 4-methoxy-indol-3-methyl glucosinolate (1MeO-I3M-GS/4MeO-I3M-GS) than wild-type plants (Figure 4.2B). Furthermore, *ref3-2 fah1-2 C4H:F5H* plants accumulate even higher levels of glucosinolates, reaching ~10 times that of wild-type glucosinolate levels (Figure 4.2B). It is yet to be determined if the hyperaccumulation of glucosinolates in *ref3-2 fah1-2 C4H:F5H* contributes to its severe dwarf phenotype.

4.4.3 Multiple mutants in phenylpropanoid pathway genes overexpressing C4H:F5H exhibit severe growth impairment

The severe growth defect of *ref3-2 fah1-2 C4H:F5H* is not the first instance of dwarfing due to *F5H* overexpression. Anderson *et al*. (2015) [54] also obtained dwarf plants when *C4H:F5H* was introduced into a *cadd* genetic background and the effect was exacerbated when a *cadc cadd* background was utilized. At the time, it was suggested that the severe dwarfism of *cadc cadd C4H:F5H* might be due to its sinapaldehyde-rich lignin composition [54]. In contrast, *ref3-2 fah1-2 C4H:F5H* should deposit sinapyl alcohol-derived S-lignin, which is not detrimental for growth [49], therefore it seems unlikely that lignin composition alone can be the cause of dwarfism in *ref3-2 fah1-2 C4H:F5H*. In order to test if the effect of *F5H* overexpression is unique to these genetic backgrounds or it is an outcome in all low lignin mutants, we introduced *C4H:F5H* into mutants of phenylalanine ammonia lyase isoforms 1 and 2 (*pal1 pal2*), and cinnamoyl-CoA reductase (*ccr1*) [31,32]. The rationale for selecting these mutants was that they do not have severe growth defects and their monolignol ratios are the same as that of wild-type plants [31,32]. PAL is at the entry point of the phenylpropanoid pathway, hence the *pal1 pal2* mutant provides a system where overall carbon flux to the pathway is reduced. CCR1 catalyzes the reduction of cinnamoyl-CoA esters to cinnamaldehydes, the precursors for monolignols, thus allowing us to test the impact of the reduction of flux through the pathway further downstream. Knocking out *F5H* had no effect on the growth phenotype of *ccr1* and *pal1 pal2*, whereas the overexpression of *F5H* resulted in dwarf plants in all cases (Figure 4.1). It is noteworthy, that the effect in *pal1 pal2* was not as severe as that of the other mutant backgrounds (Figure 4.1C).

4.4.4 Exogenous coniferyl alcohol rescues the dwarf phenotype of *ref3-2 C4H:F5H* and *cadc cadd C4H:F5H*

The results obtained above indicate that *F5H* overexpression in mutants with reduced phenylpropanoids leads to dwarfing. Besides the re-direction of flux from G- to S- monomers, a common factor among these plants may be that coniferaldehyde and coniferyl alcohol pools are already low and are further depleted by *F5H* overexpression. Moreover, the growth of plants that cannot synthesize coniferaldehyde such as the *caffeic acid O-methyltransferase* (*comt*) *caffeoyl CoA 3-O-methyl transferase* (*ccoaomt*) double mutant stalls at the cotyledon stage [60]. These

observations led us to the hypothesis that the depletion of coniferyl alcohol or derivatives thereof, causes the dwarfism of *ref3-2, ccr1, pal1 pal2* and *cadc cadd* carrying the *C4H:F5H* construct. To test this hypothesis, we planted *cadc cadd fah1-2 C4H:F5H* in media supplemented with coniferyl alcohol and coniferaldehyde as a control with similar structure. As seen in Figure 4.3, the growth of wild-type seedlings is negatively affected by the coniferaldehyde treatment and the same was observed in both mutants carrying *C4H:F5H*. In contrast, wild-type seedlings grown in coniferyl alcohol grow similarly to the mock treated seedlings, whereas *cadc cadd fah1-2 C4H:F5H* seedling grew significantly larger than controls.

To test if the effect of coniferyl alcohol on the growth of *C4H:F5H* plants was specific or could be attained by supplementing the media with any monolignol, we planted *ref3-2 fah1-2 C4H:F5H* and *cadc cadd fah1-2 C4H:F5H* on media containing either *p*-coumaryl alcohol or sinapyl alcohol. Figure 4.4 shows that the positive effect on growth is unique to coniferyl alcohol.

4.4.5 Coniferyl alcohol-derived compounds in *cadc cadd C4H:F5H*

In order to identify the coniferyl alcohol-derived compound(s) that restore growth in *cadc cadd fah1-2 C4H:F5H* we fed homozygous seedlings in liquid media with [$^{12}C_6$] or [$^{13}C_6$] coniferyl alcohol to identify coniferyl-alcohol derived molecules and compared these results with an untreated control. This analysis was performed on an automated pipeline developed in-house that uses raw LC/MS data from samples fed with labeled and unlabeled precursors (Simpson, unpublished). The algorithm is then capable of identifying metabolite features with the same retention time but that have a pre-determined mass shift between the labeled and unlabeled samples and in this way identify metabolite features that are derived from a particular precursor, in this case coniferyl alcohol. In the previous experiment plants were grown continuously on media containing coniferyl alcohol, while in the labeling experiments plants were transferred to media containing coniferyl alcohol for 24 and 48 hours. Growth differences were visible 24 hours after coniferyl alcohol treatment and by 48 hours the majority of *cadc cadd fah1-2 C4H:F5H* plants grown in coniferyl alcohol had developed a new leaf (data not included). This experiment allowed for the identification of 510 metabolite features that are derived from coniferyl alcohol (Figure 4.5). By performing pairwise analysis between the untreated and coniferyl alcohol treated samples we identified 10 coniferyl alcohol-derived metabolite features (FDR < 0.05) that accumulated at higher levels compared to the mock treated samples in both time points (Figure 4.5). Comparisons

with in-house datasets and published data [70], provided putative identities for some of the metabolite features and as seen in Table 4.2, we identified fragments corresponding to the same molecule in our analysis. Among the identified candidates we were able to find diglycosylated coniferyl alcohol, a coniferyl alcohol dimer and a glycosylated dimer of coniferyl alcohol (Table 4.2, Figure 4.6).

4.4.6 Inducible *F5H* overexpression leads to dwarfism in *ref3* and *cadc cadd*

The phenotypes obtained by the overexpression of *F5H* in multiple phenylpropanoid mutants and the fact that these phenotypes can be alleviated by providing exogenous coniferyl alcohol led us to the hypothesis that it is a phenomenon with a common basis. The severe phenotype of these plants complicates mechanistic studies, therefore we generated genetic material in which the dwarfing process was conditional. By generating transgenic lines with inducible promoters we expected to capture the changes in metabolism and gene transcription leading to, or interfering with, normal growth. In order to capture the initial changes in metabolism and gene transcription that allow for normal growth, we generated dexamethasone inducible lines in which plants with compromised growth could grow normally. As mentioned earlier, the *ccoaomt comt* mutant exhibits severely impaired growth and it is blocked in the synthesis of coniferyl alcohol. A functional copy of *COMT* under the control of an inducible promoter was introduced into the *ccoaomt comt* mutant in order to restore its growth. In parallel, we tested the hypothesis that by adding a wild-type copy of *C4H* under the control of an inducible promoter to *ref3-2 fah1-2 C4H:F5H* we would also be able to restore the growth of these plants. As seen in Figures 4.7A and 4.8A the induction of *C4H* or *COMT* in their respective mutant backgrounds led to only mild restoration in growth.

As a complementary approach, we used genetic backgrounds that do not exhibit severe growth defects, but in which overexpression of *F5H* leads to dwarfing, namely *cadc cadd fah1-2* and *ref3-2 fah1-2.* We introduced *F5H* under the control of an inducible promoter in an attempt to induce dwarfing. In these experiments, the stunted growth phenotype of *cadc cadd fah1-2 C4H:F5H* was robustly recapitulated by the inducible lines upon dexamethasone treatment (Figure 7B and 10). The effect on plant growth was severe in seedlings grown on media supplemented with dexamethasone (Figures 4.7B). In soil-grown plants the effect was apparent by the third day in newly developing leaves, overtime, leaves which expanded prior to dexamethasone treatment

exhibited curling of their petioles and stalled development (Figure 4.10). The *ref3-2 fah1-2 F5Hin* line displayed partial dwarfing upon continuous application of dexamethasone (Figure 4.9). Figure 4.9, shows *F5H* induction in *ref3-2* had its strongest negative effect in stem height. Unfortunately, the *ref3-2* mutation in this line causes it to be infertile and as a result, homozygous mutants are not readily identifiable. All experiments with this line must be conducted with segregating populations, limiting the types of experiments that can be conducted. To circumvent this problem, we generated an F5H-inducible line using a weaker mutant allele of *C4H* (*ref3-3*) as a background. These plants are fertile and also exhibit impaired growth upon dexamethasone treatment (Figure 4.7C).

4.5 Discussion

Reports on coniferyl alcohol-derived molecules affecting plant growth date back to the 1970s [61]; however, there are only isolated reports on this topic and they provide limited data to support the role of these molecules in plant growth. Furthermore, if this phenomenon is true, it is unknown how these molecules are synthesized, sensed and how widespread the role of these growth regulators is among plant species. The two derivatives of coniferyl alcohol that have been reported as having hormone-like properties are dihydroconiferyl alcohol and dehydrodiconiferyl alcohol [15,61–65]. Dihydroconiferyl alcohol (DCA) extracted from lettuce cotyledons acts synergistically with giberellic acid to promote hypocotyl elongation [61]. DCA has also been extracted from maple syrup and shown to have a cytokinin-like effect by promoting cell division in soybean callus [62]. Dehydrodiconiferyl alcohol glucosides (DCGs) from Vinca rosea promote cell division and are found in higher concentrations in tissues undergoing rapid cell division compared to tissues of slower growth [15,63].

Work on this area has been impeded by the interconnectivity of these molecules' synthesis and that of lignin biosynthesis, because genetic manipulations that alter the content of these molecules' precursor also affects lignin levels. As a consequence, detrimental effects on plant growth arising from genetic manipulations of the phenylpropanoid pathway are often ascribed to alterations in lignin content alone, and this may have obscured the role of DCA and DCGs as plant growth regulators. In an attempt to study the relationship between lignin content and composition, we discovered that the overexpression of *F5H* in lignin-deficient mutants, namely *ref3-2, cadc cadd, ccr1*, and *pal1 pal2* leads to a range of dwarf phenotypes. This result was unexpected due to

the absence of growth phenotypes when *F5H* is overexpressed in wild-type *Arabidopsis thaliana* and other plant species [49–53]. Besides the overexpression of *F5H* may be able to deplete coniferyl alcohol pools, one common factor among these plants is that they also have a metabolic block that interferes with coniferyl alcohol synthesis. Our results and previous reports on growth-promoting activities of coniferyl alcohol-derived molecules [15,61–65] led us to test the hypothesis that the growth deficiency in the *F5H* overexpressing plants is caused by the depletion of coniferyl alcohol and/or derivatives thereof.

4.5.1 Growth deficiency can be induced by F5H overexpression in phenylpropanoid mutants

As mentioned earlier, plants overexpressing F5H deposit primarily S lignin and do not exhibit a negative growth phenotype [49–53]. Nevertheless, it is possible that the compromised growth of *ref3-2, cadc cadd, ccr1,* and *pal1 pal2* plants overexpressing *F5H* may be the result of a combination decreased lignin content together with high S lignin deposition that may have a combinatorial effect that leads to defective secondary cell walls and dwarfism. In this scenario, the amount of lignin content in each of these mutants should correlate to the severity of their growth phenotype. Previous reports show that *cadc cadd, ccr1, pal1 pal2* and *ref3-2*, retain 70%, 63%, 35%, and 20% of wild-type lignin content, respectively [24,31,32,54]. As shown in Figure 4.1, the most severe phenotypes by the addition of *C4H:F5H*, were obtained in *ref3-2* and *cadc cadd* genetic backgrounds. These plants are on opposite ends of the spectrum in lignin reduction, therefore the reduction in total lignin content does not correlate well with the severity of the growth phenotype observed. Altogether, these data suggest that the dwarf phenotype in *cadc cadd, ccr1, pal1 pal2* and *ref3-2* plants overexpressing *F5H* is due to other alterations in phenylpropanoid metabolism, but not necessarily lignin.

The phenylpropanoid pathway is under high regulation at the transcriptional and posttranslational level. Therefore, in addition to considering the redirections in phenylpropanoid flux by the manipulation of enzymes in the pathway, interpretation of the results obtained by overexpressing *F5H* in *cadc cadd, ccr1, pal1 pal2* and *ref3*-2 plants requires the consideration of regulatory elements that could be contributing to the results. For example, the increase in cinnamic acid esters in *ref3-2 fah1-2 C4H:F5H* in comparison to *ref3-2* (Figure 4.2A) may be indicative of an increase in PAL activity arising from the overexpression of *F5H*. Consistent with this

observation, RNA-seq analysis of *C4H:F5H* plants (Wang, thesis) show a decrease in the transcript levels of *KFB01* and *KFB39* which encode Kelch domain F-box proteins involved in PAL ubiqutin-mediated degradation [66,67]. The *pal1 pal2 fah1-2 C4H:F5H* mutant exhibited the mildest effect in growth of the mutant series here presented (Figure 4.1C). Although *pal1pal2* mutants display perturbations in phenylpropanoid metabolism such as a *transparent testa* phenotype, PAL3 and PAL4 can partially sustain phenylpropanoid biosynthesis, as *pal1 pal2* mutants retain ~30% of PAL activity [32]. Together with the expression of PAL3 and PAL4, a decrease in PAL degradation in *pal1 pal2 fah1-2 C4H:F5H* may explain why these plants exhibit only mild effects in growth when *F5H* is overexpressed. PAL activity assays would be necessary in the future to test this hypothesis.

4.5.2 Coniferyl alcohol rescues the growth of plants carrying *C4H:F5H*

CCoAOMT catalyzes the methylation of caffeoyl CoA to feruloyl CoA, whereas COMT catalyzes the methylation of 5-hydroxyconiferyl aldehyde and 5-hydroxyconiferyl alcohol to sinapaldehyde and sinapyl alcohol, respectively. Mutants in *comt* or *ccoaomt* do not exhibit growth defects, but knocking out or downregulating *comt* and *ccoaomt* simultaneously leads to severe growth defects in *Arabidopsis* and *Medicago* [60,68]. Although feruloyl CoA is the precursor for G and S monolignols, *ccoaomt* plants can still synthesize both of these lignin precursors [60]. Biochemical characterization of COMT suggests that in the absence of CCoAOMT, COMT can compensate for its catalytic activity and in this way sustain coniferyl and sinapyl alcohol synthesis [69]. In addition to exhibiting impaired growth, the *ccoaomt comt* mutant and *ref3-2, cadc cadd, ccr1*, and *pal1 pal2* overexpressing *F5H* share as a common factor a block above coniferyl alcohol synthesis. As mentioned earlier, dihydroconiferyl alcohol and dehydrodiconiferyl alcohol glucosides are derivatives of coniferyl alcohol that have been reported as having effects on plant growth [15,61–65]. Altogether these data led us to the hypothesis that the phenotypes observed in *ref3-2, cadc cadd, ccr1*, and *pal1 pal2* overexpressing *F5H* are due to a decrease in coniferyl alcohol-derivatives. To test this, we grew segregating populations of *ref3-2 fah1-2 C4H:F5H* and *cadc cadd fah1-2 C4H:F5H* on media containing either coniferyl alcohol or *p*-coumaryl alcohol and sinapyl alcohol as controls. Coniferyl alcohol was the only treatment that promoted growth in both genotypes (Figure 4.4). This result further supports the hypothesis that it is not a deficiency

in lignin that leads to the growth phenotypes of these mutants, since supplying other monolignols did not restore growth.

4.5.3 Coniferyl alcohol-derived compounds in ^{13}C-coniferyl alcohol fed samples *cadc cadd fah1-2 C4H:F5H*

In order to identify the coniferyl alcohol-derived compound(s) that restore growth in *cadc cadd fah1-2 C4H:F5H,* we fed seedlings with [$^{12}C_6$] or [$^{13}C_6$] coniferyl alcohol to identify coniferyl-alcohol derived molecules and compared these results with an untreated control. This experiment allowed for the identification of 10 unique metabolite features that increased in samples treated with coniferyl alcohol in both time points and that are derived from coniferyl alcohol (Figure 4.5). Further analysis suggests these features correspond to five unique molecules. Three metabolite features correspond to diglycosylated coniferyl alcohol, a dimer of coniferyl alcohol and a glycosylated coniferyl alcohol dimer (Table 4.2). Future experiments require LC/MS-MS to provide a better fragmentation pattern that would help with identification of the two unknown metabolites. Further, in order to determine the stereochemistry of the dimers identified in this study it would be necessary to perform NMR analysis. The results here presented are consistent with role of DCA and DCGs and/or compounds like them, in plant growth, although growth assays would be required to determine biological activity. We did not identify DCA in our dataset, but we cannot discard the possibility that DCA is rapidly metabolized into DCGs and that is why we could not identify this compound.

4.5.4 Growth impairment by induction of F5H

We generated inducible lines in which we can stall plant growth by induction of *F5H.* The short time frame in which we observed negative effects on growth in these lines is consistent with previous results indicating that coniferyl alcohol-derived compounds are acting in a hormone-like manner. The *F5H* inducible lines present valuable material in which to perform time course experiments to monitor the changes in candidate compounds and transcription overtime that would help us elucidate the mechanism by which coniferyl alcohol and/or derivatives affect growth.

4.6 References

1. Taiz L, Zeiger E: *Plant Physiology*. Sinauer Associates, Inc.; 2006.
2. Wasternack C: Jasmonates: An update on biosynthesis, signal transduction and action in plant stress response, growth and development. *Ann Bot* 2007, 100:681–697.
3. Zhang Y, Li X: Salicylic acid: Biosynthesis, perception, and contributions to plant immunity. *Curr Opin Plant Biol* 2019, 50:29–36.
4. Fujioka S, Yokota T: Biosynthesis and metabolism of brassinosteroids. *Annu Rev Plant Biol* 2003, 54:137–164.
5. Waters MT, Gutjahr C, Bennett T, Nelson DC: Strigolactone signaling and evolution. *Annu Rev Plant Biol* 2017, 68:291–322.
6. Hartmann T: From waste products to ecochemicals: Fifty years research of plant secondary metabolism. *Phytochemistry* 2007, 68:2831–2846.
7. Erb M, Kliebenstein DJ: Plant secondary metabolites as defenses, regulators, and primary metabolites: The blurred functional trichotomy. *Plant Physiol* 2020, 184:39–52.
8. Lynch JH, Qian Y, Guo L, Maoz I, Huang XQ, Garcia AS, Louie G, Bowman ME, Noel JP, Morgan JA, et al.: Modulation of auxin formation by the cytosolic phenylalanine biosynthetic pathway. *Nat Chem Biol* 2020, 16:850–856.
9. Halkier BA, Gershenzon J: Biology and biochemistry of glucosinolates. *Annu Rev Plant Biol* 2006, 57:303–333.
10. Malinovsky FG, Thomsen MLF, Nintemann SJ, Jagd LM, Bourgine B, Burow M, Kliebenstein DJ: An evolutionarily young defense metabolite influences the root growth of plants via the ancient TOR signaling pathway. *Elife* 2017, 6:1–24.
11. Brown DE, Rashotte AM, Murphy AS, Normanly J, Tague BW, Peer WA, Taiz L, Muday GK: Flavonoids act as negative regulators of auxin transport *in vivo* in Arabidopsis. *Plant Physiol* 2001, 126:524–535.
12. Mo Y, Nagel C, Taylor LP: Biochemical complementation of chalcone synthase mutants defines a role for flavonols in functional pollen. *Proc Natl Acad Sci U S A* 1992, 89:7213–7217.
13. Muhlemann JK, Younts TLB, Muday GK: Flavonols control pollen tube growth and integrity by regulating ROS homeostasis during high-temperature stress. *Proc Natl Acad Sci U S A* 2018, 115:E11188–E11197.

14. Watkins JM, Chapman JM, Muday GK: Abscisic acid-induced reactive oxygen species are modulated by flavonols to control stomata aperture. *Plant Physiol* 2017, 175:1807–1825.
15. Binns AN, Chen RH, Wood HN, Lynn DG: Cell division promoting activity of naturally occurring dehydrodiconiferyl glucosides: Do cell wall components control cell division? *Proc Natl Acad Sci* 1987, 84:980–984.
16. Turner SR, Somerville CR: Collapsed xylem phenotype of Arabidopsis identifies mutants deficient in cellulose deposition in the secondary cell wall. *Plant Cell* 1997, 9:689–701.
17. Zhong R, Taylor JJ, Ye Z-H: Disruption of interfascicular fiber differentiation in an Arabidopsis mutant. *Plant Cell* 1997, 9:2159–2170.
18. Himmel ME, Ding S, Johnson DK, Adney WS, Nimlos MR, Brady JW, Foust TD: Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* 2007, 315:804-807.
19. Mottiar Y, Vanholme R, Boerjan W, Ralph J, Mansfield SD: Designer lignins: Harnessing the plasticity of lignification. *Curr Opin Biotechnol* 2016, 37:190–200.
20. Wang P, Dudareva N, Morgan JA, Chapple C: Genetic manipulation of lignocellulosic biomass for bioenergy. *Curr Opin Chem Biol* 2015, 29:32–39.
21. De Meester B, de Vries L, Özparpucu M, Gierlinger N, Corneillie S, Pallidis A, Goeminne G, Morreel K, De Bruyne M, De Rycke R, *et al*.: Vessel-specific reintroduction of CINNAMOYL COA REDUCTASE 1 (CCR1) in dwarfed *ccr1* mutants restores vessel and xylary fiber integrity and increases biomass. *Plant Physiol* 2018, 176:611–633.
22. Muro-Villanueva F, Mao X, Chapple C: Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr Opin Biotechnol* 2019, 56:202–208.
23. Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M: Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell* 2004, 16:1446–1465.
24. Schilmiller AL, Stout J, Weng JK, Humphreys J, Ruegger MO, Chapple C: Mutations in the *cinnamate 4-hydroxylase* gene impact metabolism, growth and development in Arabidopsis. *Plant J* 2009, 60:771–782.

25. Vanholme R, Cesarino I, Rataj K, Xiao Y, Sundin L, Goeminne G, Kim H, Cross J, Morreel K, Araujo P, et al.: Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science* 2013, 341:1103–1107.
26. Ha CM, Escamilla-trevino L, Carlos J, Yarce S, Kim H, Ralph J, Chen F, Dixon RA: An essential role of caffeoyl shikimate esterase in monolignol biosynthesis in *Medicago truncatula*. *Plant J* 2016, 86: 363-375.
27. Leplé J-C, Dauwe R, Morreel K, Storme V, Lapierre C, Pollet B, Naumann A, Kang K-Y, Kim H, Ruel K, et al.: Downregulation of cinnamoyl-Coenzyme A reductase in poplar : Multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell* 2007, 19:3669–3691.
28. Voelker SL, Lachenbruch B, Meinzer FC, Strauss SH: Reduced wood stiffness and strength , and altered stem form , in young antisense 4CL transgenic poplars with reduced lignin contents. *New Phytol* 2011, 189:1096–1109.
29. Wang JP, Matthews ML, Williams CM, Shi R, Yang C, Tunlaya-Anukit S, Chen H-C, Li Q, Liu J, Lin C-Y, et al.: Improving wood properties for wood utilization through multi-omics integration in lignin biosynthesis. *Nat Commun* 2018, 9:1579.
30. Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-pettenati J, Boudet A: Down-regulation of *Cinnamoyl-CoA Reductase* induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 1998, 13:71–83.
31. Goujon T, Ferret V, Mila I, Pollet B, Ruel K, Burlat V, Joseleau J-P, Barrière Y, Lapierre C, Jouanin L: Down-regulation of the *AtCCR1* gene in *Arabidopsis thaliana*: Effects on phenotype , lignins and cell wall degradability. *Planta* 2003, 217:218-228.
32. Huang J, Gu M, Lai Z, Fan B, Shi K, Zhou Y-H, Yu J-Q, Chen Z: Functional analysis of the Arabidopsis PAL gene family in plant growth , development , and response to environmental stress. *Plant Physiol* 2010, 153:1526–1538.
33. Li Y, Kim JI, Pysh L, Chapple C: Four isoforms of Arabidopsis 4-coumarate : CoA ligase have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol* 2015, 169:2409–2421.

34. Shadle G, Chen F, Reddy MSS, Jackson L, Nakashima J, Dixon RA: Down-regulation of *Hydroxycinnamoyl CoA: Shikimate Hydroxycinnamoyl Transferase* in transgenic alfalfa affects lignification , development and forage quality. *Phytochemistry* 2007, 68:1521–1529.
35. Chapple CCS, Vogt T, Ellis BE, Somerville CR: An Arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell* 1992, 4:1413–1424.
36. Goujon T, Sibout R, Pollet B, Maba B, Nussaume L, Bechtold N, Lu F, Ralph J, Mila I, Barrière Y, et al.: A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase impacts lignins and sinapoyl esters. *Plant Mol Biol* 2003, 51:973–989.
37. Fu C, Mielenz JR, Xiao X, Ge Y, Hamilton CY, Rodriguez M, Chen F, Foston M, Ragauskas A, Bouton J, et al.: Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proc Natl Acad Sci* 2011, 108:3803–3808.
38. Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier MT, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inzé D, et al.: Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol* 1996, 112:1479-1490.
39. Sibout R, Eudes A, Mouille G, Pollet B, Lapierre C, Jouanin L, Séguin A: *CINNAMYL ALCOHOL DEHYDROGENASE-C* and *-D* are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. *Plant Cell* 2005, 17:2059–2076.
40. Jones L, Ennos AR, Turner SR: Cloning and characterization of *irregular xylem4* (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*. *Plant J* 2001, 26:205–216.
41. Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C: Changes in secondary metabolism and deposition of an unusual in the *ref8* mutant of Arabidopsis. *Plant J* 2002, 30:47-59.
42. Hamann T: The plant cell wall integrity maintenance mechanism - A case study of a cell wall plasma membrane signaling network. *Phytochemistry* 2015, 112:100–109.
43. Decreux A, Messiaen J: Wall-associated kinase WAK1 onteracts with cell wall pectins in a calcium-induced conformation. *Plant Cell Physiol* 2005, 46:268–278.

44. Brutus A, Sicilia F, Macone A, Cervone F, Lorenzo G De: A domain swap approach reveals a role of the plant wall-associated kinase 1 (WAK1) as a receptor of oligogalacturonides. *Proc Natl Acad Sci U S A* 2010, 107:9452-9457.
45. Hématy K, Sado PE, Van Tuinen A, Rochange S, Desnos T, Balzergue S, Pelletier S, Renou JP, Höfte H: A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. *Curr Biol* 2007, 17:922–931.
46. Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, et al.: Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* 2014, 509:376–380.
47. Besseau S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B: Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* 2007, 19:148–162.
48. Gallego-Giraldo L, Escamilla-Trevino L, Jackson LA, Dixon RA: Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc Natl Acad Sci* 2011, 108:20814–20819.
49. Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C: Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in Arabidopsis. *Proc Natl Acad Sci U S A* 1998, 95:6619–6623.
50. Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C: Modified lignin in tobacco and poplar plants over-expressing the Arabidopsis gene encoding ferulate 5-hydroxylase. *Plant J* 2000, 22:223–234.
51. Huntley SK, Ellis D, Gilbert M, Chapple C, Mansfield SD: Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *J Agric Food Chem* 2003, 51:6178–6183.
52. Reddy MSS, Chen F, Shadle G, Jackson L, Aljoe H, Dixon RA: Targeted down-regulation of cytochrome P450 enzymes for forage quality improvement in alfalfa (*Medicago sativa L.*). *Proc Natl Acad Sci* 2005, 102:16573–16578.
53. Stewart JJ, Akiyama T, Chapple C, Ralph J, Mansfield SD: The effects on lignin structure of overexpression of ferulate 5-hydroxylase in hybrid poplar. *Plant Physiol* 2009, 150:621–635.

54. Anderson NA, Tobimatsu Y, Ciesielski PN, Ximenes E, Ralph J, Donohoe BS, Ladisch M, Chapple C: Manipulation of guaiacyl and syringyl monomer biosynthesis in an Arabidopsis cinnamyl alcohol dehydrogenase mutant results in atypical lignin biosynthesis and modified cell wall structure. *Plant Cell* 2015, 27:2195–2209.
55. Craft J, Samalova M, Baroux C, Townley H, Martinez A, Jepson I, Tsiantis M, Moore I: New pOp/LhG4 vectors for stringent glucocorticoid-dependent transgene expression in Arabidopsis. *Plant J* 2005, 41:899–918.
56. Clough SJ, Bent AF: Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J* 1998, 16:735–743.
57. Bonawitz ND, Soltau WL, Blatchley MR, Powers BL, Hurlock AK, Seals LA, Weng JK, Stout J, Chapple C: REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in Arabidopsis. *J Biol Chem* 2012, 287:5434–5445.
58. Li X, Bonawitz ND, Weng J-K, Chapple C: The growth reduction associated with repressed lignin biosynthesis in *Arabidopsis thaliana* is independent of flavonoids. *Plant Cell* 2010, 22:1620–1632.
59. Kim JI, Zhang X, Pascuzzi PE, Liu CJ, Chapple C: Glucosinolate and phenylpropanoid biosynthesis are linked by proteasome-dependent degradation of PAL. *New Phytol* 2020, 225:154–168.
60. Do CT, Pollet B, Thévenin J, Sibout R, Denoue D, Barrière Y, Lapierre C, Jouanin L: Both caffeoyl Coenzyme A 3-*O*-methyltransferase 1 and caffeic acid *O*-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. *Planta* 2007, 226:1117–1129.
61. Shibata K, Kubota T, Kamisaka S: Isolation and chemical identification of a lettuce cotyledon factor, a synergist of the gibberellin action in inducing lettuce hypocotyl elongation. *Plant Cell Physiol* 1974, 15:191-194.
62. Lee TS, Purse JG, Pryce RJ, Horgan R, Wareing PF: Dihydroconiferyl alcohol- A cell division factor from *Acer* species. *Planta* 1981, 152, 571–577.
63. Lynn DG, Chen RH, Manning KS, Wood HN: structural characterization of endogenous factors from Vinca rosea crown gall tumors that promote cell division of tobacco cells. *Proc Natl Acad Sci U S A* 1987, 84:615–619.

64. Teutonico RA, Dudley MW, Orr3 JD, Lynn DG, Binns AN: Activity and accumulation of naturally occuring dehydrodiconiferyl glucosides: Do cell wall components control cell division? *Plant Physiol* 1991, 97:288–297.
65. Tamagnone L: Inhibition of phenolic acid metabolism results in precocious cell death and altered cell morphology in leaves of transgenic tobacco plants. *Plant Cell* 1998, 10:1801–1816.
66. Zhang X, Gou M, Liu CJ: Arabidopsis Kelch repeat F-Box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. *Plant Cell* 2013, 25:4994–5010.
67. Zhang X, Gou M, Guo C, Yang H, Liu CJ: Down-regulation of Kelch domain-containing F-box protein in Arabidopsis enhances the production of (poly)phenols and tolerance to ultraviolet radiation. *Plant Physiol* 2015, 167:337–350.
68. Zhou R, Jackson L, Shadle G, Nakashima J, Temple S, Chen F, Dixon RA: Distinct cinnamoyl CoA reductases involved in parallel routes to lignin in *Medicago truncatula*. *Proc Natl Acad Sci U S A* 2010, 107:17803–17808.
69. Parvathi K, Chen F, Guo D, Blount JW, Dixon RA: Substrate preferences of *O*-methyltransferases in alfalfa suggest new pathways for 3-*O*-methylation of monolignols. *Plant J* 2001, 25:193–202.
70. Dima O, Morreel K, Vanholme B, Kim H, Ralph J, Boerjan W: Small glycosylated lignin oligomers are stored in Arabidopsis leaf vacuoles. *Plant Cell* 2015, 27:695–710.

Table 4.1. List of primer used in this project.

Primer name	Sequence (5'-> 3')	Usage
CC5409	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT CATG GGT TCA ACG GCA GAG	*COMT* amplification with attB adaptors fwd
CC5410	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTCTTA GAG CTT CTT GAG TAA CTC	*COMT* amplification with attB adaptors rv
CC5412	GGG GAC AAG TTT GTA CAA AAA AGC AGG CTT CATG GAG TCT TCT ATA TCA CAA AC	*F5H* amplification with attB adaptors fwd
CC5413	GGG GAC CAC TTT GTA CAA GAA AGC TGG GTCTTA AAG AGC ACA GAT GAG G	*F5H* amplification with attB adaptors rv

Table 4.2. Putative identification of metabolite features identified by LC/MS from seedlings fed with [$^{12}C_6$] or [$^{13}C_6$] coniferyl alcohol.
Metabolite features are identified by their exact mass (M) followed by their retention time (T).

Metabolite Feature	Putative Identification
M563T371	Diglycosylated coniferyl alcohol
M358T423	Coniferyl alcohol dimer
M605T437	Unknown
M180T439	Coniferyl alcohol
M358T439	Glycosylated coniferyl alcohol
M775T439	Acetate adduct of a glycosylated coniferyl alcohol dimer
M195T506	Unknown
M399T506	Unknown
M400T506	Unknown
M829T506	Unknown

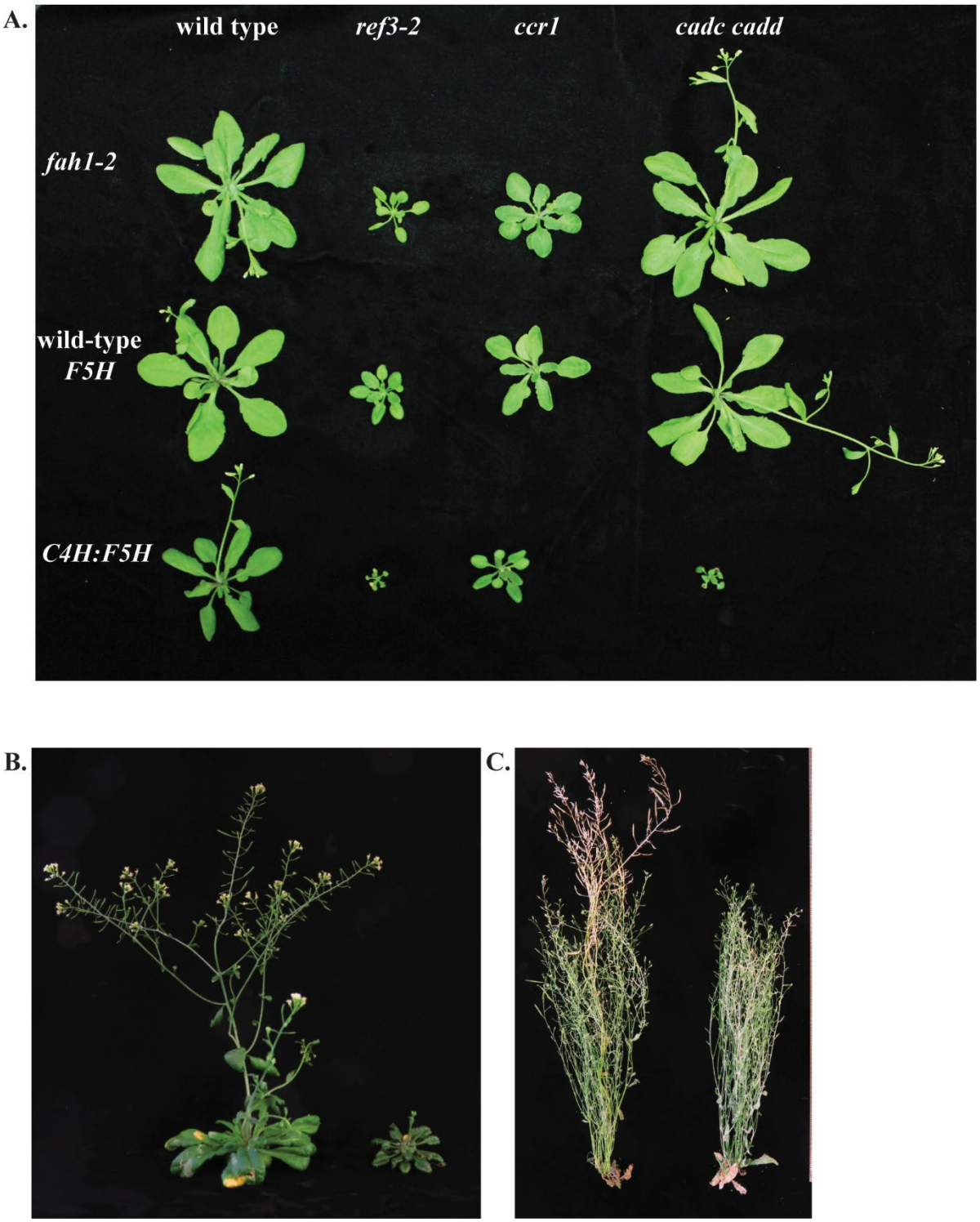


Figure 4.1. Representative photograph of F5H knockouts and overexpressors.
(A) 3-week-old wild-type, *ref3-2, ccr1,* and *cadc cadd* plants carrying a mutation in F5H (*fah1-2*) or overexpressing F5H (*C4H:F5H*). (B) *ccr1* (left) and *ccr1 C4H:F5H* (right) phenotypes at maturity. (C) *pal1 pal2* (left) and *pal1 pal2 C4H:F5H* (right) phenotypes at maturity.

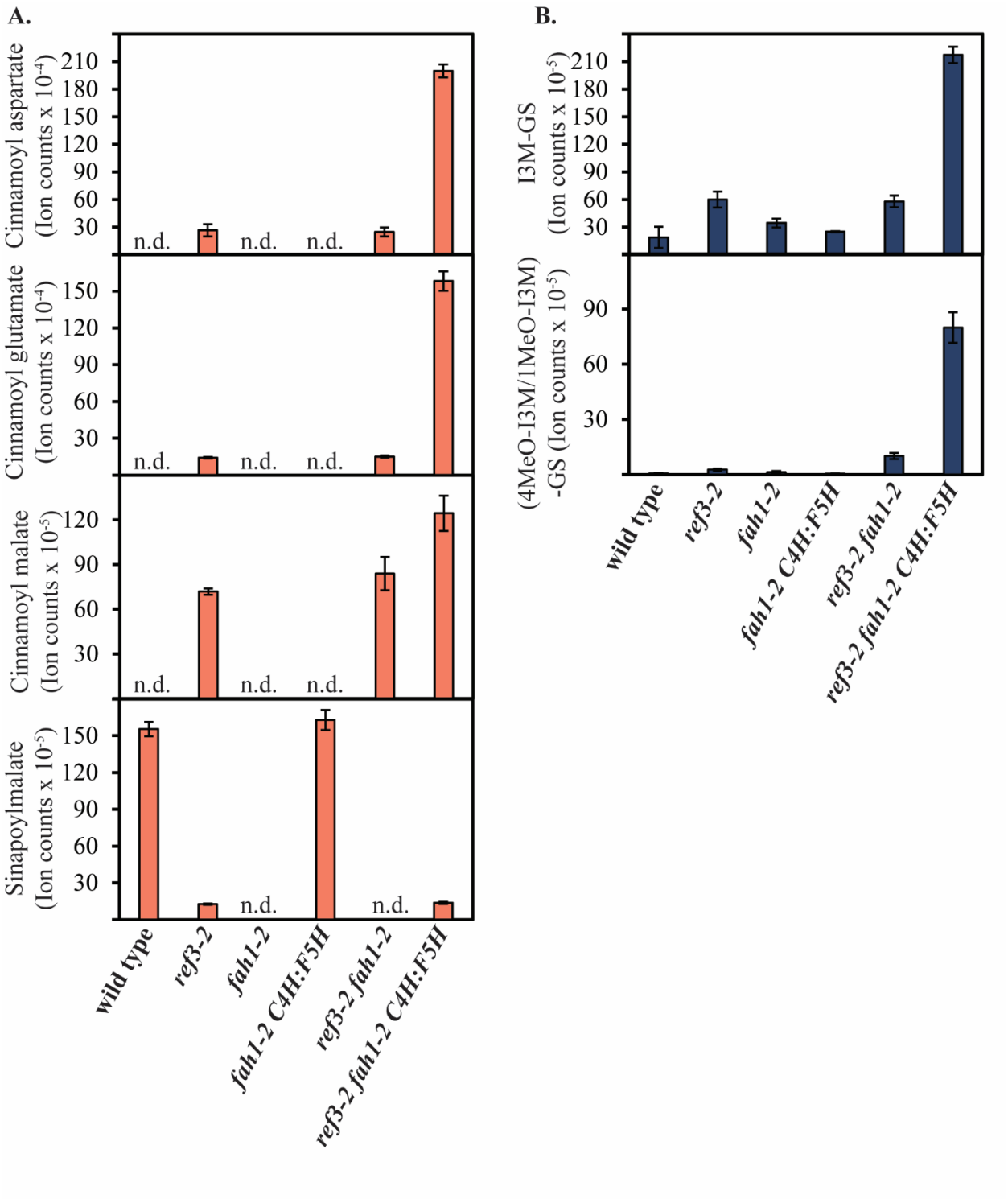


Figure 4.2. LC/MS analysis of methanol-soluble metabolites of 3 weeks-old plants.
(A) Phenylpropanoids: cinnamoyl aspartate, cinnamoyl glutamate, cinnamoyl malate, and sinapoylmalate. (B) Glucosinolates: I3M-GS: indol-3-ylmethyl glucosinolate. and (4MeO-I3M/1MeO-I3M)-GS: 1-methoxy-indol-3-ylmethyl glucosinolate. Error bars represent standard deviation among biological replicates (n=3). n.d.= not detected.

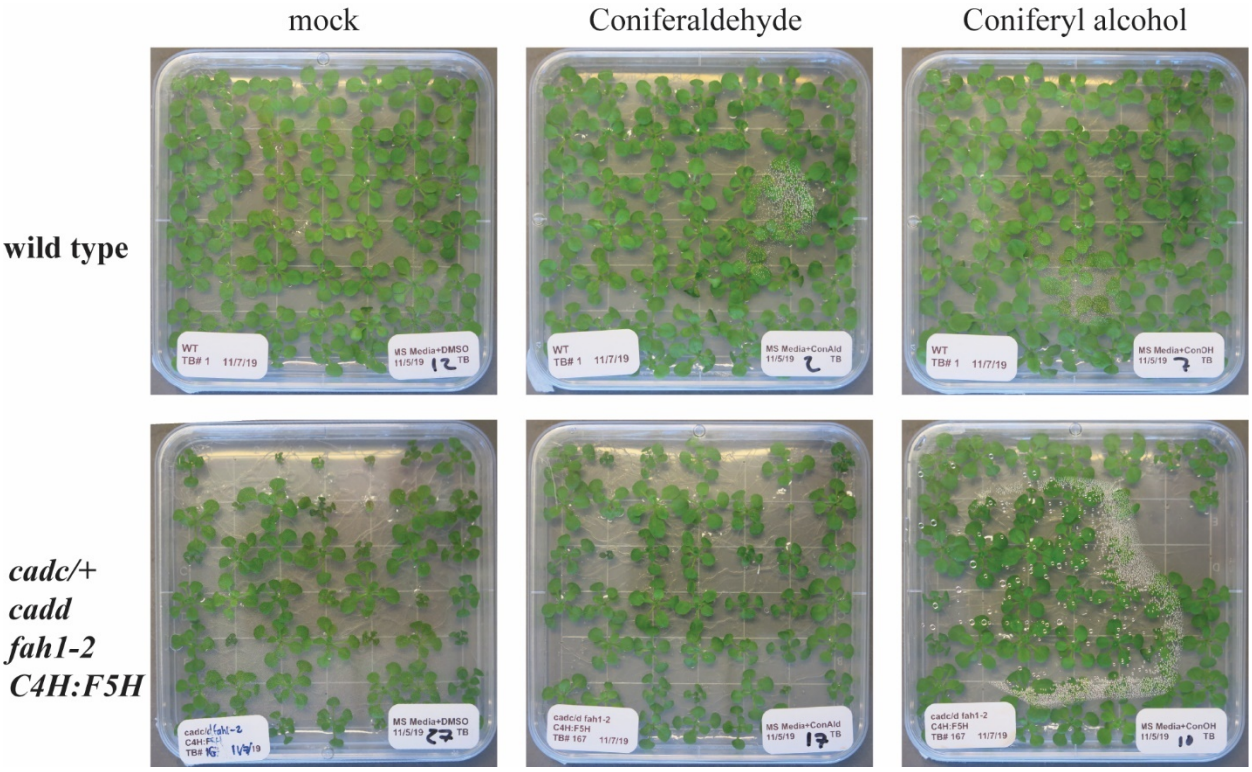


Figure 4.3. Representative photograph of plants grown in media supplemented with phenylpropanoids.
Wild type and a segregating population of *cadc/+ cadd fah1-2 C4H:F5H* grown in media supplemented with 100 µM coniferaldehyde or coniferyl alcohol in 0.03% (v/v) aqueous DMSO. The mock treatment was an equivalent concentration of DMSO. Experiment and photographs by: Tyler Bouse.

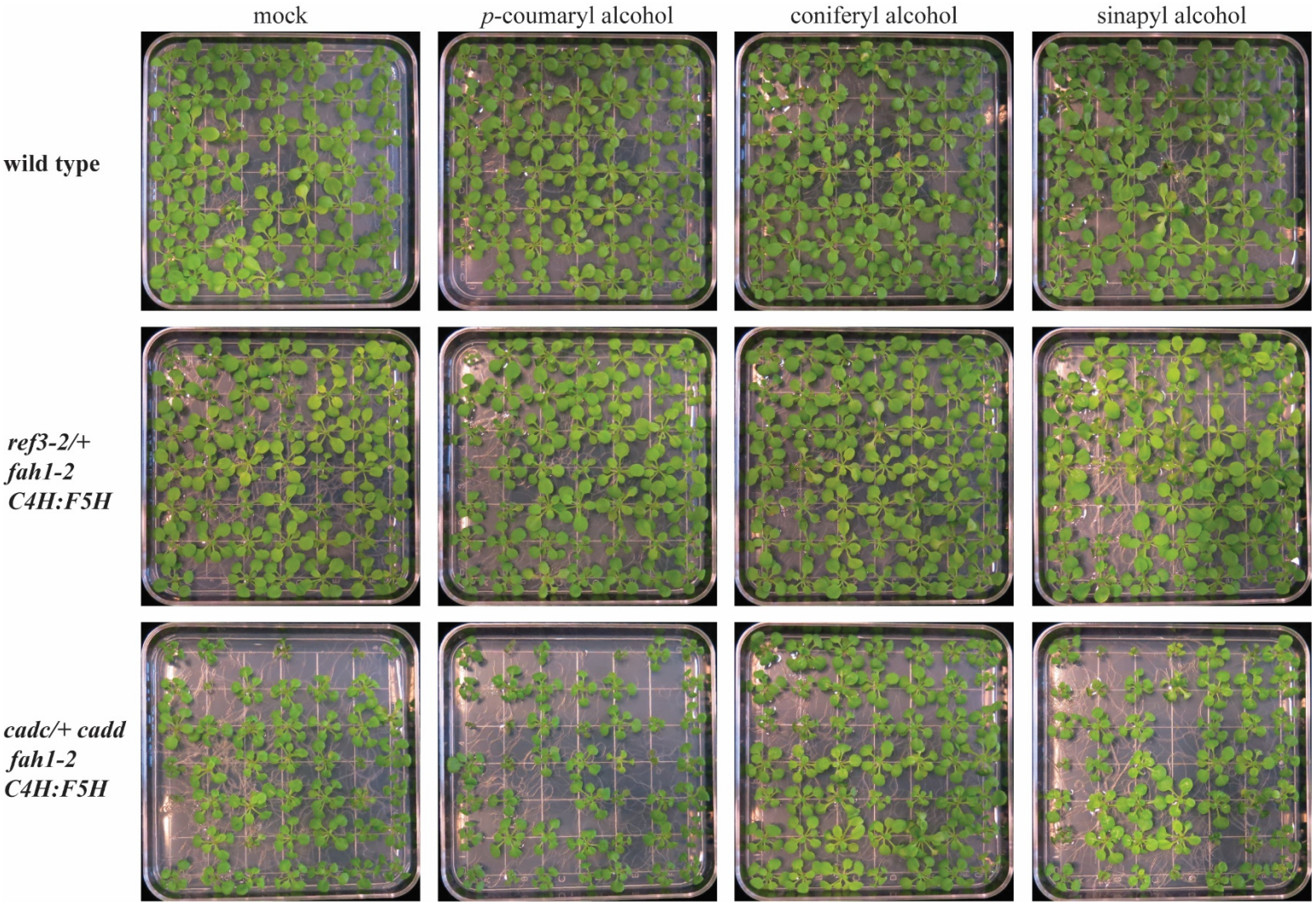


Figure 4.4. Representative photograph of plants grown in media supplemented with hydroxycinnamyl alcohols.
Wild type and segregating populations of *ref3-2/+ fah1-2 C4H:F5H* and *cadc/+ cadd fah1-2 C4H:F5H* grown in media supplemented with 200 µM *p*-coumaryl, coniferyl or sinapyl alcohol in 0.06% (v/v) aqueous DMSO. The mock treatment was an equivalent concentration of DMSO. Experiment and photographs by: Zhiwei Luo.

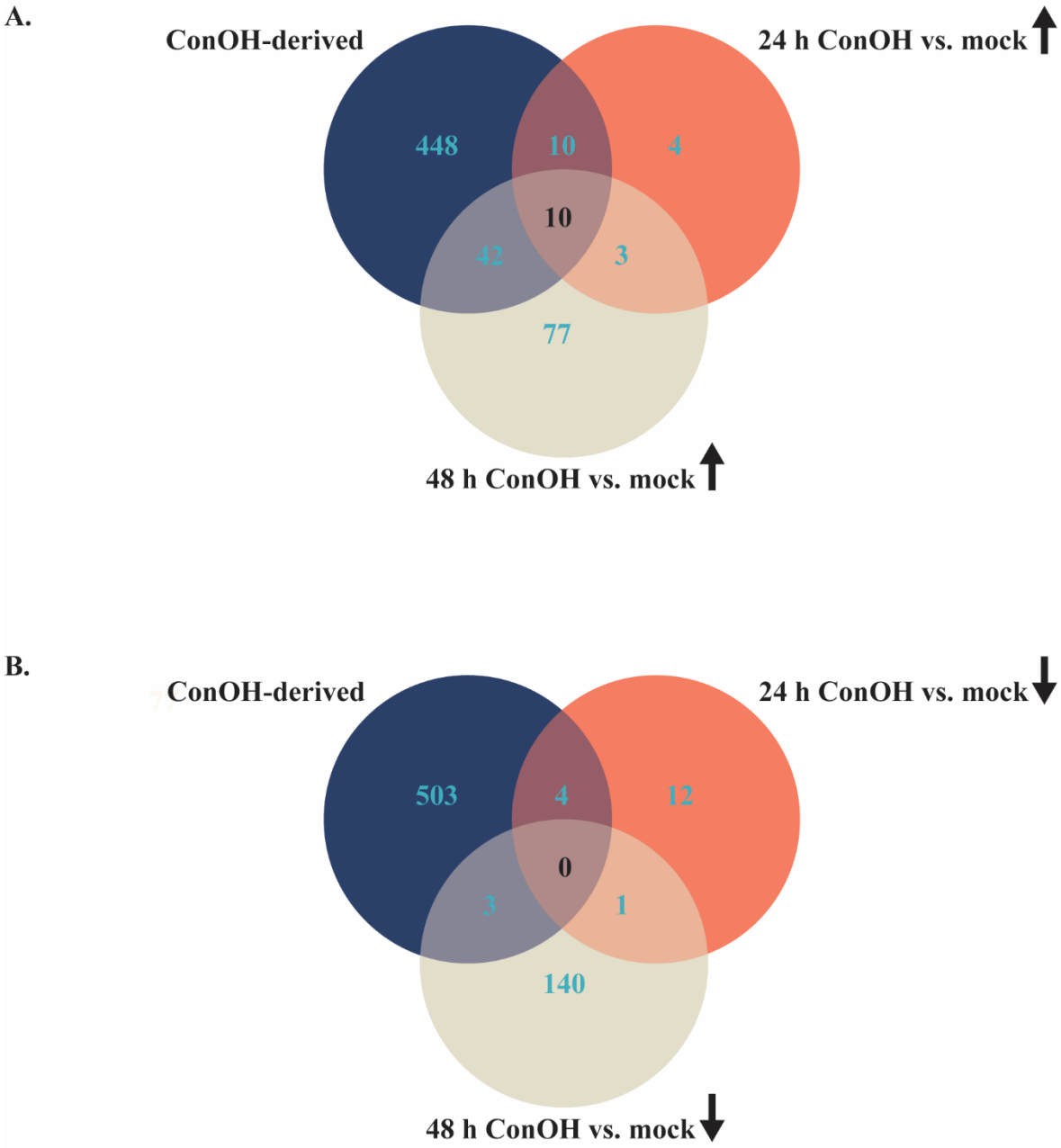


Figure 4.5. Venn diagrams depicting the overlap between coniferyl alcohol-derived metabolite features identified by LC/MS and changes in abundance of these metabolite features over time in *cadc cadd fah1-2 C4H:F5H* seedlings.
Blue: Metabolite features identified after 24 h and 48 h by feeding *cadc cadd fah1-2 C4H:F5H* seedlings 250 µM [$^{12}C_6$] or [$^{13}C_6$] coniferyl alcohol (ConOH-derived). Salmon: Pairwise comparison between metabolite features from seedlings grown in media containing 250 µM [$^{12}C_6$] coniferyl alcohol or an equivalent volume of DMSO as a mock treatment for 24 h (24 h ConOH vs. mock). Bone: Pairwise comparison between metabolite features from seedlings grown in media containing 250 µM [$^{12}C_6$] coniferyl alcohol or an equivalent volume of DMSO as a mock treatment for 48 h (48h ConOH vs. mock). (A) Metabolite features with increased abundance. (B) Metabolite features with decreased abundance.

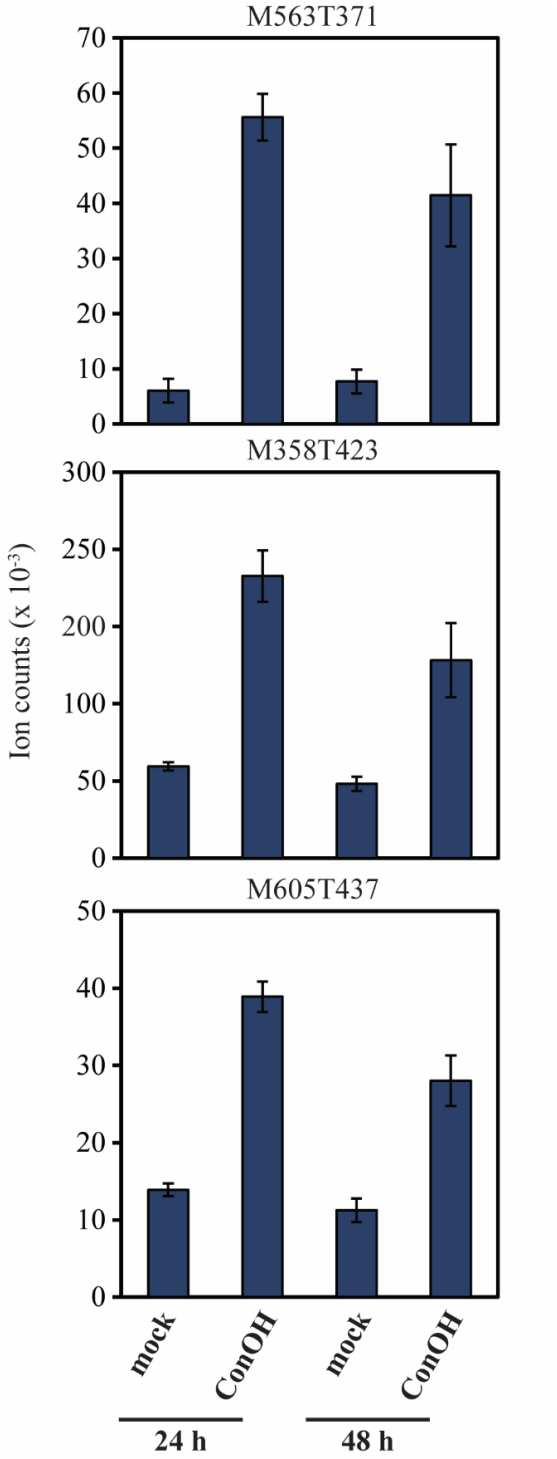


Figure 4.6. Relative abundance of coniferyl alcohol-derived metabolite features with increased abundance in *cadc cadd fah1-2 C4H:F5H* seedlings measured by LC/MS.
Grown in media supplemented with 250 µM coniferyl alcohol (ConOH) or an equivalent volume of DMSO as a mock treatment. Error bars represent standard deviation among biological replicates (*n*=3).

Figure 4.6 continued

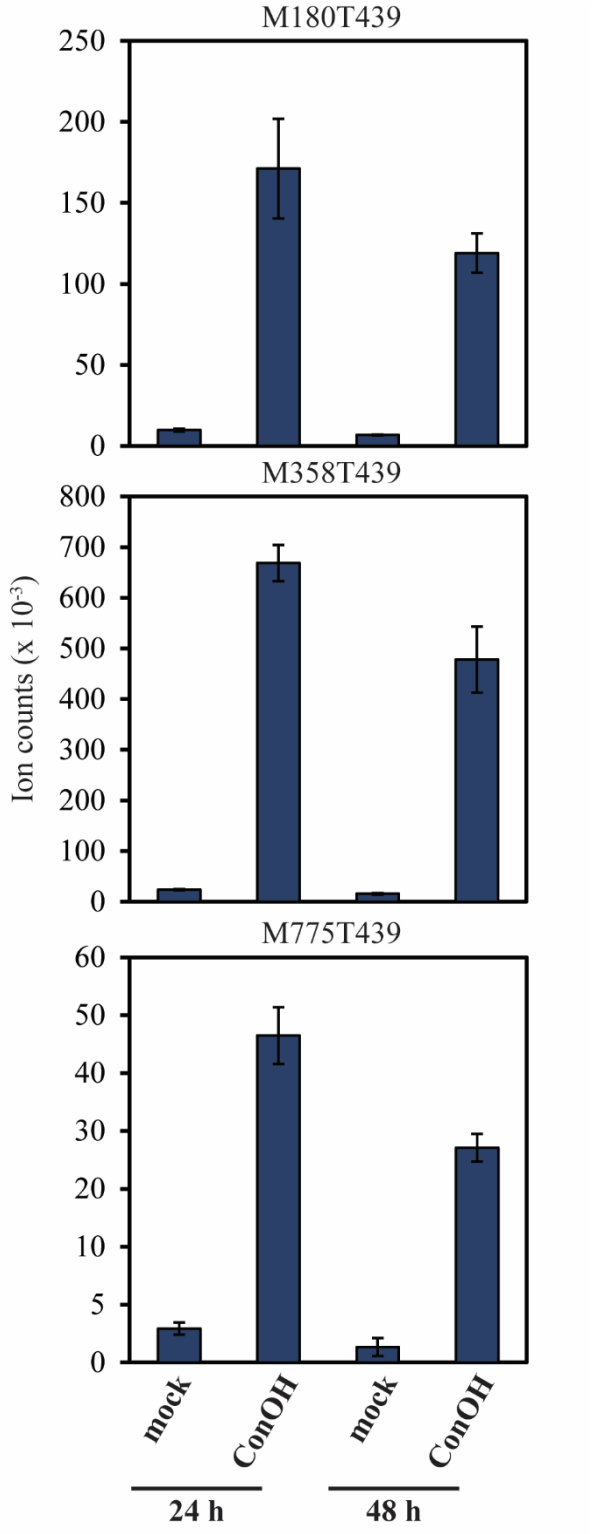
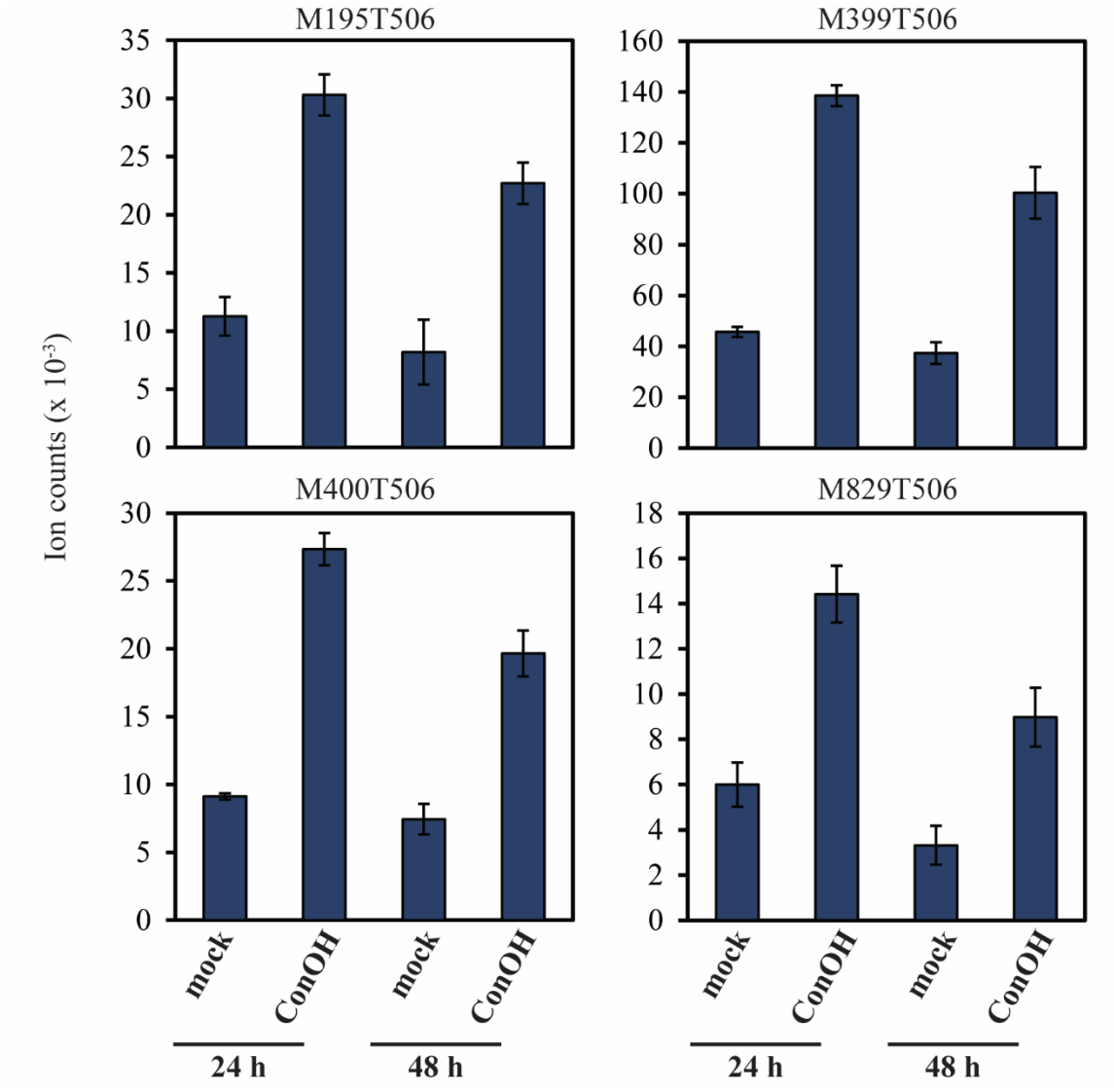


Figure 4.6 continued



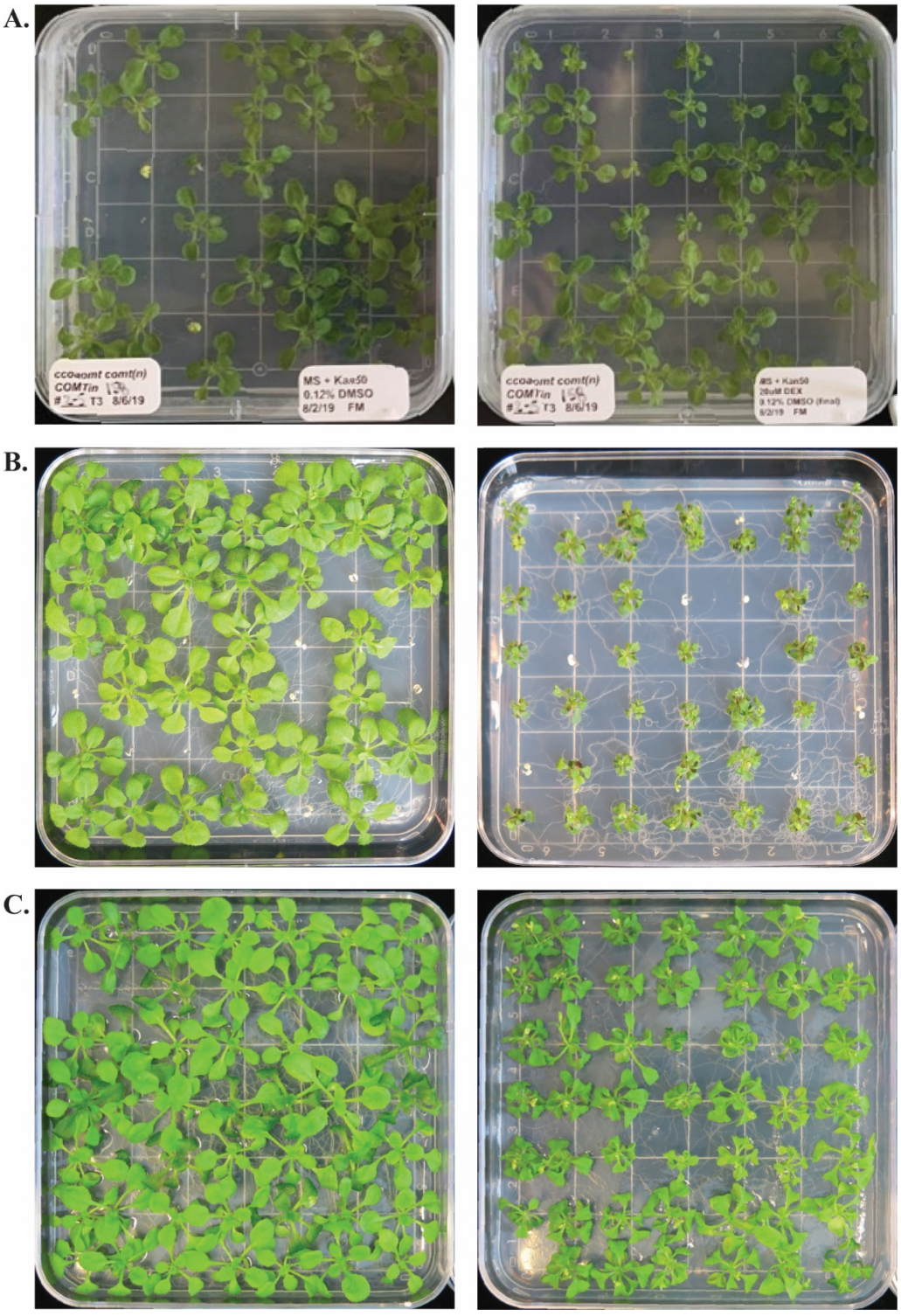


Figure 4.7. Inducible lines grown in the absence or presence of dexamethasone.
(A) Segregating population of *ccoaomt comt/+ COMTin,* mock- (left) and dex- treated (right). (B) *cadc cadd fah1-2 F5Hin,* mock- (left) and dex- treated (right). (C) *ref3-3 F5Hin,* mock- (left) and dex- treated (right).

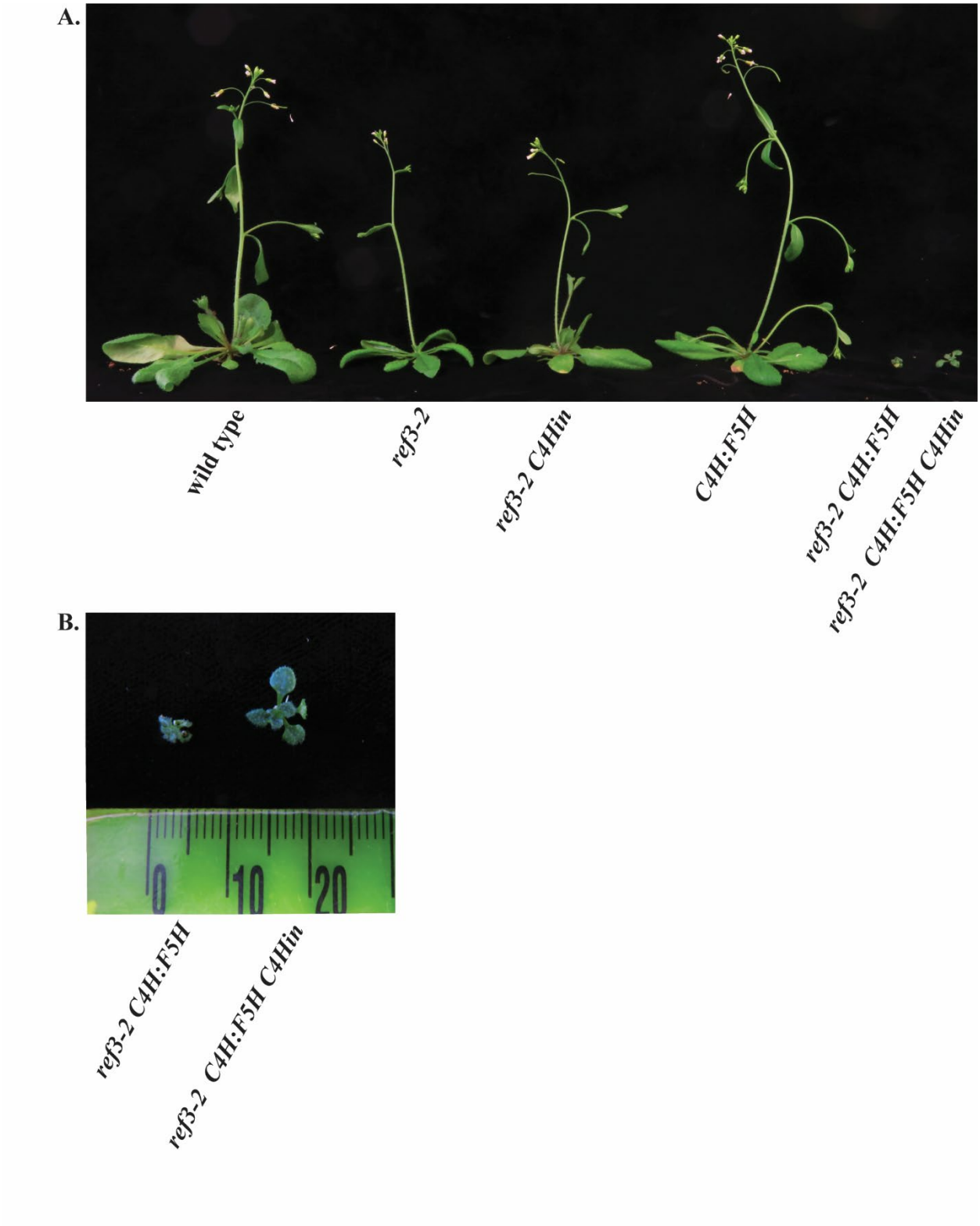


Figure 4.8. Representative photograph of *C4Hin* plants continuously treated with dexamethasone at four weeks of age.
(A) Phenotype comparison of *ref3-2 C4H:F5H C4Hin* to controls. (B) Phenotype comparison of *ref3-2 C4H:F5H* and *ref3-2 C4H:F5H C4Hin.*

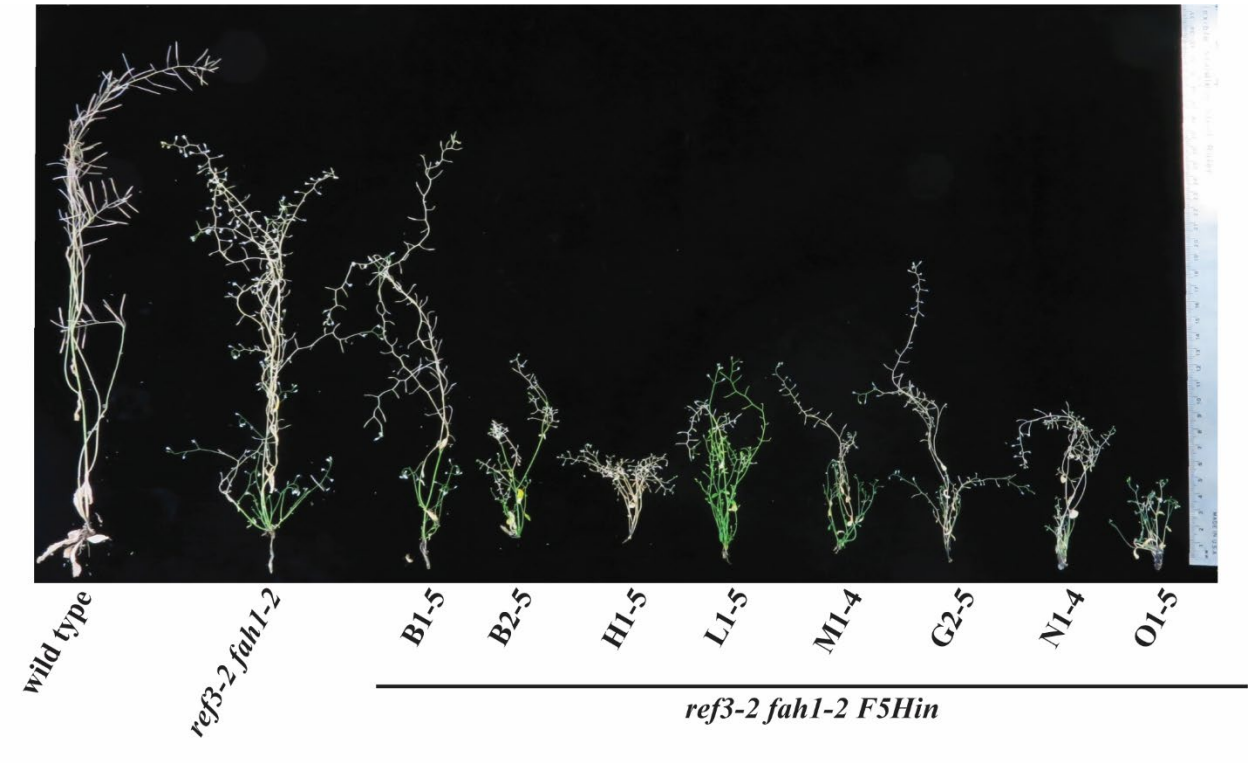


Figure 4.9. Representative photograph of *ref3-2 fah1-2 F5Hin* mature plants continuously treated with dexamethasone.

cadc cadd fah1-2

cadc/+ cadd fah1-2 F5Hin

Figure 4.10. Growth phenotype of *cadc cadd fah1-2 F5Hin* plants.
Left: *cadc cadd fah1-2.* Right: *cadc cadd fah1-2 F5Hin* line segregating for the *F5Hin* transgene treated with dexamethasone.