INVESTIGATING THE FUNCTIONAL ROLE OF MED5 AND CDK8 IN ARABIDOPSIS MEDIATOR COMPLEX

by

Xiangying Mao

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

August 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF DISSERTATION APPROVAL

Dr. Clint Chapple, Co-Chair

    Department of Biochemistry

Dr. Vikki Weake, Co-Chair

    Department of Biochemistry

Dr. Barbara Golden

    Department of Biochemistry

Dr. Tesfaye Mengiste

    Department of Botany and Plant Pathology

Approved by:

    Dr. Andrew Mesecar

        Head of the Departmental Graduate Program

For my parents Mr. Mao and Ms. Xiang,

and my fiancé Dr. Si

ACKNOWLEDGMENTS

I am so grateful for the opportunity we have had to build our friendship and supported each other over the years. It is you that make West Lafayette a place full of sweet memories.

I also want to thank my family and closest friends back in China, who make me believe that I deserve and will always be loved. I owe a debt of gratitude to my parents and closest friends who took long trips to here and showed their support as always.

Finally, I would give the most special appreciation to my fiancé Dr. Si for his love, trust and patience. It is wonderful to meet you, and I hope we could have endless luck to support each other no matter where we will be.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABA	abscisic acid
ABCB21	ATP-binding cassette subfamily B 21
ABCG40	ABC transporter G family member 40
APC8	anaphase-promoting complex subunit 8
BAM	binary alignment map
bp	base pair
C4H	cinnamic acid 4-hydroxylase
CBF	C-repeat binding factor
CDK8	cyclin-dependent kinase 8
ChIP	chromatin immunoprecipitation
COB	cobra
Col-0	Columbia-0 ecotype
CPM	count per million
cryo-EM	cryo-electron microscopy
cycC	C-type cyclin
DNA	deoxyribonucleic acid
DFRC	derivatization followed by reductive cleavage
DJC66	DNA J protein C66
DOF	DNA-binding with one finger
DREB2A	dehydration-responsive element binding protein 2A
EIL1	ethylene-insensitive 3-like 1
EIN3	ethylene-insensitive 3
EOD-FR	end-of-day far-red
ET	ethylene
FDR	False Discovery Rate
Fe	iron
FIT	Fe-deficiency induced transcription factor
FPKM	fragments per kilobase per million

GO	gene ontology
H3K27me3	trimethylation mark at Lys-27 in histone H3
HCT	hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase
HPLC	high performance liquid chromatography
HSD	honest significant difference
IAA	indole-3-acetic acid
ICS1	isochorismate synthase 1
JA	jasmonic acid
KFB	kelch-domain-containing F-box protein
KIX	kinase-inducible domain-interacting
MED	mediator subunit
MR	mutual rank
NPQ	non-photochemical quenching
NPR1	Arabidopsis nonexpressor of PR genes
ORA59	octadecanoid-responsive Arabidopsis AP2/ERF 59
PAL	phenylalanine ammonia lyase
PCA	principle component analysis
PDF1.2	plant defensin 1.2
PFT1	phytochrome and flowering time 1
phy	phytochrome
PIC	pre-initiation complex
PIF	phytochrome interacting factor
PKU1	plastid protein kinase with unknown function
PMEI	pectin methylesterase inhibitor
Pol II	RNA polymerase II
PR	pathogenesis-related
ref	reduced epidermal fluorescence
RNA	ribonucleic acid
RNAi	RNA interference
RNA-seq	mRNA sequencing
SA	salicylic acid

SAS   shade avoidance syndrome

SAUR   small auxin up RNA

SD   standard deviation

SICER   spatial clustering for identification of ChIP-Enriched regions

SID2   salicylic acid induction deficient 2

SREBP-1c sterol-regulatory-element-binding protein-1c

TF   transcription factor

TGA   thioglycolic acid

TSS   transcription start site

UV   ultraviolet

WIN1   wax inducer 1

WRI1   WRINKLED1

ABSTRACT

Author: Mao, Xiangying. Ph.D.
Institution: Purdue University
Degree Received: August 2019
Title: Investigating the Functional Role of MED5 and CDK8 in Arabidopsis Mediator Complex
Major Professor: Clint Chapple

The Mediator (Med) complex comprises about 30 subunits and is a transcriptional co-regulator in eukaryotic systems. The core Mediator complex, consisting of the head, middle and tail modules, functions as a bridge between transcription factors and basal transcription machinery, whereas the CDK8 kinase module can attenuate Mediator's ability to function as either a co-activator or co-repressor. Many Arabidopsis Mediator subunit has been functionally characterized, which reveals critical roles of Mediator in many aspects of plant growth and development, responses to biotic and abiotic stimuli, and metabolic homeostasis. Traditional genetic and biochemical approaches laid the foundation for our understanding of Mediator function, but recent transcriptomic and metabolomic studies have provided deeper insights into how specific subunits cooperate in the regulation of plant metabolism. In Chapter 1, we highlight recent developments in the investigation of Mediator and plant metabolism, with emphasis on the large-scale biology studies of *med* mutants.

We previously found that MED5, an Arabidopsis Mediator tail subunit, is required for maintaining phenylpropanoid homeostasis. A semi-dominant mutation (*reduced epidermal fluorescence 4-3*, *ref4-3*) that causes a single amino acid substitution in MED5b functions as a strong suppressor of the pathway, leading to decreased soluble phenylpropanoid accumulation, reduced lignin content and dwarfism. In contrast, loss of MED5a and MED5b (*med5*) results in increased levels of phenylpropanoids. In Chapter 2, we present our finding that *ref4-3* requires

CDK8, a Mediator kinase module subunit, to repress plant growth even though the repression of phenylpropanoid metabolism in *ref4-3* is CDK8-independent. Transcriptome profiling revealed that salicylic acid (SA) biosynthesis genes are up-regulated in a CDK8-dependent manner in *ref4-3,* resulting in hyper-accumulation of SA and up-regulation of SA response genes. Both growth repression and hyper-accumulation of SA in *ref4-3* require CDK8 with intact kinase activity, but these SA phenotypes are not connected with dwarfing. In contrast, mRNA-sequencing (RNA-seq) analysis revealed the up-regulation of a DNA J protein-encoding gene in *ref4-3*, the elimination of which partially suppresses dwarfing. Together, our study reveals genetic interactions between Mediator tail and kinase module subunits and enhances our understanding of dwarfing in phenylpropanoid pathway mutants.

In Chapter 3, we characterize other phenotypes of *med5* and *ref4-3*, and find that in addition to the up-regulated phenylpropanoid metabolism, *med5* show other interesting phenotypes including hypocotyl and petiole elongation as well as accelerated flowering, all of which are known collectively as the shade avoidance syndrome (SAS), suggesting that MED5 antagonize shade avoidance in wild-type plants. In contrast, the constitutive *ref4-3* mutant protein inhibits the process, and the stunted growth of *ref4-3* mutants is substantially alleviated by the light treatment that triggers SAS. Moreover, *ref4-3* mimics the loss-of-function *med5* mutants in maintaining abscisic acid (ABA) levels under both normal and drought growth conditions. The phenotypic characterization of *med5* mutants extend our understanding of the role of Mediator in SAS and ABA signaling, providing further insight into the physiological and metabolic responses that require MED5.

In Chapter 4, we explore the function of MED5 and CDK8 in gene expression regulation by investigating the effect of mutations in Mediator including *med5*, *ref4-3*, *cdk8-1* and *ref4-3*

cdk8-1 on genome-wide Pol II distribution. We find that loss of MED5 results in loss of Pol II occupancy at many target genes. In contrast, many genes show enriched Pol II levels in *ref4-3*, some of which overlap with those showing reduced Pol II occupancy in *med5*. In addition, Pol II occupancy is significantly reduced when CDK8 is disrupted in *ref4-3*. Our results help to narrow down the direct gene targets of MED5 and identify genes that may be closely related to the growth deficiency observed in *ref4-3* plants, providing a critical foundation to elucidate the molecular function of Mediator in transcription regulation.

STATEMENT OF PUBLISHED AND COLLABORATIVE WORK

Chapter 1 is an invited review for publication in Journal of Experimental Botany, and the submitted manuscript is still under review. The article was written by Xiangying Mao with editorial contribution from Clint Chapple and Vikki Weake. Chapter 2 is a pre-copyedited, author-produced version of an article accepted for publication in *New Phytologist* following peer review. The version of record, Mao, X., Kim, J. I., Wheeler, M. T., Heintzelman, A. K., Weake, V. M., & Chapple, C. (2019). Mutation of Mediator subunit *CDK8* counteracts the stunted growth and salicylic acid hyper-accumulation phenotypes of an Arabidopsis *MED5* mutant. *New Phytologist*, is available online at https://nph.onlinelibrary.wiley.com/doi/full/10.1111/nph.15741. The mutant *ref4-3 yuc6-1D* was generated by Jeong Im Kim, and the site-directed mutagenesis was performed by Mitchell Wheeler and Anne Heintzelman. The article was written by Xiangying Mao and edited by Clint Chapple and Vikki Weake. In Chapter 3, the shade avoidance responses in *med5* mutants including flowering time, hypocotyl and petiole lengths were first characterized by Nick Bonawitz and reanalyzed by Mitchell Wheeler.

CHAPTER 1.   MEDIATOR FUNCTION IN PLANT METABOLISM REVEALED BY LARGE-SCALE BIOLOGY

1.1   Introduction

The Mediator (Med) complex comprises about 30 subunits and is a transcriptional co-regulator in eukaryotic systems. Based on a variety of structural studies, the core Mediator complex, consisting of the head, middle and tail modules, plays a critical role in bridging the interaction between the basal transcription machinery and the various transcription factors required for transcriptional initiation (Malik and Roeder, 2010; Asturias et al., 1999; Tsai et al., 2014). An additional kinase module can reversibly associate with the core Mediator complex and modulate the interaction of Mediator with RNA polymerase II (Pol II), thereby attenuating Mediator's ability to act as either a co-activator or co-repressor (Hengartner et al., 1998; Borggrefe et al., 2002). Genome-wide studies in yeast and humans have demonstrated that Mediator also controls aspects of transcriptional elongation and termination in addition to its key roles in transcription initiation (Conaway and Conaway, 2013; Mukundan and Ansari, 2011).

Although many Mediator subunits are conserved in eukaryotes, the identification of the entire complement of plant Mediator subunits was originally hampered by the limited sequence similarity between the tail module subunits and their orthologs in non-plant species. It was over a decade after the purification of the human and yeast Mediator complex until plant Mediator was first isolated from *Arabidopsis thaliana* cell suspension cultures using chromatography followed by anti-MED6 or anti-MED2 immunoprecipitation (Bäckström et al., 2007). In addition to 22 highly conserved MED subunits, comparative genomics approaches facilitated the identification and annotation of

other, more divergent subunits in plants. For example, MED2, MED3, MED5 and MED30 were identified via short 'signature sequence motifs' in homologous subunits and MED1 and MED26 were found using Hidden Markov Model-based conserved motif prediction (Bourbon, 2008; Mathur et al., 2011). Together these studies revealed that each yeast and human MED subunit has at least one homolog in plant kingdom, demonstrating that the overall Mediator structure is conserved in eukaryotes. Since then, many Arabidopsis Mediator subunit mutants have been isolated, and their functional characterization has revealed critical roles for Mediator in many aspects of plant growth and development, responses to biotic and abiotic stimuli, and metabolic homeostasis (Mathur et al., 2011; Dolan and Chapple, 2017). Interestingly, although mutations in many Mediator subunits are not lethal, disruption of individual plant Mediator subunits often leads to global changes in gene expression. As a result of these changes in gene expression, many *med* mutants display phenotypes such as perturbed growth, development and stress responses (Yang et al., 2016; Dolan and Chapple, 2017). The viability and distinctive phenotypes of *med* mutants in plants such as Arabidopsis provides an ideal system to characterize the mechanisms by which Mediator regulates biological processes.

Compared to alterations in development that can be observed visually, or alterations in stress responses that can be evaluated at the level of gene expression, metabolic changes in plant *med* mutants are more difficult to measure. Thus, the function of Mediator in plant metabolism has only been pursued in recent years. Based on these studies, we now know that a number of Mediator subunits are required for the regulation of plant metabolism including MED5, MED16 and MED23 for the phenylpropanoid pathway (Stout et al., 2008; Dolan et al., 2017), MED16 for cellulose biosynthesis (Sorek et al., 2015), MED16 and

MED25 for iron homeostasis (Yang et al., 2014; Zhang et al., 2014), and MED15 and CDK8 for lipid biosynthesis (Kim et al., 2016; Zhu et al., 2014; Kong and Chang, 2018) (Figure 1.1). Genome-wide transcriptome analyses have revealed that those Mediator subunits are not only necessary for regulated expression of enzyme-encoding genes but also for the expression of genes involved in transcriptional or post-translational regulation of metabolic pathway genes and enzymes (Yang et al., 2014; Dolan et al., 2017). In addition, metabolite profiling of *med* mutants has provided insight on the molecular function of Mediator in plant responses to biotic and abiotic stimuli, which often involve substantial metabolic reprogramming. In this review, we highlight the function of Mediator in plant metabolism revealed by large-scale biology, including untargeted metabolomic analysis and transcriptomic studies.

1.2   Metabolic profiling reveals interactions between Mediator subunits

To investigate the downstream gene targets of individual Mediator subunits, and thereby infer the metabolic pathways impacted by Mediator-dependent regulation, several groups have performed genome-wide transcriptome analysis in various Arabidopsis *med* mutants (Table 1.1). Gene Ontology (GO) term analysis of the mis-regulated genes in *med* mutants including *med18*, *med7*, *med2*, *med5*, *med15*, *med16* and *cdk8* has identified a consistent enrichment of genes involved in plant primary or secondary metabolism (Table 1.1), suggesting that those subunits are required for the normal regulation and biosynthesis of one or more groups of plant metabolites. The substantial transcriptional reprogramming in *med* mutants suggests that wide-spread metabolic changes might also occur in these plants. Indeed, a recent liquid chromatography – mass spectrometry based untargeted

metabolic profiling of twelve Arabidopsis *med* mutants showed that loss of single MED subunits led to substantial metabolic reprogramming (Davoine et al., 2017). Hierarchical clustering of the metabolite profiles in each mutant revealed that MED22a, MED19a, MED17, MED18, MED19a and MED22a fall into one cluster, whereas MED2, MED3 and MED5 form another group, consistent with known localization of these subunits in the Mediator head and tail modules (Davoine et al., 2017). In contrast, although MED25 is assigned to the Mediator tail module (Tsai et al., 2014), *med25* clustered with *med17*, *med18*, *med19a* and *med22a* in the metabolomic analysis (Davoine et al., 2017), suggesting that MED25 genetically interacts with those head module subunits. Indeed, MED25 and MED18 are both required for pathogen defense (Kidd et al., 2009; Fallath et al., 2017). In addition, oxylipin-containing galactolipids, a group of wounding-induced compounds (Stelmach et al., 2001; Buseman et al., 2006), accumulate in both *med25* and *med18* mutant plants (Davoine et al., 2017). Like MED25, MED23 did not cluster with other Mediator tail subunits in the metabolite analysis, even though the cryo-EM structure of human Mediator complex suggests it resides in the tail (Tsai et al., 2014). It is possible that the inclusion of the dehydration-responsive element binding protein 2A (DREB2A) transcription factor mutant may have altered the clustering of the *med* mutant profiles in this study (Davoine et al., 2017). Moreover, a previous study showed that another tail subunit, MED5, and MED23 genetically interact with each other to regulate phenylpropanoid metabolism (Dolan et al., 2017). At the transcriptional level, the gene expression profiles of *med5* and *med23* were strongly correlated (Dolan and Chapple, 2018), but whether those two subunits share overlapping functions in the transcriptional regulation of metabolic processes other than phenylpropanoid metabolism remains to be

tested. Given that the structural analysis of plant Mediator complex may prove challenging, functional metabolomics could provide a surrogate approach to reveal the genetic and possibly physical interactions between different subunits.

1.3 Mediator tail module subunits function in the regulation of plant cell wall biosynthesis

The plant primary cell wall, which consists of cellulose, hemicellulose, pectic polysaccharides, and a variety of cell wall proteins, is a flexible extracellular matrix that controls plant cell wall morphogenesis (Cosgrove, 2005). In addition to these components, the plant secondary cell wall contains lignin and is therefore more rigid compared, providing structural support for plant growth and development (Boerjan et al., 2003; Zhong and Ye, 2007). Several studies implicate Mediator in playing a key role in regulating both primary and secondary plant cell wall formation. For example, MED16 was identified as a genetic suppressor of mutations in *COBRA (COB)* (Sorek et al., 2015), which encodes a glycosylphosphatidylinositol anchored protein that is necessary for plant cell expansion and cellulose biosynthesis (Schindelman et al., 2001). The *cobra* mutant suffers from reduced cellulose production and retarded growth (Schindelman et al., 2001), which can be partially suppressed by disruption of MED16 (Sorek et al., 2015). A comparison of the gene expression profiles of *cob-6* and *med16 cob-6* identified two *Pectin Methylesterase Inhibitors* (*PMEI8* and *PMEI9*) that were up-regulated *med16 cob-6*, and overexpression of those two *PMEI*s partially suppressed the *cobra* phenotype (Sorek et al., 2015). These data suggest that cellulose deficiency can be compensated for by increased pectin esterification and that Mediator may be required for plants to respond to disruptions in cell wall integrity.

In addition to its role in regulating cellulose biosynthesis for primary cell wall formation, Mediator is required to maintain phenylpropanoid homeostasis, which in turn is critical for lignin biosynthesis. In both wild type and lignin-deficient mutants, the levels of phenylpropanoids are limited by MED5a and MED5b, two paralogs that reside in the tail module of Mediator (Bonawitz et al., 2012; Kim et al., 2015; Bonawitz et al., 2014; Anderson et al., 2015). In contrast, *reduced epidermal fluorescence 4* (*ref4-3*), a semi-dominant MED5 mutant characterized by a single amino acid substitution in MED5b, displays down-regulated phenylpropanoids as well as stunted growth, a phenotype frequently observed in lignin-deficient mutants (Ruegger and Chapple, 2001; Stout et al., 2008; Bonawitz et al., 2012). To investigate the mechanism by which MED5 contributes to phenylpropanoid homeostasis, a *ref4-3* suppressor screen was performed, and demonstrated that disruption of three other Mediator tail module subunits MED2, MED16 and MED23 is sufficient to suppress the stunted growth and/or down-regulated phenylpropanoids in *ref4-3* (Dolan et al., 2017). Unlike MED16 and MED23, loss of MED2 rescued only the dwarfism of *ref4-3*, but not the low phenylpropanoid phenotype (Dolan et al., 2017). More recently, a reverse genetic study showed that disruption of *CDK8*, a kinase module subunit of Mediator, and *MED12*, another kinase module subunit required for the kinase activity of CDK8, was sufficient to rescue the stunted growth of *ref4-3* but again not its low lignin content (Mao et al., 2019). Hence, identification of Mediator subunits as *ref4-3* suppressors reveals that MED5 genetically interacts with other MED subunits from both tail and kinase modules to repress phenylpropanoid metabolism and plant growth, and that the relationship between lignin deficiency and dwarfism can be disentangled genetically.

Taken together, traditional genetics studies coupled with RNA-seq and whole-genome sequencing analyses revealed that multiple MED subunits from the tail module are required for cell wall biosynthesis. At the same time, the substantial transcriptional reprogramming in *med* mutants makes it difficult to differentiate the genes that are responsible for the mis-regulated cell wall metabolism from the ones that are induced by it, a relationship that needs to be clarified in future studies.

1.4 Mediator is required for the regulation of plant metabolic pathways through direct and indirect mechanisms

As a universal transcription co-regulator, Mediator is required for the normal expression of many genes involved in plant metabolism, including those directly involved in metabolite biosynthesis as well as those function in transcriptional or post-translational regulation of the pathways (Figure 1.2). At the molecular level, an individual Mediator subunit can be required for normal expression of appropriate transcription factors, as well as the downstream biosynthetic genes involved in a given process. For instance, Arabidopsis MED15 is required for the activation of lipid biosynthesis in seedlings and mature seeds (Kim et al., 2016), where it governs the transcription of lipid metabolism-related genes through two distinct mechanisms (Kim et al., 2016). First, MED15 is required to activate the expression of *WRINKLED1* (*WRI1*), a plant-specific transcription factor critical for fatty acid biosynthesis and glycolysis (Maeo et al., 2009). Second, MED15 is recruited to the transcription starts sites of lipid metabolism-related genes by WRI1, where together they activate target gene expression (Kim et al., 2016). Thus, WRI1-targeted genes are activated by the MED15-containing Mediator complex at multiple levels, revealing the complexity of lipid biosynthesis regulation.

In addition to being involved in the regulation of biosynthetic genes or their corresponding transcriptional activators, Mediator is required for the regulation of genes involved in post-translational regulation of biosynthetic enzymes. For example, phenylalanine ammonia lyase (PAL), the first enzyme of the phenylpropanoid biosynthetic pathway, can be targeted for degradation via the ubiquitin-proteasome pathway by Kelch-domain F-box (KFB) proteins including KFB1, KFB20, KFB39 and KFB50 (Zhang et al., 2013a, 2015). Although loss of MED5 leads to modest up-regulation of phenylpropanoid biosynthetic genes including *PAL*s, it results in much more substantial up-regulation of the F-box encoding genes *KFB39* and *KFB50* (Dolan et al., 2017). These data indicate that transcriptional regulation of proteolytic turnover may be a key mechanism by which MED5 down-regulates phenylpropanoid metabolism.

Combinations of these mechanisms may be used to regulated metabolic pathway-related genes by multiple Mediator subunits. For example, Iron (Fe) uptake is a strictly balanced process in plants, given its critical role in reduction-oxidation reactions and toxicity when over-accumulated (Briat et al., 2010; Rout and Sahoo, 2015). The Arabidopsis tail module subunits MED16 and MED25 are both required for maintaining iron homeostasis (Yang et al., 2014; Zhang et al., 2014). Although the activation of *Fe-deficiency induced transcription factor* (FIT) and FIT-targeted genes depends on MED16 and MED25 (Yang et al., 2014), only MED16 can interact with FIT (Zhang et al., 2014). In contrast, MED25 interacts with ethylene-insensitive 3 (EIN3) and ethylene-insensitive 3-like 1 (EIL1) (Yang et al., 2014), two ethylene signaling transcription factors that can stabilize FIT and thereby trigger the full activation of FIT-targeted genes (Lingam et al., 2011). The mechanism of FIT activation by two distinct Mediator subunits from the same

module is reminiscent of the Wnt–β-catenin pathway in metazoans, which requires the interaction between MED12 and β-catenin (Kim et al., 2006) as well as the repressed degradation of β-catenin through CDK8 (Morris et al., 2008; Firestein et al., 2008; Zhao et al., 2013a). Therefore, Mediator can activate its downstream targets by functioning in the post-translational regulation of the corresponding transcription factors as well as directly involved in transcriptional regulation.

1.5 The function of mediator in metabolism is well conserved across kingdoms and within plant lineages

Some Mediator subunits play similar roles in plants, animals and fungi by regulating processes such as development, lipid metabolism and pathogen defense (Wang and Chen, 2004; Loncle et al., 2007; Gillmor et al., 2009; Taubert et al., 2006; Kim et al., 2015; Kidd et al., 2009; Yang et al., 2004). For example, recent studies have revealed that CDK8 and MED15 play a conserved role in lipid metabolism across kingdoms, and their function is likely conserved in other plant species. CDK8 is one of the most conserved Mediator subunits across the kingdoms (Figure 1.3). In Arabidopsis and common wheat (*Triticum aestivum L.),* CDK8 interacts with WAX INDUCER1 (WIN1), an activator for wax biosynthesis, functioning as a positive regulator in plant lipid metabolism (Zhu et al., 2014; Kong and Chang, 2018). The role of CDK8 in the activation of plant lipid biosynthesis contrasts with findings in Drosophila and mammals in which CDK8 represses lipid accumulation (Zhao et al., 2012). Interestingly, in mammals CDK8 phosphorylates the lipid metabolism-related transcription factors sterol-regulatory-element-binding protein-1c (SREBP-1c), thereby decreasing its stability and down-regulating lipid biosynthesis (Kong and Chang, 2018; Zhao et al., 2012). In wheat, however,

phosphorylation of WIN1 by CDK8 stimulates its transactivation activity, indicating that the stability of transcription factors in plant lipid metabolism is not always negatively regulated by CDK8 phosphorylation. Alternatively, CDK8 may decrease the stability of WIN1 and trigger its rapid turnover, activating WIN1-targeted genes through an "activation by destruction" mechanism (Geng et al., 2012), a phenomenon which was previously observed in an Arabidopsis MED25 study (Inigo et al., 2012).

As mentioned previously, the tail module subunit MED15 is involved in lipid metabolism in yeast, animals and plants (Taubert et al., 2006; Yang et al., 2006; Thakur et al., 2009; Kim et al., 2016). Notably, although the interacting partners vary, a kinase-inducible domain-interacting (KIX) domain resides in every characterized MED15 and is required for its association with lipid-specific transcription activators (Taubert et al., 2006; Yang et al., 2006; Thakur et al., 2009; Kim et al., 2016). In Arabidopsis, MED15 interacts with WRI1 through the KIX domain in MED15 and thereby activates lipid metabolism-related genes (Kim et al., 2016). Because of the universal role of WRI1 in plant lipid metabolism (Cernac and Benning, 2004; Liu et al., 2010; Shen et al., 2010; Ye et al., 2018) and the conserved KIX domain embedded in MED15 across the plant kingdom (Thakur et al., 2013), the interaction between WRI1 and MED15 may be a generalized mechanism for activating lipid biosynthetic genes.

1.6   Concluding remarks

Characterization of the Mediator complex in plant metabolism originated from genetic exploration of well-studied metabolic pathways and have expanded rapidly by next-generation sequencing and multi-omics analyses. Although high-throughput

transcriptome, interactome and metabolome studies have enabled us to investigate the function of Mediator from a global perspective, the massive amounts of data obtained from genome-wide analyses pose challenges as well. First, compared to the abundant gene expression profiles of different *med* mutants, information on the interaction between different MED subunits in plant Mediator is still lacking, and interactions between Mediator and transcription factors have been explored to only a limited extent (Ou et al., 2011; Kumar et al., 2018). Similarly, metabolomic studies are hindered by ambiguously annotated metabolite databases even in model plants such as Arabidopsis (Heyman and Dubery, 2016). Second, although Mediator is involved in the regulation of many plant metabolism-related genes, few studies have focused on the direct targets of plant Mediator and the mechanism by which specific MED subunits, especially those from the tail module, recruit the Mediator to target genes. Tools to enable the genome-wide analysis of Mediator localization are required to shed light on the molecular mechanisms used by Mediator to control gene expression in plants. Last, the majority of studies on plant Mediator have been focused on Arabidopsis. Since the high sequence similarity between orthologous subunits within the plant kingdom suggests a well-conserved function of Mediator in plant metabolism (Mathur et al., 2011), examining *med* mutants in other plant species, especially those that exhibit specialized secondary metabolite production, may provide critical insight into Mediator function in plant metabolism.

Table 1.1 Overview of the available transcriptome datasets in Arabidopsis

Module	Subunit	Transcriptome analysis
Head	MED6	
	MED8	
	MED11	
	MED17[a]	
	MED18[a]	Lai et al., 2014[b]; Davoine et al., 2017
	MED19[a]	
	MED20	
	MED22[a]	
	MED28	
	MED30	
Middle	MED4[a]	
	MED7	Kumar et al., 2018[b]
	MED9	
	MED10	
	MED21	
	MED31	
Tail	MED2[a]	Dolan et al., 2017[b]
	MED3[a]	
	MED5[a]	Bonawitz et al., 2014[b]; Dolan et al., 2017[b]
	MED14	Zhang et al., 2013
	MED15	Canet et al., 2012[b]
	MED16	Zhang et al., 2012; Zhang et al., 2013; Hemsley et al., 2014; Yang et al., 2014; Zhang et al., 2014; Sorek et al., 2015; Dolan et al., 2017[b]
	MED23[a]	Dolan et al., 2017
	MED25[a]	Cevik et al., 2012; Yang et al., 2014; Davoine et al., 2017
Kinase	MED12	Gillmor et al., 2014
	MED13	Gillmor et al., 2014
	CDK8	Zhu et al., 2014; Mao et al., 2019[b]
	cycC	

[a] loss-of-function mutant of this MED subunit is included in the non-targeted metabolome analysis (Davoine et al., 2017).
[b] Genes mis-regulated in this *med* mutant are enriched for genes involved in primary or secondary metabolism.
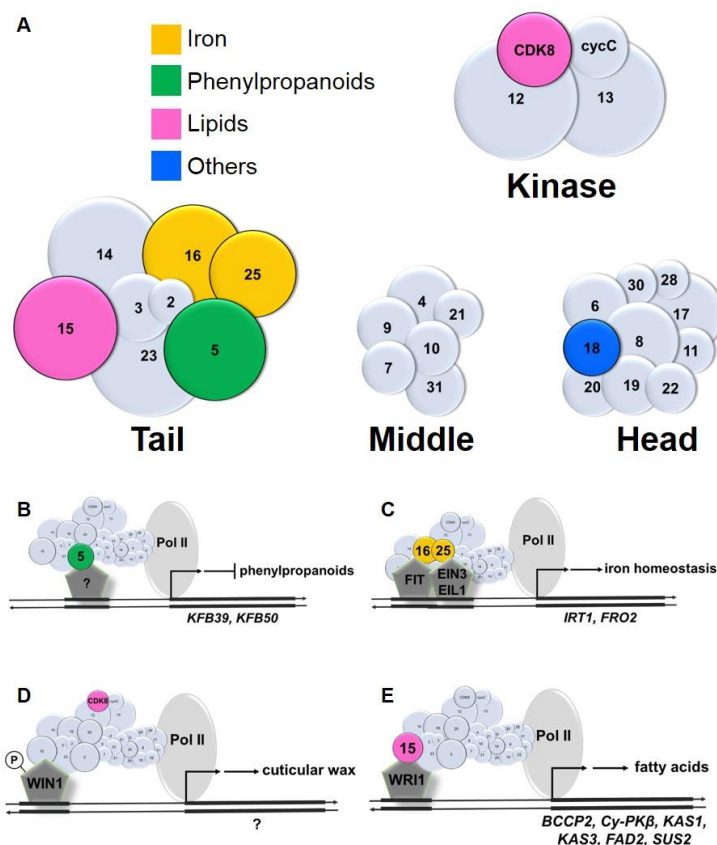
Figure 1.1 Multiple Mediator subunits from different modules are involved in plant metabolism

(A) Arabidopsis MED subunits have been assigned to head, middle, tail and Cdk8 kinase modules according to a connection map from *Saccharomyces cerevisiae* (*S. cerevisiae*). The size of each subunit depicted in the model corresponds to its molecular weight. The subunits with characterized functions in plant metabolism are colored in different colors to represent different groups of metabolites.

(B) Arabidopsis MED5 is required for the normal expression of two genes encoding Kelch-domain-containing F-box proteins (KFBs) KFB39 and KFB50 that target phenylalanine ammonia lyase (PAL), the first enzyme of phenylpropanoid pathway for degradation. This is likely the mechanism by which MED5 is involved in the negative regulation of phenylpropanoid metabolism.

(C) In Arabidopsis, MED16 and MED25 are both required for maintaining iron homeostasis. MED16 interacts with the transcription factor Fe-deficiency induced transcription factor (FIT)- and activates FIT-targeted genes. In addition, MED25 interacts with ethylene-insensitive 3 (EIN3) and ethylene-insensitive 3-like 1 (EIL1), two ethylene signaling transcription factors that can stabilize FIT and thereby trigger the full activation of FIT-targeted genes including *iron-regulated transporter 1* (*IRT1*) and *ferric chelate reductase 2* (*FRO2*).

(D) In Arabidopsis and common wheat, CDK8 interacts with WAX INDUCER1 (WIN1), a conserved activator for wax biosynthesis. WIN1 is a substrate of CDK8 in common wheat, and the CDK8-dependent phosphorylation of WIN1 is required for cuticular wax biosynthesis, however, the downstream gene targets of WIN1/CDK8 remain to be determined.

(E) Arabidopsis MED15 can interact with WRINKLED1 (WRI1) and co-activate WRI1-targeted genes, which thereby activates biosynthesis of fatty acids.
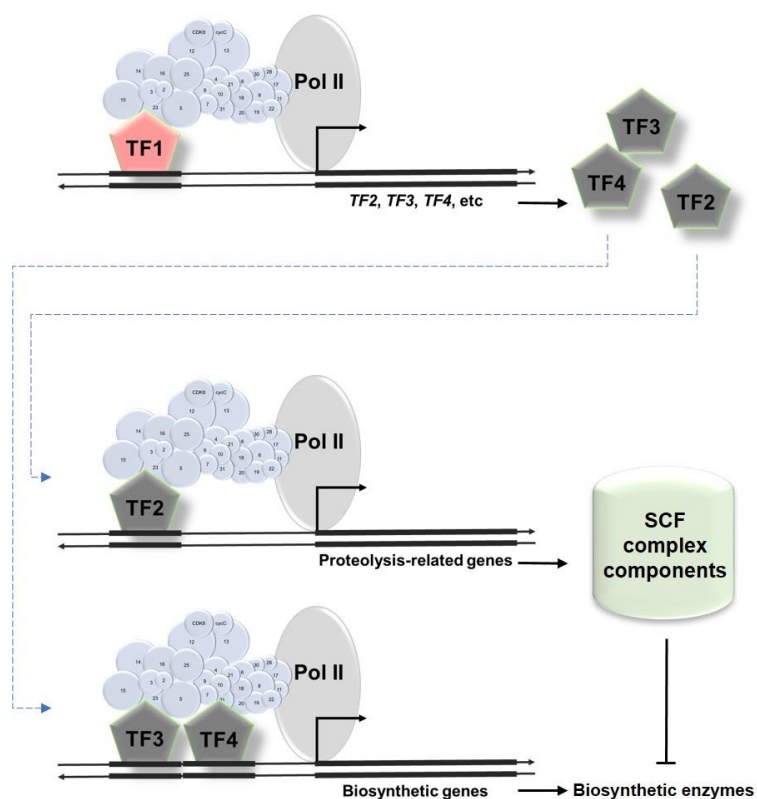
Figure 1.2 Plant metabolic pathways can be regulated by the Mediator complex at multiple levels

Mediator (blue) bridges the interaction between transcription factors (TFs) and RNA Polymerase II (Pol II) and thereby maintains the normal expression of downstream genes. First, together with some TFs (i.e., TF1), Mediator can regulate the transcription of genes encoding the TFs (grey) (top). Second, Mediator and some TFs (i.e., TF3 and TF4) can co-activate biosynthetic genes (bottom). Moreover, together with some other TFs (i.e., TF2), Mediator can co-activate the genes coding for the components of Skp1-cullin-F-box (SCF) complex, lead to proteasome-mediated turnover of the biosynthetic enzymes and thereby down-regulate the metabolite biosynthesis (middle). The indirect regulatory mechanism at the post-translational level has been observed in Arabidopsis that MED5 is required for the normal expression of genes encoding Kelch-domain-containing F-box proteins (KFBs) that target the first biosynthetic enzyme of phenylpropanoid pathway for degradation, and thereby limits phenylpropanoid biosynthesis.

Module	Subunit	AGI	S. cerevisiae	H. sapiens	Z. mays	O. sativa
Head	MED6	At3g21350	25	39	61	59
	MED8	At2g03070	17	23	63	63
	MED11	At3g01435	32	47	85	79
	MED17	At5g20170	27	29	64	63
	MED18	At2g22370	24	38	86	86
	MED19a	At5g12230	13	27	82	70
	MED20a	At4g09070	13	34	67	72
	MED20b	At2g28020	27	37	70	70
	MED22a	At1g16430	28	38	83	80
	MED22b	At1g07950	28	40	82	78
	MED28	At3g52860		39	71	69
	MED30	At5g63480		30	56	56
Middle	MED4	At5g02850	26	28	62	64
	MED7a	At5g03220	42	42	64	80
	MED9	At1g55080	24	23	45	40
	MED10a	At1g26660	41	41	74	68
	MED10b	At5g41910	38	40	75	78
	MED21	At4g04780	40	41	54	61
	MED31	At5g19910	22	37	55	61
Tail	MED2	At1g11760	16	25	66	67
	MED3	At3g09180	30	13	59	67
	MED5a	At3g23590	24	19	69	55
	MED5b	At2g48110	6	16	66	53
	MED14	At3g04740	17	31	69	67
	MED15a	At1g15780	32	21	58	59
	MED16	At4g04920	5	24	72	74
	MED23	At1g23230		30	77	76
	MED25	At1g25540		33	62	61
Kinase	MED12	At4g00450	25	25	63	61
	MED13	At1g55325	20	29	59	61
	CDK8	At5g63610	40	53	79	78
	cycCa	At5g48630	41	51	67	77
	cycCb	At5g48640	44	50	69	79

% amino acid sequence similarity compared to *A. thaliana*

Lowest within species   Highest within species

Figure 1.3 Mediator subunits are conserved within the plant kingdom

The protein sequences of Arabidopsis Mediator subunits were compared with their orthologs in *Saccharomyces cerevisiae* (*S. cerevisiae*), *Homo sapiens* (*H. sapiens*), *Zea mays* (*Z. mays*) and *Oryza sativa* (*O. sativa*). The numbers within the heatmap indicate sequence similarities (%) between Arabidopsis Mediator subunits and their orthologs in other species. Within each column, the intensities of the heatmap represent the levels of conservation of the subunits within each species.

CHAPTER 2.   MUTATION OF MEDIATOR SUBUNIT CDK8 COUNTERACTS THE STUNTED GROWTH AND SALICYLIC ACID HYPER-ACCUMULATION PHENOTYPES OF AN ARABIDOPSIS MED5 MUTANT

2.1   Introduction

The multiprotein Mediator complex comprises about 30 subunits and serves as an integrative hub for transcription regulation in eukaryotic systems (Malik and Roeder, 2010). The core Mediator complex has been subdivided into head, middle and tail domains and functions as a bridge between transcription factors and the basal transcription machinery (Asturias et al., 1999; Tsai et al., 2014). The CDK8 kinase module, which reversibly associates with the core Mediator complex, differentially regulates Mediator's activity as either a co-activator or co-repressor (Hengartner et al., 1998; Borggrefe et al., 2002). Recently, cryo-electron microscopy (cryo-EM) structures of yeast and human Mediator revealed that the association between the kinase module and the core Mediator complex is predominantly achieved through the interaction between MED13 (kinase module) and MED19 (middle module) (Tsai et al., 2013, 2014). The high-resolution cryo-EM maps not only demonstrate the interfaces between different modules of Mediator which are critical for proper transcriptional regulation (Nozawa et al., 2017; Tsai et al., 2017), but suggest an overall conserved structure of Mediator across different eukaryotic systems as well.

As in other eukaryotes (Conaway and Conaway, 2011), the Mediator complex plays a role in many aspects of plant life, including growth, development and responses to stress (Dolan and Chapple, 2017). Despite the critical nature of the complex overall, disruption of some Mediator subunit (MED) genes is not lethal in plants, and in many cases leads to

distinctive phenotypes (Yang et al., 2016; Dolan and Chapple, 2017). The ability to knock out specific subunits and study the resulting phenotypes suggests that plants can be valuable eukaryotic systems to mechanistically characterize Mediator and its involvement in plant-specific biological processes.

In addition to its role in growth and development, recent studies have demonstrated that Mediator is required for the normal regulation of secondary metabolism in Arabidopsis. Specifically, MED5a and MED5b, two MED tail subunits, are required to maintain phenylpropanoid homeostasis (Bonawitz et al., 2012). Three *reduced epidermal fluorescence 4* mutants (*ref4-1, ref4-2* and *ref4-3*) characterized by single amino acid substitutions in MED5b (D647N for *ref4-1* and *ref4-2* and G383S for *ref4-3*) were isolated as strong repressors of the phenylpropanoid pathway, indicated by decreased soluble phenylpropanoid metabolite accumulation, reduced lignin content and dwarfism (Stout et al., 2008). In contrast, disruption of *MED5a* and *MED5b* (*med5a/5b*) results in the hyper-accumulation of phenylpropanoids (Bonawitz et al., 2012), indicating that MED5 plays a widespread role in homeostatic repression of phenylpropanoid biosynthesis.

A *ref4-3* suppressor screen identified three tail subunits of Mediator, MED2, MED16 and MED23, that are required for the repressive action of *ref4-3* upon phenylpropanoid metabolism and plant growth (Stout et al., 2008; Dolan et al., 2017). Disruption of either MED16 or MED23 restores soluble phenylpropanoid accumulation and growth in *ref4-3* background, whereas loss of MED2 rescues only the dwarfism of *ref4-3* (Dolan et al., 2017). Transcriptome analysis of *ref4-3* revealed that genes encoding the enzymes in the phenylpropanoid pathway display only modest changes in expression. In contrast, negative regulators of phenylpropanoid metabolism are up-regulated compared

to wild type to an extent that is positively correlated with the level of soluble phenylpropanoid restoration in each of the suppressors (Dolan et al., 2017).

In the original *ref4-3* suppressor screen, we isolated multiple alleles of *med23* and *med16*, but only a single allele of *med2*, suggesting that the screen might not have been saturated. In addition to the tail module subunits, the dissociable CDK8 kinase module can regulate the activity of the core Mediator complex during transcription. Although CDK8 is generally recognized as a negative regulator of transcription in yeast (Kuchin et al., 1995; Rickert et al., 1999; Gonzalez et al., 2014), studies in mammalian systems indicate that CDK8 contributes to both transcriptional activation and repression (Knuesel et al., 2009; Nemet et al., 2014). Investigations in Arabidopsis revealed that CDK8 is necessary for floral organ development (Wang and Chen, 2004), mitochondrial retrograde signaling (Ng et al., 2013), pathogen defense (Zhu et al., 2014) and auxin signaling (Ito et al., 2016). Considering that CDK8 can activate down-stream gene targets in a Mediator-dependent fashion, and in *ref4-3*, negative regulators of the phenylpropanoid pathway show elevated steady state mRNA levels, we tested the hypothesis that CDK8 is required for *ref4-3* to repress phenylpropanoid metabolism and plant growth.

Here, we report that *MED5* genetically interacts with *CDK8* in Arabidopsis. Our data indicate that CDK8, and specifically its kinase activity, is required for *ref4-3* to repress plant growth. In contrast, the lignin content of *ref4-3 cdk8-1* remained low compared to wild type, indicating that low lignin content is not the cause of dwarfing in *ref4-3*. Although the phytohormone salicylic acid (SA) is hyper-accumulated in *ref4-3* and this phenotype can be suppressed by elimination of CDK8 kinase activity, blocking SA biosynthesis is not sufficient to rescue the stunted growth of *ref4-3*. In contrast, disruption of a gene encoding

a plastid-targeted DNAJ protein that is upregulated in *ref4-3* partially suppresses this growth phenotype. Together, our data demonstrate that growth inhibition, suppression of phenylpropanoid metabolism and hyper-accumulation of SA can be genetically separated in *ref4-3* mutants, and that chloroplast localized chaperones might play an unexpected role in regulating plant growth.

2.2   Results

2.2.1   Loss of CDK8 rescues stunted growth in ref4 mutants

To test whether *ref4-3* requires CDK8 to repress plant growth, we generated *ref4-3 cdk8-1* double mutants by crossing *cdk8* T-DNA insertion mutants (SALK_138675, *cdk8-1* and SALK_016169, *cdk8-2*) with *ref4-3*. Although *ref4-3* mutants exhibit a dwarf phenotype, both *ref4-3 cdk8-1* and *ref4-3 cdk8-2* are nearly normal in stature and rosette diameter (Figure 2.1, Figure 2.2A). Similarly, loss of CDK8 restores the growth defect in *ref4-1*, another allele identified from the previous *ref* screen (Ruegger and Chapple, 2001) (Figure 2.2B). Based on these results, we conclude that there is a genetic interaction between *CDK8* and *MED5* in Arabidopsis, and that CDK8 is required for *ref4-1* and *ref4-3* to repress plant growth.

We then tested whether the growth repression in *ref4-3* is dependent on the kinase activity of CDK8 using a transgene encoding a D176A kinase-dead version of the protein (Zhu et al., 2014; Kong and Chang, 2018) to generate *ref4-3 cdk8-1 35S:CDK8-MYC* (*ref4-3 CDK8*) and *ref4-3 cdk8-1 35S:CDK8^{D176A}-MYC* (*ref4-3 CDK8^{D176A}*) transgenic plants. After five weeks on soil, *ref4-3 CDK8* displayed stunted growth comparable to *ref4-3*, whereas *ref4-3 CDK8^{D176A}* looked identical to *ref4-3 cdk8-1* (Figure 2.3A). An anti-MYC

blot revealed that CDK8 was expressed at similar levels in *ref4-3 CDK8* and *ref4-3 CDK8*[D176A] (Figure 2.4) indicating that these phenotypes were not the result of different levels of transgene expression. The distinct growth phenotype between *ref4-3 CDK8* and *ref4-3 CDK8*[D176A] indicates that it is the kinase activity of CDK8 that is essential for growth repression in *ref4-3*.

In addition to CDK8, the Arabidopsis Mediator kinase module consists of C-type cyclin (cycC), MED12 and MED13. MED12 is necessary for CDK8 to demonstrate kinase activity (Knuesel et al., 2009). To independently evaluate whether CDK8 kinase activity is required for *ref4-3* phenotypes, we used a T-DNA insertion line in *MED12* (SALK_108241, *med12*) to generate *ref4-3 med12* double mutants. As with CDK8, we found that loss of MED12 is sufficient to rescue the stunted growth of *ref4-3* in terms of rosette size (Figure 2.3B). Notably, the resulting double mutants also exhibited a late flowering phenotype that was not observed in either *ref4-3* or *med12* single mutants under the same growth condition, indicating that the flowering phenotype is a result of a synthetic interaction between *ref4-3* and *med12*.

The restored growth of *ref4-3 cdk8-1* raised the question of whether CDK8 is required for growth repression in other lignin-deficient mutants. Among the *ref* mutants, *ref8-1*, a mutant deficient in the gene encoding the phenylpropanoid biosynthetic enzyme *p*-coumaroyl shikimate 3'-hydroxylase, is similar to *ref4-3* in that its phenylpropanoid metabolism is repressed and it exhibits stunted growth (Bonawitz et al., 2014). Moreover, loss of MED5 restores the stunted growth and transcriptional reprogramming of *ref8-1*, suggesting that the dwarfism of this other lignin-deficient mutant requires intact Mediator (Bonawitz et al., 2014). To elucidate the relationship between CDK8 and the stunted

growth of *ref8*, we generated *ref8-1 cdk8-1* mutants. Little, if any, growth restoration was observed in the resulting double mutants compared to *ref8-1* at multiple growth stages (Figure 2.5). Thus, our observations indicate that *ref8-1* leads to growth repression largely independent of CDK8.

2.2.2   CDK8 is not required for down-regulation of phenylpropanoids in *ref4-3*

Given that *ref4-3* plays a repressive role in phenylpropanoid metabolism that can be suppressed by several Mediator tail subunits (Dolan et al., 2017), we next determined whether loss of *CDK8* also suppresses the phenylpropanoid deficient phenotype of *ref4-3*. Because sinapoylmalate is localized in the upper epidermis, it can be readily visualized *in vivo* under ultraviolet (UV) light. As expected, *ref4-3* displayed a characteristic *ref* phenotype compared to wild type and *cdk8-1* (Figure 2.6A), indicating a decreased level of sinapoylmalate in the mutant. *ref4-3 cdk8-1* was similarly red under UV light even though the growth phenotype of *ref4-3* had been reversed (Figure 2.6A). The *ref* phenotype was also observed in *ref4-3 cdk8-2* and *ref4-1 cdk8-1* (Figure 2.7A and 2.7B). High performance liquid chromatography (HPLC) analysis confirmed these observations (Figure 2.6B), indicating that CDK8 is not required for *ref4-3* to repress sinapoylmalate biosynthesis. Similarly, in *ref4-3*, total lignin content was reduced 40% compared to wild-type and *cdk8-1* plants and in spite of the strong growth restoration seen in *ref4-3 cdk8-1*, their total lignin content remained low (Fig. 2.6C and 2.6D). Consistent with the dispensable role of CDK8 in down-regulation of phenylpropanoid metabolism in *ref4-3*, expression of either wild-type or kinase-dead *CDK8* in *ref4-3 cdk8-1* background did not cause any difference in sinapoylmalate accumulation compared to *ref4-3 cdk8-1* (Figure

2.7C). Taken together with the observation that loss of CDK8 largely rescues the stunted growth of *ref4-3*, these findings indicate that the dwarfism of *ref4-3* is independent of its restricted phenylpropanoid metabolism.

2.2.3 The stunted growth of ref4-3 is not dependent on the phosphorylation event introduced by the G383S mutation

Because the G383S mutation in the *ref4-3* allele introduces a potential phosphorylation site, we wondered if the defects in growth and phenylpropanoid metabolism observed in *ref4-3* plants could result from ectopic/hyper-phosphorylation of MED5 by one or more kinases, possibly including CDK8 (Stout et al., 2008). We also considered whether the increased side-chain size of S383 in *ref4-3* could itself lead to these phenotypes, independent of phosphorylation status. To distinguish between these possibilities, we generated a series of *med5* constructs in which various site-directed mutants were expressed under the control of the *CINNAMATE 4-HYDROXYLASE* (*C4H*) promoter (Bonawitz et al., 2012) such that the transgenes would be expressed in cells involved in phenylpropanoid metabolism. *MED5b* transgene was expressed at similar level across all different transgenic lines, none of which was less than the expression of *MED5b* in wild-type plants (Figure 2.8). We then assayed transgenic *med5a/5b* double mutant plants carrying these constructs for sinapoylmalate and lignin content. MED5a and MED5b share semi-redundant function in repression of phenylpropanoid metabolism and *med5a/5b* double mutants have increased sinapoylmalate and lignin content compared to wild type (Bonawitz et al., 2012); thus, expression of *C4H:MED5b* constructs with wild-type function should restore levels to that of the single *med5a* mutant alone. Indeed, the control *MED5b*[G383G] transgenic displayed normal growth and accumulated sinapoylmalate and

lignin content similar to that in *med5a* (Figure 2.9). In contrast, the *ref4-3* mimic, *MED5b*^{G383S}, showed similar dwarf phenotypes as compared with *ref4-3* and accumulated less phenylpropanoids compared to *med5a* (Figure 2.9). Expression of *MED5b*^{G383T} containing an alternative phosphorylation site, or the phospho-mimics *MED5b*^{G383D} and *MED5b*^{G383E}, also resulted in dwarfing and reduced phenylpropanoid levels (Figure 2.9). In contrast, expression of the non-phosphorylatable *MED5b*^{G383A} and *MED5b*^{G383V} had differing effects on plant growth and phenylpropanoid metabolism; *MED5b*^{G383A} was slightly dwarf and showed similar sinapoylmalate levels compared to the *MED5b*^{G383G} control, but *MED5b*^{G383V} plants were more stunted and showed sinapoylmalate levels that were intermediate between the *MED5b*^{G383G} control and *MED5b*^{G383S} (Figure 2.9A and 2.9B). Because neither A nor V can be phosphorylated, but these showed different plant growth phenotypes and sinapoylmalate levels, we conclude that the G383 residue is important to the function of MED5b, and the increased side-chain size at position 383 caused by substitution of S for G is likely responsible for the plant growth and phenylpropanoid phenotypes associated with *ref4-3*. In contrast to sinapoylmalate, lignin levels were reduced in both *MED5b*^{G383A} and *MED5b*^{G383V} (Figure 2.9C). Thus, the reduced phenylpropanoid accumulation in *ref4-3* is likely independent of the phosphorylation event introduced by the G383S mutation. We note that although the D647N mutation in *ref4-1* does not introduce a novel phosphorylation in MED5b, *ref4-1* mutants can also be rescued by loss of CDK8, further suggesting that the *ref4* phenotypes are not dependent upon CDK8-mediated phosphorylation of MED5b.

2.2.4   Disruption of *CDK8* partially rescues the transcriptional reprogramming of the *ref4-3* mutants

Our observations on *ref4-3 cdk8-1* indicate that the phenotypes of *ref4-3* plants can be genetically separated, and that dwarfism in these plants may result from aberrant gene expression in biological processes other than lignin biosynthesis. we performed messenger RNA sequencing (RNA-seq) using three-week-old rosettes of *ref4-3 cdk8-1* together with wild type, *cdk8-1*, *ref4-1* and *ref4-3*. Principle component analysis (PCA) revealed a clear clustering of samples by genotype (Figure 2.10). We next determined the differentially expressed gene set in each mutant compared to wild type, and performed a gene ontology (GO) term analysis focused on biological processes. Compared to *ref4-3*, *ref4-1* is a weaker allele in terms of reduced phenylpropanoid accumulation and stunted growth (Figure 2.11A). In *ref4-1*, 2927 genes were differentially expressed (Figure 2.11B). More substantial gene expression changes were observed in *ref4-3*, which included 7770 mis-regulated genes, representing more than one-third of the expressed genes (count per million (CPM) reads > 1 in 3 or more samples) (Figure 2.11B). We noticed that over 90% of the differentially expressed genes in *ref4-1* were also mis-regulated in *ref4-3* but with larger fold change (Figure 2.11B and 2.11C). This finding not only suggests that the point mutations in *ref4-1* and *ref4-3* lead to a widespread transcriptional reprogramming by similar mechanisms, but also indicates that our RNA-seq analysis captures subtle differences in gene expression between alleles. Consistent with our previous observation (Dolan et al., 2017), GO term analysis of the up-regulated genes in *ref4-3* showed an enrichment of genes involved in different stress responses and transcription regulation (Table 2.2). Up-regulated genes in *ref4-1* were also enriched for stress responses except for response to UV-B and salt stress, and transcriptional regulation (Table 2.2). In contrast

to the up-regulated genes, genes that were down-regulated in *ref4-1* and *ref4-3* were enriched for those involved in photosynthesis (Table 2.3). In *ref4-3*, genes related to ribosome biogenesis and cytokinin response were also enriched among the down-regulated genes (Table 2.3).

We identified 4053 genes that were mis-regulated in *cdk8-1*, 60% of which displayed reduced expression compared to wild type (Figure 2.12). Among the genes up-regulated in *cdk8-1*, only a limited number of GO terms were enriched, namely response to light, photosynthesis, and microtubule-based movement (Table 2.2). In contrast, transcripts related to defense response were significantly over-represented among the down-regulated genes (Table 2.3), which is consistent with the reported function of CDK8 in biotic stress responses (Zhu et al., 2014).

Although *ref4-3 cdk8-1* displays wild-type growth, a significant number of genes remained mis-regulated in the double mutant, including 3767 up-regulated genes and 4537 down-regulated genes when compared to wild type. Transcripts associated with response to water deprivation and abscisic acid (ABA) were the most significantly enriched among up-regulated genes (Table 2.2), whereas the down-regulated genes were enriched for those involved in defense responses (Table 2.3). Consistent with their *ref* phenotypes, many phenylpropanoid biosynthetic genes were down-regulated in *ref4-3*, and most of them were not rescued in *ref4-3 cdk8-1* (Figure 2.13).

Our phenotypic analysis revealed that *ref4-3* requires CDK8 to repress plant growth but not phenylpropanoid metabolism. To identify the genes that are associated with the dwarf phenotype of *ref4-3*, we focused on the genes that displayed altered expression in *ref4-3* compared to wild type, but whose expression was rescued in *ref4-3 cdk8-1* (Figure

2.14A). In total, 73 genes were significantly down-regulated in *ref4-3* compared to wild type and displayed at least a partially restored expression in the absence of CDK8 (FDR < 0.05, absolute value of $\log_2$FC > 1) (Figure 2.14B, Table 2.4). GO term analysis revealed that within this gene set, genes associated with regulation of organ growth, photosynthesis and auxin-related signaling pathway were over-represented (Table 2.5). In contrast, 378 genes were significantly up-regulated in *ref4-3* compared to wild type, the abnormal expression of which was at least partially alleviated in *ref4-3 cdk8-1* (Figure 2.14C, Table 2.6). The most significantly enriched GO categories within this gene set included suberin biosynthesis, lipid transport and defense responses (Table 2.7).

2.2.5   SA biosynthesis and signaling are activated in *ref4-3* but not in *ref4-3 cdk8-1*

GO term analysis revealed that genes involved in defense responses, especially those respond to SA, were up-regulated in *ref4-3* in a CDK8-dependent manner (Table 2.7). Previous studies proposed that hyper-activated SA biosynthesis and signaling leads to dwarfism of hydroxycinnamoyl CoA: shikimate hydroxycinnamoyl transferase (HCT)-RNA interference (RNAi) transgenics in both Arabidopsis and alfalfa (Gallego-Giraldo et al., 2011a, 2011b). Thus, we wondered if the aberrant activation of SA signaling could be the cause of dwarfism in *ref4-3*. In *ref4-3*, genes encoding proteins involved in SA biosynthesis and storage (Figure 2.15A) (Ward et al., 1991; Yang et al., 1997; D'Maris Amick Dempsey et al., 2011) were significantly increased in expression compared to wild type and *cdk8-1* (Figure 2.15B). SA signaling marker genes were also up-regulated in *ref4-3* (Figure 2.15C). Loss of CDK8 in *ref4-3* eliminated the up-regulation of SA biosynthetic and signaling genes (Figure 2.15B and 2.15C), indicating that *ref4-3* requires CDK8 to

activate SA signaling. Consistent with these observations, there was an enhanced accumulation of both free SA and SA conjugates in *ref4-3*, which was blocked by loss of *CDK8* (Figure 2.15D and 2.15E). Moreover, both free SA and SA conjugates were hyper-accumulated in *ref4-3 CDK8*, whereas *ref4-3 CDK8*^{D176A} plants accumulated wild-type level of SA (Figure 2.15D and 2.15E), indicating that over-accumulation of SA in *ref4-3* is dependent on the kinase activity of CDK8. Taken together, our RNA-seq analysis and SA measurement demonstrate that CDK8 and its kinase activity is necessary for the hyper-accumulation of SA in *ref4-3*.

To test whether SA accumulation in *ref4-3* leads to the growth defects in these plants, we crossed *ref4-3* with *salicylic acid induction deficient 2* (*sid2-4*), a mutant defective in isochorismate synthase 1. We found that the growth phenotype of *ref4-3 sid2-4* was unchanged relative to *ref4-3* (Figure 2.15F), even though HPLC analyses revealed that both free SA and total SA levels were reduced to below wild-type levels in *ref4-3 sid2-4* (Figure 2.15G). We also used a mutant line with disruption in *NPR1* (CS_3726, *npr1-1*) (Cao et al., 1997), an essential regulator of SA signaling, to generate a *ref4-3 npr1-1* double mutant. Whereas *npr1-1* mutants displayed wild-type growth, *ref4-3 npr1-1* was indistinguishable from *ref4-3* (Figure 2.16). These data indicate that SA accumulation is not the cause of dwarfing in the mutant.

2.2.6   Enhanced auxin accumulation is not sufficient to restore the stunted growth of *ref4-3*

Our GO term analysis suggested that auxin signaling is perturbed in *ref4-3* and rescued in *ref4-3 cdk8-1* (Table 2.5). Multiple genes involved in auxin signaling, including indole-3-acetic acid (IAA) induced genes such as *IAA1*, *IAA7* and *IAA29*, as well as small

auxin up RNA (SAUR) genes including *SAUR20*, *SAUR22* and *AT5G18010* were down-regulated in *ref4-3* compared to wild type (Figure 2.17A). In contrast, disruption of *CDK8* resulted in up-regulation of all the genes mentioned above, which is consistent with a previous finding that CDK8 kinase module plays a repressive role in auxin transcriptional responses (Ito et al., 2016). Except for *IAA29,* which was up-regulated in *ref4-3 cdk8-1* compared to wild type, the other five genes displayed wild-type expression in the absence of CDK8 in the *ref4-3* background, suggesting that *ref4-3* represses auxin signaling in a CDK8-dependent fashion.

Given that auxin plays a critical role in plant growth (Teale et al., 2006), we next aimed to determine whether repressed auxin signaling contributes to the dwarfism of *ref4-3*. A previous study demonstrated that YUCCA6 (YUC6) functions in tryptophan-dependent auxin biosynthesis, and the dominant mutant *yuc6-1D* is sufficient to cause hyperaccumulation of auxin (Kim et al., 2007). We therefore constructed a *ref4-3 yuc6-1D* double mutant and evaluated its growth phenotype. Although introduction of *yuc6-1D* into *ref4-3* led to activation of the auxin-responsive gene *At4g02520* (Figure 2.17B) (Smith et al., 2003) as well as high-auxin developmental phenotypes including elongated petioles and narrow leaves (Figure 2.17C), *ref4-3 yuc6-1D* was as dwarf as *ref4-3* (Figure 2.17D), indicating that repressed auxin signaling in *ref4-3* is probably not the leading cause for its stunted growth.

2.2.7 Disruption of a DNA J PROTEIN C66 (DJC66) partially restores the growth deficiency of *ref4-3*

Among the genes that showed greatest mis-regulation in *ref4-3*, *DJC66*, a gene encoding a small J-domain containing protein, was up-regulated more than 23-fold in *ref4-*

3 compared to wild type, and its expression was partially rescued in all *ref4-3* suppressors including *ref4-3 cdk8-1* (Figure 2.18A and 2.18B). While DJC66 has not been functionally characterized, it was proposed to be critical for leaf growth because of its interaction with anaphase-promoting complex subunit 8 (APC8) (Arabidopsis Interactome Mapping Consortium, 2011; Schulz et al., 2014), a protein involved in cell cycle progression (Eloy et al., 2015). To test whether DJC66 is required for the dwarfism of *ref4-3*, we crossed *ref4-3* to a T-DNA insertion line in *DJC66* (SALK_149745C, *djc66*) and generated *ref4-3 djc66* double mutants. Compared to *ref4-3*, *ref4-3 djc66* displayed modest but significant growth restoration (Fig. 2.18C, 2.18D, 2.18E and 2.18F), suggesting that the stunted growth of *ref4-3* is partially dependent on DJC66. Moreover, like *ref4-3*, *ref4-3 djc66* had lower levels of sinapoylmalate and lignin compared to wild type, indicating that the protein is not involved in the suppression of phenylpropanoid metabolism and may function specifically in the dwarfing phenotype of the mutant (Fig. 2.18G and 2.18H). Taken together, our data reveal that DJC66 is a novel suppressor that partially suppresses the stunted growth of *ref4-3*. Further, unlike all previous *ref4-3* suppressors, DJC66 presumably functions independent of Mediator's role in transcriptional regulation because it is localized to the plastid, rather than the nucleus (Chiu et al., 2013).

2.3   Discussion

In this study, we used the Arabidopsis *ref4-3* mutant to examine the function of MED5 in the context of Mediator. *ref4-3* carries a missense mutation in *MED5b*, and exhibits dwarfism and reduced phenylpropanoids (Stout et al., 2008; Bonawitz et al., 2012). Loss of MED5 leads to increased phenylpropanoid accumulation in an otherwise wild-type

genetic background (Bonawitz et al., 2012), and disruption of *MED5a* and *MED5b* can restore the phenylpropanoid-deficient phenotypes of other *ref* mutants (Anderson et al., 2015; Kim et al., 2015). Thus, our data suggest that the proteins encoded by semi-dominant *ref4* alleles mimic the action of wild-type MED5 in homeostatic repression of phenylpropanoid biosynthesis, and thus provide genetic tools that are complementary to biochemical approaches to investigate the interaction between MED5 and other transcriptional regulators.

We previously reported that loss of Mediator tail module subunits MED2, MED16 or MED23 relieves the growth defects of *ref4-3* (Dolan et al., 2017). Here, we show that loss of CDK8, a kinase module subunit, has a similar effect. Unlike disruption of MED16 or MED23, loss of CDK8 does not restore the restricted lignin biosynthesis in *ref4-3*, which again demonstrates that the stunted growth and reduced lignin content of *ref4-3* can be genetically disentangled as was found for MED2 (Dolan et al., 2017).

The identification of CDK8 as a novel *ref4-3* suppressor also provides new evidence for the functional/genetic, and potentially physical, interaction between the tail and kinase modules of Mediator. Arabidopsis CDK8 functions together with MED25 to activate the pathogen defense marker gene *PDF1.2* (Zhu et al., 2014), and physical interaction between MED5 and the kinase module has been suggested by several studies in mammalian cells (Ito et al., 2002; Knuesel et al., 2009). Our study suggests that the interaction between MED5 and CDK8 may be preserved in the Arabidopsis Mediator complex. Alternatively, the genetic interaction between CDK8 and MED5 may reflect a functional but indirect interaction between these two subunits. Recent cryo-EM structures of yeast Mediator complex revealed that the CDK8 kinase module can reversibly associate

with the head and middle module through the interaction between MED13 (kinase module) and MED19 (middle module) (Tsai et al., 2013), whereas MED5, embedded in the tail, is located distal to those two modules (Tsai et al., 2014). Given that the overall structure of Mediator is conserved in eukaryotic systems (Tsai et al., 2014), the available high-resolution map of yeast Mediator (Tsai et al., 2013, 2014) suggests that MED5 and CDK8 do not physically interact with each other. Nevertheless, the potential physical interaction between different Mediator subunits in plants still needs to be evaluated by future studies.

Although some genetic studies have shown that CDK8 has functions independent of its kinase activity (Zhu et al., 2014), our data demonstrate that CDK8 kinase activity is required for both growth deficiency and increased SA accumulation in *ref4-3*, consistent with the critical role of CDK8 with intact kinase activity in retrograde signaling and stress response (Ng et al., 2013; Zhu et al., 2014). Because CDK8 is dispensable for normal phenylpropanoid accumulation whereas MED5 is critical for this process, it is unlikely that wild-type MED5 is a general substrate of CDK8, and that the phosphorylation of MED5 by CDK8 is required for phenylpropanoid homeostasis.

The interacting partners and/or substrates of Arabidopsis CDK8 remain to be identified, but in other eukaryotes include the C-terminal domain of Pol II, histone proteins, individual Mediator tail subunits including MED2 and MED3 (Hallberg et al., 2004; Gonzalez et al., 2014) and various transcription factors (Rzymski et al., 2015; Lee et al., 2016). Notably, a recent study in common wheat revealed that CDK8 can phosphorylate the transcription factor wax inducer 1 (TaWIN1), which thereby activates TaWIN1-targeted genes and promotes very-long-chain aldehyde biosynthesis (Kong and Chang, 2018). The identification of TaWIN1 as a target of CDK8 suggests that besides substrates

of CDK8 common to all eukaryotes, CDK8 may phosphorylate plant-specific transcription factors, possibly including those that are necessary for growth inhibition of *ref4-3*.

Many plant hormones including SA and auxin play critical roles in the cross-talk between growth and immunity (Kazan and Manners, 2009; Huot et al., 2014). Although the stunted *ref4-3* mutant hyper-accumulates SA, our data suggest that the SA content of the mutant is unrelated to its dwarfism. Similarly, dwarfism of *ref8-1*, another lignin-deficient mutant, is also independent of its SA accumulation (Bonawitz et al., 2014). Thus, hyper-accumulation of SA is not a universal mechanism underpinning dwarfism in lignin-deficient mutants. Moreover, our data further show that the repressed growth in *ref4-3* is likely independent of auxin signaling. Together, we conclude that perturbation of hormone signaling is not the underlying cause for dwarfism associated with lignin deficiency.

Although *ref4-3* and *ref8-1* show multiple similarities including repressed phenylpropanoid metabolism, significant changes in their transcriptome and growth deficiency independent of SA, CDK8 is a suppressor of *ref4-3*, but not of *ref8-1*. In fact, while multiple MED subunits were identified as suppressors of *ref4-3* (Dolan et al., 2017), MED5 is the only characterized suppressor that can fully restore the growth of *ref8-1* (Bonawitz et al., 2014). The difference between *ref4-3* and *ref8-1,* as well as previously identified low-lignin mutants, indicates that multiple mechanisms exist for dwarfing in plants that co-occur with perturbed phenylpropanoid metabolic phenotypes. Specifically, the dwarfism of *ref4-3* may result from abnormal transcriptional reprogramming achieved by mutated MED5b itself, whereas the stunted growth of *ref8-1* is due to restricted flux through phenylpropanoid pathway or an abnormal response triggered by over-

accumulation of phenylpropanoid pathway intermediates that requires wild-type MED5 for perception.

Our study raises the possibility that an alternative mechanism involving chaperone pathways might be involved in *ref4-3* associated dwarfism. We identified *DJC66*, encoding a co-chaperone DnaJ protein, as a highly-upregulated gene in *ref4*-3 that was partially rescued by loss of CDK8. Similar to elimination of CDK8, loss of DJC66 suppresses the stunted growth of *ref4-3* but does not affect phenylpropanoid biosynthesis. DJC66 interacts with the anaphase-promoting complex subunit APC8, suggesting its potential role in cell cycle regulation and plant growth (Schulz et al., 2014; Eloy et al., 2015). In addition, DJC66 can be targeted to chloroplasts, and its expression is significantly induced under heat and cold stresses (Chiu et al., 2013). Given that CDK8 is essential for retrograde signaling and general abiotic stress responses (Ng et al., 2013), it is likely that DJC66 functions downstream of CDK8 in growth repression of *ref4-3* (Figure 2.19). Moreover, the partial growth restoration in *ref4-3 djc66* and the fact that DJC66 is only one of the DnaJ cochaperones (Chiu et al., 2013) suggest that other DnaJ proteins may share redundant function with DJC66 and contribute to the stunted growth of *ref4-3* as well.

2.4   Methods

2.4.1   Plant material and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was the wild type in this study. Plants were cultivated at a temperature of 23°C, under a long-day photoperiod (16 hr light/8 hr dark) with a light intensity of 100 $\mu E\,m^{-2}\,s^{-1}$. Homozygous mutants used in this study were isolated based on previous reports, with the corresponding accession numbers and

primers listed in Table 2.1 (Cao et al., 1997; Kim et al., 2007; Bonawitz et al., 2012; Zhu et al., 2014).

2.4.2 Transgenic plants

Plant binary vectors (pBA-myc) that carry either *CDK8* or *CDK8*^{D176A} driven by the 35S promoter were transformed into *ref4-3 cdk8-1* mutants by *Agrobacterium tumefaciens*-mediated transformation (Clough and Bent, 1998; Zhu et al., 2014). Similarly, a series of *MED5b*^{G383*} ('*' represents amino acids G, S, T, D, E, A and V) constructs were first cloned into pCC0996, a binary vector in which transgene expression is driven by the Arabidopsis *C4H* promoter (Bonawitz et al., 2012) and transformed into *med5* mutants. Transgenic lines were selected based on their resistance to Basta. The homozygous lines identified in the T3 generation were used for phenotypic characterization. To determine the CDK8 protein levels in the selected transgenic lines, 0.5 g of two-week-old seedlings were harvested and prepared for crude protein extracts in 1 mL Tris-HCl buffer (150 mM, pH 8.0). After centrifugation, 50 µL lysate from each sample was loaded on 10% SDS-page gel and protein gel blotting was performed using anti-MYC antibody (1:1000 dilution, Sigma M4439).

2.4.3 Lignin analysis

Total lignin content was quantified using extractive free cell walls by thioglycolic acid (TGA) lignin analysis, as described previously (Li et al., 2015).

2.4.4   HPLC analysis of secondary metabolites

Sinapoylmalate content of three-week-old whole rosettes was quantified by HPLC as previously reported (Dolan et al., 2017).

2.4.5   High-throughput mRNA sequencing

Samples of wild-type, *ref4-1*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* three-week-old rosettes were harvested in triplicate with a randomized design. Each sample contained five whole rosettes of the same genotype from five individual pots. RNA extraction and whole-transcriptome sequencing were performed as previously described (Dolan et al., 2017). The RNA-seq data of this study have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002) with accession number GSE111290. The previous RNA-seq data of *ref4-3* and its suppressors have been deposited in Gene Expression Omnibus with accession number GSE95574 (Dolan et al., 2017).

2.4.6   Differential expression analysis

Count matrices for individual samples were generated for each gene using HTSeq-count (Anders et al., 2015). Differential expression analysis was performed based on the result of expressed genes with greater than one count per million (CPM) reads in at least three samples. The filtered data was then subjected to edgeR (v3.12.1) analysis using generalized linear model (GLM) approach (McCarthy et al., 2012), which was performed using statistical program R (v3.4.1). Significance testing was performed and adjusted using Benjamin-Hochberg method, reported as False Discovery Rate (FDR) with a cut-off at FDR < 0.05. Venn diagrams were created with the online tool Venny (v 2.1,

http://bioinfogp.cnb.csic.es/tools/venny/) and GO term analysis was performed using the

DAVID Bioinformatics Resource (v.6.8, https://david.ncifcrf.gov/) (Huang et al., 2008).

2.4.7   Determination of SA levels

SA extraction and detection was performed as previously described (Rozhon et al.,

2005). Both free and total SA were quantified by HPLC using 2-methoxybenzoic acid as

an internal standard.

Table 2.1 Primers used in this study

Gene	Accession number	Mutant line	Name		Sequence
CDK8	AT5G63610	SALK_138675	*cdk8-1*	LP	TTGGTCTTGGCATCGATCTAC
			cdk8-1	RP	TTGGTGAAGGCACTTATGGTC
CDK8	AT5G63610	SALK_016169	*cdk8-2*	LP	CACTATTCCGTGCTCTTCTGC
			cdk8-2	RP	TTTCCTGATCGTCGATTTTTG
MED12	AT4G00450	SALK_108241	*med12*	LP	TCCATTTTGCTTTACTGCAGG
			med12	RP	TCGATCATCCCGCTAACTATG
SID2	AT1G74710	SALK_133146	*sid2-4*	LP	TCTGGGCTCAAACACTAAAACAC
			sid2-4	RP	GAATCAGAGGTGACGTTGAAGAC
NPR1	AT1G64280	CS_3726	*npr1-1*	F	CGTGTGCTCTTCATTTCGCTGT
			npr1-1	R	GTGCGGTTCTACCTTCCAAAGTT
YUCCA6	AT5G25620		*yuc6-1D*	LP	TGGTACTAATTCAGCAAT
			yuc6-1D	RP	ACTCTACGTACATTGAAG
DJC66	AT3G13310	SALK_149745C	*djc66*	LP	TGATCTAATGGCAACGATTCC
			djc66	RP	CGCTGTAGAATCTGGCTGTTC
GSTF2	AT4G02520			F	CATCGCCCTCCACGAGAA
				R	GCTCACCGTCTTTGAGTTCGA
MED5b	AT2G48110			F	CCGATGCTAAGAGGCTATGC
				R	CGACGTTTTGATGCAGGATA
AT1G13220	AT1G13220			F	TAACGTGGCCAAAATGATGC
				R	GTTCTCCACAACCGCTTGGT

Unless specified, all mutant lines were obtained from the Arabidopsis Biological Resource Center (The Ohio State University, Columbus, OH). Left primers (LP) and right primers (RP) were used. For *npr1-1*, forward (F) and reverse (R) primers were used to amplify the region carrying the point mutation followed by restriction enzyme digestion. Expression level of *GSTF2* and *MED5b* was determined by quantitative PCR using the listed F and R primers and normalized to the reference gene *AT1G13220*.

Table 2.2 GO term analysis of the genes that are up-regulated in *ref4-1*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* compared to wild type respectively [a]

Genotype	Term	Description	Count [b]	Fold Enrichment [c]	FDR
ref4-1					
	GO:0009751	response to salicylic acid	40	2.62	3.88E-05
	GO:0009414	response to water deprivation	60	2.28	2.75E-06
	GO:0042742	defense response to bacterium	59	2.17	2.73E-05
	GO:0009737	response to abscisic acid	80	2.12	2.06E-07
	GO:0006952	defense response	102	1.84	2.08E-06
	GO:0043565	sequence-specific DNA binding	100	1.79	1.10E-05
	GO:0005576	extracellular region	227	1.57	2.09E-09
	GO:0003700	transcription factor activity, sequence-specific DNA binding	209	1.56	2.06E-08
	GO:0016301	kinase activity	129	1.48	7.16E-03
ref4-3					
	GO:0009751	response to salicylic acid	40	2.62	3.88E-05
	GO:0009414	response to water deprivation	60	2.28	2.75E-06
	GO:0042742	defense response to bacterium	59	2.17	2.73E-05
	GO:0009737	response to abscisic acid	80	2.12	2.06E-07
	GO:0006952	defense response	102	1.84	2.08E-06
	GO:0043565	sequence-specific DNA binding	100	1.79	1.10E-05
	GO:0005576	extracellular region	227	1.57	2.09E-09
	GO:0003700	transcription factor activity, sequence-specific DNA binding	209	1.56	2.06E-08
	GO:0016301	kinase activity	129	1.48	7.16E-03
cdk8-1					
	GO:0009768	photosynthesis, light harvesting in photosystem I	13	7.16	4.50E-05
	GO:0018298	protein-chromophore linkage	18	5.74	2.96E-06
	GO:0007018	microtubule-based movement	23	4.50	2.67E-06
	GO:0010114	response to red light	17	3.75	9.09E-03
	GO:0015979	photosynthesis	30	3.08	1.05E-04
	GO:0009416	response to light stimulus	39	2.72	3.32E-05
ref4-3 cdk8-1					
	GO:0007623	circadian rhythm	38	2.26	1.38E-03
	GO:0009414	response to water deprivation	90	1.96	1.19E-07
	GO:0009737	response to abscisic acid	117	1.77	2.20E-07
	GO:0003700	transcription factor activity, sequence-specific DNA binding	361	1.53	2.27E-16
	GO:0006351	transcription, DNA-templated	385	1.45	9.33E-13

Table 2.2 continued

Genotype	Term	Description	Count [b]	Fold Enrichment [c]	FDR
	GO:0006355	regulation of transcription, DNA-templated	428	1.41	7.55E-13
	GO:0003677	DNA binding	397	1.38	4.80E-10
	GO:0005634	nucleus	1536	1.12	8.84E-08

a. For Supplemental Table 1,2 4 and 6, only significant GO terms were illustrated (FDR < 0.05).
b. 'Count' represents the number of genes that were used as queries and related to the specific GO term.
c. 'Fold Enrichment' was normalized to the percentage of background genes (all expressed genes) involved in the corresponding GO term.

Table 2.3 GO term analysis of the genes that are down-regulated in *ref4-1*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* compared to wild type respectively

Genotype	Term	Description	Count	Fold Enrichment	FDR
ref4-1					
	GO:0042026	protein refolding	8	7.53	1.50E-02
	GO:0009768	photosynthesis, light harvesting in photosystem I	17	6.54	8.09E-08
	GO:0007017	microtubule-based process	14	4.94	9.57E-04
	GO:0015995	chlorophyll biosynthetic process	21	4.34	9.30E-06
	GO:0018298	protein-chromophore linkage	18	4.01	6.35E-04
	GO:0010411	xyloglucan metabolic process	15	3.74	2.21E-02
	GO:0008652	cellular amino acid biosynthetic process	29	3.56	1.79E-06
	GO:0015979	photosynthesis	44	3.16	5.07E-09
	GO:0006633	fatty acid biosynthetic process	36	2.63	1.44E-04
	GO:0009416	response to light stimulus	46	2.24	3.67E-04
	GO:0071555	cell wall organization	55	1.97	1.99E-03
	GO:0046686	response to cadmium ion	74	1.89	1.15E-04
	GO:0055114	oxidation-reduction process	198	1.51	1.39E-06
ref4-3					
	GO:0032544	plastid translation	11	4.31	1.37E-02
	GO:0009768	photosynthesis, light harvesting in photosystem I	19	4.06	3.90E-06
	GO:0015995	chlorophyll biosynthetic process	33	3.78	2.09E-11
	GO:0019253	reductive pentose-phosphate cycle	14	3.46	2.23E-02
	GO:0015979	photosynthesis	82	3.26	2.79E-25
	GO:0000162	tryptophan biosynthetic process	15	3.20	3.57E-02
	GO:0000027	ribosomal large subunit assembly	21	3.08	8.95E-04
	GO:0002181	cytoplasmic translation	29	2.84	4.16E-05
	GO:0000028	ribosomal small subunit assembly	21	2.82	6.59E-03
	GO:0015991	ATP hydrolysis coupled proton transport	19	2.79	2.72E-02
	GO:0010411	xyloglucan metabolic process	20	2.76	1.80E-02
	GO:0042254	ribosome biogenesis	81	2.76	6.77E-18
	GO:0018298	protein-chromophore linkage	22	2.72	7.84E-03
	GO:0008652	cellular amino acid biosynthetic process	36	2.45	1.03E-04
	GO:0006412	translation	222	2.41	1.15E-40
	GO:0009735	response to cytokinin	86	2.35	3.77E-13
	GO:0006633	fatty acid biosynthetic process	57	2.31	1.50E-07
	GO:0009658	chloroplast organization	55	2.08	3.27E-05

Table 2.3 continued

Genotype	Term	Description	Count	Fold Enrichment	FDR
	GO:0009416	response to light stimulus	68	1.84	2.50E-04
	GO:0009409	response to cold	98	1.62	5.85E-04
	GO:0046686	response to cadmium ion	111	1.58	3.93E-04
	GO:0055114	oxidation-reduction process	323	1.37	2.15E-07
cdk8-1					
	GO:0042343	indole glucosinolate metabolic process	11	5.21	1.16E-02
	GO:0010200	response to chitin	72	4.57	8.10E-29
	GO:0009625	response to insect	13	4.19	2.26E-02
	GO:0009617	response to bacterium	45	3.99	5.90E-14
	GO:0009626	plant-type hypersensitive response	35	3.92	6.15E-10
	GO:0098542	defense response to other organism	15	3.90	1.06E-02
	GO:0009816	defense response to bacterium, incompatible interaction	20	3.75	3.66E-04
	GO:0009620	response to fungus	32	3.68	5.41E-08
	GO:0009611	response to wounding	80	3.34	1.27E-20
	GO:0051707	response to other organism	21	3.32	1.82E-03
	GO:0009751	response to salicylic acid	58	3.08	3.05E-12
	GO:0009636	response to toxic substance	22	3.06	4.56E-03
	GO:0042542	response to hydrogen peroxide	20	3.04	1.59E-02
	GO:0050832	defense response to fungus	73	2.84	5.42E-14
	GO:0042742	defense response to bacterium	91	2.71	2.01E-16
	GO:0009753	response to jasmonic acid	51	2.70	4.79E-08
	GO:0052696	flavonoid glucuronidation	31	2.60	1.46E-03
	GO:0044550	secondary metabolite biosynthetic process	33	2.56	9.42E-04
	GO:0006952	defense response	173	2.52	2.03E-29
	GO:0009813	flavonoid biosynthetic process	37	2.44	6.02E-04
	GO:0009408	response to heat	40	2.15	6.73E-03
	GO:0006979	response to oxidative stress	69	2.13	1.89E-06
	GO:0008152	metabolic process	62	1.96	3.81E-04
	GO:0007165	signal transduction	87	1.82	4.36E-05
	GO:0006468	protein phosphorylation	157	1.70	1.99E-08
	GO:0009737	response to abscisic acid	76	1.62	3.17E-02
	GO:0009651	response to salt stress	89	1.58	1.96E-02
	GO:0055114	oxidation-reduction process	205	1.49	2.23E-06

Table 2.3 continued

Genotype	Term	Description	Count	Fold Enrichment	FDR
ref4-3 cdk8-1					
	GO:0010112	regulation of systemic acquired resistance	11	3.91	3.48E-02
	GO:0009625	response to insect	17	2.90	3.07E-02
	GO:0009617	response to bacterium	54	2.53	1.76E-09
	GO:0010200	response to chitin	72	2.42	4.92E-12
	GO:0008652	cellular amino acid biosynthetic process	37	2.29	3.60E-04
	GO:0009611	response to wounding	98	2.16	5.66E-13
	GO:0009620	response to fungus	35	2.13	6.50E-03
	GO:0009753	response to jasmonic acid	68	1.91	2.51E-05
	GO:0009735	response to cytokinin	75	1.86	1.63E-05
	GO:0042742	defense response to bacterium	116	1.82	4.18E-09
	GO:0080167	response to karrikin	53	1.81	8.24E-03
	GO:0009408	response to heat	63	1.79	1.37E-03
	GO:0042254	ribosome biogenesis	56	1.73	2.05E-02
	GO:0050832	defense response to fungus	82	1.69	5.03E-04
	GO:0006412	translation	171	1.68	1.21E-10
	GO:0046686	response to cadmium ion	129	1.66	4.95E-07
	GO:0006952	defense response	215	1.66	3.77E-13
	GO:0006979	response to oxidative stress	99	1.62	3.16E-04
	GO:0009409	response to cold	99	1.48	2.76E-02
	GO:0055114	oxidation-reduction process	339	1.30	2.12E-05

Table 2.4 A full list of the genes that are down-regulated in *ref4-3* and with restored expression in *ref4-3 cdk8-1* (FDR < 0.05, absolute value of $\log_2$FC > 1)

AGI number	$\log_2$FC (*ref4-3* vs Col-0)	$\log_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G03870	-3.44	1.22	FASCICLIN-like arabinoogalactan 9
AT1G06080	-1.24	1.23	delta 9 desaturase 1
AT1G06350	-2.06	1.43	Fatty acid desaturase family protein
AT1G07450	-2.60	2.28	NAD(P)-binding Rossmann-fold superfamily protein
AT1G11070	-1.37	4.42	
AT1G13650	-1.51	1.92	
AT1G14280	-1.65	1.11	phytochrome kinase substrate 2
AT1G15175	-1.02	1.07	other RNA
AT1G19450	-2.07	1.85	Major facilitator superfamily protein
AT1G20190	-1.40	1.54	expansin 11
AT1G21910	-1.43	1.43	Integrase-type DNA-binding superfamily protein
AT1G22380	-1.33	2.45	UDP-glucosyl transferase 85A3
AT1G25230	-1.09	1.33	Calcineurin-like metallo-phosphoesterase superfamily protein
AT1G26945	-1.30	1.71	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT1G28010	-1.47	1.16	P-glycoprotein 14
AT1G29430	-3.25	1.90	SAUR-like auxin-responsive protein family
AT1G29440	-2.54	1.50	SAUR-like auxin-responsive protein family
AT1G29450	-2.60	1.33	SAUR-like auxin-responsive protein family
AT1G29460	-2.10	1.29	SAUR-like auxin-responsive protein family
AT1G29500	-2.43	1.69	SAUR-like auxin-responsive protein family
AT1G29510	-2.50	1.19	SAUR-like auxin-responsive protein family
AT1G29910	-1.64	1.38	chlorophyll A/B binding protein 3
AT1G29920	-1.84	1.40	chlorophyll A/B-binding protein 2
AT1G33930	-1.70	1.13	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT1G35730	-2.43	2.36	pumilio 9
AT1G46120	-6.22	4.87	transposable element gene
AT1G52190	-1.25	1.24	Major facilitator superfamily protein
AT1G52430	-1.07	1.41	Ubiquitin carboxyl-terminal hydrolase-related protein

Table 2.4 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G53690	-1.91	2.41	DNA directed RNA polymerase, 7 kDa subunit
AT1G61280	-1.12	1.27	Phosphatidylinositol N-acetylglucosaminyltransferase, GPI19/PIG-P subunit
AT1G65310	-3.04	2.28	xyloglucan endotransglucosylase/hydrolase 17
AT1G67460	-1.35	1.63	Minichromosome maintenance (MCM2/3/5) family protein
AT1G73830	-1.52	2.26	BR enhanced expression 3
AT1G75750	-1.35	1.21	GAST1 protein homolog 1
AT1G76610	-2.11	2.30	Protein of unknown function, DUF617
AT1G78320	-1.64	1.61	glutathione S-transferase TAU 23
AT2G05070	-1.40	2.13	photosystem II light harvesting complex gene 2.2
AT2G05100	-1.56	2.28	photosystem II light harvesting complex gene 2.1
AT2G16367	-1.19	2.87	
AT2G18180	-1.66	1.75	Sec14p-like phosphatidylinositol transfer family protein
AT2G18969	-1.70	1.95	
AT2G21200	-2.40	1.55	SAUR-like auxin-responsive protein family
AT2G21220	-2.96	2.46	SAUR-like auxin-responsive protein family
AT2G22980	-1.28	1.22	serine carboxypeptidase-like 13
AT2G22990	-1.25	1.06	sinapoylglucose 1
AT2G24610	-1.28	1.56	cyclic nucleotide-gated channel 14
AT2G26710	-1.06	1.36	Cytochrome P450 superfamily protein
AT2G27402	-3.00	1.00	
AT2G28630	-1.26	1.29	3-ketoacyl-CoA synthase 12
AT2G29170	-2.52	1.87	NAD(P)-binding Rossmann-fold superfamily protein
AT2G37130	-1.00	1.03	Peroxidase superfamily protein
AT2G42885	-3.67	3.68	Defensin-like (DEFL) family protein
AT3G02380	-1.21	1.53	CONSTANS-like 2
AT3G03480	-1.88	1.82	acetyl CoA:(Z)-3-hexen-1-ol acetyltransferase
AT3G05900	-1.80	1.07	neurofilament protein-related
AT3G06100	-1.96	1.81	NOD26-like intrinsic protein 7;1

Table 2.4 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT3G15450	-1.47	1.80	Aluminium induced protein with YGL and LRDR motifs
AT3G21460	-1.23	1.26	Glutaredoxin family protein
AT3G23050	-1.09	1.03	indole-3-acetic acid 7
AT3G26815	-1.56	1.82	MIR169K; miRNA
AT3G26818	-1.74	1.48	MIR169M; miRNA
AT3G27620	-1.52	1.40	alternative oxidase 1C
AT3G27690	-1.43	2.32	photosystem II light harvesting complex gene 2.3
AT3G30122	-1.90	1.89	
AT3G51600	-1.07	1.26	lipid transfer protein 5
AT3G54510	-3.21	3.04	Early-responsive to dehydration stress protein (ERD4)
AT3G56230	-1.89	1.94	BTB/POZ domain-containing protein
AT3G59250	-1.49	1.47	F-box/RNI-like superfamily protein
AT3G60290	-1.80	1.20	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT3G61430	-1.36	1.04	plasma membrane intrinsic protein 1A
AT4G10160	-3.68	5.37	RING/U-box superfamily protein
AT4G12050	-2.55	3.03	Predicted AT-hook DNA-binding family protein
AT4G12690	-2.33	1.97	Plant protein of unknown function (DUF868)
AT4G16515	-1.11	1.02	
AT4G18970	-1.88	1.13	GDSL-like Lipase/Acylhydrolase superfamily protein
AT4G25780	-2.74	1.52	CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein
AT4G27310	-1.63	1.07	B-box type zinc finger family protein
AT4G27450	-1.68	1.07	Aluminium induced protein with YGL and LRDR motifs
AT4G32280	-1.21	2.27	indole-3-acetic acid inducible 29
AT4G33790	-2.24	3.90	Jojoba acyl CoA reductase-related male sterility protein
AT4G35770	-2.26	1.00	Rhodanese/Cell cycle control phosphatase superfamily protein
AT4G36105	-3.05	2.25	
AT4G38860	-1.28	1.07	SAUR-like auxin-responsive protein family

Table 2.4 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT4G39675	-1.90	2.29	
AT5G02180	-1.33	1.48	Transmembrane amino acid transporter family protein
AT5G02760	-2.37	1.46	Protein phosphatase 2C family protein
AT5G06980	-1.29	1.71	
AT5G07571	-2.43	2.25	Oleosin family protein
AT5G15180	-1.10	1.39	Peroxidase superfamily protein
AT5G16030	-1.08	1.04	
AT5G18020	-1.95	1.51	SAUR-like auxin-responsive protein family
AT5G18030	-1.12	1.01	SAUR-like auxin-responsive protein family
AT5G18050	-1.16	1.77	SAUR-like auxin-responsive protein family
AT5G19190	-1.09	1.04	
AT5G22920	-1.38	1.02	CHY-type/CTCHY-type/RING-type Zinc finger protein
AT5G25460	-1.21	1.44	Protein of unknown function, DUF642
AT5G27360	-1.49	1.15	Major facilitator superfamily protein
AT5G35777	-3.63	2.46	transposable element gene
AT5G37950	-1.63	2.17	UDP-Glycosyltransferase superfamily protein
AT5G38970	-3.15	2.22	brassinosteroid-6-oxidase 1
AT5G39860	-2.77	2.06	basic helix-loop-helix (bHLH) DNA-binding family protein
AT5G45820	-3.54	2.09	CBL-interacting protein kinase 20
AT5G49360	-1.10	1.46	beta-xylosidase 1
AT5G54145	-1.23	1.16	
AT5G54270	-1.27	1.25	light-harvesting chlorophyll B-binding protein 3
AT5G56840	-2.68	1.02	myb-like transcription factor family protein
AT5G61650	-3.68	3.68	CYCLIN P4;2
AT5G62430	-1.01	1.72	cycling DOF factor 1
AT5G67390	-2.09	1.55	

Table 2.5 Gene ontology analysis for the genes that are down-regulated in *ref4-3* and that have
restored expression in *ref4-3 cdk8-1*

Term	Description	Count	Fold Enrichment	FDR
GO:0046620	regulation of organ growth	6	95.29	3E-06
GO:0009768	photosynthesis, light harvesting in photosystem I	6	47.65	1E-04
GO:0018298	protein-chromophore linkage	6	27.58	3E-03
GO:0009926	auxin polar transport	6	19.78	1E-02
GO:0009734	auxin-activated signaling pathway	11	11.51	4E-05
GO:0009733	response to auxin	14	9.41	3E-06

Table 2.6 A full list of the genes that are up-regulated in *ref4-3* and with restored expression in *ref4-3 cdk8-1* (FDR < 0.05, absolute value of log₂FC > 1)

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G01680	1.78	-3.60	plant U-box 54
AT1G01720	2.20	-1.08	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
AT1G02310	2.28	-1.49	Glycosyl hydrolase superfamily protein
AT1G02430	2.67	-5.97	ADP-ribosylation factor D1B
AT1G02470	1.65	-1.06	Polyketide cyclase/dehydrase and lipid transport superfamily protein
AT1G02850	1.99	-1.66	beta glucosidase 11
AT1G03495	1.92	-1.69	HXXXD-type acyl-transferase family protein
AT1G04600	2.08	-2.00	myosin XI A
AT1G05100	3.25	-1.88	mitogen-activated protein kinase 18
AT1G05450	2.42	-3.20	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT1G05880	1.62	-2.99	RING/U-box superfamily protein
AT1G09080	1.80	-5.38	Heat shock protein 70 (Hsp 70) family protein
AT1G13310	2.73	-4.18	Endosomal targeting BRO1-like domain-containing protein
AT1G13470	1.09	-4.32	Protein of unknown function (DUF1262)
AT1G13520	2.05	-2.21	Protein of unknown function (DUF1262)
AT1G13550	2.62	-1.98	Protein of unknown function (DUF1262)
AT1G14080	1.94	-2.42	fucosyltransferase 6
AT1G14880	1.20	-3.74	PLANT CADMIUM RESISTANCE 1
AT1G15380	1.96	-1.10	Lactoylglutathione lyase / glyoxalase I family protein
AT1G15520	3.15	-2.81	pleiotropic drug resistance 12
AT1G15670	1.62	-1.47	Galactose oxidase/kelch repeat superfamily protein
AT1G16090	1.14	-1.28	wall associated kinase-like 7
AT1G17615	4.93	-5.16	Disease resistance protein (TIR-NBS class)
AT1G18300	1.22	-1.10	nudix hydrolase homolog 4
AT1G19020	1.23	-2.37	
AT1G19230	3.20	-3.15	Riboflavin synthase-like superfamily protein
AT1G19250	2.92	-3.37	flavin-dependent monooxygenase 1

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G20160	1.11	-2.05	Subtilisin-like serine endopeptidase family protein
AT1G21240	1.92	-3.55	wall associated kinase 3
AT1G21390	1.59	-1.43	embryo defective 2170
AT1G21525	1.40	-2.14	
AT1G21550	1.03	-1.69	Calcium-binding EF-hand family protein
AT1G22440	2.04	-1.92	Zinc-binding alcohol dehydrogenase family protein
AT1G25422	1.11	-1.82	
AT1G26240	3.80	-3.80	Proline-rich extensin-like family protein
AT1G26390	1.39	-2.16	FAD-binding Berberine family protein
AT1G26410	3.42	-1.95	FAD-binding Berberine family protein
AT1G29860	2.05	-1.02	WRKY DNA-binding protein 71
AT1G30700	1.54	-2.03	FAD-binding Berberine family protein
AT1G30850	4.44	-2.69	root hair specific 4
AT1G30900	1.28	-2.06	VACUOLAR SORTING RECEPTOR 6
AT1G31570	2.58	-2.27	transposable element gene
AT1G32950	1.18	-1.61	Subtilase family protein
AT1G32960	4.06	-4.60	Subtilase family protein
AT1G33730	1.77	-4.04	cytochrome P450, family 76, subfamily C, polypeptide 5
AT1G33840	3.81	-1.85	Protein of unknown function (DUF567)
AT1G33950	2.50	-4.70	Avirulence induced gene (AIG1) family protein
AT1G34420	1.06	-2.10	leucine-rich repeat transmembrane protein kinase family protein
AT1G35230	1.56	-3.77	arabinogalactan protein 5
AT1G35330	2.90	-2.36	RING/U-box superfamily protein
AT1G35710	1.39	-2.69	Protein kinase family protein with leucine-rich repeat domain
AT1G44130	2.16	-2.57	Eukaryotic aspartyl protease family protein
AT1G47880	2.80	-3.83	
AT1G47890	2.38	-4.66	receptor like protein 7
AT1G48470	1.17	-1.75	glutamine synthetase 1;5
AT1G48510	6.77	-2.64	Surfeit locus 1 cytochrome c oxidase biogenesis protein
AT1G49960	2.41	-2.14	Xanthine/uracil permease family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G51850	1.47	-1.44	Leucine-rich repeat protein kinase family protein
AT1G51860	2.93	-3.42	Leucine-rich repeat protein kinase family protein
AT1G53625	1.44	-1.23	
AT1G54540	3.73	-2.71	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
AT1G55790	2.72	-2.76	Domain of unknown function (DUF2431)
AT1G56120	1.09	-1.75	Leucine-rich repeat transmembrane protein kinase
AT1G57850	3.86	-3.50	Toll-Interleukin-Resistance (TIR) domain family protein
AT1G58225	2.75	-3.88	
AT1G58390	1.30	-1.02	Disease resistance protein (CC-NBS-LRR class) family
AT1G59590	1.81	-2.32	ZCF37
AT1G60470	5.50	-1.75	galactinol synthase 4
AT1G61550	1.49	-1.67	S-locus lectin protein kinase family protein
AT1G61810	1.15	-1.15	beta-glucosidase 45
AT1G61930	1.64	-1.72	Protein of unknown function, DUF584
AT1G62262	2.18	-3.34	SLAC1 homologue 4
AT1G62370	1.29	-1.29	RING/U-box superfamily protein
AT1G63840	3.22	-2.02	RING/U-box superfamily protein
AT1G64070	3.36	-2.20	Disease resistance protein (TIR-NBS-LRR class) family
AT1G64080	1.41	-1.02	
AT1G65483	2.33	-3.53	
AT1G65690	2.54	-1.92	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
AT1G66380	1.62	-2.66	myb domain protein 114
AT1G66390	4.75	-2.08	myb domain protein 90
AT1G66570	4.54	-3.50	sucrose-proton symporter 7
AT1G66725	1.80	-1.79	MIR163; miRNA
AT1G66830	2.10	-1.19	Leucine-rich repeat protein kinase family protein
AT1G66960	3.03	-5.86	Terpenoid cyclases family protein
AT1G67360	3.00	-1.17	Rubber elongation factor protein (REF)
AT1G68440	1.16	-1.85	

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT1G68570	3.16	-1.44	Major facilitator superfamily protein
AT1G68620	2.50	-3.72	alpha/beta-Hydrolases superfamily protein
AT1G68850	2.50	-1.96	Peroxidase superfamily protein
AT1G69490	2.94	-1.02	NAC-like, activated by AP3/PI
AT1G69880	2.30	-2.03	thioredoxin H-type 8
AT1G70170	3.00	-2.65	matrix metalloproteinase
AT1G70800	2.93	-1.20	Calcium-dependent lipid-binding (CaLB domain) family protein
AT1G71390	4.13	-3.87	receptor like protein 11
AT1G71530	6.21	-4.67	Protein kinase superfamily protein
AT1G72240	1.78	-1.51	
AT1G72540	1.56	-3.16	Protein kinase superfamily protein
AT1G72950	1.22	-1.36	Disease resistance protein (TIR-NBS class)
AT1G73220	5.33	-2.81	organic cation/carnitine transporter1
AT1G73805	1.39	-2.98	Calmodulin binding protein-like
AT1G73810	1.05	-1.46	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
AT1G74140	4.75	-3.01	Rhomboid-related intramembrane serine protease family protein
AT1G74460	1.18	-1.31	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G74590	2.16	-2.63	glutathione S-transferase TAU 10
AT1G74710	1.01	-1.79	ADC synthase superfamily protein
AT1G74930	1.10	-1.83	Integrase-type DNA-binding superfamily protein
AT1G75000	2.46	-2.60	GNS1/SUR4 membrane protein family
AT1G75040	3.03	-3.53	pathogenesis-related gene 5
AT1G75910	5.12	-3.37	extracellular lipase 4
AT1G76530	1.25	-1.41	Auxin efflux carrier family protein
AT1G76640	1.46	-2.35	Calcium-binding EF-hand family protein
AT1G76970	1.66	-1.72	Target of Myb protein 1
AT1G76980	1.51	-1.37	
AT2G02330	4.06	-4.06	
AT2G03360	3.01	-2.78	Glycosyltransferase family 61 protein
AT2G04090	2.44	-2.88	MATE efflux family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT2G04100	2.19	-1.39	MATE efflux family protein
AT2G04135	3.51	-1.18	transposable element gene
AT2G04450	1.52	-2.59	nudix hydrolase homolog 6
AT2G04495	1.84	-4.00	
AT2G05440	1.25	-1.12	GLYCINE RICH PROTEIN 9
AT2G11891	5.72	-2.05	
AT2G13810	1.28	-3.90	AGD2-like defense response protein 1
AT2G14610	1.47	-3.73	pathogenesis-related gene 1
AT2G14620	2.09	-1.97	xyloglucan endotransglucosylase/hydrolase 10
AT2G16895	2.23	-1.86	
AT2G18150	1.34	-1.13	Peroxidase superfamily protein
AT2G18370	1.41	-2.04	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT2G18660	1.43	-4.68	plant natriuretic peptide A
AT2G19500	4.76	-4.76	cytokinin oxidase 2
AT2G21100	3.37	-2.40	Disease resistance-responsive (dirigent-like protein) family protein
AT2G22510	1.17	-2.05	hydroxyproline-rich glycoprotein family protein
AT2G23110	4.34	-1.69	Late embryogenesis abundant protein, group 6
AT2G23540	1.46	-1.25	GDSL-like Lipase/Acylhydrolase superfamily protein
AT2G23830	3.01	-2.11	PapD-like superfamily protein
AT2G24160	1.32	-1.42	
AT2G24210	1.90	-2.22	terpene synthase 10
AT2G25260	1.82	-1.91	
AT2G25440	1.27	-3.26	receptor like protein 20
AT2G25510	1.13	-1.53	
AT2G26150	1.25	-4.32	heat shock transcription factor A2
AT2G26400	2.76	-4.55	acireductone dioxygenase 3
AT2G26480	2.27	-2.19	UDP-glucosyl transferase 76D1
AT2G27389	1.36	-2.11	
AT2G28570	1.88	-2.02	

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT2G29100	3.31	-3.15	glutamate receptor 2
AT2G29110	1.69	-3.49	glutamate receptor 2
AT2G29350	1.28	-1.98	senescence-associated gene 13
AT2G29460	1.11	-2.51	glutathione S-transferase tau 4
AT2G29470	2.03	-2.81	glutathione S-transferase tau 3
AT2G30660	1.78	-1.98	ATP-dependent caseinolytic (Clp) protease/crotonase family protein
AT2G32680	1.55	-4.28	receptor like protein 23
AT2G33080	2.01	-3.37	receptor like protein 28
AT2G33580	2.52	-1.96	Protein kinase superfamily protein
AT2G34360	1.18	-1.71	MATE efflux family protein
AT2G34500	1.79	-1.69	cytochrome P450, family 710, subfamily A, polypeptide 1
AT2G34940	1.32	-2.17	VACUOLAR SORTING RECEPTOR 5
AT2G35980	2.12	-2.08	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
AT2G36580	2.00	-1.40	Pyruvate kinase family protein
AT2G37360	2.07	-2.37	ABC-2 type transporter family protein
AT2G37710	1.60	-1.32	receptor lectin kinase
AT2G37750	1.79	-1.28	
AT2G39350	1.55	-1.32	ABC-2 type transporter family protein
AT2G40095	1.29	-1.92	Alpha/beta hydrolase related protein
AT2G40370	3.16	-1.97	laccase 5
AT2G40740	2.49	-1.62	WRKY DNA-binding protein 55
AT2G41230	1.55	-1.85	
AT2G41850	2.80	-3.16	polygalacturonase abscission zone A. thaliana
AT2G42140	6.20	-2.81	VQ motif-containing protein
AT2G43570	2.32	-2.90	chitinase, putative
AT2G43840	3.30	-1.97	UDP-glycosyltransferase 74 F1
AT2G44240	5.54	-8.32	Protein of Unknown Function (DUF239)
AT2G44290	1.85	-1.83	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT2G44400	3.52	-2.17	Cysteine/Histidine-rich C1 domain family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT2G44910	1.21	-1.29	homeobox-leucine zipper protein 4
AT2G45220	1.84	-1.62	Plant invertase/pectin methylesterase inhibitor superfamily
AT2G45760	1.89	-2.70	BON association protein 2
AT2G45840	2.22	-1.99	Arabidopsis thaliana protein of unknown function (DUF821)
AT2G45900	1.16	-1.03	Phosphatidylinositol N-acetyglucosaminlytransferase subunit P-related
AT2G46270	1.62	-1.19	G-box binding factor 3
AT2G47130	1.12	-2.84	NAD(P)-binding Rossmann-fold superfamily protein
AT2G47460	3.28	-1.33	myb domain protein 12
AT2G47550	1.71	-1.83	Plant invertase/pectin methylesterase inhibitor superfamily
AT2G47950	4.94	-3.23	
AT2G48130	2.42	-2.09	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT2G48140	2.72	-1.83	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT3G01420	2.64	-2.25	Peroxidase superfamily protein
AT3G02610	1.04	-1.43	Plant stearoyl-acyl-carrier-protein desaturase family protein
AT3G03650	3.61	-2.46	Exostosin family protein
AT3G04000	2.97	-1.10	NAD(P)-binding Rossmann-fold superfamily protein
AT3G04050	2.33	-2.10	Pyruvate kinase family protein
AT3G04060	1.40	-1.05	NAC domain containing protein 46
AT3G04070	1.18	-1.44	NAC domain containing protein 47
AT3G04181	1.49	-1.42	
AT3G04300	2.69	-3.74	RmlC-like cupins superfamily protein
AT3G05650	1.42	-1.26	receptor like protein 32
AT3G06520	1.03	-1.21	agenet domain-containing protein
AT3G07600	2.20	-2.77	Heavy metal transport/detoxification superfamily protein
AT3G07970	2.10	-1.17	Pectin lyase-like superfamily protein
AT3G08040	1.51	-1.22	MATE efflux family protein
AT3G08860	2.55	-1.71	PYRIMIDINE 4

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT3G09220	1.52	-1.36	laccase 7
AT3G09790	4.44	-4.44	ubiquitin 8
AT3G09960	1.76	-4.33	Calcineurin-like metallo-phosphoesterase superfamily protein
AT3G11010	1.67	-2.74	receptor like protein 34
AT3G11080	1.25	-1.46	receptor like protein 35
AT3G11260	1.66	-1.54	WUSCHEL related homeobox 5
AT3G11402	1.08	-1.82	Cysteine/Histidine-rich C1 domain family protein
AT3G11430	2.41	-2.27	glycerol-3-phosphate acyltransferase 5
AT3G12220	1.91	-2.58	serine carboxypeptidase-like 16
AT3G12230	5.57	-3.82	serine carboxypeptidase-like 14
AT3G12830	1.83	-1.23	SAUR-like auxin-responsive protein family
AT3G12910	5.37	-2.97	NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
AT3G13090	1.31	-2.15	multidrug resistance-associated protein 8
AT3G13100	1.30	-1.73	multidrug resistance-associated protein 7
AT3G13130	7.47	-2.59	
AT3G13277	1.23	-1.38	other RNA
AT3G13433	2.56	-1.98	
AT3G13610	2.31	-3.81	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT3G14460	1.77	-3.87	LRR and NB-ARC domains-containing disease resistance protein
AT3G15536	2.57	-4.85	
AT3G18250	1.51	-2.22	Putative membrane lipoprotein
AT3G21520	2.00	-2.65	DUF679 domain membrane protein 1
AT3G21560	2.57	-1.08	UDP-Glycosyltransferase superfamily protein
AT3G21710	2.21	-1.22	
AT3G22060	2.72	-1.68	Receptor-like protein kinase-related family protein
AT3G22600	2.25	-2.07	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT3G22620	3.09	-1.95	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT3G22840	6.94	-1.31	Chlorophyll A-B binding family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT3G22910	1.62	-2.99	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
AT3G23010	1.06	-1.13	receptor like protein 36
AT3G23110	1.56	-1.83	receptor like protein 37
AT3G23120	1.18	-1.57	receptor like protein 38
AT3G23240	1.14	-1.23	ethylene response factor 1
AT3G24900	2.89	-3.69	receptor like protein 39
AT3G25010	2.06	-3.40	receptor like protein 41
AT3G25190	1.35	-1.61	Vacuolar iron transporter (VIT) family protein
AT3G25510	1.08	-2.61	disease resistance protein (TIR-NBS-LRR class), putative
AT3G25610	1.35	-1.58	ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein
AT3G25640	1.62	-1.25	Protein of unknown function, DUF617
AT3G25795	2.67	-3.22	other RNA
AT3G26210	1.19	-2.44	cytochrome P450, family 71, subfamily B, polypeptide 23
AT3G27490	6.61	-4.09	Cysteine/Histidine-rich C1 domain family protein
AT3G28510	2.21	-4.22	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT3G29110	6.39	-2.26	Terpenoid cyclases/Protein prenyltransferases superfamily protein
AT3G29250	1.65	-2.43	NAD(P)-binding Rossmann-fold superfamily protein
AT3G29590	1.70	-1.25	HXXXD-type acyl-transferase family protein
AT3G30120	1.85	-2.59	
AT3G42110	5.31	-5.31	transposable element gene
AT3G44326	2.05	-3.97	F-box family protein
AT3G44350	1.65	-2.19	NAC domain containing protein 61
AT3G44540	2.12	-1.77	fatty acid reductase 4
AT3G44550	1.24	-2.40	fatty acid reductase 5
AT3G44560	1.53	-3.51	fatty acid reductase 8
AT3G45290	2.93	-2.27	Seven transmembrane MLO family protein
AT3G45330	1.86	-2.17	Concanavalin A-like lectin protein kinase family protein
AT3G45760	5.01	-5.01	Nucleotidyltransferase family protein

Table 2.6 continued

AGI number	log₂FC (*ref4-3* vs Col-0)	log₂FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT3G45860	1.28	-1.44	cysteine-rich RLK (RECEPTOR-like protein kinase) 4
AT3G46080	1.14	-2.20	C2H2-type zinc finger family protein
AT3G46090	2.05	-4.70	C2H2 and C2HC zinc fingers superfamily protein
AT3G47040	4.86	-4.86	Glycosyl hydrolase family protein
AT3G47050	3.59	-1.67	Glycosyl hydrolase family protein
AT3G47480	1.11	-3.82	Calcium-binding EF-hand family protein
AT3G48080	1.09	-1.10	alpha/beta-Hydrolases superfamily protein
AT3G48630	1.88	-3.47	
AT3G48650	1.16	-3.47	
AT3G48850	1.44	-2.19	phosphate transporter 3;2
AT3G48920	1.77	-1.99	myb domain protein 45
AT3G49120	1.07	-1.25	peroxidase CB
AT3G49780	1.18	-1.26	phytosulfokine 4 precursor
AT3G49950	1.26	-2.07	GRAS family transcription factor
AT3G50400	3.88	-3.01	GDSL-like Lipase/Acylhydrolase superfamily protein
AT3G50470	1.28	-2.36	homolog of RPW8 3
AT3G50480	1.60	-2.83	homolog of RPW8 4
AT3G51330	1.32	-2.09	Eukaryotic aspartyl protease family protein
AT3G51860	1.33	-1.42	cation exchanger 3
AT3G53150	2.89	-5.97	UDP-glucosyl transferase 73D1
AT3G53510	2.45	-2.76	ABC-2 type transporter family protein
AT3G53980	3.23	-1.65	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT3G55090	4.73	-3.69	ABC-2 type transporter family protein
AT3G55970	1.22	-2.14	jasmonate-regulated gene 21
AT3G57240	2.07	-5.18	beta-1,3-glucanase 3
AT3G57260	1.26	-3.80	beta-1,3-glucanase 2
AT3G57510	3.70	-6.48	Pectin lyase-like superfamily protein
AT3G57700	1.20	-1.42	Protein kinase superfamily protein
AT3G57950	1.89	-1.68	

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT3G58550	2.29	-1.81	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT3G59710	1.56	-1.20	NAD(P)-binding Rossmann-fold superfamily protein
AT3G59930	6.35	-1.59	
AT3G60140	3.52	-1.26	Glycosyl hydrolase superfamily protein
AT3G60170	4.44	-2.69	transposable element gene
AT3G60420	1.43	-1.47	Phosphoglycerate mutase family protein
AT3G60470	3.11	-5.58	Plant protein of unknown function (DUF247)
AT3G60966	1.44	-2.41	RING/U-box superfamily protein
AT3G61198	1.41	-2.21	other RNA
AT4G00700	2.52	-3.79	C2 calcium/lipid-binding plant phosphoribosyltransferase family protein
AT4G00870	2.97	-1.34	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT4G01380	1.77	-6.20	plastocyanin-like domain-containing protein
AT4G03450	1.70	-5.44	Ankyrin repeat family protein
AT4G03540	3.61	-2.43	Uncharacterised protein family (UPF0497)
AT4G03950	4.44	-4.44	Nucleotide/sugar transporter family protein
AT4G04490	1.58	-2.71	cysteine-rich RLK (RECEPTOR-like protein kinase) 36
AT4G04500	2.83	-3.55	cysteine-rich RLK (RECEPTOR-like protein kinase) 37
AT4G04510	1.64	-4.57	cysteine-rich RLK (RECEPTOR-like protein kinase) 38
AT4G04760	1.32	-1.39	Major facilitator superfamily protein
AT4G05540	5.01	-5.29	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT4G09455	5.22	-5.22	transposable element gene
AT4G09770	1.61	-2.12	TRAF-like family protein
AT4G10500	2.67	-4.50	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT4G10860	5.59	-5.08	
AT4G11070	2.30	-3.46	WRKY family transcription factor
AT4G11170	2.01	-2.63	Disease resistance protein (TIR-NBS-LRR class) family
AT4G11650	2.13	-1.88	osmotin 34
AT4G11655	1.89	-2.82	Uncharacterised protein family (UPF0497)

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT4G11890	1.31	-2.41	Protein kinase superfamily protein
AT4G12580	3.08	-1.16	
AT4G13395	1.83	-3.29	ROTUNDIFOLIA like 12
AT4G13890	2.86	-4.00	Pyridoxal phosphate (PLP)-dependent transferases superfamily protein
AT4G13900	1.51	-1.49	
AT4G14090	1.54	-1.45	UDP-Glycosyltransferase superfamily protein
AT4G14390	1.34	-2.49	Ankyrin repeat family protein
AT4G14630	2.57	-1.14	germin-like protein 9
AT4G14640	1.74	-1.61	calmodulin 8
AT4G15417	1.82	-2.86	RNAse II-like 1
AT4G16600	1.42	-2.18	Nucleotide-diphospho-sugar transferases superfamily protein
AT4G17215	1.50	-1.10	Pollen Ole e 1 allergen and extensin family protein
AT4G17280	1.06	-1.08	Auxin-responsive family protein
AT4G17660	1.38	-2.63	Protein kinase superfamily protein
AT4G18360	1.20	-1.57	Aldolase-type TIM barrel family protein
AT4G18870	5.79	-3.26	E2F/DP family winged-helix DNA-binding domain
AT4G18980	1.46	-1.89	AtS40-3
AT4G19430	1.30	-2.79	
AT4G20110	1.09	-1.25	VACUOLAR SORTING RECEPTOR 7
AT4G21230	1.36	-1.24	cysteine-rich RLK (RECEPTOR-like protein kinase) 27
AT4G21840	1.04	-3.60	methionine sulfoxide reductase B8
AT4G21926	1.42	-1.28	
AT4G22070	2.49	-1.66	WRKY DNA-binding protein 31
AT4G22505	1.24	-1.33	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT4G22960	1.37	-1.30	Protein of unknown function (DUF544)
AT4G23140	1.03	-4.07	cysteine-rich RLK (RECEPTOR-like protein kinase) 6
AT4G23150	2.71	-3.87	cysteine-rich RLK (RECEPTOR-like protein kinase) 7
AT4G23160	1.56	-2.49	cysteine-rich RLK (RECEPTOR-like protein kinase) 8

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT4G23230	1.00	-1.36	cysteine-rich RLK (RECEPTOR-like protein kinase) 15
AT4G23310	2.54	-3.38	cysteine-rich RLK (RECEPTOR-like protein kinase) 23
AT4G23320	1.44	-1.48	cysteine-rich RLK (RECEPTOR-like protein kinase) 24
AT4G23610	1.13	-2.03	Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein family
AT4G24000	2.04	-1.31	cellulose synthase like G2
AT4G24450	1.10	-1.17	phosphoglucan, water dikinase
AT4G25350	2.06	-1.38	EXS (ERD1/XPR1/SYG1) family protein
AT4G26150	1.55	-1.65	cytokinin-responsive gata factor 1
AT4G28110	3.54	-2.00	myb domain protein 41
AT4G28390	1.39	-1.23	ADP/ATP carrier 3
AT4G28703	2.44	-1.25	RmlC-like cupins superfamily protein
AT4G28790	2.32	-1.98	basic helix-loop-helix (bHLH) DNA-binding superfamily protein
AT4G30460	1.11	-1.05	glycine-rich protein
AT4G30640	1.12	-1.67	RNI-like superfamily protein
AT4G32205	1.77	-1.33	transposable element gene
AT4G33550	3.57	-1.56	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT4G34135	2.51	-1.34	UDP-glucosyltransferase 73B2
AT4G34380	3.11	-2.45	Transducin/WD40 repeat-like superfamily protein
AT4G35165	2.82	-3.85	Protein of unknown function (DUF1278)
AT4G35380	2.45	-2.49	SEC7-like guanine nucleotide exchange family protein
AT4G36430	1.46	-1.16	Peroxidase superfamily protein
AT4G36610	1.82	-1.36	alpha/beta-Hydrolases superfamily protein
AT4G37370	1.25	-1.41	cytochrome P450, family 81, subfamily D, polypeptide 8
AT4G37400	2.89	-1.14	cytochrome P450, family 81, subfamily F, polypeptide 3
AT4G38080	1.68	-1.73	hydroxyproline-rich glycoprotein family protein
AT4G38560	1.18	-2.98	Arabidopsis phospholipase-like protein (PEARLI 4) family
AT4G39670	1.77	-1.94	Glycolipid transfer protein (GLTP) family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT4G39830	1.06	-2.36	Cupredoxin superfamily protein
AT5G01100	1.03	-1.65	O-fucosyltransferase family protein
AT5G03090	4.18	-4.45	
AT5G05390	2.48	-2.01	laccase 12
AT5G06230	1.97	-1.16	TRICHOME BIREFRINGENCE-LIKE 9
AT5G06730	2.81	-2.36	Peroxidase superfamily protein
AT5G07310	1.57	-2.67	Integrase-type DNA-binding superfamily protein
AT5G07760	1.37	-3.72	formin homology 2 domain-containing protein / FH2 domain-containing protein
AT5G07780	1.95	-1.96	Actin-binding FH2 (formin homology 2) family protein
AT5G07990	2.97	-1.52	Cytochrome P450 superfamily protein
AT5G08250	2.85	-4.01	Cytochrome P450 superfamily protein
AT5G09290	1.16	-3.49	Inositol monophosphatase family protein
AT5G09480	1.59	-1.06	hydroxyproline-rich glycoprotein family protein
AT5G09520	3.28	-2.57	hydroxyproline-rich glycoprotein family protein
AT5G09530	3.08	-1.69	hydroxyproline-rich glycoprotein family protein
AT5G10380	1.32	-2.56	RING/U-box superfamily protein
AT5G10625	2.25	-1.71	
AT5G10695	1.33	-1.57	
AT5G10760	1.35	-3.88	Eukaryotic aspartyl protease family protein
AT5G11210	2.13	-2.04	glutamate receptor 2.5
AT5G11920	1.75	-3.60	6-&1-fructan exohydrolase
AT5G13190	1.02	-1.31	
AT5G13200	1.44	-1.56	GRAM domain family protein
AT5G13320	1.63	-2.35	Auxin-responsive GH3 family protein
AT5G13330	1.10	-1.20	related to AP2 6l
AT5G13580	3.89	-2.50	ABC-2 type transporter family protein
AT5G13900	3.27	-3.26	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT5G14130	1.57	-2.11	Peroxidase superfamily protein
AT5G14650	1.90	-3.32	Pectin lyase-like superfamily protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT5G16900	2.62	-2.91	Leucine-rich repeat protein kinase family protein
AT5G17030	4.53	-2.34	UDP-glucosyl transferase 78D3
AT5G17125	8.26	-5.75	transposable element gene
AT5G18350	2.81	-2.71	Disease resistance protein (TIR-NBS-LRR class) family
AT5G19410	2.72	-2.85	ABC-2 type transporter family protein
AT5G19490	1.40	-1.48	Histone superfamily protein
AT5G19880	4.28	-1.94	Peroxidase superfamily protein
AT5G20860	2.91	-1.81	Plant invertase/pectin methylesterase inhibitor superfamily
AT5G22380	1.63	-2.90	NAC domain containing protein 90
AT5G22420	8.57	-6.05	fatty acid reductase 7
AT5G22490	1.50	-4.11	O-acyltransferase (WSD1-like) family protein
AT5G22545	1.18	-2.90	
AT5G22560	3.24	-5.22	Plant protein of unknown function (DUF247)
AT5G23000	2.07	-1.28	myb domain protein 37
AT5G23160	1.37	-2.14	
AT5G23190	1.95	-1.85	cytochrome P450, family 86, subfamily B, polypeptide 1
AT5G23990	5.57	-2.55	ferric reduction oxidase 5
AT5G24110	2.32	-2.80	WRKY DNA-binding protein 30
AT5G24200	2.01	-4.64	alpha/beta-Hydrolases superfamily protein
AT5G24205	1.76	-1.53	other RNA
AT5G24206	1.54	-1.33	other RNA
AT5G24210	1.61	-2.45	alpha/beta-Hydrolases superfamily protein
AT5G24530	1.07	-2.64	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein
AT5G24540	1.75	-4.78	beta glucosidase 31
AT5G25260	1.47	-4.04	SPFH/Band 7/PHB domain-containing membrane-associated protein family
AT5G25770	1.34	-1.26	alpha/beta-Hydrolases superfamily protein
AT5G25970	1.09	-1.11	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
AT5G26170	1.65	-3.16	WRKY DNA-binding protein 50
AT5G26310	2.12	-2.78	UDP-Glycosyltransferase superfamily protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT5G26690	1.99	-2.62	Heavy metal transport/detoxification superfamily protein
AT5G26920	1.29	-1.82	Cam-binding protein 60-like G
AT5G27420	1.72	-2.27	carbon/nitrogen insensitive 1
AT5G35580	5.91	-2.33	Protein kinase superfamily protein
AT5G36970	3.88	-1.72	NDR1/HIN1-like 25
AT5G37990	1.82	-1.58	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
AT5G38020	1.28	-1.12	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
AT5G38030	1.46	-1.32	MATE efflux family protein
AT5G38250	1.33	-3.89	Protein kinase family protein
AT5G38350	1.67	-6.29	Disease resistance protein (NBS-LRR class) family
AT5G38910	6.66	-6.66	RmlC-like cupins superfamily protein
AT5G39100	3.42	-3.18	germin-like protein 6
AT5G39110	4.47	-2.30	RmlC-like cupins superfamily protein
AT5G39120	3.63	-2.63	RmlC-like cupins superfamily protein
AT5G39130	3.90	-5.88	RmlC-like cupins superfamily protein
AT5G39720	1.74	-1.43	avirulence induced gene 2 like protein
AT5G40010	1.90	-4.77	AAA-ATPase 1
AT5G40690	1.65	-1.88	
AT5G40990	2.22	-2.51	GDSL lipase 1
AT5G41550	1.15	-1.33	Disease resistance protein (TIR-NBS-LRR class) family
AT5G41610	1.21	-1.08	cation/H+ exchanger 18
AT5G42050	1.66	-1.03	DCD (Development and Cell Death) domain protein
AT5G44460	2.25	-4.72	calmodulin like 43
AT5G44920	2.13	-1.83	Toll-Interleukin-Resistance (TIR) domain family protein
AT5G44990	2.52	-3.08	Glutathione S-transferase family protein
AT5G45000	1.46	-1.52	Disease resistance protein (TIR-NBS-LRR class) family
AT5G46960	6.82	-3.42	Plant invertase/pectin methylesterase inhibitor superfamily protein
AT5G48290	1.54	-1.89	Heavy metal transport/detoxification superfamily protein
AT5G48400	4.17	-4.04	Glutamate receptor family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT5G48410	2.63	-1.63	glutamate receptor 1.3
AT5G48540	1.80	-2.73	receptor-like protein kinase-related family protein
AT5G49680	1.17	-1.41	Golgi-body localisation protein domain; RNA pol II promoter Fmp27 protein domain
AT5G50140	4.26	-4.26	Ankyrin repeat family protein
AT5G50200	1.24	-1.07	nitrate transmembrane transporters
AT5G50260	2.34	-3.30	Cysteine proteinases superfamily protein
AT5G52390	2.18	-2.02	PAR1 protein
AT5G52760	1.50	-2.27	Copper transport protein family
AT5G52770	4.75	-3.00	Copper transport protein family
AT5G53110	1.64	-1.97	RING/U-box superfamily protein
AT5G53320	1.02	-1.14	Leucine-rich repeat protein kinase family protein
AT5G54060	1.90	-1.28	UDP-glucose:flavonoid 3-o-glucosyltransferase
AT5G54062	3.66	-3.88	
AT5G54610	2.58	-1.55	ankyrin
AT5G54700	2.32	-2.09	Ankyrin repeat family protein
AT5G55170	1.07	-1.35	small ubiquitin-like modifier 3
AT5G55410	1.76	-4.45	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT5G55450	2.68	-2.44	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT5G55460	1.00	-3.33	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT5G56050	1.20	-2.57	
AT5G56300	3.07	-1.29	gibberellic acid methyltransferase 2
AT5G56795	3.24	-2.20	metallothionein 1B
AT5G58860	1.55	-1.11	cytochrome P450, family 86, subfamily A, polypeptide 1
AT5G59330	1.57	-1.48	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
AT5G59490	2.66	-2.32	Haloacid dehalogenase-like hydrolase (HAD) superfamily protein
AT5G59660	1.18	-1.09	Leucine-rich repeat protein kinase family protein

Table 2.6 continued

AGI number	log$_2$FC (*ref4-3* vs Col-0)	log$_2$FC (*ref4-3 cdk8-1* vs *ref4-3*)	short description
AT5G59670	1.22	-1.67	Leucine-rich repeat protein kinase family protein
AT5G59820	1.43	-2.44	C2H2-type zinc finger family protein
AT5G60280	1.12	-2.45	Concanavalin A-like lectin protein kinase family protein
AT5G62150	1.64	-1.41	peptidoglycan-binding LysM domain-containing protein
AT5G62770	1.03	-1.81	Protein of unknown function (DUF1645)
AT5G63225	3.61	-2.58	Carbohydrate-binding X8 domain superfamily protein
AT5G63560	1.24	-1.59	HXXXD-type acyl-transferase family protein
AT5G64000	1.97	-3.87	Inositol monophosphatase family protein
AT5G64810	1.14	-4.52	WRKY DNA-binding protein 51
AT5G65090	1.42	-1.19	DNAse I-like superfamily protein
AT5G66150	1.14	-1.61	Glycosyl hydrolase family 38 protein
AT5G66690	5.70	-3.52	UDP-Glycosyltransferase superfamily protein
AT5G67310	1.16	-2.78	cytochrome P450, family 81, subfamily G, polypeptide 1
AT5G67340	1.30	-2.00	ARM repeat superfamily protein
AT5G67450	1.74	-2.72	zinc-finger protein 1

Table 2.7 Gene ontology analysis for the genes that are up-regulated in *ref4-3* and that have restored expression in *ref4-3 cdk8-1*

| Term	Description	Count	Fold Enrichment	FDR
GO:0010345	suberin biosynthetic process	10	21.86	2.2E-07
GO:0051707	response to other organism	13	10.03	5.4E-06
GO:0009627	systemic acquired resistance	10	9.60	9.4E-04
GO:0006869	lipid transport	14	6.26	4.4E-04
GO:0042742	defense response to bacterium	23	3.34	2.3E-03
GO:0006952	defense response	40	2.85	1.0E-05
GO:0007165	signal transduction	27	2.76	8.1E-03
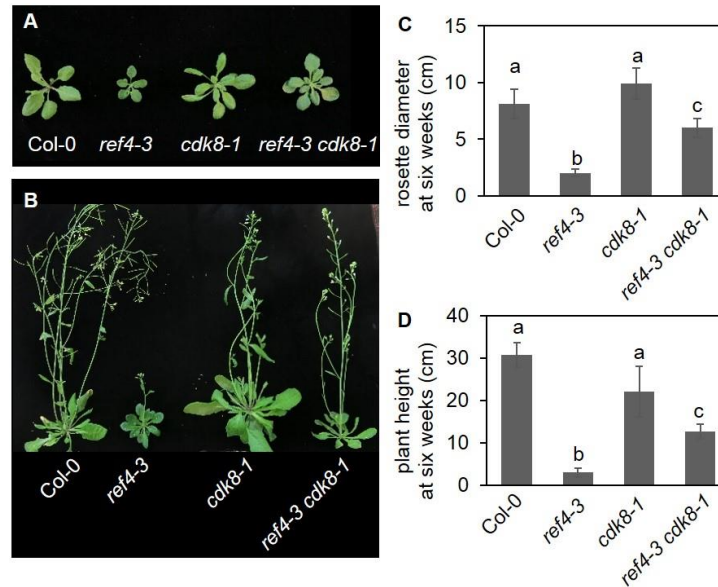
Figure 2.1 CDK8 is required for *ref4-3* to repress plant growth

(A-B) Representative photographs of *ref4-3 cdk8-1* compared to wild-type *Arabidopsis thaliana* (Columbia-0, Col-0), *ref4-3* and *cdk8-1*. *cdk8-1* is a T-DNA insertion line of CDK8, a subunit of the Mediator kinase module. Soil-grown plants were compared three weeks (A) or six weeks (B) after planting.

(C-D) Height (C) and rosette diameter (D) measurement of *ref4-3 cdk8-1* together with wild type, *ref4-3* and *cdk8-1* after growth on soil for six weeks. Data represent mean ± standard deviation (SD) (n=10). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to c to represent difference between groups.
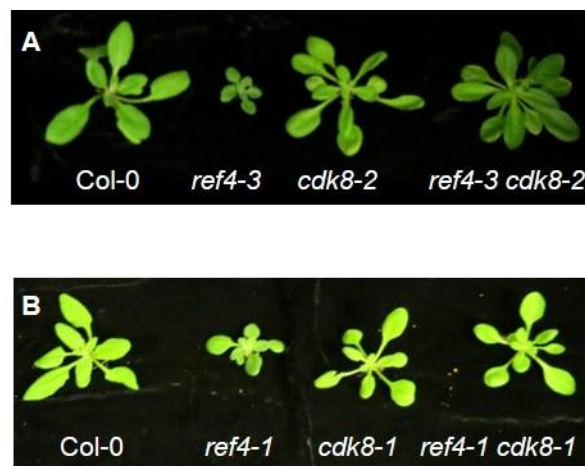
Figure 2.2 CDK8 is required for growth repression in *ref4* mutants

(A) Representative photograph of wild type, *ref4-3*, *cdk8-2* and *ref4-1 cdk8-2*. Plants were compared three weeks after planting.
(B) Representative photograph of *ref4-1 cdk8-1* compared to wild type, *ref4-1* and *cdk8-1*. Soil-grown plants were compared three weeks after planting.
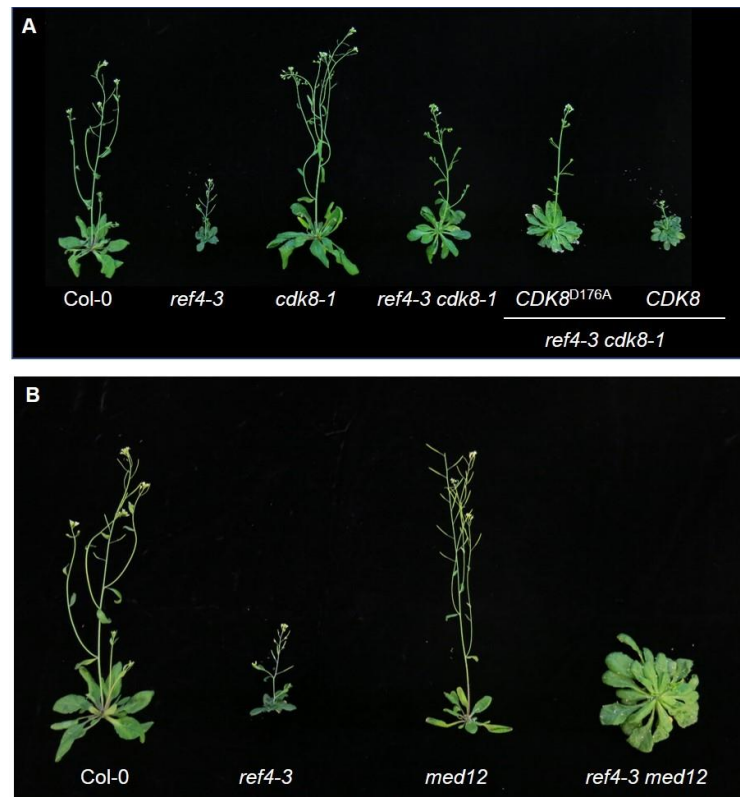
Figure 2.3 Elimination of CDK8 kinase activity is sufficient to suppress the dwarfism of *ref4-3*

(A) Five-week-old soil-grown *Arabidopsis thaliana* transgenic lines overexpressing *CDK8* in a *ref4-3 cdk8-1* background (*ref4-3 cdk8-1 CDK8*) together wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* respectively. *CDK8*^{D176A} indicates a kinase-dead version of CDK8 which carries a D to A mutation at residue 176.
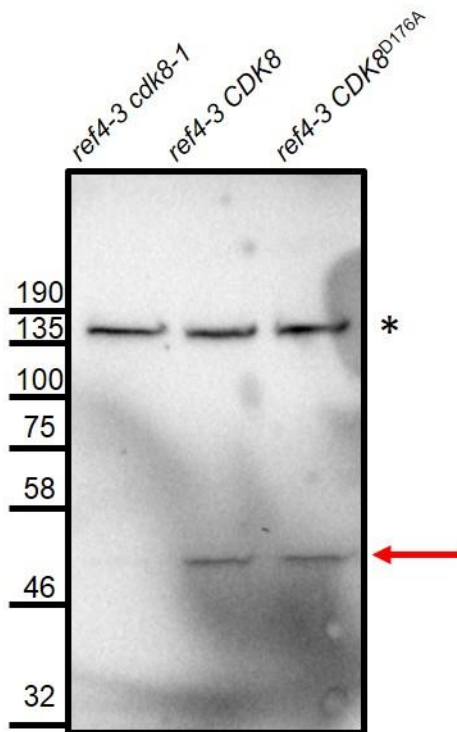(B) Five-week-old soil-grown *ref4-3 med12* compared to wild type, *ref4-3* and *med12* respectively.

Figure 2.4 Wild-type and kinase-dead CDK8 are expressed at similar levels in the *ref4-3 cdk8-1* background

Crude protein extracts were prepared from 2-week-old seedlings of *ref4-3 cdk8-1*, *ref4-3 CDK8* and *ref4-3 CDK8*[D176A] respectively. Anti-MYC antibodies recognize a band corresponding to the size of MYC-tagged CDK8 (50 kDa) only in *CDK8* transgenic plants (red arrow), whereas a non-specific band of higher molecular weight can be detected in all three samples ('*').
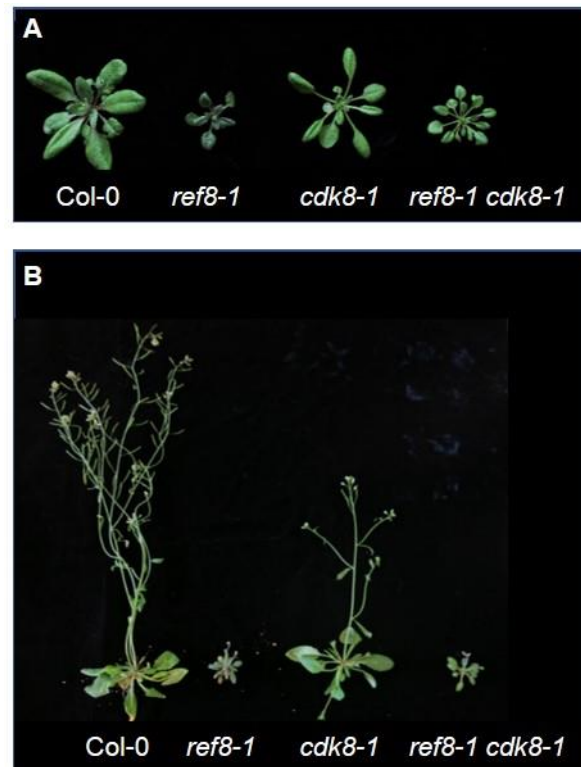
Figure 2.5 CDK8 is not necessary for *ref8-1* to repress plant growth

(A-B) Three-week-old (A) and five-week-old (B) soil-grown *ref8-1 cdk8-1* compared to wild type, *ref8-1* and *cdk8-1*.
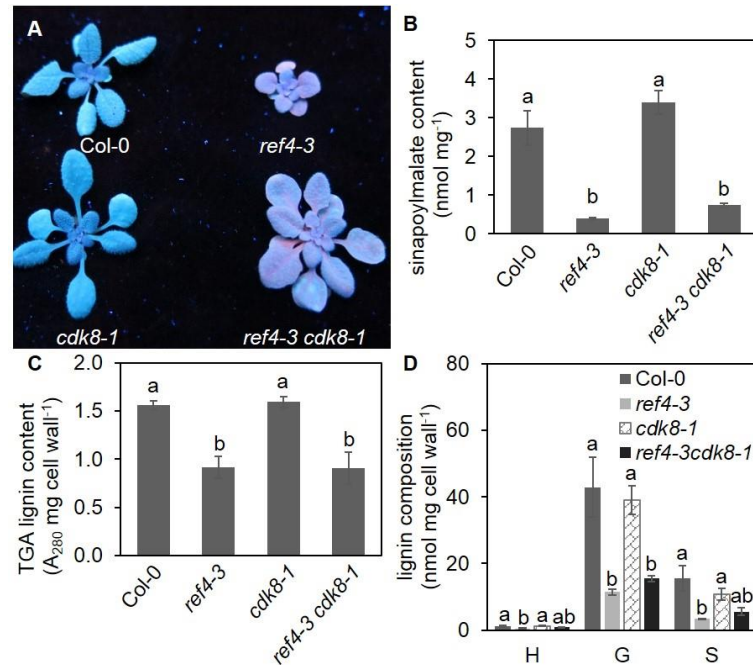
Figure 2.6 *ref4-3* represses phenylpropanoid metabolism independent of CDK8 in *Arabidopsis thaliana*

(A) Representative photograph of wild type, *ref4-3, cdk8-1* and *ref4-3 cdk8-1* under ultraviolet (UV) light. Plants were compared three weeks after planting.

(B) Sinapoylmalate content of three-week-old plants from each genotype determined by high-performance liquid chromatography (HPLC).

(C) Total lignin content in seven-week-old stem tissues quantified by thioglycolic acid (TGA) lignin analysis.

(D) Lignin monomer composition in seven-week-old stem tissues determined by the derivatization followed by reductive cleavage (DFRC) method. The p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin subunit contents were quantified and normalized to the weight of dried cell wall samples.

For panels B-D, data represent mean ± SD (n=3). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to b.
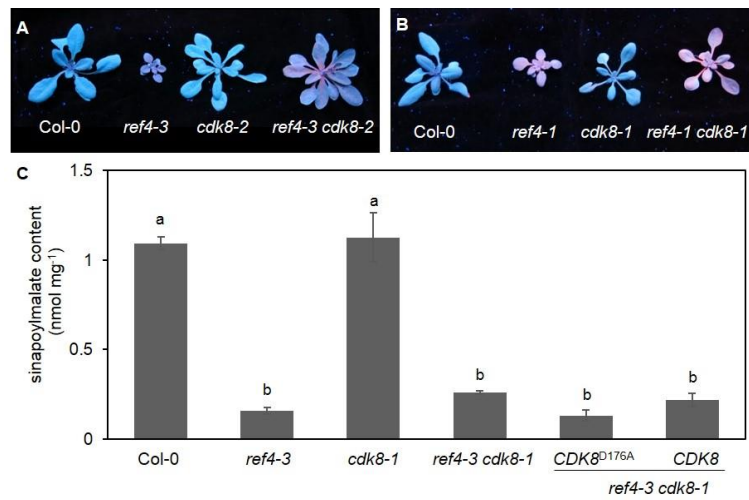
Figure 2.7 CDK8 is dispensable for reduced phenylpropanoid accumulation in *ref4* mutants

(A) Representative photograph of wild type, *ref4-3*, *cdk8-2* and *ref4-1 cdk8-2* under UV light. Plants were compared three weeks after planting.

(B) Representative photograph of *ref4-1 cdk8-1* compared to wild type, *ref4-1* and *cdk8-1* under UV light. Soil-grown plants were compared three weeks after planting.

(C) Sinapoylmalate content of three-week-old wild-type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1, ref4-3 cdk8-1 CDK8* and *ref4-3 cdk8-1 CDK8* [D176A]. Data represent mean ± SD (n=3). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to b.
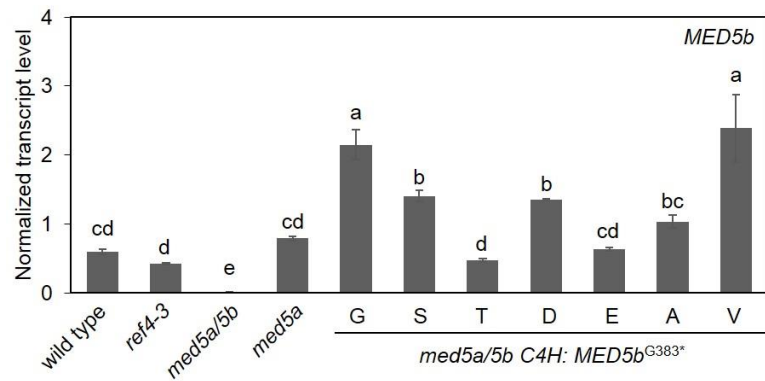
Figure 2.8 *MED5b* transgene is expressed at similar level in the selected transgenic mutants, all of which are comparable or more than expression of *MED5b* in wild type

Expression of *MED5b* was normalized to the reference gene *At1g13220* in three-week-old wild type, *ref4-3*, *med5a/5b*, *med5a* and *MED5b*[G383*] mutants, determined by quantitative PCR analysis. Data represent mean ± SD (n=3). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to e.
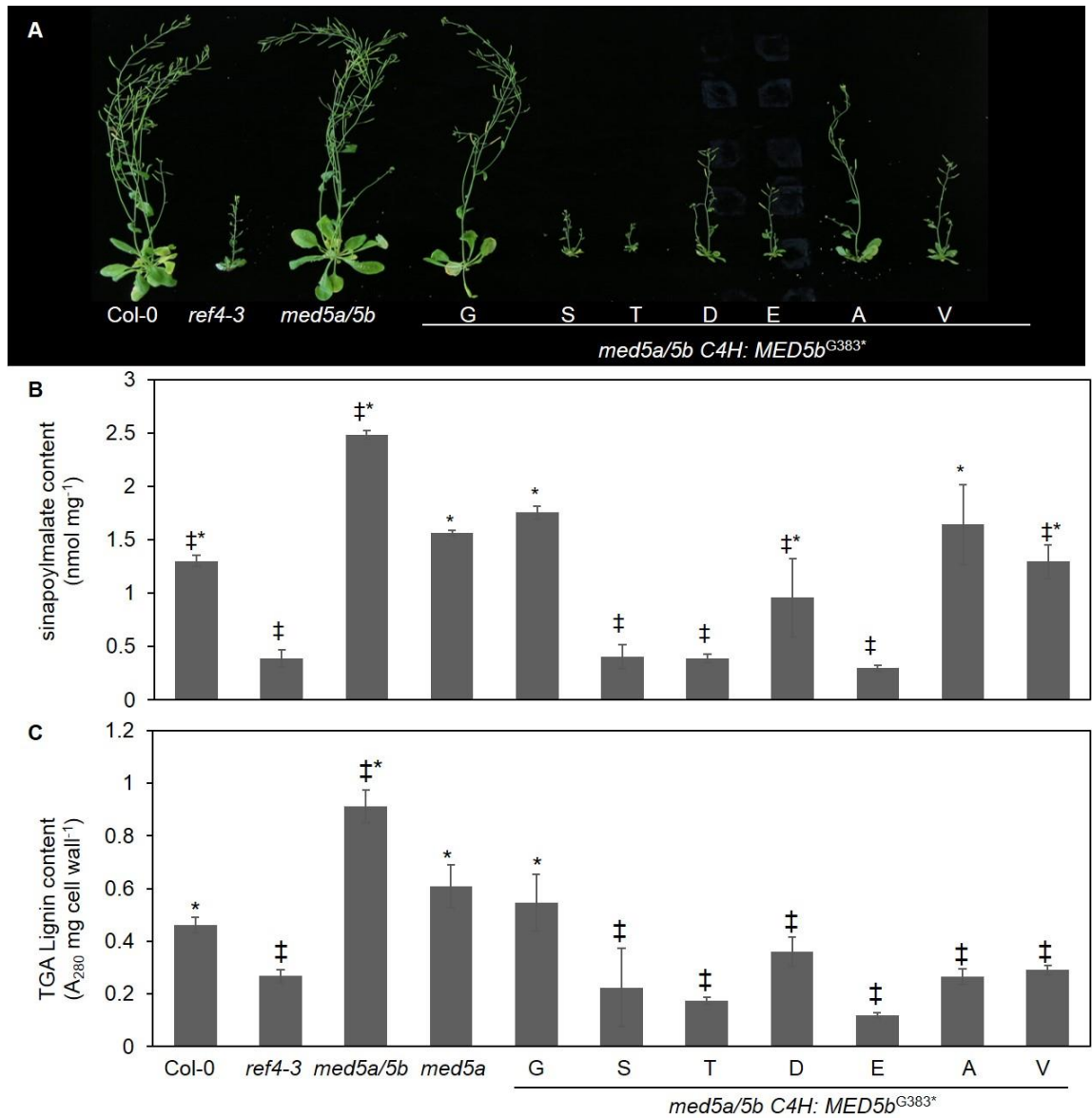
Figure 2.9 The stunted growth and reduced phenylpropanoids of *ref4-3* is not dependent on the phosphorylation event introduced by the G383S mutation

(A) Representative photograph of wild type, *ref4-3*, *med5a/5b*, *med5a* and *C4H* promoter-driven site-directed *MED5b* mutants at G383 site in *med5a/5b* background (*med5a/5b C4H: MED5b*^{G383*}). Plants were compared six weeks after planting. '*' represents the amino acid substitution including G, S, T, D, E, A and V.
(B) Quantification of sinapoylmalate content in three-week-old wild type, *ref4-3*, *med5a/5b*, *med5a* and *MED5b*^{G383*} mutants.
(C) Quantification of lignin content in seven-week-old wild type, *ref4-3*, *med5a/5b*, *med5a* and *MED5b*^{G383*} transgenics.
For panels B-C, data represent the mean ± SD (n = 3). ‡ and * indicate $p < 0.05$ (Dunnett's test) when compared to *MED5b*^{G383G} and *MED5b*^{G383S}, respectively.
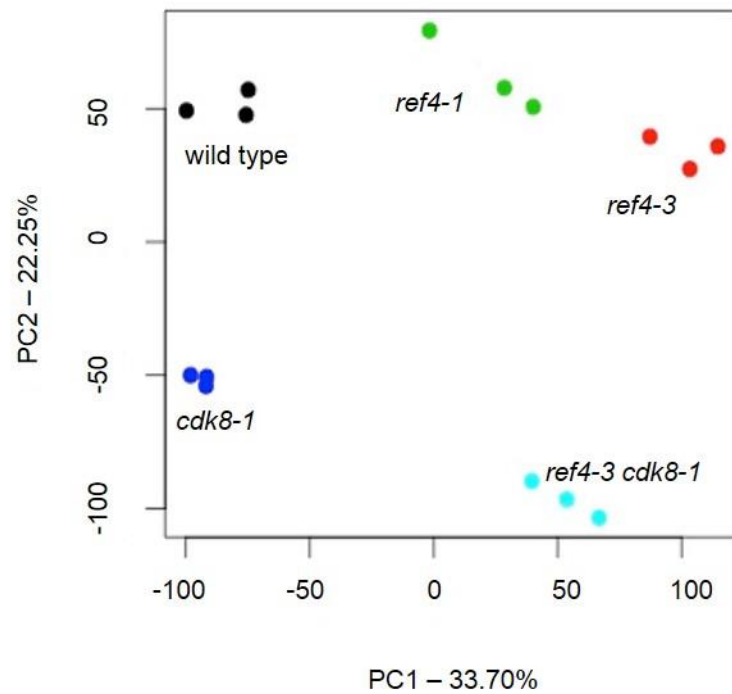
Figure 2.10 PCA of the RNA-seq samples

PCA reveals the samples of the same genotype clustered together, whereas the samples of different genotypes were separated based on the distance between each sample along the first two principal components. Each dot represents one sample, and the dots with the same color represent the samples of the same genotype.
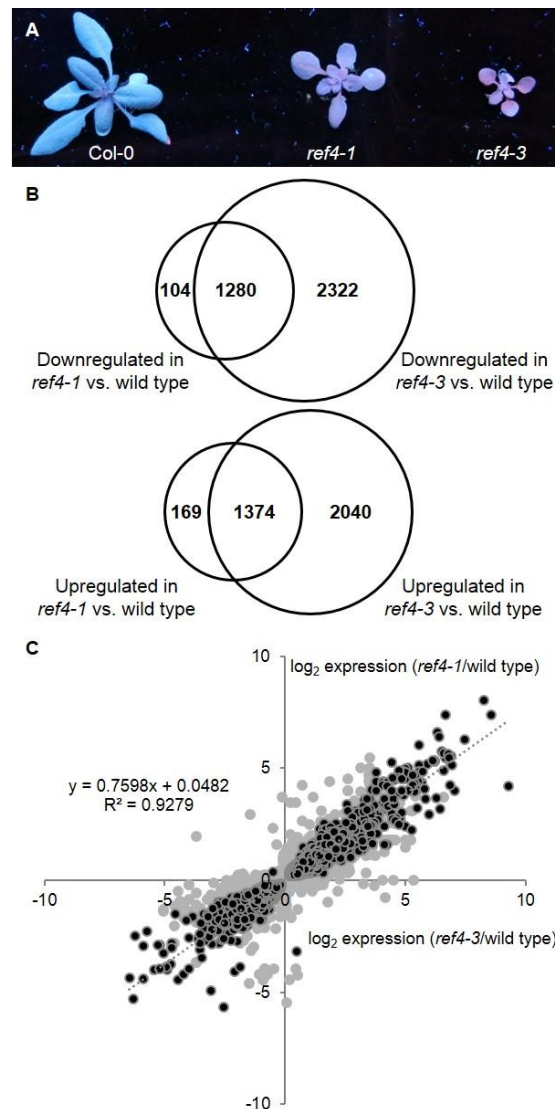
Figure 2.11 Transcriptional reprogramming in *ref4* mutants reflects the severity of alleles

(A) Representative photograph of wild-type, *ref4-1* and *ref4-3* plants under UV light. Plants were compared three weeks after planting.

(B) Venn diagrams representing the number of genes with significant decreased (top) or increased (bottom) expression in *ref4-1* or *ref4-3* compared to wild type (false discovery rate FDR < 0.05).

(C) Scatter plot representing the genes significantly mis-regulated in *ref4-1* and/or *ref4-3* compared to wild type (FDR < 0.05). Each point represents one gene, with the fold change of its expression level (counts per million reads, CPM) in *ref4-3* compared to wild type (x axis) and in *ref4-1* compared to wild type (y axis) respectively. The genes mis-regulated in both *ref4-1* and *ref4-3* are depicted as black points, while the genes mis-regulated only in *ref4-1* or *ref4-3* were depicted as grey points. The formula and the coefficient value (R^2) in the plot represent the correlation between the expression of the genes being mis-regulated in both *ref4-1* and *ref4-3*.
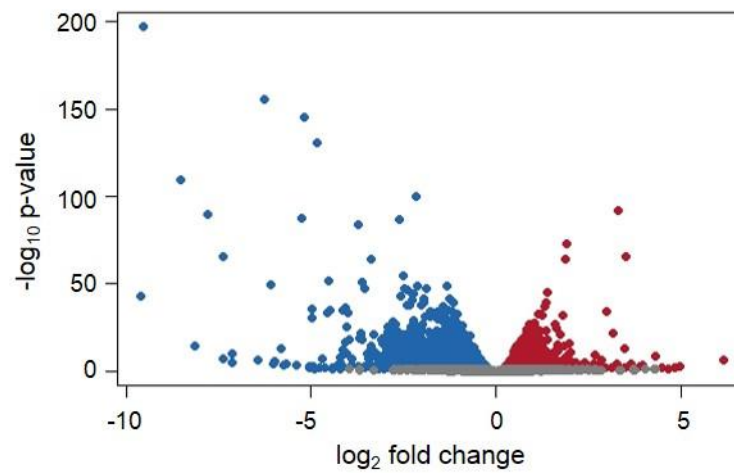
Figure 2.12 Comparison between wild type and *cdk8* mutants

A volcano plot representing genes that are differentially expressed in *cdk8-1* compared to wild type using RNA-seq analysis. For each gene represented by a dot, the x axis shows the fold change compared to wild type (log$_2$FC), whereas the y axis shows the negative value of log$_{10}$(false discovery rate) ($-$ log$_{10}$[FDR]). The genes that are significantly differentially expressed compared to wild type (FDR < 0.05) are highlighted in blue (down-regulated) or red (up-regulated), while the others are colored in grey.

Figure 2.13 Phenylpropanoid biosynthetic genes are generally repressed in *ref4-3* and *ref4-3 cdk8-1*

The expression level (fragments per kilobase per million, FPKM) of phenylpropanoid biosynthetic genes were determined by RNA-seq analysis. Data represent mean ± SD (n=3). '*' indicates FDR < 0.05 compared to wild type using edgeR analysis.

Figure 2.14 Disruption of *CDK8* rescues gene expression changes in the *ref4-3* mutant

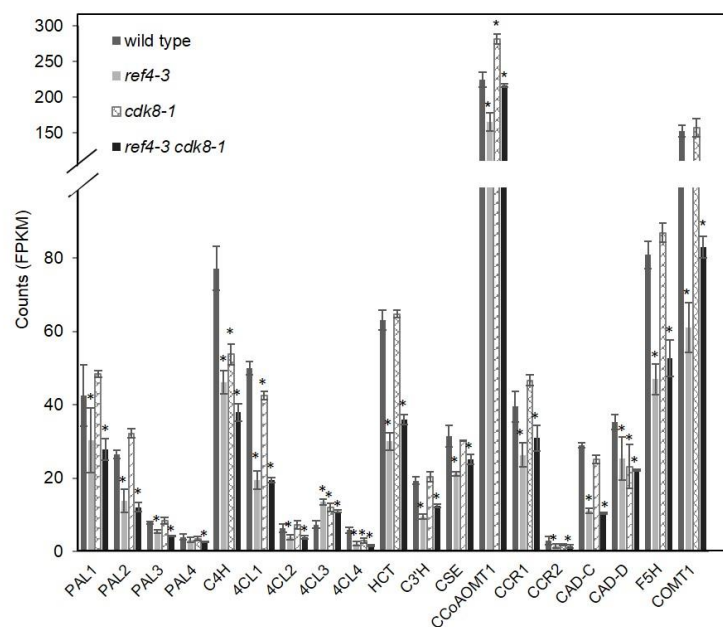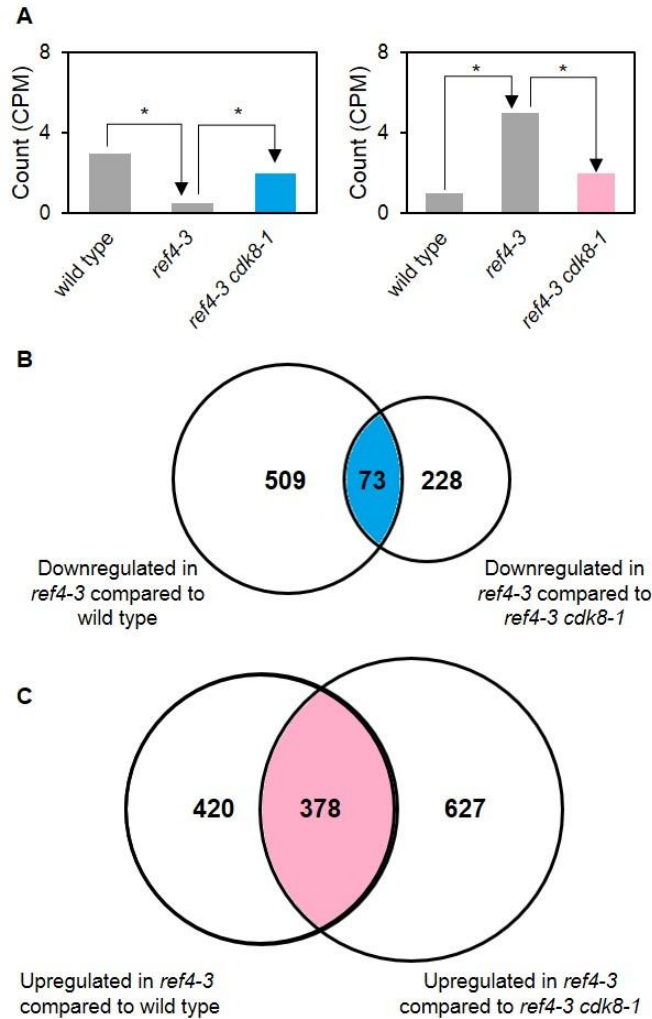(A) idealized histograms demonstrating the criteria for growth-related gene targets of interest. The potential gene targets should either be down-regulated in *ref4-3* compared to wild type with at least partial restoration of expression in *ref4-3 cdk8-1* compared to *ref4-3* (left), or up-regulated in *ref4-3* compared to wild type and at least partially repressed in *ref4-3 cdk8-1* compared to *ref4-3* (right). '*' represents the significant difference of gene expression in two genotypes (FDR <0.05, absolute value of $\log_2$FC > 1).

(B) The number of genes with significantly decreased expression in *ref4-3* compared to wild type is represented by the left Venn diagram, while the number of genes with significantly decreased expression in *ref4-3* compared to *ref4-3 cdk8-1* is represented by the right Venn diagram (FDR < 0.05, absolute value of $\log_2$FC > 1).

(C) The number of genes with significantly increased expression in *ref4-3* compared to wild type is represented by the left Venn diagram, while the number of genes with significantly decreased expression in *ref4-3* compared to *ref4-3 cdk8-1* is represented by the right Venn diagram (FDR < 0.05, absolute value of $\log_2$FC > 1).

For panels B-C, the overlapping region represents the genes that fit the criteria in the left histogram and the right one in (A) respectively.
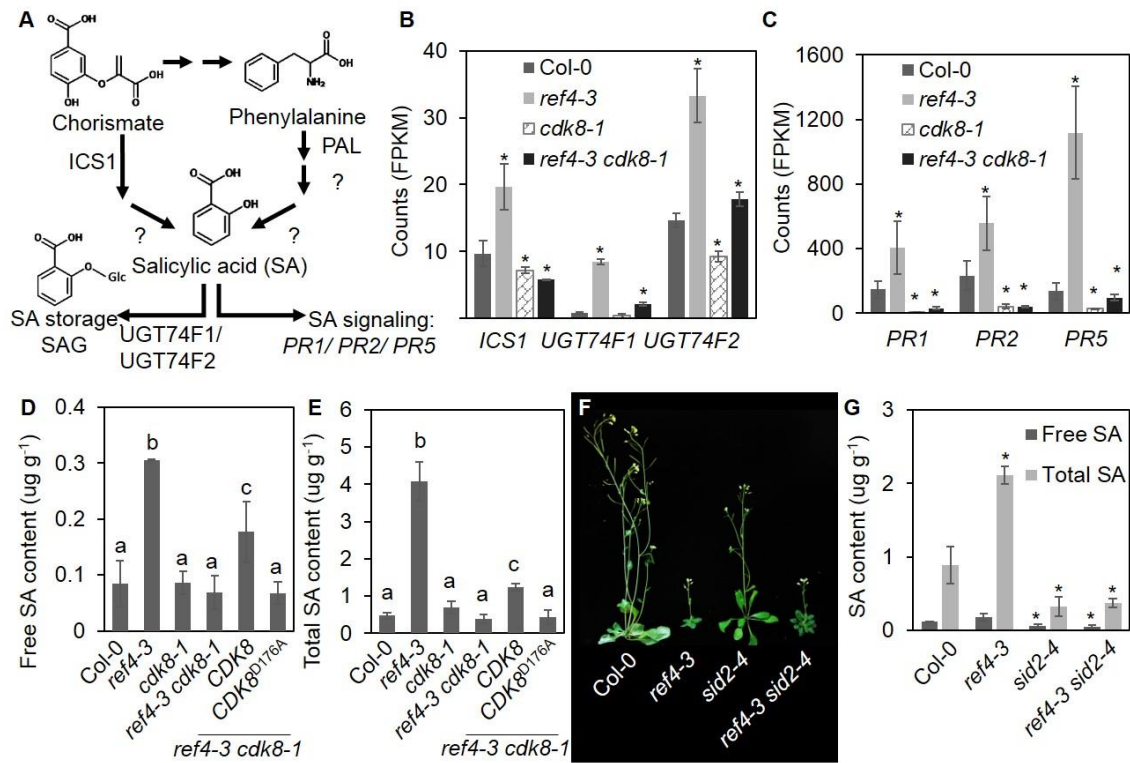
Figure 2.15 Hyper-accumulation of SA in *ref4-3* is dependent on CDK8 in *Arabidopsis thaliana*, but it is not the major cause of dwarfing in *ref4-3*

(A) The SA biosynthesis and signaling pathways. SA is synthesized from the precursor chorismate via isochorismate synthase 1 (ICS1). Phenylalanine ammonia lyase (PAL) catalyzes the first step in a less significant SA biosynthetic pathway which has yet to be fully elucidated. SA can either be converted to its glucoside form for storage by UDP-glucose dependent glucosyltransferases (UGT) UGT74F1 or UGT74F2, or serve as signal molecules for plant development and stress responses. The *pathogenesis-related* (*PR*) genes including *PR1*, *PR2* and *PR5* are marker genes for SA signaling.

(B-C) Expression level (fragments per kilobase per million, FPKM) of SA biosynthetic genes (B) and SA signaling marker genes (C) in wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1*, determined by RNA-seq analysis. Data represent mean ± SD (n=3). '*' indicates FDR < 0.05 compared to wild type.

(D-E) Free SA (D) and total SA (E) in wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1, ref4-3 cdk8-1 CDK8* and *ref4-3 cdk8-1 CDK8* [D176A] quantified by HPLC using fluorescence detection. Rosettes from three-week-old plants were used to perform quantification. Data represent mean ± SD (n=3). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to c.

(F) Five-week-old soil-grown *ref4-3 sid2-4* compared to wild type, *ref4-3* and *sid2-4* respectively.

(G) Free SA and total SA in wild-type, *ref4-3*, *sid2-4* and *ref4-3 sid2-4* quantified by HPLC using fluorescence detection. Rosettes from three-week-old plants were used to perform quantification. Data represent mean ± SD (n=3). '*' indicates $p < 0.05$ compared to wild type according to Student's t-test.

Figure 2.16 The stunted growth of *ref4-3* is independent of NPR1

Representative photograph of *ref4-3 npr1-1* compared to wild type, *npr1-1* and *ref4-3*. Soil-grown plants were compared five weeks after planting.
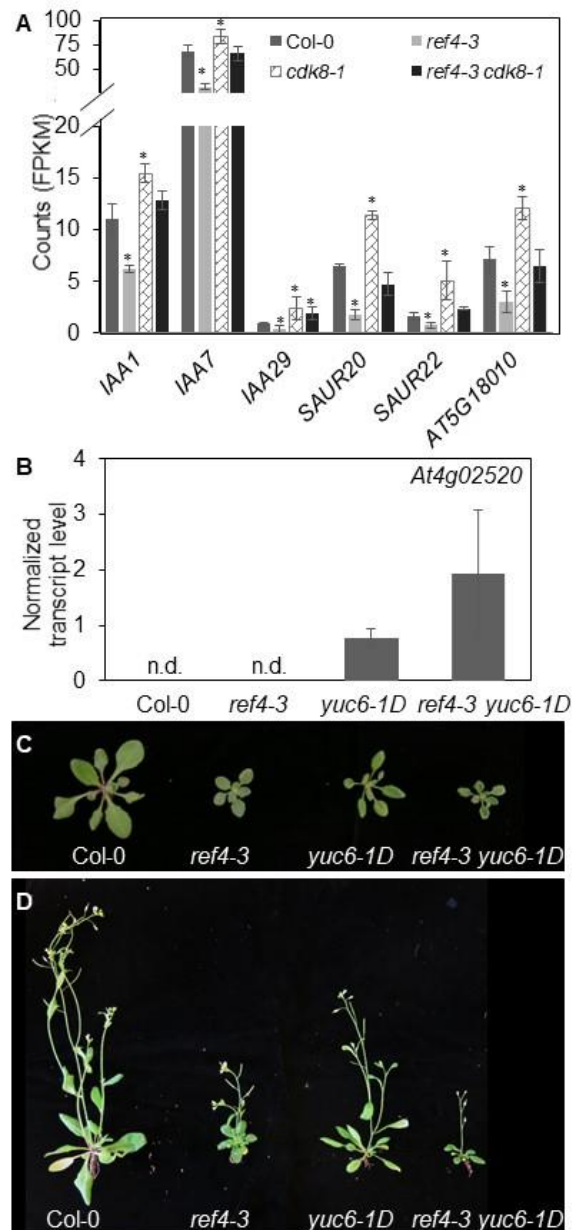
Figure 2.17 Enhanced auxin accumulation does not restore the stunted growth of *ref4-3*

(A) Expression level (FPKM) of major auxin signaling genes in wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1*, determined by RNA-seq analysis. Data represent mean ± SD (n=3). '*' indicates FDR < 0.05 compared to wild type using EdgeR analysis.
(B) Expression of *At4g02520* normalized to the reference gene *At1g13220* in wild type, *ref4-3*, *yuc6-1D* and *ref4-3 yuc6-1D*, determined by quantitative PCR analysis. Data represent mean ± SD (n=3). The expression of *At4g02520* is not detectable (n.d.) in wild type and *ref4-3*.
(C-D) Representative photographs of three-week-old (C) and five-week-old (D) soil-grown *ref4-3 yuc6-1D* compared to wild type, *ref4-3* and *yuc6-1D*.
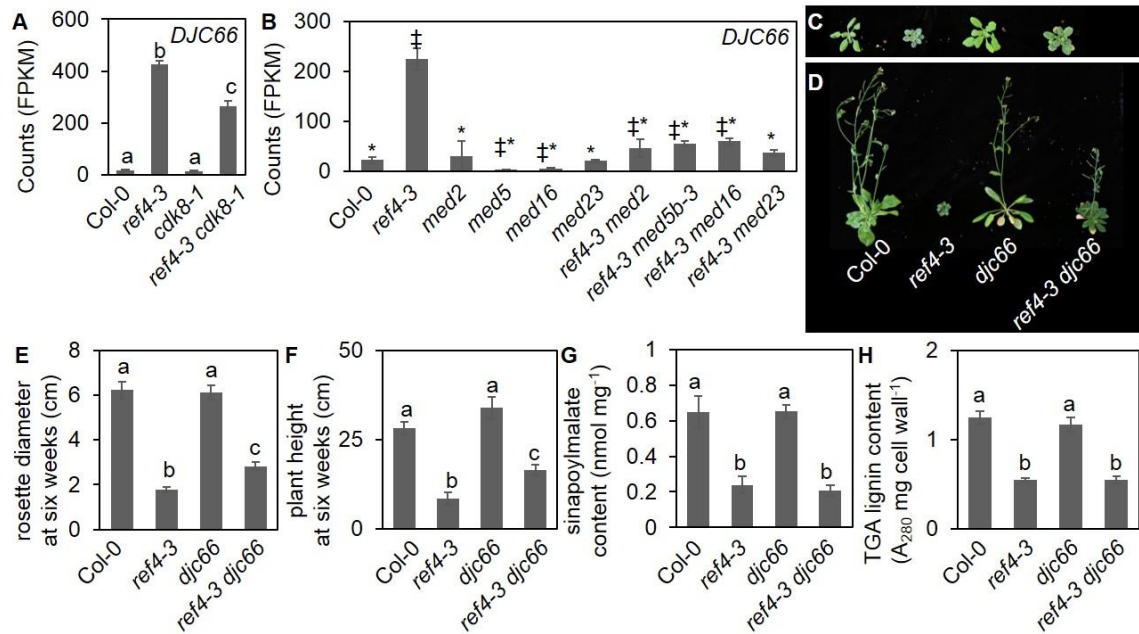
Figure 2.18 Disruption of *DJC66* partially restores the dwarfism of *ref4-3*

(A) Expression level (FPKM) of *DJC66* in wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1*. Data represent mean ± SD (n=3). Statistically significant differences (FDR < 0.05) were indicated by a to c.

(B) Expression level (FPKM) of *DJC66* in wild type, *ref4-3*, *med* T-DNA lines and *med ref4-3* double mutants determined by a previous RNA-seq analysis (Dolan et al., 2017). Data represent mean ± SD (n = 3). ‡ and * indicate FDR < 0.05 (EdgeR analysis) when compared to Col-0 and *ref4-3*, respectively.

(C-D) Three-week-old (C) and six-week-old (D) soil-grown *ref4-3 djc66* together with wild type, *ref4-3* and *djc66*.

(E-F) Height (E) and rosette diameter (F) measurement of *ref4-3 djc66* together with wild type, *ref4-3* and *djc66* after growth on soil for three weeks and six weeks respectively.

(G) Sinapoylmalate content of three-week-old wild-type, *ref4-3*, *djc66* and *ref4-3 djc66* plants determined by HPLC.

(H) Total lignin content in six-week-old stem tissues quantified by TGA lignin analysis.

For panel E-F, Data represent mean ± SD (n=10). For panel G-H, Data represent mean ± SD (n=3). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to c.
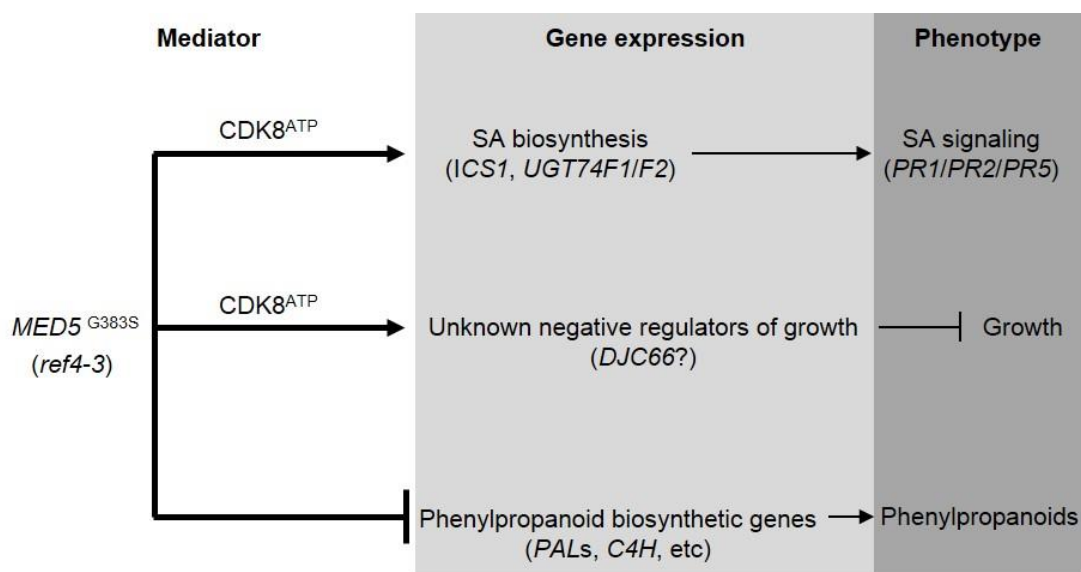
Figure 2.19 A model of the genetic interaction between *CDK8* and *ref4-3*

The findings of this study are summarized into the genetic interaction between *ref4-3* and *CDK8* in the Mediator complex (white background), mis-regulated gene targets (light grey background) and the resulting phenotypes (dark grey background). Particularly, *ref4-3* requires CDK8 with intact kinase activity (indicated by CDK8^{ATP}) to activate genes involved in SA biosynthesis and therefore leads to enhanced SA signaling. The kinase activity of CDK8 is required for growth repression of *ref4-3*; however, elimination of the kinase activity of CDK8 does not rescue the down-regulated phenylpropanoid metabolism in *ref4-3*. *DJC66* is one of the targets that are related to the dwarfism of *ref4-3*, which could be downstream of CDK8.

CHAPTER 3.   ARABIDOPSIS MED5 IS INVOLVED IN THE REGULATION OF SHADE AVOIDANCE SYNDROME AND ABA HOMEOSTASIS

3.1   Introduction

Plants growing at high density compete with one another for limited light resources. Under such conditions, the ability to detect and respond to nearby competitors is critical to maximize survival and reproductive success. Due to the absorption by chlorophyll of red, but not far-red light, light that has passed through or been reflected by plant vegetative tissues exhibits a characteristic reduction in it's red to far-red ratio (R:FR). Thus, the detection of this aspect of light quality can predict the proximity of a nearby plant (Franklin and Whitelam, 2005). In response to this threat of competition, shade-intolerant species such as Arabidopsis initiate a series of rapid morphological responses aimed at overtopping their competitors. These responses include increased elongation growth at the expense of leaf expansion, more vertical orientation of leaves, and inhibition of lateral branching (Morelli and Ruberti, 2002; Franklin, 2008). Prolonged exposure to low R:FR light indicates a failure to overtop competitors and results in the acceleration of flowering to reduce generation time at the expense of seed yield. These responses are collectively known as the shade avoidance syndrome (SAS).

The molecular mechanisms underlying SAS have been intensively explored over the past three decades, identifying the phytochrome (phy) photoreceptors as central regulators (Franklin, 2008; Roig-Villanova and Martínez-García, 2016; Sessa et al., 2018). Among the five phytochromes (phyA through phyE) encoded in the Arabidopsis genome, phyB is primarily responsible for mediating shade avoidance (Franklin and Quail, 2010),

with relatively minor contributions from phyD and phyE (Casal et al., 2013). Phytochromes can respond to changes in the ratio of R:FR and exist in either an active FR-absorbing Pfr form under high R:FR or an inactive R-absorbing Pr form under low R:FR (Roig-Villanova and Martínez-García, 2016), thus providing a sensing mechanism for changes in the ratio of R:FR. Given that activated phytochromes can target the phytochrome interacting factors (PIFs) transcription factors for ubiquitination and subsequent degradation, inactivation of one or more phytochromes stabilizes the PIFs transcription factors (Bauer, 2004; Park et al., 2004, 2012), resulting in transcriptional reprogramming that promotes the shade avoidance response (Sessa et al., 2005; Leivar et al., 2012). Besides changes in leaf morphology and flowering time, abscisic acid (ABA), a phytohormone critical for cold and drought stress responses, can be induced by low R/FR treatment (Cagnola et al., 2012; Gonzalez-Grandio et al., 2013; Holalu and Finlayson, 2017), suggesting that phytochrome-regulated shade avoidance responses may also be involved in the regulation of hormone signaling.

In addition to Phys and PIFs, Mediator (MED), a transcriptional co-regulatory complex in eukaryotes, is also involved in SAS regulation. Conserved throughout eukaryotes, Mediator is required for the transmission of information from gene specific transcription factors to the basal transcriptional machinery (Malik and Roeder, 2010). The Arabidopsis Mediator tail module subunit MED25 was originally identified as PHYTOCHROME AND FLOWERING TIME 1 (PFT1) in a screen for mutants that showed altered sensitivity of hypocotyl elongation to pulses of red and far-red light in conjunction with late flowering time phenotypes (Cerdán and Chory, 2003). Given that Mediator tail module subunits co-function in many biological processes including cold

stress responses, iron homeostasis and phenylpropanoid metabolism (Hemsley et al., 2014), other MED tail module subunits may be involved in SAS regulation as well. Interestingly, multiple *med* tail mutants including *med2*, *med14* and *med16* are susceptible to cold stress (Hemsley et al., 2014), suggesting that the Mediator complex is critical for ABA signaling, however, the function of Mediator in maintaining normal levels of ABA biosynthesis has not been investigated, and how the signals from light and those from hormones are integrated remains unknown.

In this study, we characterize SAS in a mutant with disrupted MED5a and MED5b (*med5*) and a semi-dominant *med5b* mutant (*reduced epidermal fluorescence 4*, *ref4-3*) (Stout et al., 2008; Bonawitz et al., 2012). In contrast to elevated shade avoidance response in *med5*, our data reveal that *ref4-3* mimics the function of wild-type MED5, resulting in constitutive repressed SAS under non-stressed conditions. Although shade stimulation partially suppresses the growth defects but not reduced lignin content in *ref4-3*, disruption of phyB is not sufficient to restore the stunted growth of *ref4-3*, indicating that the mechanism underlying the dwarfing is independent of phytochrome signaling. Moreover, both *med5* and *ref4-3* accumulate elevated levels of ABA under stressed and non-stressed conditions, suggesting that *ref4-3* can function as a MED5-null mutant for specific biological processes. Together, our data demonstrate that MED5 is involved in the negative regulation of SAS and ABA homeostasis, and the dwarfism of *ref4-3* can be partially rescued by shade stimulation without perturbing lignin biosynthesis.

3.2   Results

3.2.1   Arabidopsis MED5 antagonizes SAS

Multiple Mediator subunits are required for proper flowering time. While *med8*, *med13*, *med15*, *med16*, *med17*, *med18*, *med20a* and *med25* mutants show delayed flowering (Yang et al., 2016), we previously reported that loss of MED5a and MED5b (*med5*), two paralogs in the Arabidopsis Mediator tail module, results in early flowering (Dolan and Chapple, 2018). Given that *ref4-3* is a semi-dominant form of MED5 that represses the phenylpropanoid metabolism, we tested whether *ref4-3* would show a late-flowering phenotype, opposite to the early flowering observed in *med5* mutants.

Compared to wild type, *ref4-3* mutants flowered late under short-day conditions but not under long-day conditions (Figure 3.1). In contrast, *med5* flowered significantly earlier than wild type when grown in short days but not in long days (Figure 3.1). Hence, flowering time was oppositely mis-regulated in *ref4-3* and *med5* under short-day conditions. In addition to early flowering time, loss of MED5 resulted in other growth phenotypes typical of SAS including elongated petioles and hypocotyls (Figure 3.2), suggesting that MED5 antagonizes SAS (Figure 3.2). We also observed that hypocotyl length of *med5a* and *med5b* did not significantly differ from that of wild type (Figure 3.2B), suggesting that MED5a and MED5b play redundant roles in the regulation of shade avoidance response. In contrast to *med5*, *ref4-3* showed short petioles and hypocotyls (Figure 3.2), suggesting that SAS is repressed in *ref4-3*. Together with the disturbed flowering time in *ref4-3* and *med5* mutants, our results are consistent with a model in which MED5 is required for the normal regulation of SAS. Further, *ref4-3* may mimic and exaggerate the function of wild-type MED5 in SAS, leading to constitutive repressed shade avoidance responses.

3.2.2   End-of-day far-red light treatment rescues the stunted growth in *ref4-3*

Mutants with constitutively repressed SAS display multiple photomorphogenic defects including reduced rosette size and shorter petioles (Christians et al., 2012), opposite to the phenotypes observed in phytochrome-deficient mutants that have enhanced SAS (Franklin, 2003). Consistent with the observation that *ref4-3* phenocopies the mutants with constitutively repressed SAS, we reanalyzed the previous RNA-seq data (Dolan et al., 2017) and found that multiple phytochromes including *phyB*, *phyD* and *phyE* were up-regulated in *ref4-3* compared to wild type (Figure 3.3), whereas the positive regulators including *PIF*s and PIF-targeted genes displayed coordinated down-regulation in *ref4-3* mutants (Figure 3.3). Moreover, neither phytochromes nor phytochrome-targeted genes displayed drastic gene expression changes in *med5* (Figure 3.3), suggesting that MED5 may negatively affect shade avoidance responses through phytochrome-independent pathways.

Given that phytochromes are deactivated under shade conditions with enriched far-red light (Roig-Villanova and Martínez-García, 2016), to test whether the dwarfism of *ref4-3* is contributed by inhibited shade avoidance responses, we treated the plants with end-of-day far-red (EOD FR) light treatment, which included an additional 30-minute far-red light at the end of each photosynthetic period. We found that the stunted growth of *ref4-3* was partially restored under EOD-FR light treatment (Figure 3.4A – 3.4C) but phenylpropanoid accumulation in *ref4-3* remained low (Figure 3.4D), suggesting that *ref4-3* inhibits plant growth partially through repressing SAS. Moreover, in addition to disruption of CDK8, a subunit of the Mediator kinase module (Mao et al., 2019). our data demonstrate that reduced lignin content and dwarfism in *ref4-3* can be dissociated by shade stimulation as

well, buttressing our finding that the reduced lignin content is not causative for the dwarfism of *ref4-3*.

Given that phyB plays a dominant role in SAS inhibition relative to all other phytochromes in Arabidopsis (Reed et al., 1993) and that *phyB* was up-regulated in *ref4-3* (Figure 3.3), we tested whether disruption of phyB could rescue the stunted growth of *ref4-3*. To do this, we used *phyB-9*, a previously characterized phyB-null mutant (Reed et al., 1993). Although *phyB-9* showed phenotypes typical for constitutive SAS including elongated hypocotyls in both wild-type and *ref4-3* backgrounds (Figure 3.5A and 3.5C), *ref4-3 phyB-9* was as dwarf as *ref4-3* (Figure 3.5A and 3.5B, 3.5D and 3.5E), indicating that disruption of phyB was not sufficient to rescue the stunted growth of *ref4-3*, and the stunted growth of *ref4-3* is contributed by the constitutively inhibited shade avoidance responses through a phyB-independent pathway.

3.2.3   The photosynthetic machinery is disrupted in *ref4-3*

Shade avoidance responses are often negatively correlated with chlorophyll accumulation (Jackson and Prat, 1996; McCormac et al., 2001; Christians et al., 2012; Zhao et al., 2013b) which can imply the photosynthetic capacity in leaves (Croft et al., 2017). Because *ref4-3* and *med5* showed opposite shade avoidance responses, we sought to characterize the levels of photosynthesis in those two mutants.

To do this, we measured chlorophyll content and photosynthetic parameters in *ref4-3* and *med5* together with wild type. Consistent with the appearance of its dark-green rosettes, chlorophyll content in *ref4-3* was significantly higher than wild type (Figure 3.6A). In addition, non-photochemical quenching (NPQ), a key parameter for photosynthetic

machinery, was up-regulated in *ref4-3* under high light intensities (350 µmol m^{-2}s^{-1}) (Figure 3.6B), suggesting an imbalance between photosystem I and photosystem II (Wilson et al., 2006). In contrast, *med5* accumulated wild-type levels of chlorophyll, and there was no significant difference between *med5* and wild type for all the parameters we examined (Figure 3.6B and 3.6C, Table 3.1). Consistent with these phenotypes, our RNA-seq analysis revealed that the down-regulated genes in *ref4-3* compared to wild type were enriched for processes related to photosynthesis (Mao et al., 2019). In contrast, loss of MED5 resulted in little effect upon photosynthesis-related genes (Dolan et al., 2017). Hence, we conclude that MED5 is not necessary for plants to support wild-type levels of photosynthesis, and the perturbed photosynthesis in *ref4-3* is likely a neumorphic effect induced by the G383S mutation.

3.2.4 Both *ref4-3* and *med5* over-accumulate ABA under normal and drought-stressed conditions

Many plant hormones show altered accumulation in response to changes of light quality and quantity. For instance, ABA, a phytohormone critical for abiotic stress responses and seed development (Vishwakarma et al., 2017), can be rapidly induced by shade stress (Kurepin et al., 2007; Cagnola et al., 2012). We previously reported that the gene ontology (GO) term 'response to ABA' was significantly enriched for genes that were up-regulated in *ref4-3* and down-regulated in *med5* (Dolan et al., 2017), suggesting that ABA signaling is oppositely mis-regulated in *ref4-3* and in *med5*.

According to a previous RNA-seq analysis (Dolan et al., 2017), multiple ABA biosynthetic and catabolic genes (Hauser et al., 2017) were mis-regulated in *ref4-3* (Figure 3.7A). In contrast, loss of MED5 led to little, if any, mis-regulation of those genes (Figure

3.7A), suggesting that wild-type MED5 is likely to be dispensable for normal ABA biosynthesis. Moreover, multiple ABA signaling genes including *ABC transporter G family member 40* (*ABCG40*) showed up-regulated expression in *ref4-3*, however, only a few ABA signaling genes were down-regulated in *med5* (Figure 3.7A), which is consistent with the modest expression changes of the genes involved in ABA biosynthesis and degradation in *med5* mutants.

To further explore the function of MED5 in ABA signaling, we reanalyzed the RNA-seq data from *ref4-3* and *med5* mutants (Dolan et al., 2017) and compared the results with the genes mis-regulated by ABA treatment (Nemhauser et al., 2006). Notably, 80% of the genes that were mis-regulated by ABA treatment displayed similar changes in gene expression in *ref4-3* (Figure 3.7B and 3.7D), suggesting that ABA signaling is perturbed in *ref4-3* mutants, which may result from abnormal ABA accumulation. In contrast, we did not observe an opposite change in gene expression in *med5* (Figure 3.7C and 3.7E), again suggesting that ABA signaling is not substantially perturbed when MED5 is disrupted.

Consistent with the observation that gene expression changes in *ref4-3* mimics the gene expression profile of wild-type plants under ABA treatment, *ref4-3* accumulated more ABA compared to wild type under well-watered conditions (Figure 3.8A). Considering that increased ABA level can trigger changes in expression of more than 10% of genes in Arabidopsis (Nemhauser et al., 2006), the substantial transcriptome reprogramming in *ref4-3* may result, in part, from its abnormal ABA content. To our surprise, ABA levels were also higher in *med5* than wild type (Figure 3.8A). ABA levels were induced by drought conditions in wild type (McLachlan et al., 2018), and both *ref4-3* and *med5* showed

further increase in ABA levels (Figure 3.8A). Thus, wild-type MED5 is involved in the negative regulation of stress-induced ABA biosynthesis.

Given that ABA is the major signaling molecule for drought stress responses (Fujita et al., 2011), we measured leaf water loss in detached leaves to test whether *ref4-3* or *med5* display any difference in resistance towards drought stress. During the 8-hour period after detachment, *ref4-3* lost less water compared to wild-type plants (Figure 3.8B), which is likely contributed by its elevated level of endogenous ABA. In contrast, surprisingly, no significant difference of water retention was observed in *med5* (Figure 3.8B). The dissociation between ABA content and drought resistance has been previously reported (Lin et al., 2011). Considering that most ABA-responsive genes were not mis-regulated in *med5* under non-stressed conditions, the unchanged response towards drought stress in *med5* may result from its insensitivity to ABA signaling.

Collectively, our data demonstrate that MED5 is required for genes involved in maintaining wild-type levels of ABA and the G383 site in *ref4-3* is necessary for its function; however, abnormal ABA biosynthesis and signaling may be induced in *ref4-3* and *med5* through distinctive mechanisms.

3.3   Discussion

A previous characterization of *med* tail mutants revealed that loss of MED5 results in an early flowering phenotype (Dolan and Chapple, 2018). Here we performed a more detailed phenotypic study of *med5* and the semi-dominant *med5* mutant*, ref4-3*, under both short-day and long-day conditions. In addition to early flowering, we found that *med5* showed other phenotypes typical for shade avoidance syndromes including elongated

hypocotyls and petioles, suggesting that SAS was enhanced when MED5 was disrupted. In contrast, SAS was inhibited in *ref4-3*, which is consistent with the previous notion that *ref4-3* mimics and enhances the action of wild-type MED5 in maintaining phenylpropanoid homeostasis (Bonawitz et al., 2012). Although multiple Mediator subunits are required for plant normal flowering time, MED25 is the only subunit known to be critical for shade avoidance response. As opposed to *med5*, disruption of MED25 leads to repressed SAS (Cerdán and Chory, 2003), suggesting that MED25 promotes SAS. Despite the fact that MED5 and MED25 are both Mediator tail module subunits (Tsai et al., 2014), recent studies reveal that they are required for different biological processes (Bonawitz et al., 2012; Yang et al., 2014), and no genetic interaction has been observed between those two subunits (Dolan et al., 2017). Consistent with their distinctive phenotypes, genes mis-regulated in *med5* and *med25* were enriched for different gene ontologies (Dolan and Chapple, 2018; Davoine et al., 2017). In insects, MED25 predominantly functions in gene activation by interacting with strong activators (Mittler et al., 2003; Bryant and Ptashne, 2003), whereas MED24 (homologous to Arabidopsis MED5) can act as either a co-activator or a co-repressor (Stampfel et al., 2015). Thus, MED5 and MED25 may be involved in shade avoidance response through independent pathways.

In addition to inhibited shade avoidance response, *ref4-3* mutants accumulated increased level of chlorophylls compared to wild type and *med5*. Few studies have focused on the role of Mediator in plant primary metabolism, however, multiple *med* mutants including *med14* and *med16* show pale green leaves (Knight et al., 2009; Zhang et al., 2013b), suggesting that those Mediator subunits may be required for chlorophyll biosynthesis. Perturbed chlorophyll biosynthesis in *ref4-3* could be related to hyperactive

phytochrome signaling. Mutants that over-accumulate phytochromes often show dark-green leaves and accumulate increased levels of chlorophyll (Christians et al., 2012). In contrast, disruption of phytochromes causes decreased chlorophyll levels and leaves of the *phy* mutants are pale (Jackson and Prat, 1996; McCormac et al., 2001; Christians et al., 2012; Zhao et al., 2013b). Given that phytochromes positively regulate the expression of chlorophyll biosynthetic genes and chloroplast development (McCormac et al., 2001; Zhao et al., 2013b), the increased expression of phytochromes in *ref4-3* could promote chlorophyll biosynthesis synergistically.

Our targeted metabolite analysis revealed that MED5 is involved in down-regulation of ABA biosynthesis under both normal or drought conditions. In spite of the hyper-accumulation of ABA in *med5* mutants, *med5* lost water at a speed comparable to wild type, and many ABA-activated or repressed genes (Nemhauser et al., 2006) were normally expressed in *med5* mutants. Hence, our data are consistent with a model that MED5 is required for ABA-induced drought stress responses and transcriptional reprogramming, which cannot be invoked by increased accumulation of ABA when MED5 is disrupted. Except for down-regulation of an ABA hydroxylase-encoding gene *CYP707A3* (Umezawa et al., 2006), most ABA biosynthetic and catabolic genes showed wild-type expression in *med5*, therefore, the over-accumulated ABA in *med5* is likely to be induced by a feedback regulation of ABA biosynthesis and presumably achieved by reduced ABA catabolism.

Many Mediator subunits are involved in the regulation of abiotic stress responses including those towards drought and cold (Samanta and Thakur, 2015), both of which depend on normal ABA accumulation (Vishwakarma et al., 2017). In Arabidopsis, loss of

MED25 leads to enhanced drought tolerance, (Elfving et al., 2011) whereas *cdk8* mutants are hyper-sensitive to drought (Zhu et al., 2014). In addition, MED2, MED14 and MED16 are required for cold resistance (Hemsley et al., 2014). The abnormal abiotic stress response in *med* mutants has been attributed to either mis-regulated stress-induced gene expression, or to changes in plant cell wall structure (Elfving et al., 2011; Msanne et al., 2011; Hemsley et al., 2014; Zhu et al., 2014). Nevertheless, whether ABA biosynthesis is perturbed in various *med* mutants and thereby contributes to the observed phenotypes remains to be investigated.

In the dwarfing mutant *ref4-3*, the level of ABA was significantly enhanced, consistent with the constant correlation between ABA over-accumulation and stunted growth (Kleinow et al., 2009; Chung et al., 2014). We also noticed that many genes that were mis-regulated in response to ABA treatment showed similar expression changes in *ref4-3*. In contrast, most of those genes showed wild-type expression in *med5*, despite that ABA was equivalently enhanced in both *ref4-3* and *med5*. Given that *med5* are of normal growth (Stout et al., 2008), our data suggest that the dwarfism of *ref4-3* may be induced by hyper-activated ABA signal transduction. Alternatively, hyper-accumulation of ABA in *ref4-3* may be associated with its collapsed xylems (Stout et al., 2008; Feng et al., 2012) and occurs independently of its dwarfing phenotype.

Given the opposite flowering and photomorphogenesis phenotypes in *ref4-3* and *med5*, we conclude that *ref4-3* is a dominant *med5* mutant in the context of shade avoidance response, consistent with our previous finding in phenylpropanoid pathway (Stout et al., 2008; Bonawitz et al., 2012). Unlike the phenylpropanoid pathway in which many biosynthetic genes and regulatory genes were mis-regulated in *ref4-3* and *med5* in opposite

directions (Dolan et al., 2017), we did not observe a similar pattern of gene expression for shade avoidance response, suggesting that different sets of genes are involved in the down-regulated shade avoidance responses in *ref4-3* and the up-regulated SAS in *med5*. Alternatively, *ref4-3* and *med5* may differentially regulate SAS at a post-translational level (Leivar and Quail, 2011), resembling the function of MED5 in regulating phenylpropanoid metabolism by targeting F-box proteins (Dolan et al., 2017).

Interestingly, the increased ABA accumulation in both *ref4-3* and *med5* suggests that *ref4-3* acts as a loss-of-function *med5* mutant and leads to up-regulated ABA biosynthesis. Based on the function of Mediator in bridging the interaction between different transcription factors and Pol II, a direct model for the corresponding phenotype is that introduction of G383S mutation disrupts the interaction between MED5 and a (co-)repressor of ABA biosynthesis, which thereby phenocopies *med5* and results in up-regulation of ABA in *ref4-3*. Because an individual Mediator subunit can differentially regulate various pathways by interacting with different transcription factors (Chen et al., 2012; Lai et al., 2014), the molecular effect of G383S mutation may largely depend on the interacting partners of MED5 under different circumstances.

3.4   Methods

3.4.1   Plant material and growth conditions

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was the wild type in this study. Unless specified, plants were grown at a temperature of 23°C, under a long-day photoperiod (16 hr light/8 hr dark) or a short-day period (8 hr light/16 hr dark) with a light intensity of 100 $\mu E\,m^{-2}\,s^{-1}$. For hypocotyl measurement, Arabidopsis seedlings were

planted on ammonia-free MS medium (Murashige and Skoog, 1962) with 1% (w/v) sucrose and cultivated under continuous light for ten days. For end-of-day far-red (EOD-FR) light treatment, plants were exposed to far-red light for 30 min at the end of every photoperiod. Homozygous mutants used in this study were isolated based on previous reports, with the corresponding accession numbers and primers listed in Table 3.2 (Neff et al., 1998; Bonawitz et al., 2012).

3.4.2 Measurement of plant growth

Plant heights, rosette diameters, hypocotyl and petiole lengths under different growth conditions were quantified using ImageJ software (Abramoff et al., 2004).

3.4.3 Total lignin content quantification

Extractive free cell walls were prepared as described previously, followed by thioglycolic acid (TGA) lignin analysis (Li et al., 2015).

3.4.4 HPLC analysis of sinapoylmalate

Three-week-old whole rosettes were harvested, and sinapoylmalate levels were quantified by HPLC as previously reported (Dolan et al., 2017).

3.4.5 Reanalysis of RNA-seq data

Gene expression profiles of wild type, *ref4-3* and *med5* were obtained from previous studies (Dolan et al., 2017). The complete dataset can be accessed through NCBI's Gene Expression Omnibus (Edgar et al., 2002) with accession number GSE95574 (Dolan

et al., 2017). Differential expression analysis was performed as previously described between *med5* mutants and wild type (Dolan et al., 2017). Based on $\log_2$ fold change ($\log_2$FC) and false discovery rate (FDR) of each gene, volcano plots were created using built-in R functions (v3.4.1) and custom scripts. The scripts can be shared upon request.

3.4.6 Measurement of chlorophyll and photosynthetic capacity

For chlorophyll measurement, eight-week-old short-day cultivated plants were harvested and grinded in liquid nitrogen, followed by multiple rounds of extraction using 80% acetone until the pellet turns to white. Extracted chlorophylls were quantified using U-3900 spectrophotometer (HITACHI). For photosynthesis-related parameters, eight-week-old short-day cultivated plants were first treated with dark for 30 min, and a fully expanded leaf of each plant was subjected to different light intensities. The corresponding photosynthetic parameters were captured using the Fluorescence Monitoring System (Hansatech) run by Modfluor 32 software.

3.4.7 Determination of ABA levels

For this experiment, 20-day-old long-day cultivated plants were subjected to either water-deficit conditions by withholding water for 2 weeks (drought) or well-watered over the same period (control plants). Fully expanded leaves were harvested and frozen in liquid nitrogen, followed by ABA extraction and detection as previously described (McAdam, 2015). ABA was quantified by LC-MS using deuterated ABA [^{2}H$_6$] as an internal standard, synthesized by National Research Council of Canada.

3.4.8   Fresh weight loss of detached leaves assay

Four-week-old Arabidopsis plants cultivated under long-day conditions were used in this experiment, following previously described procedures (Zhang et al., 2007).

3.4.9   Statistical analysis

For box plots with overlaid points representing individual samples, data were plotted with gglot2 using statistical program R (v3.4.1). The lower and upper hinges correspond to the 25% quantile and the 75% quartile, and the lower and upper whiskers extend from each hinge to the smallest or largest values no more than 1.5× inter-quartile range. For bar graphs, data represent mean ± standard deviation (SD). Unless specified, the means were compared by one-way ANOVA, and statistically significant differences ($p <$ 0.05) were identified by Tukey's honest significant difference (HSD) test and indicated by different letters.

Table 3.1 Summary of photosynthesis-related parameters of wild-type, *ref4-3* and *med5* plants at different light intensities

Parameter [a]	genotype	light intensity (µmol / (m² • s))				
		25	50	120	240	350
Fv/Fm	wild type	0.84	0.84	0.85	0.85	0.85
	ref4-3	0.86	0.85	0.86	0.86	0.85
	med5	0.84	0.85	0.84	0.85	0.83
PSII	wild type	0.78	0.72	0.59	0.39	0.28
	ref4-3	0.79	0.75	0.53	0.37	0.30
	med5	0.78	0.73	0.57	0.41	0.28
NPQ	wild type	0.14	0.32	0.77	1.49	1.82
	ref4-3	0.13	0.28	1.32 [b]	2.05 [b]	2.32 [b]
	med5	0.14	0.25	1.00	1.62	1.46 [b]
ETR	wild type	8.21	15.19	29.83	38.96	41.84
	ref4-3	8.26	15.65	26.47	37.64	43.93
	med5	8.23	15.30	28.56	40.83	41.34
qP	wild type	0.95	0.90	0.81	0.60	0.45
	ref4-3	0.94	0.92	0.76	0.62	0.51
	med5	0.95	0.90	0.80	0.63	0.47
qL	wild type	0.76	0.66	0.54	0.35	0.23
	ref4-3	0.73	0.69	0.49	0.39	0.30
	med5	0.77	0.62	0.54	0.39	0.26
PQ-redox state	wild type	0.24	0.34	0.46	0.65	0.77
	ref4-3	0.27	0.31	0.51	0.61	0.70
	med5	0.23	0.38	0.46	0.61	0.74

[a] Parameters include photochemical efficiency of PSII (Fv/Fm), quantum yield of PSII electron transport (PSII), non-photochemical quenching (NPQ) coefficient, photosynthetic electron transport rate (ETR), photochemical quenching (qP), the proportion of open PSII centers (qL), and plastoquinone (PQ)-redox state.
[b] Statistically significant difference detected by student t-test between the measurement in wild type and that in *ref4-3* ($p < 0.05$) (n=4).

Table 3.2 Primers used in this study

Gene	Accession number	Mutant line	Name		Sequence
MED5a	AT3G23590	SALK_011621	*rfr1-3*	LP	TTTTATGGGCCTCTTTCGTG
			rfr1-3	RP	TTGGCATTAGTGAGCAAGCA
MED5b	AT2G48110	SALK_037472	*ref4-6*	LP	ACGGGGCATTGATAGAAAAA
			ref4-6	RP	AGGAGGGAATCGACAATGTG
MED5b	AT2G48110		*ref4-3*	F	CTTTGGTTGCCCATTGATCT
			ref4-3	R	GATTGGTTCCCCCAATTACA
PhyB	AT2G18790		*phyB-9*	F	GTGGAAGAAGCTCGACCAGGCTTG
			phyB-9	R	GCAAAACTCTTGCGTCTGTG

For *med5a* and *med5b*, left primers (LP) and right primers (RP) were used. For *ref4-3* and *phyB-9*, forward (F) and reverse (R) primers were used to amplify the regions carrying the point mutations followed by restriction enzyme digestion.
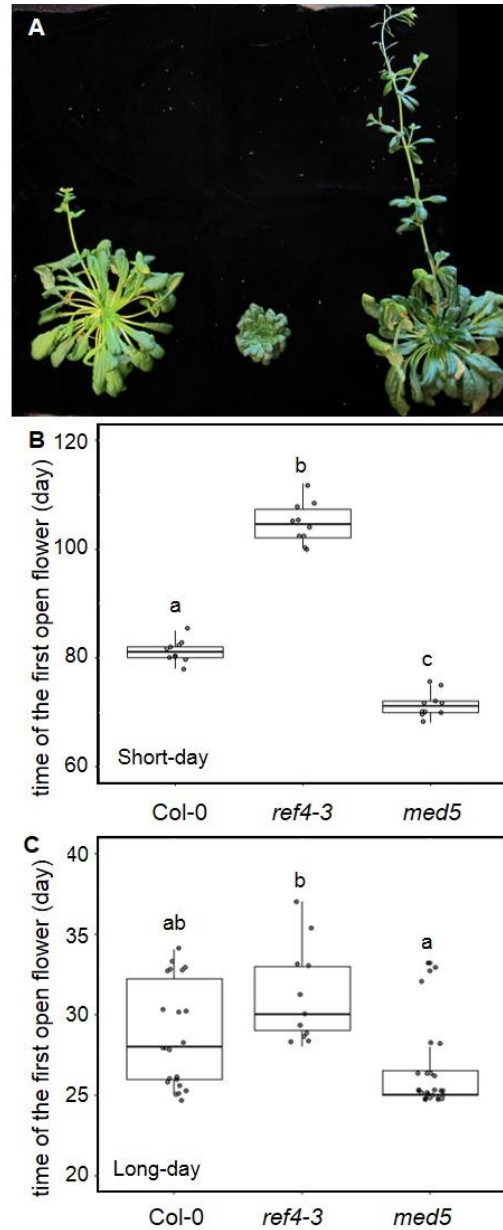
Figure 3.1 MED5a and MED5b are required for proper flowering time

(A) Representative photograph of three-month-old wild type, *ref4-3* and *med5* under short-day condition.
(B) Quantification of days to flower of wild type, *ref4-3* and *med5* under short-day condition.
(C) Quantification of days to flower of wild type, *ref4-3* and *med5* under long-day condition.
For panels B-C, data represent mean ± standard deviation (SD) (n ≥ 10). The means were compared by one-way ANOVA, and statistically significant differences (p < 0.05) were identified by Tukey's test and are indicated by a to c.
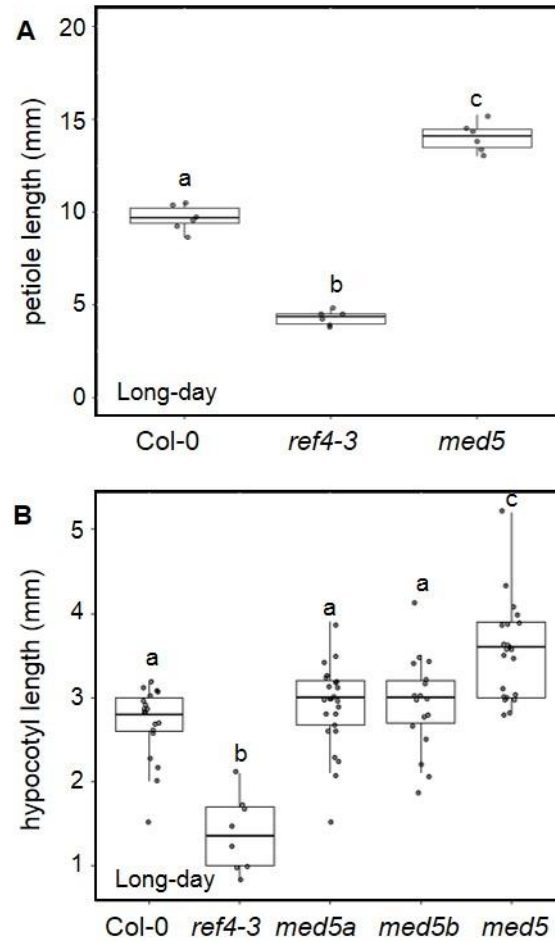
Figure 3.2 *ref4-3* and *med5* display opposite shade-avoidance response-related phenotypes

(A) Petiole length measurement of wild type, *ref4-3* and *med5* under long-day condition. Plants were grown on soil for three weeks, and six biological replicates were quantified.
(B) Hypocotyl length measurement of wild type, *ref4-3*, *med5a*, *med5b* and *med5* under long-day condition. Plants were grown on MS medium for 10 days first, and at least eight biological replicates were used for quantification.
For panels A-B, data represent mean ± standard deviation (SD) ($n \geq 6$). The means were compared by one-way ANOVA, and statistically significant differences ($p < 0.05$) were identified by Tukey's test and are indicated by a to c.
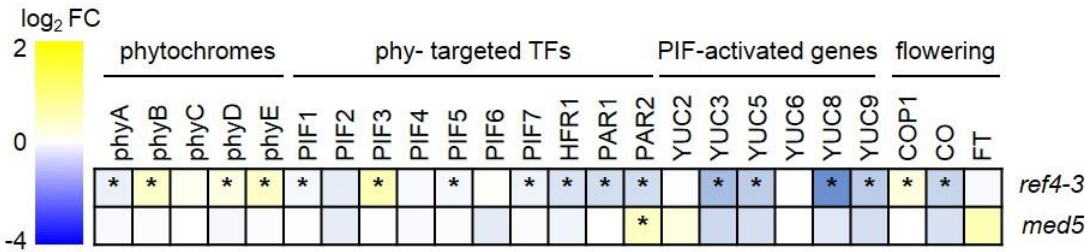
Figure 3.3 *ref4-3* has altered expression of phytochrome signaling-related genes

Log$_2$ fold change in expression of genes related to phytochrome signaling in *ref4-3* and *med5* compared to wild type (Dolan et al., 2017). Representative targets are phytochromes (phy), phytochrome-repressed transcription factors (TFs) including phytochrome interacting factors (PIFs), PIF-activated genes and markers for flowering time. '*' indicates genes with an FDR < 0.05.
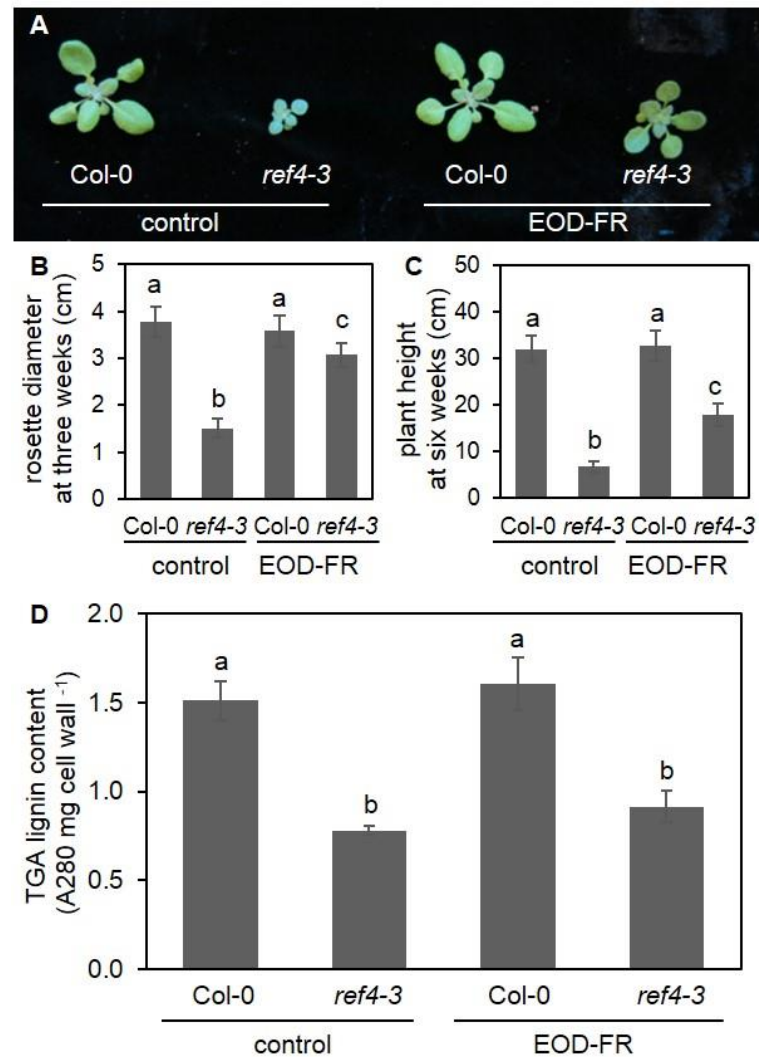
Figure 3.4 Dwarfism of *ref4-3* can be rescued by the EOD-FR light treatment

(A) Representative photograph of three-week-old wild-type and *ref4-3* plants under control light condition, or with 30 min treatment of EOD-FR light.

(B-C) Rosette diameter (B) and height (C) measurement of wild type and *ref4-3* after growth on soil for three weeks and six weeks respectively. Plants were cultivated under control light condition or with 30 min treatment of EOD-FR light. Data represent mean ± SD (n=10).

(D) Total lignin content in seven-week-old stem tissues quantified by thioglycolic acid (TGA) lignin analysis. Data represent mean ± SD (n = 3).

For panel B-D, the means were compared by one-way ANOVA, and statistically significant differences (p < 0.05) were identified by Tukey's test and are indicated by a to c.
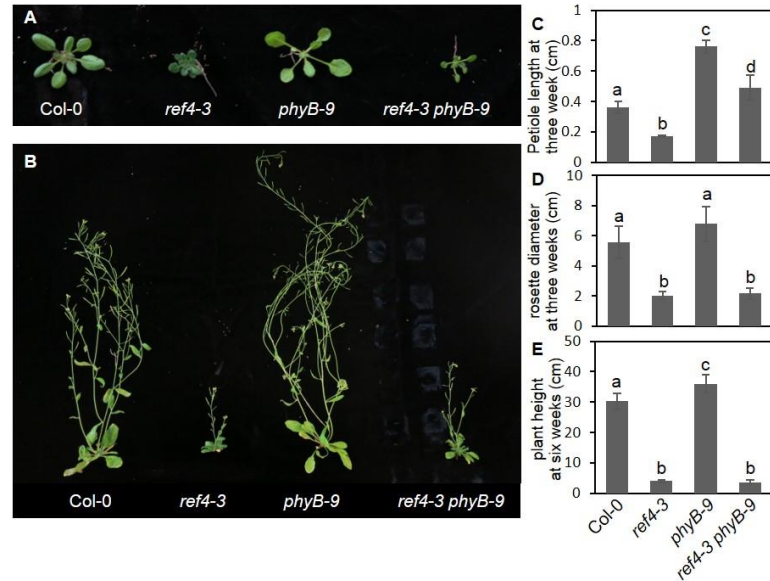
Figure 3.5 Disruption of PhyB is not sufficient to rescue the stunted growth of *ref4-3*

(A-B) Representative photograph of three-week-old (A) and six-week-old (B) wild type, *ref4-3, phyB-9* and *ref4-3 phyB-9*.
(C) Petiole length measurement of wild type, *ref4-3*, *phyB-9* and *ref4-3 phyB-9* after growth on soil for three weeks.
(D-E) Rosette diameter (D) and height (E) measurement of wild type, *ref4-3*, *phyB-9* and *ref4-3 phyB-9* after growth on soil for three weeks and six weeks respectively
For panel C-E, data represent mean ± SD (n=10). The means were compared by one-way ANOVA, and statistically significant differences (p < 0.05) were identified by Tukey's test and are indicated by a to d.
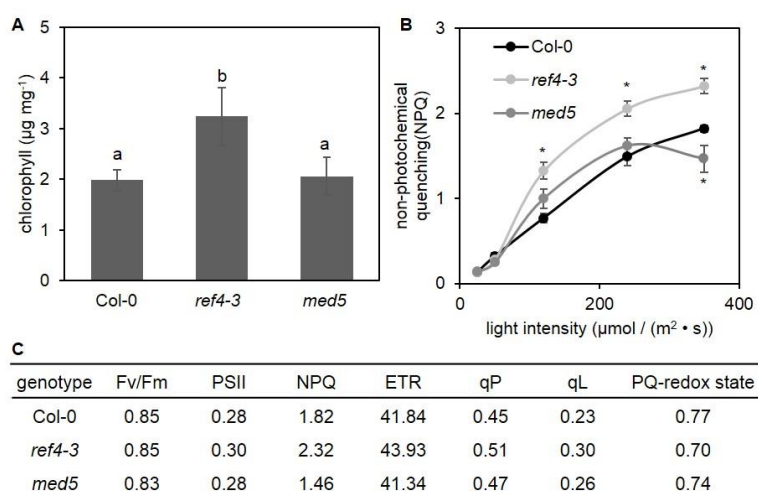
Figure 3.6 The performance of photosynthetic machinery is disrupted in *ref4-3*

(A)Total chlorophyll content in eight-week-old leaves of wild type, *ref4-3* and *med5* grown in short-day condition.

(B) Measurement of non-photochemical quenching (NPQ), the signature parameter for photosynthesis in wild type, *ref4-3* and *med5*.

(C) Summary table of photosynthesis-related parameters of wild-type, *ref4-3* and *med5* plants at 350 μmol / (m^2 • s)).
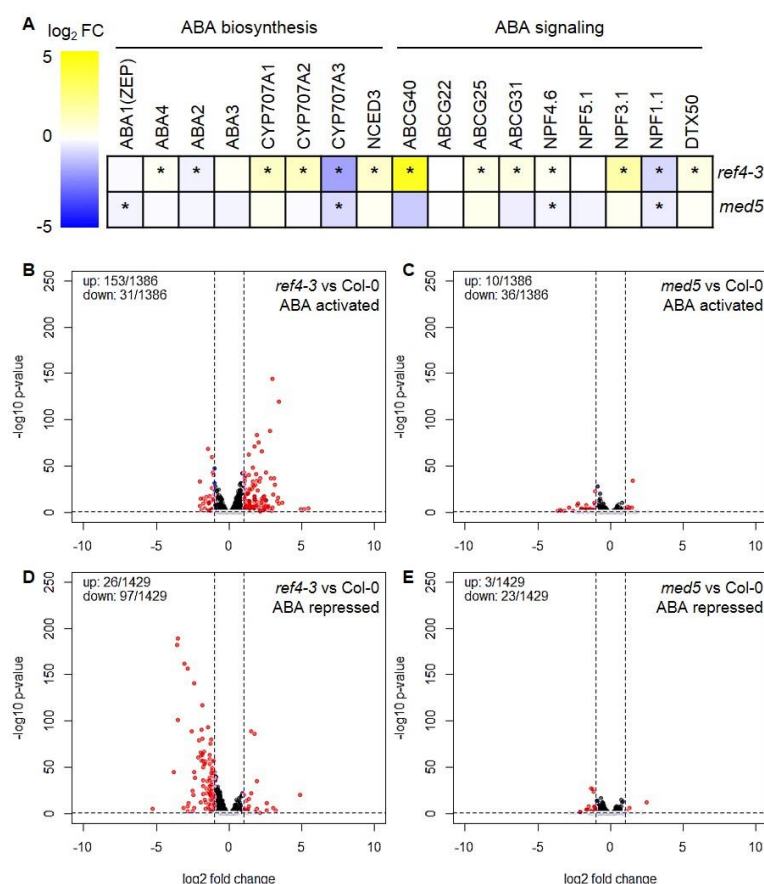
Figure 3.7 *ref4-3* has altered expression of ABA-responsive genes

(A) $\log_2$ fold change in expression of ABA biosynthetic and signaling genes in *ref4-3* and *med5* compared to wild type. '*' indicates genes with an FDR < 0.05.

(B-C) Volcano plots representing the expression level of ABA activated genes in *ref4-3* (B) and *med5* (C) compared to wild type. For each gene represented by a dot, the x axis shows the fold change in *ref4-3* or *med5* compared to wild type ($\log_2$FC), whereas the y axis shows the $\log_{10}$(false discovery rate) for each gene ($-\log_{10}$[FDR]). The genes that are significantly differentially expressed compared to wild type (FDR < 0.05) are highlighted in red, while the others are colored in grey.

(D-E) Similar to panel B-C, volcano plots representing the expression level of ABA repressed genes in *ref4-3* (D) and *med5* (E) compared to wild type.

For panel A-E, $\log_2$FC and FDR of each gene was determined by previous RNA-seq analysis (Dolan et al., 2017).
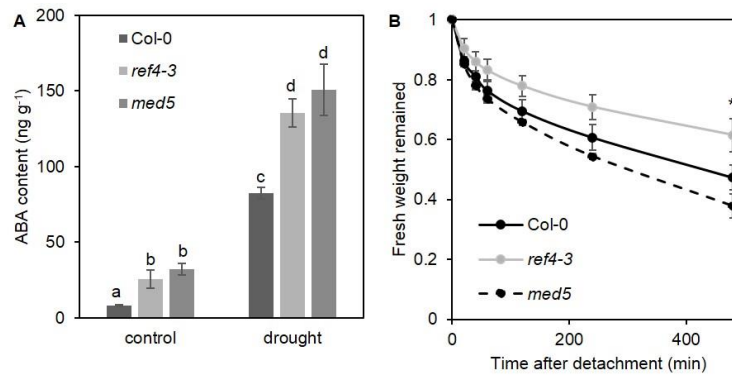
Figure 3.8 *ref4-3* is resistant to drought stress

(A) Quantification of ABA extracted from three-week-old regular-grown or drought-treated wild type, *ref4-3* and *med5*. Data represent mean ± SD (n = 3). The means were compared by one-way ANOVA, and statistically significant differences (p < 0.05) were identified by Tukey's test and are indicated by a to d.

(B) Whole rosettes from three-week-old wild type, *ref4-3* and *med5* were harvested, and recorded for changes in weight over eight-hour period. The ration between the weight at each time point and the time point 0 were reported as fresh weight remained. Data represent mean ± SD (n = 3). The means were compared by one-way ANOVA, and statistically significant differences (p < 0.05) from wild type were identified by Tukey's test and are indicated by '*'.

CHAPTER 4. COMPARATIVE ANALYSIS OF RNA POLYMERASE II OCCUPANCY CHANGES IN ARABIDOPSIS MED5 AND CDK8 MUTANTS

4.1 Introduction

The multi-subunit complex Mediator is a conserved transcription co-regulator in eukaryotes. According to structural studies, Mediator can be divided into four modules, namely the head, middle, tail and a dissociable CDK8 kinase module (Dotson et al., 2000). The tail module serves as a platform for the interaction between Mediator and various transcription factors, and once integrated, these signals can then be transmitted to general transcription machinery through contacts between the head and middle modules and RNA Polymerase II (Pol II), leading to activation or repression of downstream target expression (Meyer et al., 2010). Hence, Mediator is critical for the transcription of Pol II transcribed genes.

To elucidate the molecular function of Mediator in transcription regulation, it is critical to characterize the occupancy of Pol II and the Mediator complex at a genome-wide scale. Studies in yeast and mammals demonstrate that Mediator is required for nearly every step of transcription, including Pol II recruitment (Malik and Roeder, 2005), pre-initiation complex (PIC) formation (Reeves and Hahn, 2003), transcriptional elongation (Conaway and Conaway, 2013) and transcriptional termination (Mukundan and Ansari, 2011). The function of Mediator in Pol II recruitment has also been explored in plants at a small number of individual genes. For instance, the Arabidopsis Mediator tail module subunit MED16 is required for recruitment of Pol II to the transcription start site (TSS) of C-repeat binding factor (CBF)-targeted genes (Hemsley et al., 2014). Loss of MED16 resulted in

compromised induction of those genes under cold stress and a lack of freezing tolerance (Hemsley et al., 2014). Similarly, loss of CDK8, a kinase module subunit, led to decreased Pol II occupancy at *plant defensin 1.2* (*PDF1.2*), which accounted for reduced *PDF1.2* expression and defected biotic responses in *cdk8* mutants (Zhu et al., 2014). The above studies demonstrate that the integrity of plant Mediator complex is critical for Pol II recruitment, suggesting that the biological processes maintained by plant Mediator subunits could be valuable readouts to characterize the molecular function of Mediator.

Previous studies showed that MED5a and MED5b, a pair of orthologous Arabidopsis Mediator tail module subunits, limit phenylpropanoid biosynthesis (Stout et al., 2008; Bonawitz et al., 2012). A semi-dominant *med5* mutant named *reduced epidermal fluorescence 4-3* (*ref4-3*) carrying a single amino acid substitution in *MED5b* (G383S) was isolated as a strong suppressor of phenylpropanoid pathway, indicated by decreased soluble phenylpropanoids, reduced lignin content, and dwarfism (Stout et al., 2008), a phenotype often associated with lignin-deficient mutants (Muro-Villanueva et al., 2019). Negative regulators of phenylpropanoid biosynthesis including *KFB39* and *KFB50*, two genes encoding F-box proteins that mediate the ubiquitination and degradation of phenylalanine ammonia lyases (PALs) (Zhang et al., 2013a, 2015), were significantly mis-regulated in opposite directions in *ref4-3* and *med5* compared to wild type; however, phenylpropanoid biosynthetic genes including *PAL*s only displayed modest changes in *med5* mutants (Dolan et al., 2017). The gene expression profiles of *ref4-3* and *med5* suggest that the protein encoded by the *ref4-3* allele of MED5 is likely to enhance Pol II recruitment to negative regulators of phenylpropanoid biosynthesis, which thereby promotes their expression and leads to repressed phenylpropanoid metabolism. In addition, disruption of CDK8

suppresses the stunted growth but not the restricted phenylpropanoid metabolism in *ref4-3* (Mao et al., 2019). Investigation of the genes that were mis-regulated in *ref4-3* in a CDK8-dependent manner enabled us to identify DJC66, a DNAJ co-chaperone, as a novel suppressor of the growth defect in *ref4-3* (Mao et al., 2019). Despite this, the limited restoration of growth in *ref4-3 djc66* suggests that the stunted growth of *ref4-3* may also result from the mis-regulation of other genes. Given that the CDK8 kinase module can modulate the interaction between the core Mediator complex and Pol II and thus interfere with normal transcription (Taatjes, 2010), we proposed that CDK8 is required for abnormal Pol II recruitment to the growth-related genes in *ref4-3*.

In this study, we investigate the effect of mutations in Mediator including *med5*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* on genome-wide Pol II distribution to provide insight into the function of MED5 and CDK8 in gene expression regulation. We found that loss of MED5 resulted in loss of Pol II occupancy at many target genes. Conversely, many genes showed enriched Pol II levels in *ref4-3*, some of which overlapped with those showing reduced Pol II occupancy in *med5*. In addition, Pol II occupancy was significantly reduced when CDK8 is disrupted in *ref4-3*. Our results help to narrow down the direct gene targets of MED5 and identify genes that may be closely related to the growth deficiency observed in *ref4-3* plants.

4.2   Results

4.2.1   Pol II occupancy is highly enriched within gene bodies

To determine the changes in genome-wide Pol II occupancy caused by mutations in Mediator complex subunit genes, we performed chromatin immunoprecipitation (ChIP)-

sequencing (seq) analysis in the *med* mutants *med5*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1*, together with wild type. Chromatin was extracted from 10-day-old seedlings and precipitated with the well-characterized monoclonal antibody 4H8 that recognizes Pol II with high affinity regardless of its phosphorylation status (Ou, 2014). The C-terminal domain of the largest subunit of Pol II shows altered phosphorylation status during different stages of transcription such as the transition from initiation to elongation (Phatnani and Greenleaf, 2006); thus, our ChIP-seq data reflect total Pol II occupancy irrespective of transcription stage.

Given that Pol II recruitment is a necessary step for transcriptional initiation and reflects the transcriptional status of a gene (Gan et al., 2011; Mokry et al., 2012), we tested whether a positive correlation between Pol II occupancy and gene expression level can be captured by our Pol II ChIP-seq analysis. To do this, we classified the 33,557 annotated genes in the Arabidopsis genome into four groups according to their expression level in three-week-old wild-type plants (Mao et al., 2019): high, medium, low and silent (Figure 4.1A).We then plotted the average distribution of Pol II +/- 5 kb from the TSS of the genes in each group and found that Pol II enrichment positively correlated with expression (Figure 4.1B). In addition, Pol II occupancy was enriched downstream of the TSS, suggesting that Pol II is actively transcribing these genes.

To further characterize the distribution of Pol II, we classified genes based on length into groups of short (<1500 base pairs (bp)), medium (1500 bp to 3000 bp) or long genes (> 3000 bp) and then examined the average distribution of Pol II in these groups of genes. Regardless of gene length, Pol II was present throughout the length of the genes, buttressing our finding that Pol II occupancy was significantly enriched across gene bodies

(Figure 4.1C). Taken together, we conclude that Pol II occupancy positively correlates with gene expression level, and is highly elevated in the gene body downstream of the TSS.

4.2.2   Loss of MED5 leads to decreased Pol II occupancy at a subset of genes

To identify the genes that require MED5 for normal transcription, we compared Pol II occupancy between wild type and *med5* to identify regions with differential Pol II binding. We identified 7398 peaks of Pol II binding and found that Pol II occupancy increased in *med5* relative to wild type at only 4 regions, while 514 peaks showed reduced Pol II occupancy in *med5*. The 837 genes associated with those 514 peaks show reduced Pol II occupancy downstream of the TSS (Figure 4.2A). Consistent with the characterized function of MED5 in pathogen defense and phenylpropanoid biosynthesis (Wang et al., 2016; Stout et al., 2008; Bonawitz et al., 2012), genes exhibiting reduced Pol II occupancy in *med5* were enriched for gene ontology (GO) terms associated with biotic or abiotic stress responses, tryptophan biosynthesis and phenylpropanoid metabolism (Table 4.1).

To elucidate the relationship between the transcriptional defects due to loss of MED5 and the mis-regulated gene expression in *med5* mutants, we compared the genes with differential Pol II occupancy in *med5* with those genes showing differential gene expression based on RNA-seq analysis (Dolan et al., 2017). We found that most of the genes with decreased Pol II occupancy in *med5* also exhibited down-regulated gene expression in *med5* (Figure 4.2B). Notably, only a subset of the genes with decreased expression in *med5* also showed lower Pol II occupancy (Figure 4.2B), suggesting that these genes may represent the direct transcriptional targets of MED5. Alternatively, given that the RNA-seq data was from 3-week-old whole rosettes whereas the ChIP-seq was

performed using 10-day-old seedlings, the difference in growth stages may provide a possible explanation for the discrepancy between the two types of gene expression data.

Because the Mediator complex can be recruited to the promoters of target genes through the interaction between its tail module subunits and gene-specific transcription factors (Borggrefe and Yue, 2011), the down-regulated transcription in *med5* mutants may result from a disrupted interaction between MED5 and the corresponding transcription factors. Thus, we sought to identify the transcription factors that might be required for MED5 within Mediator to activate transcription at its target genes. To do this, we analyzed the enrichment of transcription factor binding sites at the promoter regions (-1000 bp to 200 bp relative to the TSS) of genes with lower Pol II occupancy in *med5* compared to wild type using MEME (Bailey et al., 2009). We discovered one 12-bp motif that was significantly enriched among the promoters of *med5* down-regulated genes (Figure 4.3C). This motif was most abundant in the region 300 bp upstream to 100 bp downstream from the TSS of MED5 target genes (Figure 4.3D). Interestingly, the enriched motif matched with a group of DNA-binding with one finger (DOF) transcription factors in the Arabidopsis database (O'Malley et al., 2016) (Figure 4.3C). DOF transcription factors act as negative regulators of plant flowering time (Fornara et al., 2009). Given that *dof* and *med5* mutants are both early flowering (Fornara et al., 2009; Dolan and Chapple, 2018), this finding suggests that MED5 could activate DOF-targeted genes to impact the timing of flowering in wild-type plants.

4.2.3   The *ref4-3* mutation results in increased Pol II occupancy at a subset of genes

We have previously proposed that MED5 functions as a coactivator of transcription and that the *ref4-3* mutation enhances the normal function of MED5 for a subset of its gene targets (Dolan et al., 2017). In contrast to the overall down-regulated Pol II occupancy in *med5* compared to wild type, Pol II occupancy was significantly increased at 216 of the 7151 Pol II-bound regions in *ref4-3*, whereas it was decreased at only 2 regions. The peaks with increased Pol II signals in *ref4-3* were associated with 315 genes and as before, increased Pol II occupancy was evident throughout the gene bodies (Figure 4.3A). Comparison between Pol II ChIP-seq analysis and RNA-seq analysis (Dolan et al., 2017) revealed that most of the genes with increased Pol II occupancy in *ref4-3* also showed elevated gene expression (Figure 4.3B). This suggests that up-regulated expression of these genes in *ref4-3* may stem from elevated Pol II recruitment. The genes with higher Pol II binding in *ref4-3* were enriched for GO terms associated with protein phosphorylation, regulation of phenylpropanoid metabolism and responses to various hormones namely abscisic acid (ABA), jasmonic acid (JA), salicylic acid (SA) and ethylene (ET) (Table 4.2). Most of the enriched GO categories are consistent with previously characterized phenotypes of *ref4-3*, including inhibited phenylpropanoid biosynthesis and increased SA accumulation (Stout et al., 2008; Mao et al., 2019).

To identify potential transcription factors that could be involved in recruiting Mediator in *ref4-3,* we looked for enriched sequence motifs in the promoter regions (-1000 bp upstream to 200 bp downstream relative to the TSSs) of the genes with up-regulated Pol II occupancy in *ref4-3*. Using this approach, we identified a GAGA-like motif that was significantly enriched in these promoters relative to other expressed genes (Figure 4.3C).

In Arabidopsis, the GAGA-motif binding protein basic pentacysteine 6 (BPC6) recruits polycomb-repressive complexes, conserved negative regulators that deposit a trimethylation mark at Lys-27 in histone H3 (H3K27me3) and repress expression of homeotic genes (Hecker et al., 2015). For the genes that were mis-regulated in *ref4-3*, the GAGA motifs were most abundant at the -100 to 200 bp around the TSS (Figure 4.3D). This observation is consistent with studies showing that GAGA motifs were overrepresented in Arabidopsis core promoters (Yamamoto et al., 2009; Hecker et al., 2015). Although the interaction between Mediator and polycomb-repressive complexes has not been explored in plants, studies in human cells showed that these two complexes interact (Fukasawa et al., 2012, 2015). Thus, one potential explanation for the up-regulated transcription of most *ref4-3* targets is that the semi-dominant MED5b mutant interferes with the recruitment of polycomb-repressive complexes, thereby leading to de-repression of these genes. One prediction of this model is that the genes that are upregulated in *ref4-3* mutants should be marked by H3K27me3 in wild type, and should show loss of this repressive histone modification in the *ref4-3* mutant.

4.2.4 *med5* and *ref4-3* perturb transcription of genes involved in other cellular processes in addition to phenylpropanoid biosynthesis

Previous RNA-seq analysis revealed that among the genes that were at least 2-fold mis-regulated in *med5* and *ref4-3* compared to wild type, 67% were down-regulated in *med5* and upregulated in *ref4-3* (Dolan et al., 2017), suggesting that *ref4-3* may lead to constitutive activation of genes that require MED5 for normal expression. Surprisingly, in our Pol II ChIP-seq analysis, we only found 33 genes that show increased Pol II occupancy in *ref4-3* and decreased Pol II occupancy in *med5* (Figure 4.4A, Table 4.3, Figure 4.5). The

large proportion of genes with up-regulated Pol II occupancy in *ref4-3* and unchanged Pol II occupancy in *med5* suggests that *ref4-3* may activate the expression of genes that do not usually require MED5-containing Mediator for normal transcription. In addition, given that more than 95% of the genes that showed mis-regulated Pol II occupancy in *med5* were not differentially bound by Pol II in *ref4-3*, the protein encoded by the MED5b G383S allele likely functions as a wild-type MED5 at most MED5 target genes.

Considering that MED5 is mainly required for gene activation in Arabidopsis (Dolan and Chapple, 2018), we reasoned that the genes with down-regulated Pol II occupancy in *med5* and up-regulated Pol II occupancy in *ref4-3* were mostly likely to be the direct targets of MED5. To test if any common targets of *ref4-3* and *med5* are co-expressed and thus may share similar functions, we retrieved the available co-expression data from ATTED-II database (Obayashi et al., 2018) and constructed a mutual rank (MR)-based network (Figure 4.4B, Table 4.3). Although most of the genes in the dataset were not highly co-expressed with each other as indicated by their large MR indexes, *KFB39* and *KFB50*, two negative regulators of the phenylpropanoid pathway, shared a low MR index (Figure 4B, Figure 4.5), suggesting that MED5 is involved in the transcriptional regulation of those two *KFB*s in wild-type plants. Surprisingly, *DJC66*, a gene encoding the aforementioned DNAJ cochaperone protein, is co-expressed with *KFB39* (Figure 4.4B). Given that DJC66 is not required for wild-type phenylpropanoid biosynthesis (Mao et al., 2019), the MR-based network may reveal the coordinated roles of DJC66, KFB39 and KFB50 in other cellular processes.

Another group of highly co-expressed genes were identified from the MR analysis that include *ATP-binding cassette subfamily B 21* (*ABCB21), octadecanoid-responsive*

Arabidopsis AP2/ERF 59 (*ORA59*) and *plastid protein kinase with unknown function 1* (*PKU1*) (Figure 4.4B, Table 4.3). *ABCB21* encodes an auxin importer/exporter that is regulated by auxin (Kamimoto et al., 2012). Interestingly, *ORA59* is required for auxin transport (Ståldal et al., 2012) in addition to its well-characterized function in integrating the JA and ET signaling pathways (Pre et al., 2008). The high co-expression between *ABCB21* and *ORA59* suggests that the effects of *ref4-3* and *med5* on transcription may be extended to the genes involved in auxin transport and signaling.

4.2.5 Loss of CDK8 in *ref4-3* causes an overall decrease in Pol II occupancy

Previously we identified *cdk8* as a novel suppressor of *ref4-3* that can rescue its stunted growth but not the restricted phenylpropanoid metabolism (Mao et al., 2019). To narrow down the genes that could be associated with the dwarfism of *ref4-3*, we sought to identify the genes with altered Pol II occupancy in *ref4-3* compared to wild type, but with restored Pol II occupancy in *ref4-3 cdk8-1*. Loss of CDK8 resulted in 2124 versus 10 regions with decreased versus increased Pol II occupancy in *ref4-3 cdk8-1* compared to *ref4-3*, suggesting that CDK8 is generally required for transcription activation in the *ref4-3* background. An average profile for the 4287 genes associated with those 2124 peaks revealed a substantial loss of Pol II binding in the double mutant compared to *ref4-3* (Figure 4.6A). Moreover, the vast majority of these genes showed decreased expression levels in *ref4-3 cdk8-1* compared to *ref4-3* (Figure 4.6B), suggesting that the decreased transcript levels observed in *ref4-3 cdk8-1* likely reflect reduced Pol II transcription at these genes. We also identified 315 genes that showed increased Pol II occupancy in *ref4-3* compared to wild type, and 198 of these genes had reduced Pol II occupancy in *ref4-3 cdk8-1*

compared to *ref4-3* (Figure 4.7A). An average plot of Pol II ChIP-seq signals for those genes showed elevated levels of Pol II occupancy at the TSS and throughout the gene bodies in *ref4-3*, which was restored in *ref4-3 cdk8-1* (Figure 4.7B). Additionally, loss of CDK8 in wild-type plants did not cause significant change in Pol II occupancy for these genes with increased Pol II in *ref4-3* that was rescued in *ref4-3 cdk8-1* (Figure 4.7B). Thus, we conclude that CDK8 is not necessary for Pol II recruitment at these genes.

We next compared our ChIP-seq analysis with RNA-seq expression data (Mao et al., 2019) and found that among the 198 genes with CDK8-dependent Pol II enrichment in *ref4-3*, 50 genes showed up-regulated expression in *ref4-3* and could be rescued by disruption of CDK8 (Figure 4.7C). Significantly enriched GO categories for these genes included 'response to ethylene' and 'response to salicylic acid' (Table 4.4), the latter of which is consistent with the over-accumulated SA content in *ref4-3* that is dependent on CDK8 (Mao et al., 2019). In Arabidopsis and common wheat, CDK8 interacts with the ethylene response transcription factor wax inducer 1 (Zhu et al., 2014; Kong and Chang, 2018) to activate expression of genes encoding wax metabolic enzymes (Kong and Chang, 2018). Although the function of MED5 in ethylene response has not been explored, a previous RNA-seq analysis using three-week-old rosettes showed that the genes down-regulated in both *ref4-3* and *med5* were enriched for genes involved in ethylene-activated signaling pathway (Dolan et al., 2017). Therefore, our Pol II ChIP-seq analysis performed in 10-day-old seedlings may capture an enhanced response to ethylene in *ref4-3* at an earlier stage of development, which is likely dependent on CDK8.

4.3   Discussion

The opposing and readily measurable phenylpropanoid phenotypes of *med5* and *ref4-3* plants suggest that these plants are valuable genetic tools to explore the function of MED5 in the context of the Mediator complex (Stout et al., 2008; Bonawitz et al., 2012; Dolan et al., 2017). To understand how Arabidopsis MED5 regulates Pol II-dependent transcription, we characterized the genome-wide change of Pol II occupancy in *med5* and *ref4-3* and identified genes with perturbed Pol II occupancy. Consistent with a model that MED5 functions as a transcription co-activator and *ref4-3* enhances the role of MED5 in the activation of negative regulators of phenylpropanoid metabolism and plant growth (Dolan et al., 2017), we found that loss of MED5 predominantly led to decreased Pol II occupancy compared to wild type, whereas in *ref4-3*, most of the differential binding regions exhibited increased Pol II occupancy. Most of the genes identified with increased or decreased Pol II occupancy also showed up- or down-regulated gene expression in *ref4-3* and *med5* mutants respectively, suggesting that the expression changes of those genes largely stem from perturbed Pol II transcription. Moreover, *KFB39* and *KFB50*, two previously characterized negative regulators of the phenylpropanoid metabolism (Zhang et al., 2013a, 2015) and *DJC66*, a suppressor of *ref4-3* (Mao et al., 2019) showed up-regulated Pol II occupancy in *ref4-3* and down-regulated Pol II occupancy in *med5*. Therefore, at least for the robust targets, our analysis can capture the genes that require MED5 for their normal transcription.

Although our RNA-seq and Pol II ChIP-seq studies have revealed the genes that require MED5 for proper expression or transcription, it is unclear whether these genes are direct targets of MED5-containing Mediator complex. Moreover, we do not know how

changes in MED5 in the *med5* or *ref4-3* mutants affect Mediator occupancy or distribution. Previous studies showed that loss of a single Mediator tail module subunit decreased recruitment of Mediator to its downstream genes (Ansari and Morse, 2012; Jeronimo et al., 2016; Grünberg et al., 2016), which in turn resulted in reduced Pol II recruitment, and down-regulated gene expression. Thus, our Pol II ChIP-seq analysis provides the foundation for the characterization of Mediator occupancy at a genome-wide scale. Given that MED5 is in the tail module of Mediator (Tsai et al., 2014) and that *med5* shows multiple phenotypes including reduced phenylpropanoid metabolism and early flowering (Stout et al., 2008; Bonawitz et al., 2012; Dolan and Chapple, 2018), it is likely that loss of MED5 disrupts the interaction between MED5 and the transcription factors involved in the regulation of phenylpropanoid metabolism and flowering time control, as reflected by our transcription factor binding motif analyses.

To further narrow down the list of genes that could be related to the dwarfism of *ref4-3*, we identified genes with increased Pol II occupancy in *ref4-3* compared to wild type and decreased Pol II occupancy in *ref4-3 cdk8-1* compared to *ref4-3*. Compared to the list of genes that were up-regulated in *ref4-3* in a CDK8-dependent manner (Mao et al., 2019), we found that genes involved in responses towards SA and ethylene were enriched among the overlapping upregulated targets identified from both RNA-seq and ChIP-seq analysis. Interestingly, previous RNA-seq analysis using 3-week-old rosettes showed that ethylene signaling-responsive genes were enriched for genes that were down-regulated in *ref4-3* (Dolan et al., 2017). Given that the Pol II ChIP-seq analysis was performed in 10-day-old seedlings, the difference between RNA-seq analysis and Pol II ChIP-seq analysis may result from the different growth stages of plants. The changed Pol II enrichment at

ethylene-responsive genes may reflect an abnormal activation of ethylene signaling in *ref4-3* at an early growth stage, which may contribute to the dwarfism of *ref4-3*. We also noticed that a signature motif for recruiting polycomb-repressive complexes was enriched in *ref4-3* targeted genes. Considering that polycomb-repressive complexes can place H3K27me3 marker on its target genes to repress transcription (Hecker et al., 2015) and that many genes involved in hormone biosynthesis, transport and signaling are targets of H3K27me3 (Yamamuro et al., 2016), *ref4-3* may negatively affect the recruitment of polycomb-repressive complexes, leading to hyper-activation of hormone responsive genes including those to SA and ethylene.

4.4   Methods

4.4.1   Plant materials and growth

Arabidopsis thaliana ecotype Columbia-0 (Col-0) was used as the wild type genotype in this study. Plants were cultivated on Murashige and Skoog agar plates and grown at 23°C under a long-day photoperiod (16 hr light/8 hr dark).

4.4.2   Pol II ChIP-seq library generation and sequencing

Samples of wild-type, *med5*, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* 10-day-old seedlings were harvested in triplicate with a randomized design. Each sample contained 4 grams of the same genotype from three individual pots. The harvested samples were cross-linked using 1.6 % formaldehyde by 10-min vacuum infiltration, and then immediately flash frozen in liquid nitrogen. Pol II ChIP was performed as previously described with minor modifications (Carter et al., 2018). After chromatin extraction and immunoprecipitation

using 2 μL Pol II antibodies (4H8, ab5408, Abcam) the precipitated chromatin was reverse cross-linking and purified using ChIP DNA clean & concentrator kit (D5205, Zymo Research). 5% of sheared chromatin without incubation with antibodies was used as an input control for each sample. cDNA libraries were constructed using NEXTflex ChIP-seq kit (Bioo Scientific Corporation). ChIP-seq libraries were analyzed using the bio-analyzer (Agilent) to examine the quality and quantity of each library, ensuring that the amplified genomic DNA were enriched at size 250 to 500 bp with molarity of at least 1 nM. Qualified genomic libraries were pooled and subjected to sequencing with Illumina HiSeq 2500 technology, which generated paired-end, 65 bp sequencing reads. Quality control was performed by the Purdue Genomics Core using FASTX-Toolkit (Gordon et al., 2010), followed by mapping to the Bowtie2-indexed Arabidopsis genome (TAIR10) using Tophat with default parameters (Trapnell et al., 2009) to generate Binary Alignment Map (BAM) files.

4.4.3 Statistical analysis of Pol II ChIP-seq data

For the mapped sequences, we first converted the BAM files to BED files and performed peak-calling using spatial clustering for identification of ChIP-Enriched regions (SICER) (v1.1) (Xu et al., 2014). The genome was partitioned into nonoverlapping 200 bp windows and regions with significant enrichment of Pol II against both random background and the input control were identified (FDR < 0.05). For the pair-wise differential binding test, the identified peaks from two different genotypes were subjected to ChIPComp (v1.6.0) analysis (Chen et al., 2015) using statistical program R (v3.4.1) (FDR < 0.05). Gene annotation of the identified regions with significantly different Pol II occupancy was

performed using closestBed utility of BEDtools with default settings, and the gene lists of interest were used for downstream analysis. The MEME suite (v5.0.5) was used for motif analysis (Bailey et al., 2009). First, we used MEME utility (Bailey and Elkan, 1994) to discovery significantly enriched motifs (8 bp to 12 bp) within 1000 bp upstream to 200 bp downstream of the TSS of each gene (E-value < 0.05). The sequences were scanned again with FIMO (Grant et al., 2011) to find the instances for the identified motifs, and the motif comparison tool Tomtom (Gupta et al., 2007) was used to check if the identified motifs matched with any known motifs in the Arabidopsis motif database (O'Malley et al., 2016) (E-value < 0.05).

The DAVID Bioinformatics Resource (v.6.8, https://david.ncifcrf.gov/) (Huang et al., 2008) was used to perform GO term analysis.

4.4.4 Data visualization

The BAM files of the biological triplicates were first merged using SAMtools (v1.8) (H et al., 2009), and the average plots for Pol II enrichment over the genomic regions were created using ngsplot (v2.61) (Shen et al., 2014). The normalized Pol II enrichment (reads per million mappable reads) from each sample was plotted over the genes of interest using Gviz (v3.8) (Hahne and Ivanek, 2016) with custom scripts, which can be shared upon request.

Table 4.1 GO term analysis of the genes that showed decreased Pol II occupancy in *med5* compared to wild type [a]

Term	Description	Count [b]	Fold Enrichment [c]	P-value
GO:2000762	regulation of phenylpropanoid metabolic process	4	15.78	0.001
GO:0006501	C-terminal protein lipidation	3	13.81	0.018
GO:0000162	tryptophan biosynthetic process	4	5.52	0.034
GO:0048544	recognition of pollen	5	4.06	0.033
GO:0010286	heat acclimation	5	4.06	0.033
GO:0009611	response to wounding	20	3.10	0.000
GO:0009753	response to jasmonic acid	14	2.91	0.001
GO:0009617	response to bacterium	8	2.83	0.022
GO:0006970	response to osmotic stress	10	2.61	0.014
GO:0009751	response to salicylic acid	11	2.45	0.014
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	13	2.11	0.021
GO:0009737	response to abscisic acid	22	1.73	0.017

a. For Table 4.1, 4.2 and 4.4, only significant GO terms are shown (P-value < 0.05).
b. 'Count' represents the number of genes that were used as queries and related to the specific GO term.
c. 'Fold Enrichment' was normalized to the percentage of background genes (all genes with at least one count per million reads throughout the gene body regions in at least three samples) associated with the corresponding GO term.

Table 4.2 GO term analysis of the genes that showed increased Pol II occupancy in *ref4-3* compared to wild type

Term	Description	Count	Fold Enrichment	P-value
GO:0005991	trehalose metabolic process	2	69.47	0.028
GO:0051697	protein delipidation	2	69.47	0.028
GO:2000762	regulation of phenylpropanoid metabolic process	3	29.77	0.004
GO:0043547	positive regulation of GTPase activity	3	10.42	0.033
GO:0006874	cellular calcium ion homeostasis	3	8.34	0.049
GO:0007623	circadian rhythm	7	5.46	0.002
GO:0009723	response to ethylene	7	4.38	0.005
GO:0080167	response to karrikin	6	3.90	0.019
GO:0009751	response to salicylic acid	6	3.36	0.033
GO:0009753	response to jasmonic acid	6	3.13	0.042
GO:0009737	response to abscisic acid	11	2.17	0.030
GO:0006468	protein phosphorylation	18	1.68	0.038

Table 4.3 A full list of the genes that show down-regulated Pol II occupancy in *med5* and show up-regulated Pol II occupancy in *ref4-3* [a]

AGI number	Gene name	Short description
AT2G44130 [b]	*KFB39*	Galactose oxidase/kelch repeat superfamily protein
AT3G59940 [b]	*KFB50*	Galactose oxidase/kelch repeat superfamily protein
AT3G13310 [b]	*DJC66*	Chaperone DnaJ-domain superfamily protein
AT3G62150 [c]	*ABCB21*	P-glycoprotein 21
AT1G06160 [c]	*ORA59*	octadecanoid-responsive AP2/ERF 59
AT5G61560 [c]	*PKU1*	U-box domain-containing protein kinase family protein
AT3G42880	*PRK3*	Leucine-rich repeat protein kinase family protein
AT2G05940	*RIPK*	Protein kinase superfamily protein
AT1G18910	*BTSL2*	zinc ion binding protein
AT3G50480	*HR4*	homolog of RPW8 4
AT1G78660	*GGH1*	gamma-glutamyl hydrolase 1
AT2G37710	*RLK*	receptor lectin kinase
AT5G56600	*PRF3*	profilin 3
AT3G47170		HXXXD-type acyl-transferase family protein
AT1G35515	*HOS10*	high response to osmotic stress 10
AT1G51090	*ATHMAD1*	Heavy metal transport/detoxification superfamily protein
AT2G36590	*ProT3*	proline transporter 3
AT3G62160		HXXXD-type acyl-transferase family protein
AT4G33980	*COR28*	cold-regulated gene 28
AT4G39820		Tetratricopeptide repeat (TPR)-like superfamily protein
AT1G35516		myb-like transcription factor family protein
AT1G78650	*POLD3*	DNA-directed DNA polymerase
AT2G36580		Pyruvate kinase family protein
AT2G44140		Peptidase family C54 protein
AT3G13300	*VCS*	Transducin/WD40 repeat-like superfamily protein
AT3G45240	*GRIK1*	geminivirus rep interacting kinase 1
AT3G59950		Peptidase family C54 protein
AT3G62140		NEFA-interacting nuclear protein
AT4G01750	*RGXT2*	rhamnogalacturonan xylosyltransferase 2
AT4G17098		ncRNA
AT4G17100		poly U-specific endoribonuclease-B protein
AT4G27660		hypothetical protein
AT4G39810		Polynucleotidyl transferase, ribonuclease H-like superfamily protein

a. Genes shaded in grey have available co-expression data in ATTED-II database and are included in the heatmap in Figure 4.4B.
b. Group 1 of highly co-expressed genes that are indicated in blue in Figure 4.4B.
c. Group 2 of highly co-expressed genes that are indicated in red in Figure 4.4B.

Table 4.4 GO term analysis of the genes that showed increased Pol II occupancy in *ref4-3* and restored Pol II occupancy in *ref4-3 cdk8-1*

Term	Description	Count	Fold Enrichment	P-value
GO:0009723	response to ethylene	3	9.59	0.037
GO:0009751	response to salicylic acid	3	8.58	0.045

A

Expression level	Gene number	Min (FPKM)	Max (FPKM)
high	1273	100.03	13049.57
medium	8679	10.01	99.97
low	12546	4.85E-06	10.00
silent	11059	0	0

B

High
Medium
Low
Silent

Read (CPM)

-3000   -1500   TSS   1500   3000
Genomic Region (5' -> 3')

C

Less than 1.5 kb
1.5 kb – 3 kb
3 kb or longer

Read (CPM)

-5000   -2500   TSS   2500   5000
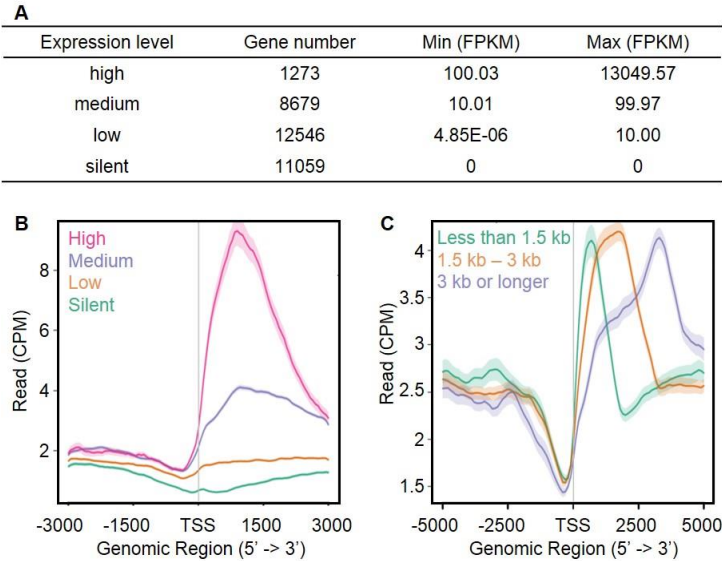Genomic Region (5' -> 3')

Figure 4.1 Pol II occupancy is highly elevated in the gene body regions of annotated genes

(A) Genes were classified into four groups including high, medium, low and silent, according to the expression level of each gene (fragments per kilobase of exon per million mapped reads (FPKM)) determined in three-week-old wild-type plants (Mao et al., 2019).

(B) Genes were classified into four groups based on their expression levels as described in panel A, and the enrichment of average Pol II signal (count per million mappable reads, CPM) in wild-type samples is plotted over a -3000 base pair (bp) to 3000 bp region with respect to the transcriptional start site (TSS).

(C) Genes were classified into three groups based on their lengths, and the enrichment of average Pol II signal (CPM) in wild-type samples is plotted over a -5000 bp to 5000 bp region with respect to the TSS.
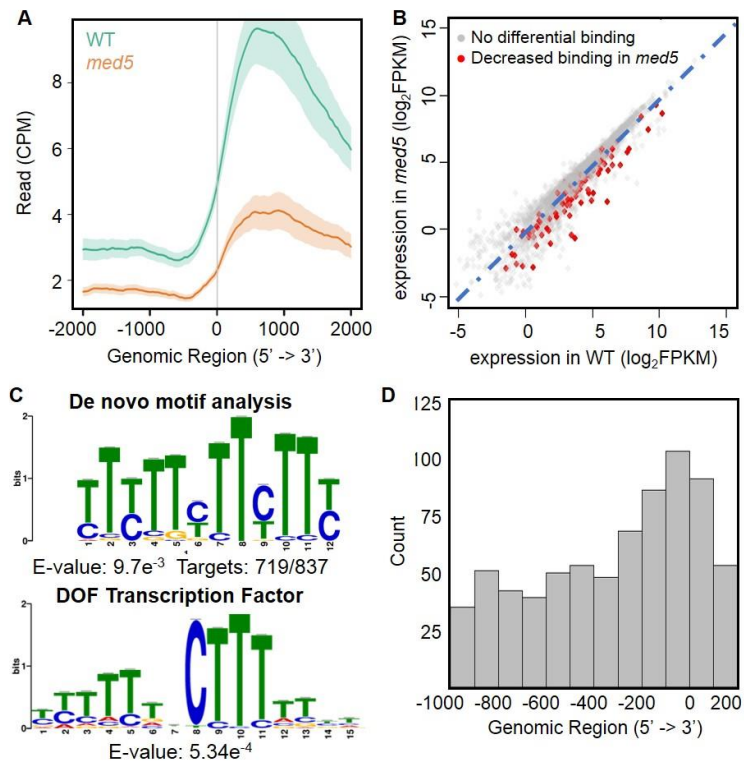
Figure 4.2 Genome-wide analysis identifies most of the differential binding sites have decreased Pol II occupancy in *med5* compared to wild type

(A) An average plot of the Pol II signals in wild type and *med5* for genes with decreased Pol II occupancy in *med5*. Pol II signals are plotted over a -2000 bp to 2000 bp region with respect to TSS.

(B) Comparison between MED5-targeted genes identified from Pol II ChIP-seq analysis and those from RNA-seq analysis. Expression levels of genes that were mis-regulated in *med5* (y-axis) are plotted against their expression levels in wild type (x-axis). The genes with decreased Pol II occupancy in *med5* are marked in red, whereas the others are marked in grey.

(C) A 12-bp motif enriched in the promoters of MED5-targeted genes and its best match in Arabidopsis known motif database (O'Malley et al., 2016). De novo motif analysis using MEME suite identified a significantly enriched motif in the promoter regions (-1000 bp to 200 bp with respect to TSS) of 835 genes with down-regulated Pol II occupancy in *med5*, among which 719 genes contained this enriched motif. Comparison between the enriched motif with known motifs in Arabidopsis database showed that the best match was the binding site for DNA-binding with one finger (DOF) transcription factors.

(D) A bar graph summarizing the occurrence (y-axis) of the enriched motif in the -1000 bp to 200 bp regions (x-axis) in MED5-targeted genes.
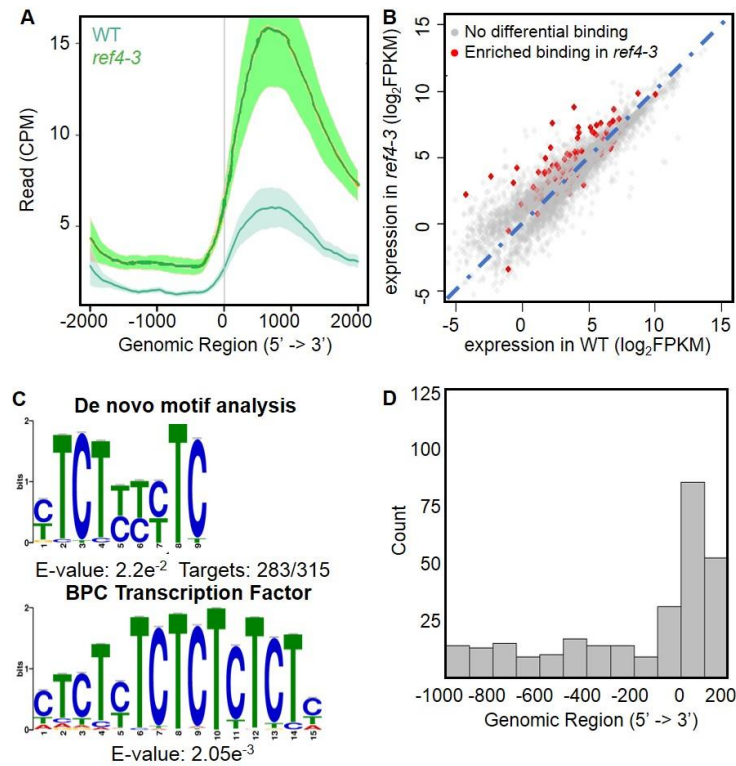
Figure 4.3 Genome-wide analysis identifies most of the differential binding sites have increased Pol II occupancy in *ref4-3* compared to wild type

(A) An average plot of the Pol II signals in wild type and *ref4-3* for genes with increased Pol II occupancy in *ref4-3*. Pol II signals are plotted over a -2000 bp to 2000 bp region with respect to TSSs.

(B) Comparison between genes targeted by *ref4-3* identified from Pol II ChIP-seq analysis and those from RNA-seq analysis. Expression levels of genes that were mis-regulated in *ref4-3* (y-axis) are plotted against their expression levels in wild type (x-axis). The genes with increased Pol II occupancy in *ref4-3* are marked in red, whereas the others are marked in grey.

(C) A 9-bp motif enriched in the promoters of genes targeted by *ref4-3* and its best match in Arabidopsis known motif database (O'Malley et al., 2016). De novo motif analysis using MEME suite identified a significantly enriched motif in the promoter regions (-1000 bp to 200 bp with respect to TSSs) of 315 genes with up-regulated Pol II occupancy in *ref4-3*, among which 283 genes contained this enriched motif. Comparison between the enriched motif with known motifs in Arabidopsis database showed that the best match was the binding site for basic pentacysteine (BPC) transcription factors.

(D) A bar graph summarizing the occurrence (y-axis) of the enriched motif in the -1000 bp to 200 bp regions (x-axis) in genes with increased Pol II recruitment in *ref4-3*.
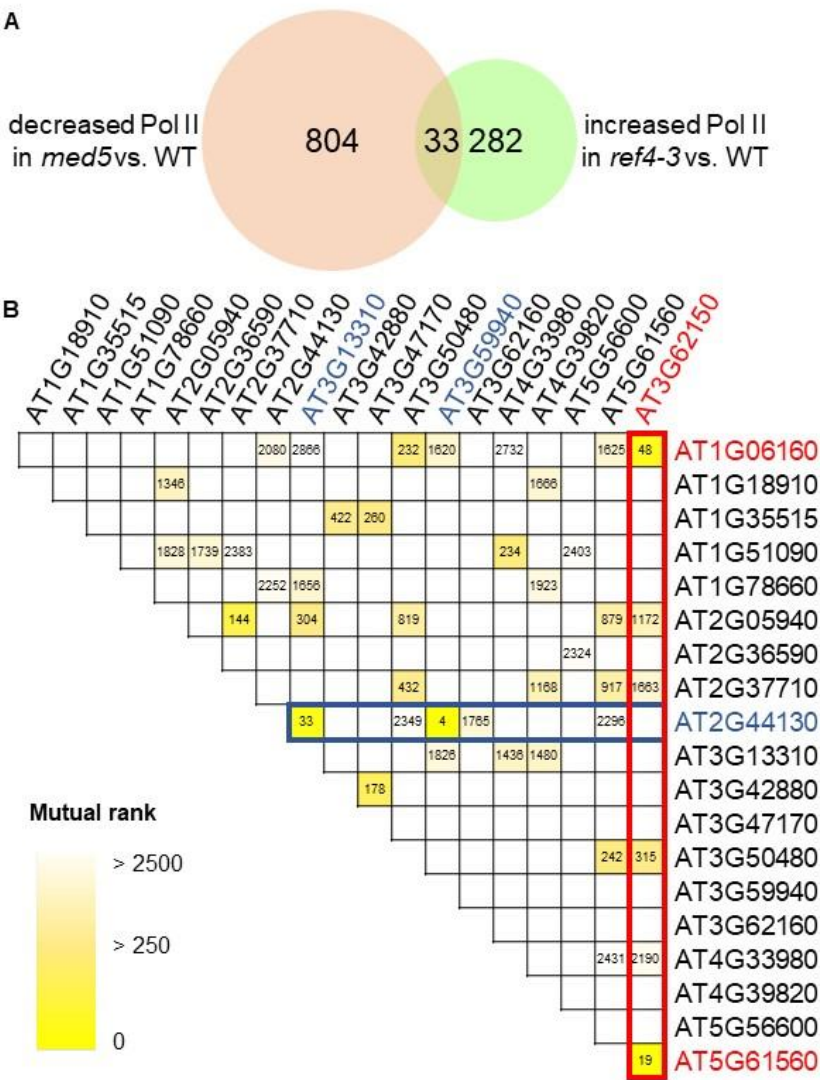
Figure 4.4 *ref4-3* and *med5* share some common targets that are highly co-expressed with each other

(A) Overlap in genes with reduced Pol II recruitment in *med5* and those with increased Pol II recruitment in *ref4-3*.
(B) A heat map depicting the correlation of the genes that are commonly mis-regulated in *med5* and *ref4-3* (Table 4.1). Genes with no co-expression data available were omitted in this plot. The numbers within the heat map indicate mutual rank (MR) scores for each pair of genes, and the MR scores larger than 2500 are omitted. Two groups of highly co-expressed genes are marked in red and blue respectively.

Figure 4.5 Pol II occupancy is significantly up-regulated in the gene body regions of *KFB39*, *KFB50* and *DJC66* in *ref4-3*, and down-regulated in *med5*
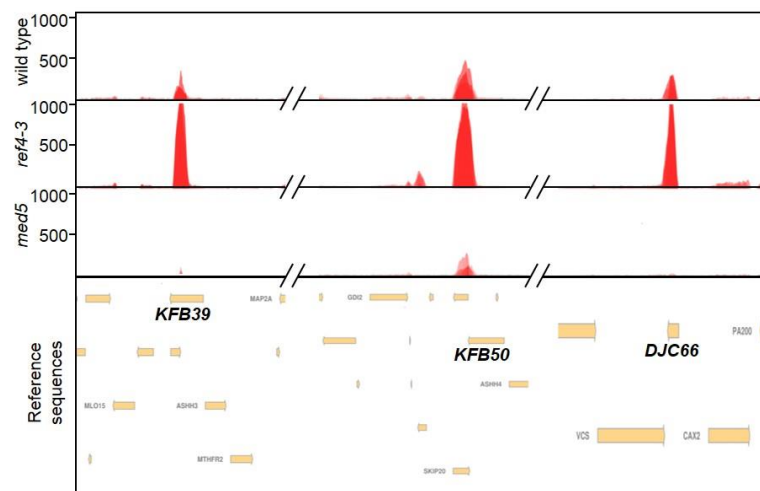
KFB39, *KFB50* and *DJC66* (bottom, yellow arrows) are three examples of the targets with mis-regulated Pol II occupancy in *ref4-3* and *med5*. Normalized Pol II signals (red) in each biological replicate are overlaid on top of each other.
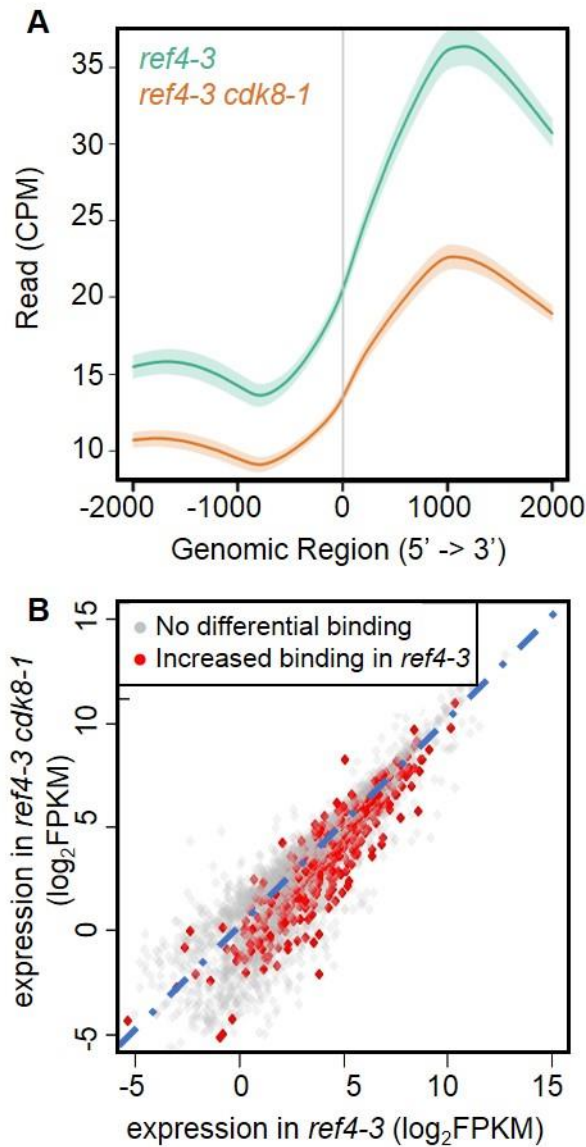
Figure 4.6 Loss of CDK8 in *ref4-3* causes an overall down-regulation for the Pol II occupancy

(A) An average plot of the Pol II signals in *ref4-3* and *ref4-3 cdk8-1* for genes with down-regulated Pol II occupancy in *ref4-3 cdk8-1* compared to *ref4-3*. Pol II signals are plotted over a -2000 bp to 2000 bp region with respect to TSSs.

(B) Overlap between the genes with reduced Pol II occupancy in *ref4-3 cdk8-1* compared to *ref4-3* and the genes with changed expression in the double mutant compared to *ref4-3*. Expression levels of genes that were differentially expressed in *ref4-3 cdk8-1* (y-axis) and *ref4-3* (x-axis) are shown in the scatter plot, and each dot represents one gene. The genes with decreased Pol II occupancy in *ref4-3 cdk8-1* compared to *ref4-3* are marked in red, whereas the others are marked in grey.

Figure 4.7 Genes with enhanced transcription in *ref4-3* and rescued in *ref4-3 cdk8-1* display similar expression level changes

(A) Overlap in genes with increased Pol II recruitment in *ref4-3* compared to wild type and those with decreased Pol II recruitment in *ref4-3 cdk8-1* compared to *ref4-3*.
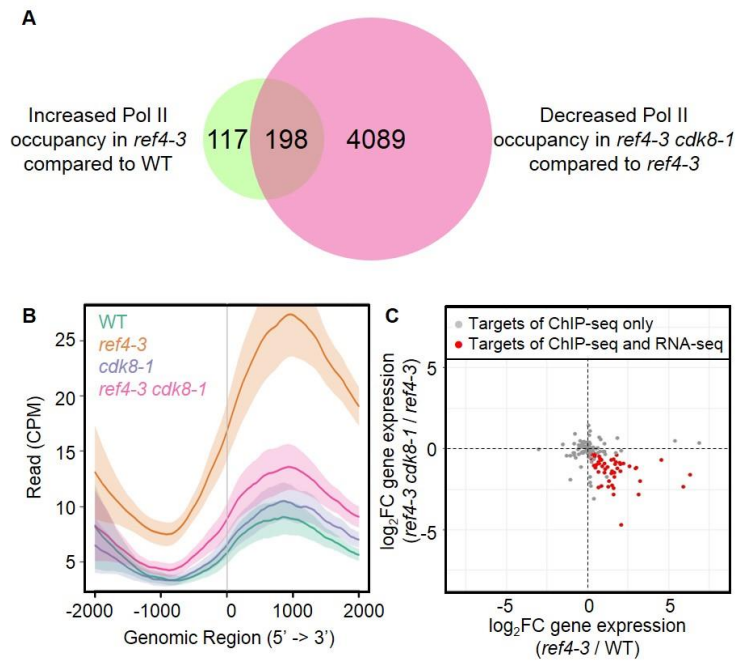
(B) An average plot of the Pol II signals in wild type, *ref4-3*, *cdk8-1* and *ref4-3 cdk8-1* for genes with increased Pol II occupancy in *ref4-3* compared to wild type and with rescued Pol II occupancy in *ref4-3 cdk8-1*. Pol II signals are plotted over a -2000 bp to 2000 bp region with respect to TSSs.

(B) Overlap between the genes that are up-regulated in *ref4-3* in a CDK8-dependent manner identified from RNA-seq and Pol II ChIP-seq. Each dot in the scatter plot represents a gene that showed up-regulated Pol II occupancy in *ref4-3* and can be rescued by disruption of CDK8. Log$_2$ fold change (FC) by comparison between the gene expression level in *ref4-3* and wild type can be read from the x-axis, whereas the log$_2$ fold change by comparison between the gene expression level in *ref4-3 cdk8-1 and ref4-3* can be read from the y-axis. The common targets identified from RNA-seq and Pol II ChIP-seq are marked in red, whereas the others are marked in grey.

REFERENCES

Abramoff, M.D., Magalhães, P.J., and Ram, S.J. (2004). Biophotonics international. *Biophotonics Int.* **11**: 36–42.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.

Anderson, N., Bonawitz, N.D., Nyffeler, K.E., and Chapple, C. (2015). Loss of ferulate 5-hydroxylase leads to Mediator-dependent inhibition of soluble phenylpropanoid biosynthesis in Arabidopsis. *Plant Physiol.*: pp.00294.2015.

Ansari, S.A. and Morse, R.H. (2012). Selective role of Mediator tail module in the transcription of highly regulated genes in yeast. *Transcription* **3**: 110–114.

Arabidopsis Interactome Mapping Consortium (2011). Evidence for network evolution in an Arabidopsis interactome map. *Science* **333**: 601–7.

Asturias, F.J., Jiang, Y.W., Myers, L.C., Gustafsson, C.M., and Kornberg, R.D. (1999). Conserved structures of mediator and RNA polymerase II holoenzyme. *Science.* **283**: 985–987.

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

Bailey, T.L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings. Int. Conf. Intell. Syst. Mol. Biol.* **2**: 28–36.

Bailey, T.L., Mikael, B., Buske Fabian A, Martin, F., Grant Charles E, Luca, C., Jingyuan, R., Li Wilfred W, and Noble William S (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**: W202–W208.

Bauer, D. (2004). Constitutive Photomorphogenesis 1 and Multiple Photoreceptors Control Degradation of Phytochrome Interacting Factor 3, a Transcription Factor Required for Light Signaling in Arabidopsis. *Plant Cell* **16**: 1433–1445.

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. *Annu. Rev. Plant Biol.* **54**: 519–546.

Bonawitz, N.D., Im Kim, J., Tobimatsu, Y., Ciesielski, P.N., Anderson, N.A., Ximenes, E., Maeda, J., Ralph, J., Donohoe, B.S., Ladisch, M., and others (2014). Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. *Nature* **509**: 376.

Bonawitz, N.D., Soltau, W.L., Blatchley, M.R., Powers, B.L., Hurlock, A.K., Seals, L.A., Weng, J.K., Stout, J., and Chapple, C. (2012). REF4 and RFR1, subunits of the transcriptional coregulatory complex mediator, are required for phenylpropanoid homeostasis in Arabidopsis. *J. Biol. Chem.* **287**: 5434–5445.

Borggrefe, T., Davis, R., Erdjument-Bromage, H., Tempst, P., and Kornberg, R.D. (2002). A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast. *J. Biol. Chem.* **277**: 44202–44207.

Borggrefe, T. and Yue, X. (2011). Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Semin. Cell Dev. Biol.* **22**: 759–768.

Bourbon, H.M. (2008). Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. *Nucleic Acids Res.* **36**: 3993–4008.

Briat, J.F., Duc, C., Ravet, K., and Gaymard, F. (2010). Ferritins and iron storage in plants. *Biochim. Biophys. Acta - Gen. Subj.* **1800**: 806–814.

Bryant, G.O. and Ptashne, M. (2003). Independent recruitment in vivo by Gal4 of two complexes required for transcription. *Mol. Cell* **11**: 1301–1309.

Buseman, C.M., Tamura, P., Sparks, A.A., Baughman, E.J., Maatta, S., Zhao, J., Roth, M.R., Esch, S.W., Shah, J., Williams, T.D., and Ruth, W. (2006). Wounding Stimulates the Accumulation of Glycerolipids Containing Oxophytodienoic Acid and Dinor-Oxophytodienoic Acid in Arabidopsis Leaves. *PLANT Physiol.* **142**: 28–39.

Cagnola, J.I., Ploschuk, E., Benech-Arnold, T., Finlayson, S.A., and Casal, J.J. (2012). Stem Transcriptome Reveals Mechanisms to Reduce the Energetic Cost of Shade-Avoidance Responses in Tomato. *PLANT Physiol.* **160**: 1110–1119.

Cao, H., Glazebrook, J., Clarke, J.D., Volko, S., and Dong, X. (1997). The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. *Cell* **88**: 57–63.

Carter, B., Bishop, B., Ho, K.K., Huang, R., Jia, W., Zhang, H., Pascuzzi, P.E., Deal, R.B., and Ogas, J. (2018). The chromatin remodelers PKL and PIE1 act in an epigenetic pathway that determines H3K27me3 homeostasis in Arabidopsis. *Plant Cell* **30**: 1337–1352.

Casal, J.J., Candia, A.N., and Sellaro, R. (2013). Light perception and signalling by phytochrome A. *J. Exp. Bot.* **65**: 2835–2845.

Cerdán, P.D. and Chory, J. (2003). Regulation of flowering time by light quality. *Nature* **423**: 881.

Cernac, A. and Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. *Plant J.* **40**: 575–585.

Chen, L., Wang, C., Qin, Z.S., and Wu, H. (2015). A novel statistical method for quantitative comparison of multiple ChIP-seq datasets. *Bioinformatics* **31**: 1889–1896.

Chen, R., Jiang, H., Li, L., Zhai, Q., Qi, L., Zhou, W., Liu, X., Li, H., Zheng, W., Sun, J., and Li, C. (2012). The Arabidopsis Mediator Subunit MED25 Differentially Regulates Jasmonate and Abscisic Acid Signaling through Interacting with the MYC2 and ABI5 Transcription Factors. *Plant Cell* **24**: 2898–2916.

Chiu, C.C., Chen, L.J., Su, P.H., and Li, H. min (2013). Evolution of Chloroplast J Proteins. *PLoS One* **8**: e70384.

Christians, M.J., Gingerich, D.J., Hua, Z., Lauer, T.D., and Vierstra, R.D. (2012). The Light-Response BTB1 and BTB2 Proteins Assemble Nuclear Ubiquitin Ligases That Modify Phytochrome B and D Signaling in Arabidopsis. *PLANT Physiol.* **160**: 118–134.

Chung, Y., Kwon, S. Il, and Choe, S. (2014). Antagonistic Regulation of Arabidopsis Growth by Brassinosteroids and Abiotic Stresses. *Mol. Cells* **37**: 795–803.

Clough, S.J. and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**: 735–743.

Conaway, R.C. and Conaway, J.W. (2011). Function and regulation of the Mediator complex. *Curr. Opin. Genet. Dev.* **21**: 225–230.

Conaway, R.C. and Conaway, J.W. (2013). The Mediator complex and transcription elongation. *Biochim. Biophys. Acta (BBA)-Gene Regul. Mech.* **1829**: 69–75.

Cosgrove, D.J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* **6**: 850–861.

Croft, H., Chen, J.M., Luo, X., Bartlett, P., Chen, B., and Staebler, R.M. (2017). Leaf chlorophyll content as a proxy for leaf photosynthetic capacity. *Glob. Chang. Biol.* **23**: 3513–3524.

D'Maris Amick Dempsey, A.C., Vlot, M.C.W., and Daniel, F.K. (2011). Salicylic acid biosynthesis and metabolism. *Arab. book/American Soc. Plant Biol.* **9**.

Davoine, C., Abreu, I.N., Khajeh, K., Blomberg, J., Kidd, B.N., Kazan, K., Schenk, P.M., Gerber, L., Nilsson, O., Moritz, T., and Björklund, S. (2017). Functional metabolomics as a tool to analyze Mediator function and structure in plants. *PLoS One* **12**: e0179640.

Dolan, W.L. and Chapple, C. (2017). Conservation and Divergence of Mediator Structure and Function: Insights from Plants. *Plant Cell Physiol.* **1**: 4–21.

Dolan, W.L. and Chapple, C. (2018). Transcriptome Analysis of Four Arabidopsis thaliana Mediator Tail Mutants Reveals Overlapping and Unique Functions in Gene Regulation. *G3 &#58; Genes|Genomes|Genetics* **8**: 3093–3108.

Dolan, W.L., Dilkes, B.P., Stout, J.M., Bonawitz, N.D., and Chapple, C. (2017). Mediator Complex Subunits MED2, MED5, MED16, and MED23 Genetically Interact in the Regulation of Phenylpropanoid Biosynthesis. *Plant Cell* **29**: 3269–3285.

Dotson, M.R., Yuan, C.X., Roeder, R.G., Myers, L.C., Gustafsson, C.M., Jiang, Y.W., Li, Y., Kornberg, R.D., and Asturias, F.J. (2000). Structural organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci.* **97**: 14307–14310.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**: 207–10.

Elfving, N., Davoine, C., Benlloch, R., Blomberg, J., Brannstrom, K., Muller, D., Nilsson, A., Ulfstedt, M., Ronne, H., Wingsle, G., Nilsson, O., and Bjorklund, S. (2011). The Arabidopsis thaliana Med25 mediator subunit integrates environmental cues to control plant development. *Proc. Natl. Acad. Sci.* **108**: 8245–8250.

Eloy, N.B., de Freitas Lima, M., Ferreira, P.C.G., and Inzé, D. (2015). The Role of the Anaphase-Promoting Complex/Cyclosome in Plant Growth. *CRC. Crit. Rev. Plant Sci.* **34**: 487–505.

Fallath, T., Kidd, B.N., Stiller, J., Davoine, C., Björklund, S., Manners, J.M., Kazan, K., and Schenk, P.M. (2017). MEDIATOR18 and MEDIATOR20 confer susceptibility to Fusarium oxysporum in Arabidopsis thaliana. *PLoS One* **12**: e0176022.

Feng, D.X., Tasset, C., Hanemian, M., Barlet, X., Hu, J., Trémousaygue, D., Deslandes, L., and Marco, Y. (2012). Biological control of bacterial wilt in Arabidopsis thaliana involves abscissic acid signalling. *New Phytol.* **194**: 1035–1045.

Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S., and others (2008). CDK8 is a colorectal cancer oncogene that regulates β-catenin activity. *Nature* **455**: 547.

Fornara, F., Panigrahi, K.C.S., Gissot, L., Sauerbrunn, N., Rühl, M., Jarillo, J.A., and Coupland, G. (2009). Arabidopsis DOF Transcription Factors act Redundantly to Reduce CONSTANS Expression and are Essential for a Photoperiodic Flowering Response. *Dev. Cell* **17**: 75–86.

Franklin, K.A. (2003). Mutant Analyses Define Multiple Roles for Phytochrome C in Arabidopsis Photomorphogenesis. *PLANT CELL ONLINE* **15**: 1981–1989.

Franklin, K.A. (2008). Shade avoidance. *New Phytol.* **179**: 930–944.

Franklin, K.A. and Quail, P.H. (2010). Phytochrome functions in Arabidopsis development. *J. Exp. Bot.* **61**: 11–24.

Franklin, K.A. and Whitelam, G.C. (2005). Phytochromes and shade-avoidance responses in plants. *Ann. Bot.* **96**: 169–175.

Fujita, Y., Fujita, M., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2011). ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J. Plant Res.* **124**: 509–525.

Fukasawa, R., Iida, S., Tsutsui, T., Hirose, Y., and Ohkuma, Y. (2015). Mediator complex cooperatively regulates transcription of retinoic acid target genes with Polycomb Repressive Complex 2 during neuronal differentiation. *J. Biochem.* **158**: 373–384.

Fukasawa, R., Tsutsui, T., Hirose, Y., Tanaka, A., and Ohkuma, Y. (2012). Mediator CDK subunits are platforms for interactions with various chromatin regulatory complexes. *J. Biochem.* **152**: 241–249.

Gallego-Giraldo, L., Escamilla-Trevino, L., Jackson, L.A., and Dixon, R.A. (2011a). Salicylic acid mediates the reduced growth of lignin down-regulated plants. *Proc. Natl. Acad. Sci.* **108**: 20814–20819.

Gallego-Giraldo, L., Jikumaru, Y., Kamiya, Y., Tang, Y., and Dixon, R.A. (2011b). Selective lignin downregulation leads to constitutive defense response expression in alfalfa (Medicago sativa L.). *New Phytol.* **190**: 627–639.

Gan, W.Q., Koehoorn, M., Davies, H.W., Demers, P.A., Tamburic, L., and Brauer, M. (2011). Long-term exposure to traffic-related air pollution and the risk of coronary heart disease hospitalization and mortality. *Environ. Health Perspect.* **119**: 501–507.

Geng, F., Wenzel, S., and Tansey, W.P. (2012). Ubiquitin and Proteasomes in Transcription. *Annu. Rev. Biochem.* **81**: 177–201.

Gillmor, C.S., Park, M.Y., Smith, M.R., Pepitone, R., Kerstetter, R.A., and Poethig, R.S. (2009). The MED12-MED13 module of Mediator regulates the timing of embryo patterning in Arabidopsis. *Development* **137**: 113–122.

Gonzalez-Grandio, E., Poza-Carrion, C., Sorzano, C.O.S., and Cubas, P. (2013). BRANCHED1 Promotes Axillary Bud Dormancy in Response to Shade in Arabidopsis. *Plant Cell* **25**: 834–850.

Gonzalez, D., Hamidi, N., Del Sol, R., Benschop, J.J., Nancy, T., Li, C., Francis, L., Tzouros, M., Krijgsveld, J., Holstege, F.C.P., and Conlan, R.S. (2014). Suppression of Mediator is regulated by Cdk8-dependent Grr1 turnover of the Med3 coactivator. *Proc. Natl. Acad. Sci.* **111**: 2500–2505.

Gordon, A., Hannon, G.J., and others (2010). Fastx-toolkit. *FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab. cshl. edu/fastx_toolkit* **5**.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**: 1017–1018.

Grünberg, S., Henikoff, S., Hahn, S., and Zentner, G.E. (2016). Mediator binding to UASs is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. *EMBO J.* **35**: 2435–2446.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. *Genome Biol.* **8**: R24.

H, L., B, H., A, W., T, F., J, R., N, H., G, M., G, A., and R, D. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Hahne, F. and Ivanek, R. (2016). Visualizing Genomic Data Using Gviz and Bioconductor BT - Statistical Genomics: Methods and Protocols. In Springer, pp. 335–351.

Hallberg, M., Polozkov, G. V., Hu, G.-Z., Beve, J., Gustafsson, C.M., Ronne, H., and Bjorklund, S. (2004). Site-specific Srb10-dependent phosphorylation of the yeast Mediator subunit Med2 regulates gene expression from the 2- m plasmid. *Proc. Natl. Acad. Sci.* **101**: 3370–3375.

Hauser, F., Li, Z., Waadt, R., and Schroeder, J.I. (2017). SnapShot: Abscisic Acid Signaling. *Cell* **171**: 1708-1708.e0.

Hecker, A., Brand, L.H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D. (2015). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant Physiol.* **168**: 1013–1024.

Hemsley, P.A., Hurst, C.H., Kaliyadasa, E., Lamb, R., Knight, M.R., De Cothi, E.A., Steele, J.F., and Knight, H. (2014). The Arabidopsis Mediator Complex Subunits MED16, MED14, and MED2 Regulate Mediator and RNA Polymerase II Recruitment to CBF-Responsive Cold-Regulated Genes. *Plant Cell* **26**: 465–484.

Hengartner, C.J., Myer, V.E., Liao, S.M., Wilson, C.J., Koh, S.S., and Young, R.A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. *Mol. Cell* **2**: 43–53.

Heyman, H.M. and Dubery, I.A. (2016). The potential of mass spectrometry imaging in plant metabolomics: a review. *Phytochem. Rev.* **15**: 297–316.

Holalu, S. V. and Finlayson, S.A. (2017). The ratio of red light to far red light alters Arabidopsis axillary bud growth and abscisic acid signalling before stem auxin changes. *J. Exp. Bot.* **68**: 943–952.

Huang, D.W., Sherman, B.T., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2008). DAVID gene ID conversion tool. *Bioinformation* **2**: 428.

Huot, B., Yao, J., Montgomery, B.L., and He, S.Y. (2014). Growth-defense tradeoffs in plants: A balancing act to optimize fitness. *Mol. Plant* **7**: 1267–1287.

Inigo, S., Giraldez, A.N., Chory, J., and Cerdan, P.D. (2012). Proteasome-Mediated Turnover of Arabidopsis MED25 Is Coupled to the Activation of FLOWERING LOCUS T Transcription. *PLANT Physiol.* **160**: 1662–1673.

Ito, J., Fukaki, H., Onoda, M., Li, L., Li, C., Tasaka, M., and Furutani, M. (2016). Auxin-dependent compositional change in Mediator in ARF7- and ARF19-mediated transcription. *Proc. Natl. Acad. Sci.* **113**: 6562–6567.

Ito, M., Okano, H.J., Darnell, R.B., and Roeder, R.G. (2002). The TRAP100 component

of the TRAP/Mediator complex is essential in broad transcriptional events and development. *EMBO J.* **21**: 3464–3475.

Jackson, S.D. and Prat, S. (1996). Control of tuberisation in potato by gibberellins and phytochrome B. *Physiol. Plant.* **98**: 407–412.

Jeronimo, C., Langelier, M.-F., Bataille, A.R., Pascal, J.M., Pugh, B.F., and Robert, F. (2016). Tail and kinase modules differently regulate core mediator recruitment and function in vivo. *Mol. Cell* **64**: 455–466.

Kamimoto, Y., Terasaka, K., Hamamoto, M., Takanashi, K., Fukuda, S., Shitan, N., et al. (2012). Arabidopsis ABCB21 is a facultative auxin importer/exporter regulated by cytoplasmic auxin concentration. *Plant Cell Physiol.* **53**: 2090–2100.

Kazan, K. and Manners, J.M. (2009). Linking development to defense: auxin in plant-pathogen interactions. *Trends Plant Sci.* **14**: 373–382.

Kidd, B.N., Edgar, C.I., Kumar, K.K., Aitken, E.A., Schenk, P.M., Manners, J.M., and Kazan, K. (2009). The Mediator Complex Subunit PFT1 Is a Key Regulator of Jasmonate-Dependent Defense in Arabidopsis. *PLANT CELL ONLINE* **21**: 2237–2252.

Kim, J.I., Sharkhuu, A., Jin, J.B., Li, P., Jeong, J.C., Baek, D., et al. (2007). yucca6, a Dominant Mutation in Arabidopsis, Affects Auxin Accumulation and Auxin-Related Phenotypes. *PLANT Physiol.* **145**: 722–735.

Kim, J.I., Dolan, W.L., Anderson, N.A., and Chapple, C. (2015). Indole Glucosinolate Biosynthesis Limits Phenylpropanoid Accumulation in Arabidopsis thaliana. *Plant Cell* **27**: 1529–1546.

Kim, M.J., Jang, I.-C., and Chua, N.-H. (2016). The Mediator Complex MED15 Subunit Mediates Activation of Downstream Lipid-Related Genes by the WRINKLED1 Transcription Factor. *Plant Physiol.* **171**: 1951–1964.

Kim, S., Xu, X., Hecht, A., and Boyer, T.G. (2006). Mediator is a transducer of Wnt/β-catenin signaling. *J. Biol. Chem.* **281**: 14066–14075.

Kleinow, T., Himbert, S., Krenz, B., Jeske, H., and Koncz, C. (2009). NAC domain transcription factor ATAF1 interacts with SNF1-related kinases and silencing of its subfamily causes severe developmental defects in Arabidopsis. *Plant Sci.* **177**: 360–370.

Knight, H., Mugford, S.G., Ülker, B., Gao, D., Thorlby, G., and Knight, M.R. (2009). Identification of SFR6, a key component in cold acclimation acting post-translationally on CBF function. *Plant J.* **58**: 97–108.

Knuesel, M.T., Meyer, K.D., Bernecky, C., and Taatjes, D.J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. *Genes Dev.* **23**: 439–451.

Kong, L. and Chang, C. (2018). Suppression of wheat TaCDK8/TaWIN1 interaction negatively affects germination of Blumeria graminis f.sp. tritici by interfering with very-long-chain aldehyde biosynthesis. *Plant Mol. Biol.* **96**: 165–178.

Kuchin, S., Yeghiayan, P., and Carlson, M. (1995). Cyclin-dependent protein kinase and cyclin homologs SSN3 and SSN8 contribute to transcriptional control in yeast. *Proc. Natl. Acad. Sci.* **92**: 4006–4010.

Kumar, V., Waseem, M., Dwivedi, N., Maji, S., Kumar, A., and Thakur, J.K. (2018). KIX domain of AtMed15a, a Mediator subunit of Arabidopsis, is required for its interaction with different proteins. *Plant Signal. Behav.* **13**: e1428514.

Kurepin, L. V., Shah, S., and Reid, D.M. (2007). Light quality regulation of endogenous levels of auxin, abscisic acid and ethylene production in petioles and leaves of wild type and ACC deaminase transgenic Brassica napus seedlings. *Plant Growth Regul.* **52**: 53–60.

Lai, Z., Schluttenhofer, C.M., Bhide, K., Shreve, J., Thimmapuram, J., Lee, S.Y., Yun, D.J., and Mengiste, T. (2014). MED18 interaction with distinct transcription factors regulates multiple plant functions. *Nat. Commun.* **5**.

Lee, T., Poss, Z.C., Taatjes, D.J., Odell, A.T., Dowell, R.D., Pelish, H.E., Shair, M.D., Tangpeerachaikul, A., Old, W.M., and Ebmeier, C.C. (2016). Identification of Mediator Kinase Substrates in Human Cells using Cortistatin A and Quantitative Phosphoproteomics. *Cell Rep.* **15**: 436–450.

Leivar, P. and Quail, P.H. (2011). PIFs: pivotal components in a cellular signaling hub. *Trends Plant Sci.* **16**: 19–28.

Leivar, P., Tepperman, J.M., Cohn, M.M., Monte, E., Al-Sady, B., Erickson, E., and Quail, P.H. (2012). Dynamic Antagonism between Phytochromes and PIF Family Basic Helix-Loop-Helix Factors Induces Selective Reciprocal Responses to Light and Shade in a Rapidly Responsive Transcriptional Network in Arabidopsis. *Plant Cell* **24**: 1398–1419.

Li, Y., Kim, J.I., Pysh, L., and Chapple, C. (2015). Four isoforms of Arabidopsis thaliana 4-coumarate: CoA ligase (4CL) have overlapping yet distinct roles in phenylpropanoid metabolism. *Plant Physiol.*: pp.00838.2015.

Lin, P.C., Pomeranz, M.C., Jikumaru, Y., Kang, S.G., Hah, C., Fujioka, S., Kamiya, Y., and Jang, J.C. (2011). The Arabidopsis tandem zinc finger protein AtTZF1 affects ABA- and GA-mediated growth, stress and gene expression responses. *Plant J.* **65**: 253–268.

Lingam, S., Mohrbacher, J., Brumbarova, T., Potuschak, T., Fink-Straube, C., Blondet, E., Genschik, P., and Bauer, P. (2011). Interaction between the bHLH Transcription Factor FIT and ETHYLENE INSENSITIVE3/ETHYLENE INSENSITIVE3-LIKE1 Reveals Molecular Linkage between the Regulation of Iron Acquisition and Ethylene Signaling in Arabidopsis. *Plant Cell* **23**: 1815–1829.

Liu, J., Hua, W., Zhan, G., Wei, F., Wang, X., Liu, G., and Wang, H. (2010). Increasing seed mass and oil content in transgenic Arabidopsis by the overexpression of wri1-like gene from Brassica napus. *Plant Physiol. Biochem.* **48**: 9–15.

Loncle, N., Boube, M., Joulia, L., Boschiero, C., Werner, M., Cribbs, D.L., and Bourbon, H.M. (2007). Distinct roles for Mediator Cdk8 module subunits in Drosophila development. *EMBO J.* **26**: 1045–1054.

Maeo, K., Tokuda, T., Ayame, A., Mitsui, N., Kawai, T., Tsukagoshi, H., Ishiguro, S., and Nakamura, K. (2009). An AP2-type transcription factor, WRINKLED1, of Arabidopsis thaliana binds to the AW-box sequence conserved among proximal upstream regions of genes involved in fatty acid synthesis. *Plant J.* **60**: 476–487.

Malik, S. and Roeder, R.G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem. Sci.* **30**: 256–263.

Malik, S. and Roeder, R.G. (2010). The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nat. Rev. Genet.* **11**: 761.

Mao, X., Kim, J.I., Wheeler, M.T., Heintzelman, A.K., Weake, V.M., and Chapple, C. (2019). Mutation of Mediator subunit CDK8 counteracts the stunted growth and salicylic acid hyperaccumulation phenotypes of an Arabidopsis MED5 mutant. *New Phytol.*: nph.15741.

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

McAdam, S.A.M. (2015). Physicochemical quantification of abscisic acid levels in plant tissues with an added internal standard by ultra-performance liquid chromatography. *Bio-protocol* **5**: 1–13.

McCarthy, D., Chen, Y., Research, G.S.-N. acids, and 2012, U. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *academic.oup.com*.

McCormac, A.C., Fischer, A., Kumar, A.M., Söll, D., and Terry, M.J. (2001). Regulation of HEMA1 expression by phytochrome and a plastid signal during de-etiolation in Arabidopsis Thaliana. *Plant J.* **25**: 549–561.

McLachlan, D.H., Pridgeon, A.J., and Hetherington, A.M. (2018). How Arabidopsis Talks to Itself about Its Water Supply. *Mol. Cell* **70**: 991–992.

Meyer, K.D., Lin, S.C., Bernecky, C., Gao, Y., and Taatjes, D.J. (2010). P53 activates transcription by directing structural shifts in Mediator. *Nat. Struct. Mol. Biol.* **17**: 753–760.

Mittler, G., Stühler, T., Santolin, L., Uhlmann, T., Kremmer, E., Lottspeich, F., Berti, L., and Meisterernst, M. (2003). A novel docking site on Mediator is critical for activation by VP16 in mammalian cells. *EMBO J.* **22**: 6494–6504.

Mokry, M., Hatzis, P., Schuijers, J., Lansu, N., Ruzius, F.P., Clevers, H., and Cuppen, E. (2012). Integrated genome-wide analysis of transcription factor occupancy, RNA polymerase II binding and steady-state RNA levels identify differentially regulated functional gene classes. *Nucleic Acids Res.* **40**: 148–158.

Morelli, G. and Ruberti, I. (2002). Light and shade in the photocontrol of Arabidopsis growth. *Trends Plant Sci.* **7**: 399–404.

Morris, E.J., Ji, J.Y., Yang, F., Di Stefano, L., Herr, A., Moon, N.S., Kwon, E.J., Haigis, K.M., Näär, A.M., and Dyson, N.J. (2008). E2F1 represses β-catenin transcription and is antagonized by both pRB and CDK8. *Nature* **455**: 552–556.

Msanne, J., Lin, J., Stone, J.M., and Awada, T. (2011). Characterization of abiotic stress-responsive Arabidopsis thaliana RD29A and RD29B genes and evaluation of transgenes. *Planta* **234**: 97–107.

Mukundan, B. and Ansari, A. (2011). Novel role for mediator complex subunit Srb5/Med18 in termination of transcription. *J. Biol. Chem.* **286**: 37053–37057.

Murashige, T. and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* **15**: 473–497.

Muro-Villanueva, F., Mao, X., and Chapple, C. (2019). Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr. Opin. Biotechnol.* **56**: 202–208.

Neff, M.M., Neff, J.D., Chory, J., and Pepper, A.E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in Arabidopsis thaliana genetics. *Plant J.* **14**: 387–392.

Nemet, J., Jelicic, B., Rubelj, I., and Sopta, M. (2014). The two faces of Cdk8, a positive/negative regulator of transcription. *Biochimie* **97**: 22–27.

Nemhauser, J.L., Hong, F., and Chory, J. (2006). Different Plant Hormones Regulate Similar Processes through Largely Nonoverlapping Transcriptional Responses. *Cell* **126**: 467–475.

Ng, S., Giraud, E., Duncan, O., Law, S.R., Wang, Y., Xu, L., et al. (2013). Cyclin-dependent kinase E1 (CDKE1) provides a cellular switch in plants between growth and stress responses. *J. Biol. Chem.* **288**: 3449–3459.

Nozawa, K., Schneider, T.R., and Cramer, P. (2017). Core Mediator structure at 3.4 Å extends model of transcription initiation complex. *Nature* **545**: 248–251.

O'Malley, R.C., Huang, S.-S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**: 1280–1292.

Obayashi, T., Aoki, Y., Tadaka, S., Kagaya, Y., and Kinoshita, K. (2018). ATTED-II in 2018: A Plant Coexpression Database Based on Investigation of the Statistical Property of the Mutual Rank Index. *Plant Cell Physiol.* **59**: e3.

Ou, B., Yin, K.Q., Liu, S.N., Yang, Y., Gu, T., Wing Hui, J.M., Zhang, L., Miao, J., Kondou, Y., Matsui, M., Gu, H.Y., and Qu, L.J. (2011). A high-throughput screening system for arabidopsis transcription factors and its application to Med25-dependent transcriptional regulation. *Mol. Plant* **4**: 546–555.

Ou, M. (2014). Searching for RNAP II on the Compact HSV-1 Genome.

Park, E., Kim, J., Lee, Y., Shin, J., Oh, E., Chung, W. Il, Jang, R.L., and Choi, G. (2004). Degradation of phytochrome interacting factor 3 in phytochrome-mediated light signaling. *Plant Cell Physiol.* **45**: 968–975.

Park, E., Park, J., Kim, J., Nagatani, A., Lagarias, J.C., and Choi, G. (2012). Phytochrome B inhibits binding of phytochrome-interacting factors to their target promoters. *Plant J.* **72**: 537–546.

Phatnani, H.P. and Greenleaf, A.L. (2006). Phosphorylation and functions of the RNA polymerase II CTD. *Genes Dev.* **20**: 2922–2936.

Pre, M., Atallah, M., Champion, A., De Vos, M., Pieterse, C.M.J., and Memelink, J. (2008). The AP2/ERF Domain Transcription Factor ORA59 Integrates Jasmonic Acid and Ethylene Signals in Plant Defense. *PLANT Physiol.* **147**: 1347–1357.

Reed, J.W., Nagpal, P., Poole, D.S., Furuya, M., and Chory, J. (1993). Mutations in the gene for the red/far-red light receptor phytochrome B alter cell elongation and physiological responses throughout Arabidopsis development. *Plant Cell* **5**: 147–157.

Reeves, W.M. and Hahn, S. (2003). Activator-independent functions of the yeast mediator sin4 complex in preinitiation complex formation and transcription reinitiation. *Mol. Cell. Biol.* **23**: 349–358.

Rickert, P., Corden, J.L., and Lees, E. (1999). Cyclin C/CDK8 and cyclin H/CDK7/p36 are biochemically distinct CTD kinases. *Oncogene* **18**: 1093–1102.

Roig-Villanova, I. and Martínez-García, J.F. (2016). Plant Responses to Vegetation Proximity: A Whole Life Avoiding Shade. *Front. Plant Sci.* **7**.

Rout, G.R. and Sahoo, S. (2015). ROLE OF IRON IN PLANT GROWTH AND METABOLISM. *Rev. Agric. Sci.* **3**: 1–24.

Rozhon, W., Petutschnig, E., Wrzaczek, M., and Jonak, C. (2005). Quantification of free and total salicylic acid in plants by solid-phase extraction and isocratic high-performance anion-exchange chromatography. *Anal. Bioanal. Chem.* **382**: 1620–1627.

Ruegger, M. and Chapple, C. (2001). Mutations that reduce sinapoylmalate accumulation in Arabidopsis thaliana define loci with diverse roles in phenylpropanoid metabolism. *Genetics* **159**: 1741–1749.

Rzymski, T., Mikula, M., Wiklik, K., and Brzózka, K. (2015). CDK8 kinase - An emerging target in targeted cancer therapy. *Biochim. Biophys. Acta - Proteins Proteomics* **1854**: 1617–1629.

Samanta, S. and Thakur, J.K. (2015). Importance of Mediator complex in the regulation and integration of diverse signaling pathways in plants. *Front. Plant Sci.* **6**.

Schindelman, G., Morikami, A., Jung, J., Baskin, T.I., Carpita, N.C., Derbyshire, P., McCann, M.C., and Benfey, P.N. (2001). COBRA encodes a putative GPI-anchored protein, which is polarly localized and necessary for oriented cell expansion in arabidopsis. *Genes Dev.* **15**: 1115–1127.

Schulz, P., Jansseune, K., Degenkolbe, T., Meret, M., Claeys, H., Skirycz, A., Teige, M., Willmitzer, L., and Hannah, M.A. (2014). Poly(ADP-Ribose)polymerase activity controls plant growth by promoting leaf cell number. *PLoS One* **9**: e90322.

Sessa, G., Carabelli, M., Possenti, M., Morelli, G., and Ruberti, I. (2018). Multiple Pathways in the Control of the Shade Avoidance Response. *Plants* **7**: 102.

Sessa, G., Carabelli, M., Sassi, M., Ciolfi, A., Possenti, M., Mittempergher, F., Becker, J., Morelli, G., and Ruberti, I. (2005). A dynamic balance between gene activation and repression regulates the shade avoidance response in Arabidopsis. *Genes Dev.* **19**: 2811–2815.

Shen, B., Allen, W.B., Zheng, P., Li, C., Glassman, K., Ranch, J., Nubel, D., and Tarczynski, M.C. (2010). Expression of ZmLEC1 and ZmWRI1 Increases Seed Oil Production in Maize. *PLANT Physiol.* **153**: 980–987.

Shen, L., Shao, N., Liu, X., and Nestler, E. (2014). Ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**: 284.

Smith, A.P., Nourizadeh, S.D., Peer, W.A., Xu, J., Bandyopadhyay, A., Murphy, A.S., and Goldsbrough, P.B. (2003). Arabidopsis AtGSTF2 is regulated by ethylene and auxin, and encodes a glutathione S-transferase that interacts with flavonoids. *Plant J.* **36**: 433–442.

Sorek, N., Szemenyei, H., Sorek, H., Landers, A., Knight, H., Bauer, S., Wemmer, D.E., and Somerville, C.R. (2015). Identification of MEDIATOR16 as the Arabidopsis COBRA suppressor MONGOOSE1. *Proc. Natl. Acad. Sci.* **112**: 16048–16053.

Ståldal, V., Cierlik, I., Landberg, K., Myrenås, M., Sundström, J.F., Eklund, D.M., Chen, S., Baylis, T., Ljung, K., and Sundberg, E. (2012). The Arabidopsis thaliana transcriptional activator STYLISH1 regulates genes affecting stamen development, cell expansion and timing of flowering. *Plant Mol. Biol.* **78**: 545–559.

Stampfel, G., Frank, O., Wienerroither, S., Reiter, F., Stark, A., Tfs, S., et al. (2015). Transcriptional regulators form diverse groups with context-dependent regulatory functions. *Nature*.

Stelmach, B.A., Müller, A., Hennig, P., Gebhardt, S., Schubert-Zsilavecz, M., and Weiler, E.W. (2001). A Novel Class of Oxylipins, sn1-O-(12-Oxophytodienoyl)-sn2-O-(hexadecatrienoyl)-monogalactosyl Diglyceride, from Arabidopsis thaliana. *J. Biol. Chem.* **276**: 12832–12838.

Stout, J., Romero-Severson, E., Ruegger, M.O., and Chapple, C. (2008). Semidominant mutations in reduced epidermal fluorescence 4 reduce phenylpropanoid content in Arabidopsis. *Genetics* **178**: 2237–2251.

Taatjes, D.J. (2010). The human Mediator complex: A versatile, genome-wide regulator of transcription. *Trends Biochem. Sci.* **35**: 315–322.

Taubert, S., Van Gilst, M.R., Hansen, M., and Yamamoto, K.R. (2006). A mediator subunit, MDT-15, integrates regulation of fatty acid metabolism by NHR-49-dependent and -independent pathways in C. elegans. *Genes Dev.* **20**: 1137–1149.

Teale, W.D., Paponov, I.A., and Palme, K. (2006). Auxin in action: Signalling, transport and the control of plant growth and development. *Nat. Rev. Mol. Cell Biol.* **7**: 847–859.

Thakur, J.K., Arthanari, H., Yang, F., Chau, K.H., Wagner, G., and Näär, A.M. (2009). Mediator subunit Gal11p/MED15 is required for fatty acid-dependent gene activation by yeast transcription Factor Oaf1p. *J. Biol. Chem.* **284**: 4422–4428.

Thakur, J.K., Yadav, A., and Yadav, G. (2013). Molecular recognition by the KIX domain and its role in gene regulation. *Nucleic Acids Res.* **42**: 2112–2125.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**: 1105–1111.

Tsai, K.-L., Tomomori-Sato, C., Sato, S., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2014). Subunit architecture and functional modular rearrangements of the transcriptional mediator complex. *Cell* **157**: 1430–1444.

Tsai, K.-L., Yu, X., Gopalan, S., Chao, T.-C., Zhang, Y., Florens, L., Washburn, M.P., Murakami, K., Conaway, R.C., Conaway, J.W., and others (2017). Mediator structure and rearrangements required for holoenzyme formation. *Nature* **544**: 196.

Tsai, K.L., Sato, S., Tomomori-Sato, C., Conaway, R.C., Conaway, J.W., and Asturias, F.J. (2013). A conserved Mediator-CDK8 kinase module association regulates Mediator-RNA polymerase II interaction. *Nat. Struct. Mol. Biol.* **20**: 611–619.

Umezawa, T., Okamoto, M., Kushiro, T., Nambara, E., Oono, Y., Seki, M., Kobayashi, M., Koshiba, T., Kamiya, Y., and Shinozaki, K. (2006). CYP707A3, a major ABA 8′-hydroxylase involved in dehydration and rehydration response in Arabidopsis thaliana. *Plant J.* **46**: 171–182.

Vishwakarma, K., Upadhyay, N., Kumar, N., Yadav, G., Singh, J., Mishra, R.K., Kumar, V., Verma, R., Upadhyay, R.G., Pandey, M., and Sharma, S. (2017). Abscisic Acid Signaling and Abiotic Stress Tolerance in Plants: A Review on Current Knowledge and Future Prospects. *Front. Plant Sci.* **08**.

Wang, C., Du, X., and Mou, Z. (2016). The Mediator Complex Subunits MED14, MED15, and MED16 Are Involved in Defense Signaling Crosstalk in Arabidopsis. *Front. Plant Sci.* **7**.

Wang, W. and Chen, X. (2004). HUA ENHANCER3 reveals a role for a cyclin-dependent protein kinase in the specification of floral organ identity in Arabidopsis. *Development* **131**: 3147–3156.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Métraux, J.-P., and Ryals, J.A. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* **3**: 1085–1094.

Wilson, K.E., Ivanov, A.G., Öquist, G., Grodzinski, B., Sarhan, F., and Huner, N.P.A. (2006). Energy balance, organellar redox status, and acclimation to environmental stress. *Can. J. Bot.* **84**: 1355–1370.

Xu, S., Grullon, S., Ge, K., and Peng, W. (2014). Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Stem Cell Transcr. Networks*: 97–111.

Yamamoto, Y.Y., Yoshitsugu, T., Sakurai, T., Seki, M., Shinozaki, K., and Obokata, J. (2009). Heterogeneity of Arabidopsis core promoters revealed by high-density TSS analysis. *Plant J.* **60**: 350–362.

Yamamuro, C., Zhu, J.-K., and Yang, Z. (2016). Epigenetic modifications and plant hormone action. *Mol. Plant* **9**: 57–70.

Yang, F., DeBeaumont, R., Zhou, S., and Naar, A.M. (2004). The activator-recruited cofactor/Mediator coactivator subunit ARC92 is a functionally important target of the VP16 transcriptional activator. *Proc. Natl. Acad. Sci.* **101**: 2339–2344.

Yang, F., Vought, B.W., Satterlee, J.S., Walker, A.K., Sun, Z.-Y.J., Watts, J.L., DeBeaumont, R., Saito, R.M., Hyberts, S.G., Yang, S., and others (2006). An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. *Nature* **442**: 700.

Yang, Y., Li, L., and Qu, L.J. (2016). Plant Mediator complex and its critical functions in transcription regulation. *J. Integr. Plant Biol.* **58**: 106–118.

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

Yang, Y., Shah, J., and Klessig, D.F. (1997). Signal perception and transduction in plant defense responses. *Genes Dev.* **11**: 1621–1639.

Ye, J., Wang, C., Sun, Y., Qu, J., Mao, H., and Chua, N.-H. (2018). Overexpression of a Transcription Factor Increases Lipid Content in a Woody Perennial Jatropha curcas. *Front. Plant Sci.* **9**.

Zhang, X., Gou, M., Guo, C., Yang, H., and Liu, C.-J. (2015). Down-regulation of Kelch domain-containing F-box protein in Arabidopsis enhances the production of (poly) phenols and tolerance to ultraviolet radiation. *Plant Physiol.* **167**: 337–350.

Zhang, X., Gou, M., and Liu, C.-J. (2013a). Arabidopsis Kelch Repeat F-Box Proteins Regulate Phenylpropanoid Biosynthesis via Controlling the Turnover of Phenylalanine Ammonia-Lyase. *Plant Cell* **25**: 4994–5010.

Zhang, X., Yao, J., Zhang, Y., Sun, Y., and Mou, Z. (2013b). The Arabidopsis Mediator complex subunits MED14/SWP and MED16/SFR6/IEN1 differentially regulate defense gene expression in plant immune responses. *Plant J.* **75**: 484–497.

Zhang, Y., Wu, H., Wang, N., Fan, H., Chen, C., Cui, Y., Liu, H., and Ling, H.Q. (2014). Mediator subunit 16 functions in the regulation of iron uptake gene expression in Arabidopsis. *New Phytol.* **203**: 770–783.

Zhao, J., Ramos, R., and Demma, M. (2013a). CDK8 regulates E2F1 transcriptional activity through S375 phosphorylation. *Oncogene* **32**: 3520–3530.

Zhao, J., Zhou, J. jun, Wang, Y. ying, Gu, J. wei, and Xie, X. zhi (2013b). Positive Regulation of Phytochrome B on Chlorophyll Biosynthesis and Chloroplast Development in Rice. *Rice Sci.* **20**: 243–248.

Zhao, X., Feng, D., Wang, Q., Abdulla, A., Xie, X.J., Zhou, J., et al. (2012). Regulation of lipogenesis by cyclin-dependent kinase 8 - Mediated control of SREBP-1. *J. Clin. Invest.* **122**: 2417–2427.

Zhong, R. and Ye, Z.H. (2007). Regulation of cell wall biosynthesis. *Curr. Opin. Plant Biol.* **10**: 564–572.

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

VITA

Xiangying (Candy) Mao

Education

Fudan University, China	Biotechnology	B.S.	2013
Purdue University, IN	Biochemistry	graduate student	2013-present
Purdue University, IN	Statistics	grad certification	2018

Employment and Professional Experience

Graduate research assistant, Department of Biochemistry, Purdue University

    Co-advisors: Dr. Clint Chapple and Dr. Vikki Weake   02/2014-present

Intern, Department of Marketing, Oncology, Bristol-Myers Squibb, Shanghai, China

    Director: Mr. Rong Yu   02/2013-06/2013

Undergrad research assistant, State Key Laboratory of Genetic Engineering, Fudan University, Shanghai, China

    PI: Dr. Hongyan Wang   03/2012 - 06/2013

Undergrad research assistant, Institute of integrated Traditional Chinese Medicine & Western Medicine, Huashan Hospital, Shanghai, China

    PI: Dr. Qin Bian   02/2011 - 03/2012

Honors and Awards

Arnold Kent Balls Award ($1000), Department of Biochemistry, Purdue University, 04/2019

Travel Award ($1000), Weiner Travel Award, Purdue University, 10/2018

Certificate of excellence in interdisciplinary research, Purdue Graduate School, 05/2018

Travel Award ($1000), American Society of Biochemistry and Molecular Biology, 01/2018

Travel Award ($775), American Society of Plant Biologists, 04/2017

Bird Stair Fellowship ($5000), Department of Biochemistry, Purdue University, Investigating the genome-wide distribution of Pol II in Arabidopsis med mutants, 02/2017

Honorable mention, Outstanding research poster, 2016 Global Climate and Energy Project (GCEP) Research Symposium, Stanford University, 11/2016

Bird Stair Fellowship ($4200), Department of Biochemistry, Purdue University, Investigating the interaction between MED5 and CDK8, 08/2015

Conferences and Presentations

RECOMB/ISCB Conference on Regulatory and Systems Genomics, New York City, NY, 12/2018. Poster entitled 'Mutation of Mediator subunit CDK8 counteracts the stunted growth in an Arabidopsis MED5 mutant'.

Midwest Chromatin and Epigenetics Meeting, West Lafayette, IN, 06/2018. Talk entitled 'Investigating the interaction between MED5 and CDK8 in Arabidopsis'.

American Society of Biochemistry and Molecular Biology, San Diego, CA, 04/2018. Talk entitled 'Investigating the interaction between MED5 and CDK8 in Arabidopsis'

PepsiCo Journey Through Science Day, New York City, NY, 09/2017. Poster entitled 'MED5 and CDK8 play a role in lignin-modification-induced dwarfing in Arabidopsis'

American Society of Plant Biologists (ASPB), Honolulu, HI, 06/2017. Poster entitled 'MED5 and CDK8 play a role in lignin-modification-induced dwarfing in Arabidopsis'

2017 Midwest ASPB Meeting, West Lafayette, IN, 02/2017. Talk entitled 'MED5 and CDK8 play a role in lignin-modification-induced dwarfing in Arabidopsis'

GCEP research Symposium, Stanford, CA, 11/2016. Poster entitled 'MED5 and CDK8 play a role in lignin-modification-induced dwarfing in Arabidopsis'

American Society of Plant Biologists (ASPB), Austin, TX, 07/2016. Poster entitled 'Investigating the interaction between MED5 and CDK8 in Arabidopsis'

Midwest Chromatin Meeting, Grand Rapids, MI, 06/2016. Poster entitled 'Investigating the interaction between MED5 and CDK8 in Arabidopsis'

Publications

X Mao, JI Kim, MT Wheeler, AK Heintzelman, VM Weake, C Chapple. (2019). Arabidopsis Mediator subunit CDK8 counteracts the stunted growth and hyperaccumulation of SA in plants carrying a mutant allele of MED5. *New Phytologist.* In press.

F Muro-Villanueva, **X Mao**, C Chapple. (2019). Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Current Opinion in Biotechnology*. 56: 202-208.

XY Mao, Q Bian, ZY Shen. (2012). Icariin exerts an osteogenetic effect in regular mice in vitro via MAPK signaling pathway. *Journal of Chinese Integrative Medicine,* 10: 1272-1278.

Manuscripts submitted / in Preparation

X Mao, VM Weake, C Chapple. (2019). The function of Mediator revealed by large-scale biology: through the lens of plant metabolism. *Journal of Experimental Botany*. Submitted.

X Mao, ND Bonawitz, MT Wheeler, VM Weake, C Chapple. (2019). Arabidopsis MED5 is involved in the regulation of shade avoidance syndrome and ABA homeostasis under non-stressed conditions. *Plant Physiology*. In preparation.

Teaching Activities

Graduate teaching assistant, Department of Biochemistry, Purdue University.

Introduction to R programming, Dr. Pete Pascuzzi, 05/2017-07/2018

Graduate mentor, Department of Biochemistry, Purdue University for:

       Gilbert Kayanja (Biochemistry rotation student, 01/2017-03/2017)

       Anne Heintzelman (REU undergraduate, 05/2017-07/2017)

       Mitchell Wheeler (Chemistry undergraduate, 08/2017-12/2018)

       Lizhi (Bobby) Chen (Biochemistry rotation student, 01/2019-02/2019)

<u>Service and Leadership</u>

Social coordinator, Grad Women in College of Agriculture, Purdue University

06/2015-present

Volunteer, Science in Schools, Purdue University

10/2014-present

Judge, Lafayette Regional Science and Engineering Fair

03/2014-present