

The role of the GIR1-AMTL1 interaction in sepal cell fate and glucosinolate accumulation

by

Lauren Apprill

B.S., Kansas State University, 2025

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Biochemistry and Molecular Biophysics

College of Arts and Sciences

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2026

Approved by:

Major Professor

Kathrin Schrick

Abstract

HD-Zip IV transcription factors are the key regulators of epidermal development in *Arabidopsis*. As the epidermis serves as the first line of defense against the environment, its proper development is vital for plant survival. One HD-Zip IV member, *Arabidopsis thaliana* MERISTEM LAYER1 (ATML1), is a known regulator of giant cell formation in sepals, a modified leaf structure protecting the developing reproductive tissues within the flower bud. GLABRA2 INTERACTING REPRESSOR1 (GIR1) is a small protein that was previously reported to play a role as a transcriptional coregulator. In this work, it is shown that GIR1 physically interacts with ATML1. Since the function of GIR1 in transcription is poorly understood, we sought to elucidate the mechanism and downstream developmental impacts of this interaction.

Protein-protein interaction analyses revealed that the GIR1-ATML1 interaction occurs through the predicted zinc finger of GIR1 and START adjacent domain (STAD) of ATML1. Additionally, GIR1 displayed interactions with TOPLESS (TPL) and TOPLESS-RELATED (TPR) proteins, which are corepressors known to recruit histone deacetylase chromatin remodelers. The complex between ATML1, GIR1, and TPL/TPR proteins was detected using co-immunoprecipitation, indicating that GIR1 acts as an adaptor protein to bridge the interactions. The *gir1* loss-of-function mutant was observed to have more giant cells in sepals, a phenotype similar to *ATML1* overexpressing lines, suggesting that GIR1 functions as a repressor to ATML1. RNA sequencing revealed several genes upregulated in *gir1* sepals compared to wild type that are involved in the glucosinolate (GSL) biosynthesis pathway. GSLs are sulfur- and nitrogen-containing secondary metabolites that serve as biotic defense compounds in the Brassicales. Chemical analysis along with mass spectrometry imaging confirmed GSL

accumulation in *gir1* sepals. Thus, GIR1 acts as a repressor to ATML1 that functions to negatively regulate giant cell formation and GSL biosynthesis in sepals.

The interaction between GIR1 and ATML1 was shown in both plant and yeast systems but less is known about the minimal requirements or biological nature of the interaction. Identification of an acidic patch in ATML1 STAD and yeast bait autoactivation assays suggest STAD's role as a transcriptional activation domain. Mutational analysis revealed residues in the acidic patch critical for the transactivation activity. In order to detect the direct interaction between STAD and GIR1, we employed a biophysical method known as surface plasmon resonance (SPR). It was confirmed that STAD can directly interact with GIR1 with a low micromolar affinity. Moreover, no effective K_D could be calculated when the interaction occurred with a *gir1* zinc finger mutant protein. Interaction of wild-type GIR1 with a STAD acidic patch mutant resulted in a two-fold decrease in affinity.

Together, these studies define a novel transcriptional repression complex involved in regulating giant cell formation and GSL accumulation in sepals. Given the agricultural importance of this compound, this work can be leveraged for engineering crops with tailored cell-specific GSL profiles.

Table of Contents

List of Figures	vii
List of Tables	ix
List of Abbreviations	x
Acknowledgements	xiii
Chapter 1 - Background and Introduction	1
Class IV homeodomain leucine-zipper (HD-Zip IV) transcription factors	1
Gene repression mechanisms	3
Adaptor protein family	5
Research objectives and questions	7
Thesis Outline	8
References	10
Chapter 2 - ATML1-GIR1-TPL transcriptional repression module controls glucosinolates and giant cells in <i>Arabidopsis thaliana</i> sepals	13
Abstract	13
Introduction	14
Results	16
The gir1 mutants display excess giant cells on sepals	16
Giant cell phenotypes reveal genetic interaction between GIR1 and ATML1	17
GIR1 complements the excess giant cell phenotype of gir1 mutants	17
Yeast 2-Hybrid and in planta assays demonstrate GIR1 interaction with ATML1 STAD... ..	19
ATML1, GIR1, and TPL/TPR proteins form a ternary complex	20
GSL pathway genes are upregulated in sepals from gir1 mutants	21
GSL levels are elevated in sepals from gir1 mutants	23
Discussion	24
Materials and Methods	27

Plant growth conditions and transformation	27
Construction of binary vectors and other plasmids	28
Y2H assays.....	29
Transient protein expression and coimmunoprecipitation	30
Phenotypic characterization and microscopy.....	31
RNA extraction and transcriptomics analysis	31
Glucosinolate extraction and analysis	32
Mass Spectrometry Imaging	32
Figures	34
References.....	50
Chapter 3 - Characterization of the direct interaction between the transactivation domain ATML1 STAD and coregulator GIR1	56
Abstract.....	56
Introduction.....	57
Results.....	59
ATML1 STAD contains an acidic patch important for transcriptional activation.	59
ATML1 STAD interacts with GIR1 directly with a low micromolar affinity.....	59
Discussion.....	60
Materials and Methods.....	62
Autoactivation yeast two-hybrid assay	62
Recombinant protein production.....	62
Trace mineral panel of aqueous samples by ICP-MS	63
Surface Plasmon Resonance	64
Figures and Tables	65
References.....	71
Chapter 4 - Conclusions and future directions.....	72
Mechanisms of ATML1-mediated gene expression regulation	72

Cell- and tissue-specific GSL biosynthesis regulation through adaptor proteins	73
Structural effects of GIR1 zinc finger	75
References	76
Appendix A - Supplemental Figures and Tables	77
References	96

List of Figures

Chapter 2:

Figure 1. <i>gir1</i> mutant displays a giant cell phenotype in sepals.	34
Figure 2. The EAR motif and cysteine residues of the Zn finger are required for rescue of the giant cell phenotype by the <i>GIR1</i> native promoter construct.....	36
Figure 3. GIR1 interacts with HD-Zip IV transcription factor ATML1 via its Zn finger motif. .	38
Figure 4. GIR1 interacts with TPL and TPRs via its EAR motif.	40
Figure 5. Differential expression analysis suggests that GIR1 acts as a repressor of ATML1 to control GSL biosynthesis genes in sepals.	43
Figure 6. GSL levels are elevated in sepals from <i>gir1</i> mutants.	47
Figure 7. Model for biological function of ATML1-GIR1-TPL/TPR complex in repression of GSL biosynthesis and giant cell growth.....	49

Chapter 3:

Figure 1. ATML1 STAD functions as a transactivation domain.	65
Figure 2. Purification of recombinantly expressed wild-type and mutant GIR1 and ATML1 STAD.	66
Figure 3. SPR analysis confirmed direct in vitro interaction between ATML1 STAD and GIR1.....	69

Appendix:

Appendix Figure 1. The <i>gir1</i> and <i>gir1;gir2</i> mutants exhibit similarly enhanced giant cell formation.	78
Appendix Figure 2. Overexpression of <i>GIR1</i> results in rescue of giant cell phenotype of <i>gir1</i> mutants.	79
Appendix Figure 3. Western blot of GIR1 proteins expressed in Yeast 2-Hybrid assays.....	81

Appendix Figure 4. Heatmap from RNA-seq data showing three biological replicates for each genotype.	83
Appendix Figure 5. RNA-seq data for aliphatic glucosinolate biosynthesis genes. Related to Figure 5.	84
Appendix Figure 6. Graphs of RNA-seq data for indolic glucosinolate biosynthesis genes and other glucosinolate pathway genes. Related to Figure 5.....	86
Appendix Figure 7. MALDI-MS imaging of the sepal epidermis indicates high 4-MSOB glucosinolate levels in the <i>gir1</i> mutants.....	88
Appendix Figure 8. Floral-specific expression of TPL, TPRs, and GIR1 in Arabidopsis.	89

List of Tables

Chapter 3:

Table 1. Metals detected in purified protein samples from ICP-MS.....	68
Table 2. Steady-state binding affinities of GIR1 to ATML1 STAD and atml1 STAD ^{D589L}	70

Appendix:

Appendix Table 1. Summary of expression of glucosinolate biosynthesis genes in <i>gir1</i> sepals..	90
Appendix Table 2. Names and chemical properties of glucosinolates described in this study. ...	92
Appendix Table 3. Oligonucleotides used in this study.	93
Appendix Table 4. L1 box elements in GSL biosynthesis gene promoters.....	95

List of Abbreviations

ACL1	ABAXIALLY CURLED LEAF1
ATML1	<i>Arabidopsis thaliana</i> MERISTEM LAYER1
BZR1	BRASSINAZOLE-RESISTANT1
Co-IP	co-immunoprecipitation
Col	Columbia
DEGs	differentially expressed genes
DTT	dithiothreitol
EAR	ethylene-responsive element binding factor-associated amphiphilic repression
EMSA	electrophoretic mobility shift assay
FC	fold-change
FDR	false discovery rate
FIE	FERTILIZATION-INDEPENDENT ENDOSPERM
GIR1	GLABRA2 INTERACTING REPRESSOR1
GIR2	GLABRA2 INTERACTING REPRESSOR2
GL2	GLABRA2
GO	gene ontology
GSL	glucosinolate
HAT	histone acetyltransferase
HDAC	histone deacetylase
HD2	Histone Deacetylase 2
HD-Zip IV	class IV homeodomain leucine-zipper
HD	homeodomain
HPLC-DAD	high performance liquid chromatography-diode-array detection
iCLIP	individual-nucleotide resolution UV-crosslinking immunoprecipitation

ICP-MS	inductively coupled plasmon mass spectrometry
IP	immunoprecipitated
kDa	kilodalton
K _D	equilibrium dissociation constant
K-Hex	kaempferol hexoside
K-Rha	kaempferol rhamnoside
LER	leaf endoplasmic reticulum
MALDI-MS	matrix-assisted laser desorption/ionization mass spectrometry
NAD	nicotinamide adenine dinucleotide
NMR	nuclear magnetic resonance
PDF2	PROTODERMAL FACTOR2
PMSF	phenylmethanesulfonylfluoride
PRC2	polycomb repressive complex 2
PTM	post translational modification
PVPP	polyvinylpolypyrrolidone
Q-Rha	quercetin rhamnoside
RNA-seq	RNA-sequencing
ROC4	RICE OUTERMOST CELL-SPECIFIC GENE 4
RPD3/HDA1	Reduced Potassium Dependence 3/Histone Deacetylase 1
RU	resonance units
SD	standard deviation
SEM	scanning electron microscope
SIED1	salt-induced EIN3/EIL1-dependent 1
SIR2	Silent Information Regulator 2
SlCycB2	<i>Solanum lycopersicum</i> B-TYPE CYCLIN
SPR	surface plasmon resonance
SSIMS	static secondary ion mass spectrometry

STAD	START adjacent domain
START	STeriodogenic Acute Regulatory protein-related lipid Transfer
TF	transcription factor
TPL	TOPLESS
TPR	TOPLESS-RELATED
UTRs	untranslated regions
WB	western blot
WT	wild type
Y2H	yeast two-hybrid
YPDA	yeast peptone dextrose adenine

Acknowledgements

I'd like to thank my mentor Dr. Kathrin Schrick for her support from the early years of my undergraduate degree through the completion of this thesis. I am grateful for her guidance as she taught me experiments, offered improvements to my writing, and advised me through the beginnings of my scientific career. I'd also like to thank the members of my committee, Dr. Timothy Durrett and Dr. Brandon Garcia, for their feedback during committee meetings and my final exam. I'd lastly like to extend my gratitude to the current and former members of Dr. Schrick's lab for creating a fun and supportive environment to achieve our research goals.

Chapter 1 - Background and Introduction

Class IV homeodomain leucine-zipper (HD-Zip IV) transcription factors

Eukaryotes contain diverse cell types based on distinct gene expression and epigenetic patterns to accommodate an organisms' complexity in structure and function¹. As such, cell-type differentiation is achieved with the successful up- and down-regulation of genes. Transcription factors (TFs) play a major role in this process due to their ability to recruit RNA polymerase to initiate transcription or regulators that can modify chromatin architecture and alter accessibility to a promoter². One universal characteristic of TF structure is a DNA binding domain, which recognizes cis-acting DNA elements upstream of target genes. TFs may also have other unique domains that serve additional purposes, such as domains that facilitate protein interactions with coregulators². These nuances in TF structure or differences in corresponding target genes has led to an immense diversity of TFs, making up around 5% of the proteome in *Arabidopsis* and 9% in humans^{3,4}. TFs are classified into families based on similarity of amino acid sequence, structure, or DNA binding activity.

Homeodomain (HD) TFs comprise a large well-conserved family found across all metazoa⁵. First discovered in *Drosophila*, each HD TF gene contains a characteristic homeobox, a 180 bp element that encodes a helix-turn-helix structure that functions as the HD DNA binding domain⁵⁻⁷. At a high level, HD TFs play important roles in developmental processes, with mutations in HD TFs resulting in ectopic expression and abnormal development⁵. The sub classes of HD TFs in plants include PHD finger, ZF-HD, WOX, BELL, KNOX, DDT, NDX, LD, PINTOX, SAWADEE, and HD-Zip⁸⁻¹⁰. TFs are classified in these families based on several parameters including sequence similarity, protein size, conserved motifs or domains, and intron

architecture¹⁰. Specifically, the HD-Zip proteins are further classified into subfamilies I-IV¹¹. Per their namesake, these proteins contain a leucine zipper, a structural motif defined by an alpha-helix with a leucine every seventh position residing on the same side of the helix¹¹. These ordered residues expose a hydrophobic region, which encourages dimerization¹⁰. The leucine zipper domain is located immediately downstream of the HD, an orientation uniquely found in plants¹⁰.

Class IV HD-Zip TFs comprise 16 members in *Arabidopsis* whose founding member was GLABRA2 (GL2), for which a loss of function mutant analysis revealed root hair and trichome defects^{12,13}. It has been shown that class IV members bind to L1 box sequences in the promoter, enabling activation or repression of target genes¹⁴. Apart from the HD and leucine zipper domain, class IV TFs contain a STeroidogenic Acute Regulatory protein-related lipid Transfer (START) domain and the START adjacent domain (STAD), which is also referred to as SAD¹⁵. While START has been characterized to serve as a lipid binding domain in plants and animals, little is known about STAD's function¹⁰.

One member of HD-Zip IV TFs, *Arabidopsis thaliana* MERISTEM LAYER1 (ATML1), initially described in 1996, was reported to be expressed exclusively in the shoot¹⁶. Subsequent studies showed ATML1's importance in regulating cell-type differentiation in the epidermis during embryogenesis and later throughout developmental stages¹⁷. The epidermis has many important functions including acting as the first line of defense against environmental stressors, being the outermost cell layer of plants, and playing key roles in regulating growth patterns^{17,18}. Thus, understanding the regulation of epidermal development is critical for proper plant development. Epidermal cells can differentiate into various cell types that exhibit specialized functions such as trichomes which aid in the prevention of insect attack or guard cells that

contribute to proper gas exchange and water upkeep^{17,18}. Intriguingly, upon an ectopic expression analysis of *ATML1*, guard cell and trichome formation was seen in inner tissue layers instead of on the outer epidermis¹⁸. Moreover, it was reported that *ATML1* acts in a dose-dependent manner to regulate giant cell formation in sepals, where the *atml1* T-DNA insertion mutant displayed fewer giant cells and overexpressing *ATML1* exhibited more¹⁹. Further, the double mutant between *ATML1* and *PROTODERMAL FACTOR2* (*PDF2*), which is functionally redundant with *ATML1*, revealed an embryo-lethal phenotype²⁰.

Gene repression mechanisms

Chemical and structural changes to genetic material without altering the DNA sequence are epigenetic mechanisms that affect gene expression²¹. For example, DNA methylation and chromatin structure modifications, can either positively or negatively modulate gene expression²¹. Due to the constant modifications of chromatin to create accessibility or reduced access for transcription machinery, chromatin is described as a dynamic structure²². DNA methylation, which typically results in decreased gene expression, can occur on cytosines through various DNA methyltransferases²³. Conversely, gene expression can be altered when histones are post-translationally modified by histone modifying enzymes affecting the compactness of the nucleosomes, consequently altering the accessibility of the transcription machinery or transcriptional activators^{21,24}. All post translational modifications (PTM) are reversible and can include acetylation, methylation, phosphorylation, ubiquitination, and sumoylation²¹. When a chromatin structure modification is initiated, complexes are commonly formed between a TF, co-regulator, and a chromatin remodeling protein²¹.

One PTM, acetylation on histone tail lysine residues by histone acetyltransferases (HAT), is correlated with activation of gene expression whereas deacetylation, by histone deacetylases (HDAC), condense chromatin and results in gene repression²². The chemical nature of this PTM is reflected by a nucleosome's positively charged histone attracted to the negatively charged DNA backbone²⁵. Thus, acetylation on a lysine neutralizes the charge and results in a looser nucleosome and the region becoming more transcriptionally active. Deacetylation reverses this phenomenon and reintroduces the positive charge, restoring the ionic association between histones and DNA and consequently repressing gene expression in that region²⁵. There are three primary classes of HDACs: Reduced Potassium Dependence 3/Histone Deacetylase 1 (RPD3/HDA1) which is highly conserved across the eukaryotes and has been widely studied, Silent Information Regulator 2 (SIR2), and Histone Deacetylase 2 (HD2)²⁶. Each class requires different cofactors that aid in catalytic activity including nicotinamide adenine dinucleotide (NAD) for SIR2 and zinc for RPD3/HDA1 and HD2²⁴. Often, deacetylases and demethylases work in tandem, sometimes through direct interaction, to regulate wide-spread cellular processes²⁴. Timely and accurate HDAC activity is vital, where improper function leads to misexpression of genes and can eventually cause different epigenetic diseases²⁵. HDACs have been a target for competitive inhibitors in therapeutics, such as the drugs SAHA (Vorinostat, Zolinza™, Merck & Co, Inc., USA) and FK228 (Romidepsin, Istodax™, Celgene Corp., USA) for treating cutaneous T-cell lymphoma²⁵. HDACs are effective at remodeling chromatin structure to modulate gene expression. However, due to their inability to bind DNA, they rely on protein interactions to target specific regions.

Short sequence motifs in proteins can sometimes function to regulate gene expression by facilitating the formation of protein complexes. The ethylene-responsive element binding factor-

associated amphiphilic repression (EAR) motif is one of the main transcriptional repression motifs that is well conserved across the plant kingdom²¹. EAR motifs were first discovered in a TFIIIA-type zinc finger protein and are characterized by the sequences LxLxL, DLN_xP, or DLN_{xx}P with x denoting any amino acid^{21,27}. Proteins containing EAR motifs have been shown to interact with a wide range of co-repressor proteins. For example, many transcription factors with EAR motifs can recruit polycomb repressive complex 2 (PRC2)²⁷. Although PRC proteins lack any specific DNA sequence targeting ability on their own, they can be brought to specific sites by the protein interactions to mediate chromatin condensation via its tri-methyltransferase activity on lysine 27 of histone H3^{27,28}. In one study, it was discovered that floral development is properly regulated by a repressive protein complex containing a PRC protein, FERTILIZATION-INDEPENDENT ENDOSPERM (FIE), the co-repressor KNUCKLES containing an EAR motif and zinc finger, and WUSCHEL, a HD TF²⁹. EAR motifs are also widely associated with mediating protein-protein interactions with the co-repressor proteins TOPLESS (TPL), TOPLESS-RELATED PROTEIN (TPR), and SAP18³⁰. TPL works by recruiting an HDAC to remodel the chromatin to a more condensed structure and less transcriptionally active³¹. In one example, a transcriptional repression complex was identified between an EAR motif-containing transcription factor related to brassinosteroid response, BRASSINAZOLE-RESISTANT1 (BZR1), TPL, and HDA19³². Thus, EAR motifs can negatively regulate target gene expression by indirectly affecting HDAC and PRC2 activity.

Adaptor protein family

A small group of three related proteins in *Arabidopsis* make up a family previously described as adaptor proteins^{33,34}. GL2-interacting repressor1 (GIR1) and GL2-interacting repressor2 (GIR2) are small, functionally redundant, and plant-specific proteins that share 65%

amino acid identity and 71% similarity³³. They were first discovered by their interaction with an HD-Zip IV TF, GL2³³. Phenotypic analysis of *gir1* and *gir2* T-DNA insertion mutants exhibited no obvious defects, however, the *gir1;gir2* double mutant displayed more root hairs similar to the *gl2* mutant phenotype³³. Expression profiling revealed that *GIR1* mRNA levels are high in the root meristem, which resembles *GL2* root expression patterns³³. Studies in commercially relevant crops have revealed important functions for their respective GIR1 homologs. In cotton, its GIR1 homolog, *GaFzl*, was found to regulate cotton fiber formation, which is analogous to trichome formation³⁵. The rice homolog, ABAXIALLY CURLED LEAF1 (*ACL1*) can interact with HD-Zip IV members ROC2, ROC5, and ROC7 to negatively regulate drought tolerance and pest resistance by increasing formation of bulliform cells, the specialized epidermal cells that contribute to water transpiration, and cuticular wax production³⁶. Lastly, the tomato GIR1 homolog, *Solanum lycopersicum* B-TYPE CYCLIN (*SlCycB2*), was observed to interact with the HD-Zip IV TF WOOLLY to positively regulate the lesser branched type I trichome formation^{37,38}. Since *SlCycB2* is thought to work at the G2/M transition during mitosis to help induce cell division, it was speculated the switch from endoreduplication to mitotic cell division in trichomes is driven by *SlCycB2*³⁸. Due to phenotypic and expression data primarily focused on root and trichome tissues, wider functions of GIR1 and GIR2 and the molecular mechanisms underlying their interaction with HD-Zip IV proteins remain unknown.

GIR1 and GIR2's namesake as adaptor proteins came from one study where GIR1 was found to interact with TPL and GL2, essentially linking the three proteins in a repression complex³⁴. It was reported that the interaction with TPL is facilitated through a putative EAR motif with the sequence LxLxL, and deletion of the EAR motif abolishes the interaction³⁴. Another important structural motif of the adaptor proteins is two CPxC sequences forming a

predicted zinc finger³⁹. Zinc fingers are commonly associated with aiding in protein-protein interactions, especially with DNA-binding proteins³⁹. Further, it was discovered the carboxy-terminus of the adaptor proteins is required for interaction with the HD-Zip IV TFs, which also fits the hypothesis that plant proteins containing zinc fingers are associated with epidermal development^{39,40}. Another study described a third member of the adaptor protein family, an evolutionarily closely related protein with high sequence similarity: salt-induced EIN3/EIL1-dependent 1 (SIED1)^{39,41}. Mutant analysis revealed the null *sied1* mutant displayed hypersensitivity and low survival rates under reactive oxygen species stress whereas overexpressing SIED1 displayed increased salt resistance and a leaf rolling phenotype^{39,41}. The wider functions of SIED1 require further investigation.

Research objectives and questions

Although the HD, leucine zipper, and START domains of HD-Zip IV TFs have been well characterized, the STAD domain and its function remain largely unstudied. Likewise, the molecular mechanisms by which ATML1 controls target gene expression are unclear. Prior work showing GIR1 interacting with GL2 to regulate trichome and root hair development³³ raises the question whether GIR1 can directly interact with other HD-Zip IV members in *Arabidopsis*, particularly given evidence from other crop species that HD-Zip IV proteins work in conjunction with GIR1 homologs to regulate epidermal cell-type differentiation³⁵⁻³⁸. Further, could GIR1 and HD-Zip IV-mediated regulation contribute more broadly to epidermal patterning in *Arabidopsis* than previously documented and what are the mechanisms used to achieve proper regulation?

The fundamental objective of this thesis is to elucidate the mechanism and developmental impact of gene expression regulation through ATML1. Specifically, I sought to discover how

ATML1 activity is restricted through its protein-protein interactions and the consequential downstream developmental effects. The key hypothesis is that ATML1 interacts with GIR1 which acts as a repressive adaptor protein, due to the presence of an EAR motif, that forms a larger transcription repression complex regulating cell-type differentiation in epidermal tissues.

Thesis Outline

The first chapter provides an introduction and background on HD TFs, specifically the HD-Zip IV family, and their importance in epidermal development in plants. It also describes transcriptional repression mechanisms of epigenetic pattern changes involving co-repressors. Lastly, background of a small adaptor protein family highlighting the structural motifs previously implicated for their protein-protein interactions with TFs and their relevance in economically important crops is discussed.

The second chapter focuses on uncovering the molecular mechanism of GIR1-ATML1 regulation and its downstream effects. First, a phenotypic analysis of giant cells in *gir1-1* mutant sepals was performed along with a rescue experiment involving *GIR1* expression with its native promoter in a *gir1-1* mutant background. Next, a series of experiments mapped the interacting regions of ATML1, GIR1, and TPL/TPRs through yeast two-hybrid and Co-IP. To confirm GIR1's repressive nature with its interaction with ATML1, an RNA sequencing experiment was executed on sepal tissues. A GO enrichment analysis prompted a quantitative chemical analysis comparing glucosinolate levels in wild type, *gir1-1*, and *atml1-3*. Lastly, a model is proposed detailing the effects of an ATML1-GIR1-TPL/TPR transcriptional repression complex.

The third chapter further investigates the interaction between GIR1 and ATML1 by elucidating the necessary components of the interaction. Yeast autoactivation assays along identification of an acidic patch in ATML1 STAD suggested it functions as a transcriptional activation domain. Recombinant GIR1 and STAD were expressed, purified, and subject to surface plasmon resonance to detect the physical interaction in vitro. The affinity of the interaction was calculated with a steady-state analysis.

Finally, the fourth chapter outlines main conclusions of this thesis drawing attention to the relevant applications, specifically the potential to contribute to modern sustainable agricultural applications. Chapter four also includes remarks concerning the work's novelty and suggests future directions to expand on these findings.

References

1. Chan, M.M., Smith, Z.D., Grosswendt, S., Kretzmer, H., Norman, T.M., Adamson, B., Jost, M., Quinn, J.J., Yang, D., Jones, M.G., et al. (2019). Molecular recording of mammalian embryogenesis. *Nature* 570, 77–82. 10.1038/s41586-019-1184-5.
2. Strader, L., Weijers, D., and Wagner, D. (2022). Plant transcription factors - being in the right place with the right company. *Curr Opin Plant Biol* 65, 102136. 10.1016/j.pbi.2021.102136.
3. Goos, H., Kinnunen, M., Salokas, K., Tan, Z., Liu, X., Yadav, L., Zhang, Q., Wei, G.H., and Varjosalo, M. (2022). Human transcription factor protein interaction networks. *Nat Commun* 13, 766. 10.1038/s41467-022-28341-5.
4. Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., et al. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. 10.1126/science.290.5499.2105.
5. Gehring, W.J., Affolter, M., and Burglin, T. (1994). Homeodomain proteins. *Annu Rev Biochem* 63, 487–526. 10.1146/annurev.bi.63.070194.002415.
6. Carrasco, A.E., McGinnis, W., Gehring, W.J., and De Robertis, E.M. (1984). Cloning of an *X. laevis* gene expressed during early embryogenesis coding for a peptide region homologous to *Drosophila* homeotic genes. *Cell* 37, 409–414. 10.1016/0092-8674(84)90371-4.
7. Gehring, W.J., Muller, M., Affolter, M., Percival-Smith, A., Billeter, M., Qian, Y.Q., Otting, G., and Wuthrich, K. (1990). The structure of the homeodomain and its functional implications. *Trends Genet.* 6, 323–329.
8. Vollbrecht, E., Veit, B., Sinha, N., and Hake, S. (1991). The developmental gene Knotted-1 is a member of a maize homeobox gene family. *Nature* 350, 241–243. 10.1038/350241a0.
9. Schena, M., and Davis, R.W. (1992). HD-Zip proteins: members of an Arabidopsis homeodomain protein superfamily. *Proc Natl Acad Sci U S A* 89, 3894–3898. 10.1073/pnas.89.9.3894.
10. Capella, M., P.A. Ribone, A.L. Arce, and R.L. Chan (2016). Homeodomain-leucine zipper transcription factors: structural features of these proteins, unique to plants, in *Plant transcription factors*. Elsevier, 113–126.
11. Elhiti, M., and Stasolla, C. (2009). Structure and function of homodomain-leucine zipper (HD-Zip) proteins. *Plant Signal Behav* 4, 86–88. 10.4161/psb.4.2.7692.
12. Di Cristina, M., Sessa, G., Dolan, L., Linstead, P., Baima, S., Ruberti, I., and Morelli, G. (1996). The Arabidopsis *Athb-10* (GLABRA2) is an HD-Zip protein required for regulation of root hair development. *Plant J.* 10, 393–402.
13. Rerie, W.G., Feldmann, K.A., and Marks, M.D. (1994). The GLABRA2 gene encodes a homeo domain protein required for normal trichome development in Arabidopsis. *Genes Dev.* 8, 1388–1399.
14. Abe, M., Takahashi, T., and Komeda, Y. (2001). Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J.* 26, 487–494.
15. Ponting, C.P., and Aravind, L. (1999). START: a lipid-binding domain in StAR, HD-ZIP and signalling proteins. *Trends Biochem.Sci.* 24, 130–132.

16. Lu, P., Porat, R., Nadeau, J.A., and O'Neill, S.D. (1996). Identification of a meristem L1 layer-specific gene in *Arabidopsis* that is expressed during embryonic pattern formation and defines a new class of homeobox genes. *The Plant cell* 8, 2155–2168.
17. Iida, H., and Takada, S. (2021). A Quarter Century History of ATML1 Gene Research. *Plants (Basel)* 10. 10.3390/plants10020290.
18. Schrick, K., Ahmad, B., and Nguyen, H.V. (2023). HD-Zip IV transcription factors: Drivers of epidermal cell fate integrate metabolic signals. *Curr Opin Plant Biol* 75, 102417. 10.1016/j.pbi.2023.102417.
19. Meyer, H.M., Teles, J., Formosa-Jordan, P., Refahi, Y., San-Bento, R., Ingram, G., Jonsson, H., Locke, J.C., and Roeder, A.H. (2017). Fluctuations of the transcription factor ATML1 generate the pattern of giant cells in the *Arabidopsis* sepal. *Elife* 6, e191131. 10.7554/eLife.19131.
20. Ogawa, E., Yamada, Y., Sezaki, N., Kosaka, S., Kondo, H., Kamata, N., Abe, M., Komeda, Y., and Takahashi, T. (2015). ATML1 and PDF2 Play a Redundant and Essential Role in *Arabidopsis* Embryo Development. *Plant Cell Physiol* 56, 1183–1192. 10.1093/pcp/pcv045.
21. Kagale, S., and Rozwadowski, K. (2011). EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. *Epigenetics* 6, 141–146. 10.4161/epi.6.2.13627.
22. Bowen, A.J., Gonzalez, D., Mullins, J.G., Bhatt, A.M., Martinez, A., and Conlan, R.S. (2010). PAH-domain-specific interactions of the *Arabidopsis* transcription coregulator SIN3-LIKE1 (SNL1) with telomere-binding protein 1 and ALWAYS EARLY2 Myb-DNA binding factors. *J Mol Biol* 395, 937–949. 10.1016/j.jmb.2009.11.065.
23. Feng, S., Jacobsen, S.E., and Reik, W. (2010). Epigenetic reprogramming in plant and animal development. *Science* 330, 622–627. 10.1126/science.1190614.
24. Liu, X., Yang, S., Zhao, M., Luo, M., Yu, C.W., Chen, C.Y., Tai, R., and Wu, K. (2014). Transcriptional repression by histone deacetylases in plants. *Mol Plant* 7, 764–772. 10.1093/mp/ssu033.
25. Park, S.Y., and Kim, J.S. (2020). A short guide to histone deacetylases including recent progress on class II enzymes. *Exp Mol Med* 52, 204–212. 10.1038/s12276-020-0382-4.
26. Hollender, C., and Liu, Z. (2008). Histone deacetylase genes in *Arabidopsis* development. *J Integr Plant Biol* 50, 875–885. 10.1111/j.1744-7909.2008.00704.x.
27. Baile, F., Merini, W., Hidalgo, I., and Calonje, M. (2021). EAR domain-containing transcription factors trigger PRC2-mediated chromatin marking in *Arabidopsis*. *Plant Cell* 33, 2701–2715. 10.1093/plcell/koab139.
28. Zhou, Y., Romero-Campero, F.J., Gomez-Zambrano, A., Turck, F., and Calonje, M. (2017). H2A monoubiquitination in *Arabidopsis thaliana* is generally independent of LHP1 and PRC2 activity. *Genome Biol* 18, 69. 10.1186/s13059-017-1197-z.
29. Sun, B., Zhou, Y., Cai, J., Shang, E., Yamaguchi, N., Xiao, J., Looi, L.S., Wee, W.Y., Gao, X., Wagner, D., and Ito, T. (2019). Integration of Transcriptional Repression and Polycomb-Mediated Silencing of WUSCHEL in Floral Meristems. *Plant Cell* 31, 1488–1505. 10.1105/tpc.18.00450.
30. Lee, M.S., An, J.H., and Cho, H.T. (2016). Biological and molecular functions of two EAR motifs of *Arabidopsis* IAA7. *J Plant Biol* 59, 24–32. 10.1007/s12374-016-0453-1.
31. Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science* 312, 1520–1523. 10.1126/science.1123841.

32. Kim, H., Shim, D., Moon, S., Lee, J., Bae, W., Choi, H., Kim, K., and Ryu, H. (2019). Transcriptional network regulation of the brassinosteroid signaling pathway by the BES1-TPL-HDA19 co-repressor complex. *Planta* *250*, 1371–1377. 10.1007/s00425-019-03233-Z.
33. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. I. Interaction with the repressor GLABRA2 and regulation of root hair development. *Biochem Biophys Res Commun* *488*, 547–553. 10.1016/j.bbrc.2017.05.084.
34. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. II. Interaction with the co-repressor TOPLESS and promotion of histone deacetylation of target chromatin. *Biochem Biophys Res Commun* *488*, 609–613. 10.1016/j.bbrc.2017.05.085.
35. Feng, X., Cheng, H., Zuo, D., Zhang, Y., Wang, Q., Liu, K., Ashraf, J., Yang, Q., Li, S., Chen, X., and Song, G. (2019). Fine mapping and identification of the fuzzless gene GaFzl in DPL972 (*Gossypium arboreum*). *Theor Appl Genet* *132*, 2169–2179. 10.1007/s00122-019-03330-3.
36. Tao, Z., Zhu, L., Li, H., Sun, B., Liu, X., Li, D., Hu, W., Wang, S., Miao, X., and Shi, Z. (2024). ACL1-ROC4/5 complex reveals a common mechanism in rice response to brown planthopper infestation and drought. *Nat Commun* *15*, 8107. 10.1038/s41467-024-52436-w.
37. Yang, C., Gao, Y., Gao, S., Yu, G., Xiong, C., Chang, J., Li, H., and Ye, Z. (2015). Transcriptome profile analysis of cell proliferation molecular processes during multicellular trichome formation induced by tomato Wov gene in tobacco. *BMC Genomics* *16*, 868. 10.1186/s12864-015-2099-7.
38. Yang, C., Li, H., Zhang, J., Luo, Z., Gong, P., Zhang, C., Li, J., Wang, T., Zhang, Y., Lu, Y., and Ye, Z. (2011). A regulatory gene induces trichome formation and embryo lethality in tomato. *Proc Natl Acad Sci U S A* *108*, 11836–11841. 10.1073/pnas.1100532108.
39. Zhang, Q., Hong, T., Su, T., Li, Q., Xiang, A., Chen, X., He, G., Zhao, F., and Wu, R. (2025). Genome-wide identification and functional characterization of CLG family genes reveal likely roles in epidermal development in Arabidopsis. *Plant Cell Rep* *44*, 32. 10.1007/s00299-024-03421-6.
40. Han, G., Li, Y., Qiao, Z., Wang, C., Zhao, Y., Guo, J., Chen, M., and Wang, B. (2021). Advances in the Regulation of Epidermal Cell Development by C2H2 Zinc Finger Proteins in Plants. *Front Plant Sci* *12*, 754512. 10.3389/fpls.2021.754512.
41. Peng, J., Li, Z., Wen, X., Li, W., Shi, H., Yang, L., Zhu, H., and Guo, H. (2014). Salt-induced stabilization of EIN3/EIL1 confers salinity tolerance by deterring ROS accumulation in Arabidopsis. *PLoS Genet* *10*, e1004664. 10.1371/journal.pgen.1004664.

Chapter 2 - ATML1-GIR1-TPL transcriptional repression module controls glucosinolates and giant cells in *Arabidopsis thaliana* sepals

Abstract

Glucosinolates (GSLs) are sulfur- and nitrogen-containing secondary metabolites that serve as defense compounds in *Arabidopsis* and other members of the Brassicales. Although the enzymatic pathway that produces GSLs is well-studied, the upstream mechanisms that control their tissue-specific synthesis are poorly understood. We identified a novel repression module that transcriptionally regulates GSL levels in sepals, the modified leaves that protect the reproductive tissues within the floral bud. GLABRA2 (GL2) INTERACTING REPRESSOR1 (GIR1) interacts with *Arabidopsis thaliana* MERISTEM LAYER1 (ATML1), an HD-Zip IV transcription factor known to be required for giant cell formation in the sepal epidermis. This interaction requires the predicted zinc finger of GIR1, and the C-terminal START adjacent domain (STAD) of ATML1. The *gir1* loss-of-function mutants exhibit excess giant cells, in contrast to *atml1* mutants which lack giant cells, supporting the role of GIR1 as a negative regulator of ATML1. We confirmed that GIR1 interacts with TOPLESS (TPL) and TOPLESS-RELATED (TPR) corepressors that recruit chromatin remodelers, and coimmunoprecipitation demonstrated that GIR1 acts as an adaptor protein between ATML1 and TPL/TPR. RNA sequencing revealed that numerous genes involved in GSL biosynthesis, including the key transcriptional regulator MYB29, are upregulated in *gir1* mutants. Consistent with the transcriptomic data, chemical analysis revealed that *gir1* mutants display elevated GSL levels in sepals. Mass spectrometry imaging confirmed high GSL accumulation in *gir1* sepals compared to wild type and *atml1*. Overall, our findings uncover a previously unrecognized link between

cell expansion and GSL metabolism and suggest strategies for engineering plants with cell-type specific GSL profiles.

Introduction

GSLs are secondary metabolites that naturally occur in the plant model *Arabidopsis thaliana* and other members of the Brassicales including agriculturally important oilseed crops such as canola, as well as cruciferous vegetables including broccoli, kale, and cabbage¹. On the plant side, GSLs serve as important defense compounds against herbivores and pathogens^{2,3}. Relevant to the human diet, GSLs contribute to the pungent flavors and health-promoting properties of cruciferous vegetables^{4,5}. Epidemiological studies have identified GSL byproducts, such as isothiocyanates, as bioactive derivatives possessing anti-inflammatory, antioxidant and anticarcinogenic properties⁶⁻⁹. The hydrolysis of GSLs via myrosinase occurs when plant cells are damaged^{10,11}, or through catabolic breakdown by gut microbiota^{12,13}.

In *Arabidopsis*, the GSL biosynthesis pathway begins with amino acid chain elongation from methionine for aliphatic GSLs or from tryptophan for indolic GSLs. A series of enzyme reactions convert the amino acid derivatives to form the core GSL scaffold. Structural diversity arises from side-chain modifications catalyzed by various hydroxylases, methyltransferases, and oxygenases^{14,15}. Overall pathway architecture is conserved across the Brassicales with lineage-specific expansions and neo-functionalizations determining biological activities of specific GSLs^{5,16,17}. Transcriptional control of the pathway is primarily governed by specific R2R3-MYB transcription factors. MYB28, MYB29, and MYB76 regulate genes of the aliphatic GSL pathway, while MYB34, MYB51, and MYB122 regulate the indolic pathway¹⁸⁻²⁰. Environmental

and developmental cues are integrated in response to jasmonic acid, and interactions with bHLH and MYC transcription factors modulate the pathway under stress²¹.

While the GSL biosynthesis pathway and its transcriptional control are extensively studied⁴, the epidermal cell-type regulation and upstream molecular mechanisms that negatively control their synthesis are poorly understood. GSL biosynthesis is thought to mainly occur in the vasculature²², and GTR transporters mediate redistribution to specialized sulfur-rich storage cells in the phloem²³⁻²⁴. However, there is evidence for nonrandom distribution of GSLs in the leaf epidermis²⁵. *Arabidopsis thaliana* MERISTEM 1 (ATML1) is a member of the homeodomain leucine-zipper (HD-Zip) IV transcription factor family that is expressed in the outermost layer (L1) of vegetative and floral organs and is critical for formation of the epidermis²⁶. It functions redundantly with its close paralog PROTODERMAL FACTOR2 (PDF2) to positively regulate the expression of L1-specific genes during epidermal differentiation^{27,28}. In *Arabidopsis* sepals, ATML1 overexpression promotes the formation of giant cells that nearly cover the entire sepal whereas *atml1* loss-of-function results in loss of giant cells^{29,30}.

In the present study, we identify a novel transcriptional repressor of the GSL pathway that regulates members of the class IV homeodomain leucine zipper (HD-Zip IV) family. GIR1 (GLABRA2 (GL2)-interacting repressor) was initially reported to physically interact with the HD-Zip IV transcription factor GL2³¹. GIR1 contains an ethylene-responsive element binding factor-associated amphiphilic repression (EAR) motif required for TOPLESS (TPL) corepressor recruitment³². We demonstrate that GIR1 directly interacts with ATML1 and modulates its activity at the post-translational level. GIR1 contains C-terminal cysteine motifs that are predicted to form a C4 zinc finger motif required for its interaction with ATML1. This protein-protein interaction occurs through the HD-Zip IV START Adjacent Domain, a C-terminal

domain of unknown function. Genetic analysis reveals that GIR1 controls giant cell formation in an opposing manner to ATML1. Our protein interaction studies confirmed that the GIR1 EAR motif interacts with TPL and further uncovered that it can interact with other TPL family members, TOPLESS-RELATED (TPR) proteins. We propose a model in which GIR1 engages the HD-Zip IV TF ATML1, and the TPL/TPR corepressors simultaneously to facilitate transcriptional repression of genes of the GSL biosynthesis pathway.

Results

The *gir1* mutants display excess giant cells on sepals

GIR1 and *GIR1*-like genes encode small proteins of ~100 amino acids that are thought to function in transcriptional repression^{31,32}, but whose molecular and biological activities are not clearly defined. Tomato and tobacco *GIR1* homologs have roles in trichome development³³⁻³⁵. In cotton, *GIR1*-related members are implicated in fiber elongation^{36,37}. While performing genetic crosses with *gir1* mutants from *Arabidopsis*, we discovered an overproduction of giant cells on their sepals (**Fig. 1A,B**). Giant cells are elongated epidermal cells, visible by their large size and outward bulging. Sepals, the outermost floral organs, protect the reproductive tissues from the environment and control flower opening³⁸. *GIR2*, a paralog of *GIR1*, shares high nucleotide and predicted amino acid similarity to *GIR1*³¹. However, no obvious phenotypes were observed in sepals of *gir2* null mutants compared to the control (**Appendix Figure 1, Fig. 1C**). We constructed *gir1;gir2* double mutants and found that their giant cell phenotypes were indistinguishable from that of *gir1* single mutants (**Appendix Figure 1, Fig. 1C**). These data show that *GIR1* but not *GIR2* negatively regulates giant cell formation on sepals.

Giant cell phenotypes reveal genetic interaction between GIR1 and ATML1

The *atml1* loss-of-function mutants were previously reported to exhibit defective giant cell formation³⁰. The giant cell overproduction we observed in *gir1* mutants is similar to that observed in lines overexpressing *ATML1*²⁹. Thus, *GIR1* appears to function antagonistically to *ATML1*. To further test the relationship between the *gir1* and *atml1* mutations, we performed genetic crosses and found that *gir1;atml1* double mutants contain fewer giant cells, indicating that *atml1* is genetically epistatic to *gir1* (**Fig. 1B**). We quantified the number of giant cells along with giant cell widths and uncovered significant increases in both metrics in the *gir1* mutants in comparison to wild type and *atml1*, whereas this phenotype was lost in the *gir1;atml1* double mutant (**Fig. 1C**). Thus, *gir1* excess giant cell formation depends on *ATML1* activity. These results suggest that the balance between *ATML1* and *GIR1* activity is important in determining giant cell fate within the sepal.

GIR1 complements the excess giant cell phenotype of *gir1* mutants

Next, a complementation experiment was conducted to verify that the giant cell phenotype is due to the loss of *GIR1* function in the *gir1* mutant background. We generated transgenic lines carrying a genomic fragment containing the native promoter and coding region of wild-type *GIR1* (**Fig. 2A, 2B**). Observation and quantification of giant cells versus small cells on sepals confirmed rescue of the *gir1* mutant phenotype (**Fig. 2C, 2D**). From sequence alignments and AlphaFold models³⁹, we identified conserved motifs which we targeted for mutational analysis (**Figs. 2A, 3A**). The EAR motif is a transcriptional repression motif consisting of LxLxL (x = any amino acid)^{40,41}. The EAR motif of GIR1, located within the N-terminus, was shown to physically interact with the transcriptional corepressor TPL³². Conserved

Cys motifs (CPxC), located within the C-terminus of GIR1, are predicted to form a zinc finger comprised of four Cys residues.

We utilized Cys to Ala mutagenesis to test whether the zinc finger is required for *GIR1* regulation of giant cell formation. Transgenic lines carrying the genomic fragment with the *gir1*^{C71A} or *gir1*^{C87A} mutations failed to complement the *gir1* giant cell phenotype (**Figs. 2A, 2C, 2D**). Similarly, *gir1*^{ΔEAR} transgenic lines carrying a genomic fragment harboring a deletion of the EAR motif failed to rescue the *gir1* giant cell phenotype (**Figs. 2A, 2C, 2D**). Since the EAR motif is required to recruit the TPL corepressor, we asked whether a translational fusion of TPL to the *gir1*^{ΔEAR} mutant protein restores the function of GIR1. The TPL C-terminal 944 amino acid coding region was fused to the C-terminus of the *gir1*^{ΔEAR} genomic sequence (**Fig. 2B**). Phenotypic characterization showed that the *proGIR1:gir1*^{ΔEAR}:*TPL* transgene complemented the giant cell phenotype, similar to the wild-type *pro:GIR1:GIR1* (**Figs. 2C, 2D**). This result suggests that GIR1 recruits TPL through the EAR motif which is critical for *GIR1* function.

Next, we repeated the complementation experiment with *gir1* mutants utilizing the cauliflower mosaic virus strong constitutive 35S promoter (**Appendix Figure 2A**). The *GIR1* overexpression similarly showed complementation of *gir1* giant cell phenotype, and quantification revealed a trend in the reduction of giant cell formation in the *35S:GIR1* lines (**Appendix Figure 2B, 2C**). In contrast, 35S overexpression lines carrying the *gir1*^{ΔEAR}, *gir1*^{C71A} or *gir1*^{C87A} mutations showed giant cell phenotypes indistinguishable from the *gir1* mutant (**Appendix Figure 2B, 2C**). Overall, these experiments show that *GIR1* requires the EAR motif and the Cys residues within the predicted zinc finger to effectively suppress giant cell overproduction in sepals.

Yeast 2-Hybrid and in planta assays demonstrate GIR1 interaction with ATML1 STAD

We utilized two separate in vivo assays to test the interaction between GIR1 and ATML1, considering the predicted structural features of each protein. Aside from the EAR motif and predicted zinc finger, the GIR1 protein appears highly disordered (**Fig. 3A**). ATML1, a member of the HD-Zip IV TF family, contains a homeodomain (HD) DNA binding domain and a leucine zipper dimerization domain, as well as a steroidogenic acute regulatory (StAR) protein-related lipid transfer (START) and a START adjacent domain (STAD) (**Fig. 3B**).

Using Y2H assays, we confirmed interaction between GIR1 and ATML1 and mapped the site of interaction to STAD for ATML1 (**Fig. 3C**). Our Y2H assays with putative STAD mutants confirmed the importance of D589 as well as E633, for interaction with GIR1 (**Fig. 3C**). We used Y2H assays to determine which features of GIR1 are critical for ATML1 interaction. Mutants affecting the N-terminal EAR motif (*gir1^{ΔEAR}*) were able to interact with ATML1 while C-terminal cysteine (CPxC) motif mutants with alanine substitutions (*gir1^{C71A}*, *gir1^{C87A}*) failed to interact with ATML1 (**Fig. 3C**).

The GIR1-ATML1 protein interactions were validated by co-immunoprecipitation (Co-IP) using transient expression of HA- and GFP-tagged proteins in *N. benthamiana*. Co-IP analysis revealed that EGFP:GIR1, when used as bait, strongly interacted with HA:ATML1 (**Figs. 3D, 3E**). Furthermore, specific mutations mapping to STAD abolished the interaction, consistent with the Y2H analysis. The HA:atml1^{F578A} and HA:atml1^{D589L} mutants showed a weaker and no detectable interaction, respectively (**Fig. 3D**). Consistent with the Y2H results, the *gir1* mutants affecting the N-terminal EAR motif (*gir1^{ΔEAR}*) interacted with ATML1 in Co-

IP, while C-terminal cysteine (CPxC) motif mutants *gir1*^{C71A} and *gir1*^{C87A} failed to interact with ATML1 (**Fig. 3E**). Taken together, the Co-IP experiments confirmed the critical importance of the Cys residues (C71A, C87A) of GIR1, and the STAD residues (F587A, D589L) of ATML1 for interaction (**Fig. 3F**).

ATML1, GIR1, and TPL/TPR proteins form a ternary complex

A previous study showed that GIR1 binds the TPL corepressor, suggesting its role as an adaptor protein³². We conducted Co-IP experiments with EGFP-tagged *gir1* mutant proteins and HA-tagged TPL to address the functional requirements for this interaction. Our results show that the N-terminal EAR motif is necessary for TPL binding while the C-terminal Cys motifs (Cys71, Cys87) are dispensable for this interaction (**Fig. 4A**). Since TPL is a member of a small gene family, we used Co-IP to address whether several TPR proteins are also able to bind with GIR1. Our results show that four family members (TPR1, TPR2, TPR3, TPR4) bind to GIR1, similar to the established GIR1-TPL interaction (**Fig. 4B**). We validated the GIR1-TPL/TPR interaction results with Y2H assays and additionally confirmed that *gir1*^{ΔEAR} mutant fails to interact with these corepressor proteins, although both mutant and wild-type proteins show similar levels of expression (**Fig. 4C, Appendix Figure 3**). Another corepressor protein, SAP18, appears to not interact with GIR1, and served as a negative control, along with the empty vector (**Fig. 4C**).

Since GIR1 interacts with both ATML1 (**Fig. 3**) and TPL/TPR proteins, we investigated whether GIR1 can interact with both proteins simultaneously thereby forming a tripartite complex. HA:ATML1 was co-expressed with EGFP-tagged TPR2 and TPR4 in *N. benthamiana* either in the absence or presence of EGFP:GIR1. Immunoprecipitation using anti-HA magnetic beads showed that HA-ATML1 pulled down TRP2 and TPR4 in the presence of GIR1 (**Fig. 4D**). In contrast a significantly attenuated interaction was observed between ATML1 and TPR2/TPR4

when they were co-expressed in the absence of GIR1. Taken together, our results suggest that GIR1 facilitates the formation of a ternary complex through simultaneous interactions with ATML1 and TPL/TPR family members such as TPR2 and TPR4 (**Fig. 4E**).

GSL pathway genes are upregulated in sepals from *gir1* mutants

To understand the molecular basis of the excess giant cell phenotype in *gir1* mutants, we performed an RNA-sequencing (RNA-seq) experiment on sepal tissue from stage 12 flowers of wild type, *gir1*, and *atml1* mutants (**Datasets S1-S6**). As expected, the RNA-seq data showed that transcript levels of *GIR1* and *ATML1* were reduced in the corresponding mutants in comparison to wild type (**Fig. 5A**). Our transcriptome analysis revealed that overall, the *gir1* mutants exhibit more upregulated genes relative to wild type while *atml1* mutants exhibit more downregulated genes (**Figs. 5B**). The upregulation of genes in the *gir1* mutant in comparison to both wild type and *atml1*, as visualized in a volcano plot (**Fig. 5C**) and heat map (**Fig. 5D, Appendix Figure 4**), provides support for the hypothesis that *GIR1* functions as a repressor of the giant cell fate, while *ATML1* functions as an activator of giant cell fate.

Gene ontology (GO) and overrepresentation analysis of the RNA-seq data revealed the upregulation of 13 GSL pathway genes in *gir1* mutants (**Figs. 5E, 5F**). The upregulated genes included *MYB29*, a transcription factor gene implicated in aliphatic GSL production⁴², and *MAM1*, an aliphatic GSL biosynthesis gene, both of which were upregulated >10-fold in *gir1* mutants (**Figs. 5F, 5G, 5H**). Other genes significantly upregulated in *gir1* included more aliphatic GSL biosynthesis genes (*MAM3*, *GSTU20*, *IPMI SSU3*), as well as genes implicated in indolic GSL biosynthesis (*CYP79B2*, *CYP79B3*), sulfur metabolism (*AKN2*) and GSL catabolism (*ESP*) (**Fig. 5H**). It is noteworthy that the *MYB29* paralog, *MYB28*, also showed an upregulated trend in *gir1* mutants (**Fig. 5F**).

In addition to the genes detected by GO analysis, further scanning of the RNA-seq data uncovered numerous other GSL pathway genes that were significantly upregulated or showed an upregulated trend in *gir1* mutants in comparison to wild type and *atml1*: a total of 29 GSL-related genes were upregulated in *gir1*: 21 aliphatic ([Appendix Figure 5, Appendix Table 1](#)) and 9 indolic GSL biosynthesis genes, two of which are also active in the aliphatic pathway ([Appendix Figure 6, Appendix Table 1](#)). The benzenic GSL pathway genes were not differentially expressed in *gir1* mutants, whereas one of the ER body – GSL-related genes, *NAII*, was shown to be significantly upregulated ([Appendix Figure 6, Appendix Table 1](#)). *MYB29* and several other GSL-related genes were significantly downregulated or showed a decreased trend in *atml1* mutants, consistent with our hypothesis that *ATML1* is a positive regulator of GSL biosynthesis, while *GIR1* acts as a repressor. Although we found evidence for *GIR1*-mediated transcriptional regulation of the aliphatic pathway through *MYB29* and to a lesser extent, *MYB28*, regulators of the indolic pathway, *MYB34* and *MYB51*, showed no differential expression ([Appendix Figure 6](#)), and transcript levels of *MYB76* and *MYB122* were not detected in the RNA-seq data ([Datasets S1-S6](#)).

The GO analysis revealed that genes related to 2-oxocarboxylic acid metabolism were also highly enriched in the *gir1* mutant to a similar extent in enrichment as the GSL biosynthesis genes ([Fig. 5E](#)). Also referred to as 2-keto acid, 2-oxocarboxylic acid is a key intermediate in amino acid synthesis. In the initial steps of aliphatic GSL biosynthesis, methionine undergoes a transamination with the aminotransferase BCAT4 to produce 2-oxocarboxylic acid⁴³. A series of reactions elongate the methionine side chain before entering the core structure formation of the pathway. Upregulated genes involved in 2-oxocarboxylic acid metabolism indicate an increase in

elongated chain methionine, consistent with the increased production of aliphatic GSLs in *gir1* mutant sepals.

GSL levels are elevated in sepals from *gir1* mutants

To address whether upregulation of GSL biosynthesis pathway genes leads to changes in GSL production, we collected sepals from wild type, *gir1*, and *atml1* mutants. We also collected bud tissues lacking sepals, which were harvested concurrently at stage 12, like the sepals, to serve as a control. Extracts were prepared for high performance liquid chromatography-diode-array detection (HPLC-DAD) analysis ([Fig. 6A](#)). Quantification of the data revealed that total *gir1* GSL levels were consistent between sepal and buds lacking sepals. In contrast, total GSL levels were lower in buds lacking sepals compared to wild type or *atml1* ([Fig. 6B](#), [Appendix Table 2](#)).

Overall, five GSLs (4MSOB, I3M, 3MSOP, 8MOSOO, 5MSOP) were detected in sepals and three additional minor compounds (4MTB, 4OHI3M, 7MSOH) were detected in buds lacking sepals. Four aliphatic GSL compounds (4MSOB, 3MSOP, 8MOSOO, 5MSOP) and one indolic GSL compound (I3M) were significantly increased about two-fold in *gir1* mutant sepals in comparison to wild-type sepals ([Fig. 6B](#), [6C](#), [Appendix Table 1](#)). In contrast, *atml1* mutant sepals exhibited a reduction in GSL levels for three of the aliphatic compounds (4MSOB, 3MSOP, 5MSOP). In bud tissues lacking sepals, we failed to detect a notable difference in GSL profiles between the mutants and wild type ([Fig. 6C](#), [6D](#)). Altogether, the quantitative differences in GSL profiles of *gir1* and *atml1* mutants appear to be restricted to sepals. Using matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) imaging of sepals, we confirmed that GSL levels are elevated in *gir1* sepals compared to wild type and *atml1* ([Appendix Figure 7](#)).

We superimposed the results from our transcriptome analysis and GSL profiles on a current model of the GSL biosynthesis pathway from *Arabidopsis* (Fig. 6E). In this model, genes upregulated in *gir1* encode enzymes that are predicted to contribute to the increased accumulation of GSL compounds observed in *gir1*. Therefore, the nearly global upregulation of GSL gene expression can explain the increased levels of GSLs detected in sepals from *gir1* mutants.

Discussion

We propose a model for the role of the ATML1-GIR1-TPL/TPR tripartite complex in the repression of GSL biosynthesis as well as the repression of giant cell elongation growth in Fig. 7. GIR1 interacts with ATML1 through its zinc finger and ATML1-STAD, and this interaction is required for GIR1 activity. Could a detailed molecular understanding of the binding interface identify novel ways to control the GIR1-ATML1 interaction that in turn dictates GSL accumulation in the sepal, or in other specified tissues? It is noteworthy that the *MYB29* paralog, *MYB28*, showed an upregulated trend (Fig. 5F) whereas the mRNA levels of other described transcriptional regulators of the GSL pathway were not significantly altered in sepals, suggesting that the ATML1-GIR1-TPL/TPR module could play an important role in transcriptional regulation of the pathway in a tissue- or cell-specific manner.

Building on this model, GIR1 contains C-terminal cysteine motifs that are predicted to form a C4 zinc finger motif critical for its interaction with ATML1 and other HD-Zip IV transcription factors. This protein-protein interaction occurs through STAD, a domain of unknown function. While the function of STAD is not reported, we show here that this domain is required for interaction with GIR1. Mutant analysis suggests that GIR1 controls the ratios for

normal giant cell and small cells in sepals. Additionally, the GIR1 EAR motif is not only required to interact with TPL but also with other TPL family members, TPR proteins.

GIR1 negatively regulates giant cell elongation in sepals, but other adaptor protein family members *gir2* and *sied1* do not exhibit giant cell phenotypes (Fig. 1C, Appendix Figure 1E). However, could these proteins have a similar role to GIR1 in GSL regulation? Adaptor proteins may exhibit GSL regulation similar to GIR1 or act to regulate GSLs in a tissue-specific manner. GSL accumulation and composition varies widely across plant tissues so small differences in transcriptional regulation of biosynthesis or transport may be in part from specific adaptor protein activity, which fine tunes this regulatory network. Establishing the plant-wide adaptor protein-mediated GSL regulation may reveal specificity in diverse epidermal tissues, such as leaves, roots, and seeds, which have agronomic relevance in Brassica oilseed crops including camelina and pennycress.

This broader view is further supported by work in other species. ABAXIALLY CURLED LEAF1 (ACL1), a rice ortholog of GIR1, drives bulliform cell division and cell expansion, resulting in abaxial leaf curling when overexpressed⁵⁰. This shows that GIR1-like adaptors can influence epidermal development by altering both cell proliferation and enlargement in diverse species. A recent investigation utilizing proximity labeling and protein-protein interaction assays demonstrated that ACL1 directly interacts with RICE OUTERMOST CELL-SPECIFIC GENE 4 (ROC4) and ROC5 transcription factors, thereby negatively modulating their activities through corepressor recruitment and/or interfering with their homo- or heterodimerization⁵¹. This is consistent with our findings that GIR1 bridges the interaction between ATML1 and TPL/TPR corepressors, consequentially resulting in decreased ATML1 target gene activation.

The *gir1* mutant displays two phenotypes: (i) excess giant cell formation in sepals, and (ii) elevated GSLs in sepals, but not other floral tissues. These findings raise an important question about the relationship between GSLs and cell expansion. Giant cell morphology is described as elongated, bulging cells from the epidermis. Since bulging cells may be first subject to insect attack, one possibility is that GSLs accumulate in giant cells providing a selective advantage to prevent insect herbivory. On the other hand, it has been reported that GSL derivatives can act as signaling molecules to influence cell growth⁴⁴. This is consistent with the idea that regulation of GSL accumulation may consequently control endoreduplication and cell size. The signal resolution from our MALDI-MS imaging experiments did not permit us to distinguish whether GSLs preferentially reside in giant cells but instead provided clues that GSL accumulate non-uniformly across the sepal. A spatial understanding of the GSL pathway in *gir1* mutants will aid in deciphering how the GIR1-ATML1 regulatory module can be leveraged to modify GSL profiles in crops.

Alternative pathways may provide means to regulate GSL homeostasis through ATML1. A recent study implicated leaf endoplasmic reticulum (LER) bodies accumulate β -glucosidases to hydrolyze GSLs for defense purposes, and genetic analyses showed that ATML1 drives LER body formation⁴⁵. Therefore, ATML1-mediated regulation of LER bodies could control timely response to biotic attack by shifting the balance of storage and hydrolysis of GSLs. Functionally, linking ATML1-regulated LER body formation and GSL derivative levels provides a connection between plant developmental patterns and chemical defense where the plant can localize enzymes needed for GSL bioactivation in the cellular context in which activation occurs.

More generally, recent evidence suggests that GSL biosynthesis is not only controlled at the transcriptional level, but also through light cues. The light-signaling TF involved in GSL

biosynthesis, HY5 was identified to directly interact with histone deacetylases⁴⁶. Similarly, GIR1 and the other adaptor proteins are also regulated by light due to the light responsiveness and circadian rhythm elements in their promoters and diurnal fluctuations in transcript levels⁴⁷⁻⁴⁹. Although ATML1 is not itself light-regulated, its target gene expression can be influenced by light due to its association with light-regulated adaptor proteins. Rather than acting as a developmental switch, the temporal regulation of GSL biosynthesis through GIR1 adds an additional layer of complexity to ATML1-driven regulation.

GSL secondary metabolites display versatile applications. On one hand, GSLs and their derivatives possess pesticidal and antimicrobial activities^{2,3}, and therefore high levels of GSLs are advantageous in floral tissues which are especially vulnerable to herbivory and pathogen attack. On the other hand, high levels of GSLs in camelina seed meals are detrimental for cattle feeds^{52,53}. The regulatory mechanism that we have identified in *Arabidopsis* addresses these opposing needs: Our preliminary data indicate that the *gir1* GSL overproduction mutant displays high levels of these compounds specifically in the sepals, but not in seeds.

Materials and Methods

Plant growth conditions and transformation

Arabidopsis thaliana plants were of the Columbia (Col-0) ecotype. The T-DNA insertion lines *gir1-1* (SALK_089888C) and *gir2-1* (SALK_048394) were previously described³¹ and provided by Vitaly Citovsky, and *atml1-3* (SALK_033408) was previously described³⁰ and provided by Adrienne Roeder. Plants were grown in BM6 (Berger, Canada) soil supplemented with vermiculite and perlite (Hummert International, Topeka, Kansas) at a 28:4:1 ratio. Sown seeds were stratified at 4°C for 4-6 days and transferred to 23°C under continuous light. Plants

for GSL analysis were grown in diurnal 14 h light/10 h dark conditions. The *gir1* plants were transformed by floral dip using the *Agrobacterium tumefaciens* strain GV3101 (pMP90)⁵⁴. Transformants were selected on 0.8% agar medium supplemented with 1 mM KNO₃ and 20 µg/mL Hygromycin B for native promoter constructs and 80 µg/mL kanamycin for CaMV 35S promoter constructs. At least ten transformants per line were transferred to the soil. T1 plants were screened for a T2 segregation ratio of 3:1 of the relevant antibiotic resistance or EGFP expression. For each transgene, at least three to four transformants were characterized, and at least one representative homozygous line, verified by sequencing, was selected for further study.

Construction of binary vectors and other plasmids

The cDNA sequences for *ATML1* and *GIR1*, were PCR-amplified and cloned using the pENTR/D-TOPO Cloning Kit (Invitrogen, Carlsbad, CA). pENTR/D-TOPO plasmids for *gir1* mutants (*ΔEAR*, *R23L*, *R30E*, *R23L;R30L*, *C71A*, *C87A*) and *atml1* mutants (*K547E*, *F578A*, *D589L*, *N593A*, *N627A*, *E633L*, *K733R*) were created from wild-type templates using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs). Gateway cloning using LR Clonase II (Invitrogen) was used to construct the following: The 35S promoter constructs were generated by transferring pENTR/D-TOPO or pDONR207 sequences to pK7WGF2⁵⁵ for EGFP fusions and pEarleyGate 201⁵⁶ for hemagglutinin (3xHA) tag fusions; and the Y2H constructs were generated by transferring pENTR/D-TOPO sequences to pDEST32 and pDEST22 plasmids (ProQuest, Invitrogen). PCR amplification with Q5 High Fidelity DNA Polymerase (New England Biolabs), gel purification with Nucleospin Gel and PCR Clean-up (Machery-Nagel), and NEBuilder HiFi DNA Assembly (New England Biolabs) into the *Hind*III and *Kpn*I linearized binary vector SR54, a derivative of pCAMBIA1300, was used to generate the *GIR1* promoter constructs. For *proGIR1:GIR1*, a ~5 kb genomic fragment containing the *GIR1*

promoter, 5' UTR, gene body, 3' UTR, and 340 bp downstream sequence was taken. For the *proGIR1:gir1^{ΔEAR}* construct, the genomic region was amplified as two segments: an upstream fragment prior to the EAR motif and a downstream fragment with the remainder of the sequence. For the *proGIR1:gir1^{C71A}* and *proGIR1:gir1^{C87A}* constructs, the genomic region was similarly amplified in two segments to incorporate the respective Cys to Ala substitution mutations. A four-fragment DNA assembly was performed to generate the *proGIR1:gir1^{ΔEAR}:TPL* construct. The first segment was identical to first segment for the *proGIR1:gir1^{ΔEAR}* construct. The second segment contained the region downstream of the EAR motif through the last amino acid of *GIR1*. The third fragment contained the TPL C-terminal domain (944 residues). The fourth fragment contained the *GIR1* 3'-UTR and the 340 bp downstream sequence. The oligonucleotide primers utilized for plasmid construction are listed in [Appendix Table 1](#).

Y2H assays

The pDEST32 bait and pDEST22 prey vectors were transformed in opposite mating type strains Y2HGold and Y187 as described previously⁵⁸ using standard LiAc/PEG transformation. To assay the bait-prey interactions, mating was carried out on yeast peptone dextrose adenine (YPDA) medium to produce diploids containing both the prey and bait vectors. Following incubation for 1-2 days, diploid cells were selected on -L-W medium. Diploids were grown in liquid -L-W media to an OD600 of 1.0. From the normalized cultures, four serial dilutions were prepared and placed on permissive and selective media using a 48-pin multiplex plating tool ("Frogger," Dankar, Inc.) to assay reporter gene expression indicating interactions. Bait and prey vectors lacking inserts were used as negative controls. Confirmation of bait protein expression involved the extraction of total yeast proteins using the NaOH/trichloroacetic acid method.

Transient protein expression and coimmunoprecipitation

Agrobacterium cultures harboring bait, prey, or p19⁵⁹ plasmids were grown overnight. Cell mixtures with equal amounts of bait and prey plasmids (final OD₆₀₀ = 0.4) were combined with cells harboring p19 (final OD₆₀₀ = 0.2) and were pelleted and resuspended in 1 mL of infiltration media (10 mM MES [pH 5.6], 10 mM MgCl₂, 150 μM acetosyringone, 6 mg/ml glucose). Cells were incubated in the dark for 1 h. Leaves from 4–6-week-old *N. benthamiana* plants grown under 14 h light /10 h dark cycles were infiltrated on the abaxial side with a 1 ml needleless syringe. EGFP expression was assessed after 48 h by epifluorescence microscopy, and expressing regions were cut, weighed, flash-frozen, and stored at -80°C. Frozen tissue (0.5 g) was ground in liquid nitrogen, and resuspended in 2.5 mL extraction buffer (150 mM NaCl, 50 mM Tris-HCl [pH 7.5], 5% [v/v] glycerol, 1 mM EDTA, 0.5% [v/v] Triton X-100, 1 mM dithiothreitol (DTT), 1X protease inhibitor (Sigma Aldrich, P9599), 1 mM phenylmethanesulfonylfluoride (PMSF), 2% [w/v] polyvinylpyrrolidone (PVPP). Lysates were vortexed for 1 min and incubated on ice for 30 min before pelleting at 20K × g for 10 min at 4°C. Lysates (100 μL) were heated for 5 min at 99°C, combined with 100 μl of 2X Laemmli buffer and stored at -20°C for use as input controls. Cleared lysates were incubated with 15 μL of GFP-Trap magnetic agarose beads (Chromotek) or anti-HA magnetic beads (ThermoFisher Scientific, 88836) for 60 min at 4°C with continuous end-to-end rotation. Beads were pelleted and supernatants were removed, followed by multiple 5 min washes (150 mM NaCl, 50 mM Tris [pH 7.5], 5% [v/v] glycerol, 1 mM EDTA, 0.5% [v/v] Triton X-100, 1 mM DTT, 1X protease inhibitor, 1mM PMSF) with continuous end-to-end rotation. Beads were eluted in 100 μL Laemmli buffer and resolved by SDS-PAGE followed by western blot with rabbit anti-GFP (1:2000, Invitrogen A-11122) or rat anti-HA (1:5000, Roche, 11867431001) primary antibodies

and goat anti-rabbit (1:2000; Enzo, ADI-SAB-300) or anti-rat (1:3000; Millipore, AP136P) secondary antibodies. SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher Scientific) was used for chemiluminescent detection with an Azure c300 imager (Azure Biosystems).

Phenotypic characterization and microscopy

Intact flower buds or sepals were dissected using fine forceps and transferred to double stick tape for examination. Images were captured with a Leica DFC295 digital camera mounted on a Leica M125 fluorescence stereomicroscope. Expression of EGFP:GIR1 and EGFP:gir1^{ΔEAR} was analyzed using the GFP2 filter. Scanning electron microscope (SEM) images were recorded on an SNE-Alpha high-resolution tabletop SEM with a 5-axis XYZRT stage at 5 nm resolution (Nanoimages, Lafayette, CA). Giant cell quantification was performed by marking giant cells based on characteristic morphological features followed by measuring the widths using ImageJ.

RNA extraction and transcriptomics analysis

Stage 12 unopened flowers were dissected to obtain 200 sepals (from 50 buds) for each genotype in triplicate. Samples were frozen in liquid nitrogen prior to RNA extraction using the QIAGEN RNeasy Plant Mini Kit (Venlo, Netherlands). RNA quality evaluation, cDNA library construction, and Illumina sequencing were performed by Novogene (Sacramento, CA). Genes were identified as differentially expressed if their q (adjusted p value) ≤ 0.05 and $\log_2(\text{fold change}) \geq 1$. Mainstream hierarchical clustering grouped genes by their FPKM values. Gene ontology and KEGG enrichment analyses were performed for corresponding genes with an adjusted p value < 0.05 .

Glucosinolate extraction and analysis

GSL content in all genotypes within all tissues was measured as previously described⁶⁰. Immediately after tissue collection each sample was placed in 1 mL of 100% methanol. Samples were homogenized for 3 min in a paint shaker, centrifuged, and 400 μ L of supernatants were transferred to a 96-well filter plate with DEAE sephadex. The filter plate was washed with water, 90% methanol and water again. Sephadex-bound GSLs were eluted after an overnight incubation with 110 μ L of sulfatase. Individual desulfo-GSLs within each sample were separated and detected by HPLC-DAD, identified, and quantified by comparison to standard curves from purified desulfo-GSLs⁶¹.

Mass Spectrometry Imaging

Sepals were dissected under an optical microscope, and the abaxial side of the sepal was mounted facing up on a glass slide using a double-sided conductive carbon tape. Three sepals from each genotype were placed on the same slide to minimize batch effects. 9-aminoacridine (9-AA; 15 mg/mL in MeOH: H₂O (30:70, v/v) was applied as a matrix using a TM sprayer (HTX Technologies, Chapel Hill, NC), and Au was sputter-coated for 20 s at 40 mA (Cressington 108; Ted Pella, Redding, CA) to provide surface conductivity. MALDI-MS imaging was performed using a QExactive HF Orbitrap MS (Thermo Scientific, San Jose, CA) with a MALDI source (Spectroglyph, Kennewick, WA) equipped with a 349-nm laser (Explorer One; Spectra Physics, Milpitas, CA). Data were collected in negative-ion mode for the m/z range of 400–500 with a mass resolution of 120,000 at m/z 200 to target the GSLs signals. Raster steps of 20 μ m were used, considering the small size of sepals. Raw data were converted to imzML files using ImageInsight (Spectroglyph), and imzML files are loaded into the MSiReader (North Carolina

State University; Raleigh, NC) software⁶² to generate MS images, and METASPACE⁶³ was used for annotation with a false discovery rate (FDR) threshold of 10% and a mass tolerance of ± 3 ppm.

Figures

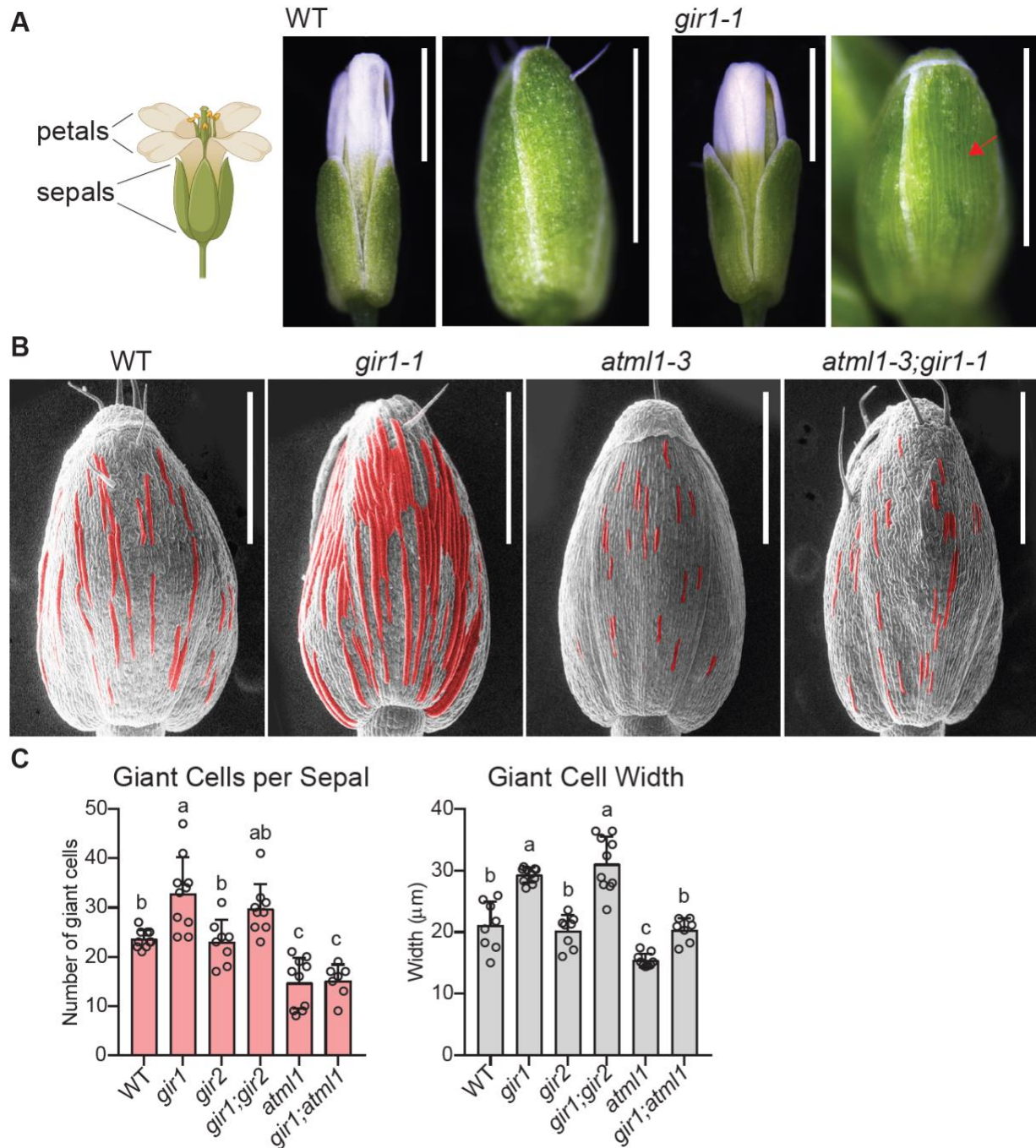


Figure 1. *gir1* mutant displays a giant cell phenotype in sepals.

(A) Illustration of *Arabidopsis thaliana* flower indicating anatomical positions of four petals and three out of four sepals. Stereo microscope images show wild-type (WT) flowers (stage 14) and

buds (stage 12)⁶⁴ in comparison to *gir1* mutants, which display an increase in giant cell formation (arrow). Bar = 1 mm

(B) Comparison of giant cell phenotype on sepals from *gir1* versus *atml1* mutant. Enhanced giant cell formation is suppressed in the *gir1;atml1* double mutant, suggesting that *atml1* is epistatic to *gir1*. SEM images of stage 12 flowers with giant cells falsely colored red. Images were taken at stage 12 (closed flower). Bar = 500 μ m.

(C) Quantification of giant cell numbers per sepal and giant cell width for WT in comparison to *gir1*, *gir2*, *gir1;gir2*, *atml1*, and *gir1;atml1*. Standard deviations are indicated for n=7-10 biological replicates. Significant differences were assessed by one-way ANOVA, Tukey's test, and are indicated by letters (P value < 0.05). Images of *gir2* and *gir1;gir2* mutant buds are shown in [Appendix Figure 1](#).

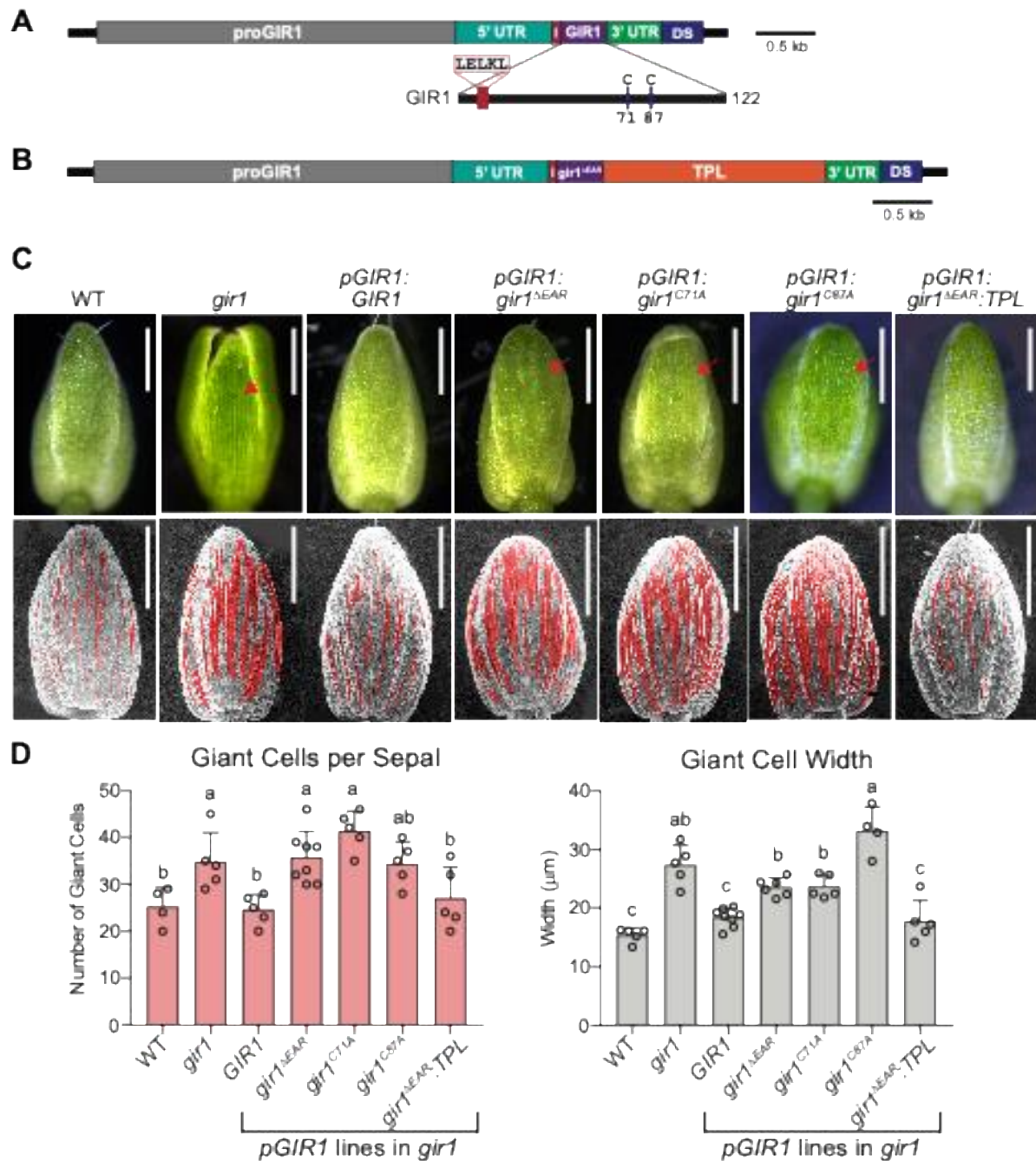


Figure 2. The EAR motif and cysteine residues of the Zn finger are required for rescue of the giant cell phenotype by the *GIR1* native promoter construct.

(A) and (B) Schematics of the transgenic native promoter constructs for *GIR1* wild-type and mutants affecting the function of the EAR motif and predicted Zn finger.

(A) The *GIR1* native promoter (*pGIR1*) was used to drive expression of *GIR1*, and the *gir1*^{*ΔEAR*} and *gir1*^{*C68A*} mutants, and **(B)** expression of C-terminal fusion of TPL to *gir1*^{*ΔEAR*}.

(C) Stereo microscopy images of flower buds (top) and matching scanning electron microscopy images (bottom) with genotypes indicated. Giant cells are false-colored red. Images were taken at stage 12 (closed flower). Bar = 1 mm. Sepal giant cell phenotypic analysis of *gir1* mutant and transgenic lines with constructs shown in **(A)** and **(B)**. The *GIR1* and *gir1*^{*ΔEAR*}-*TPL* constructs rescued *gir1* phenotype while *gir1* mutants (*gir1*^{*ΔEAR*}, *gir1*^{*C71A*}, and *gir1*^{*C87A*}) failed to rescue. Bar = 500 μm.

(D) Number of giant cells along with average giant cell width per sepal were quantitated for the genotypes shown in **(C)**. Standard deviations are indicated for n=7-10 biological replicates. Significant differences were assessed by one-way ANOVA, Tukey's test, and are indicated by letters (P value < 0.05).

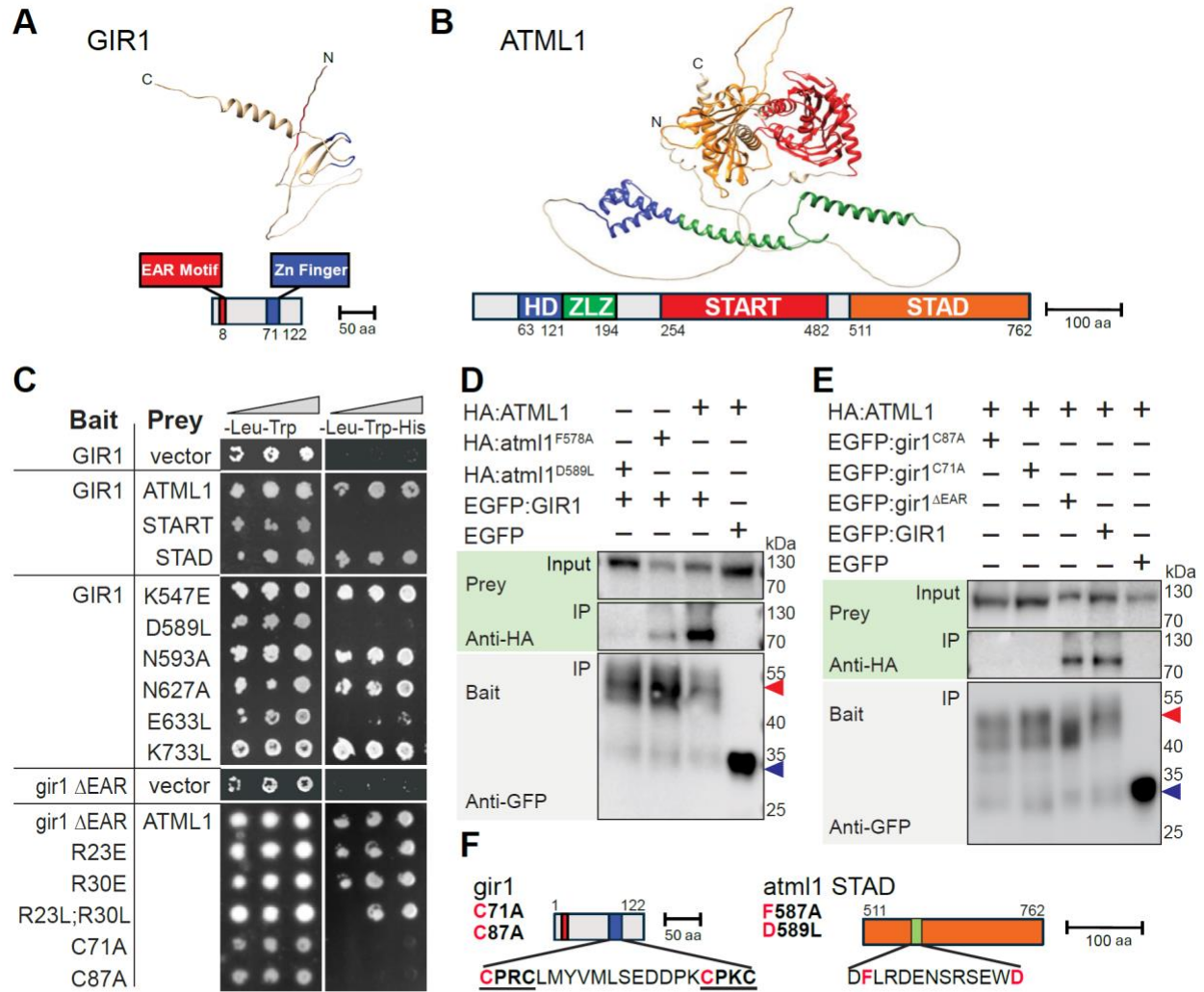


Figure 3. GIR1 interacts with HD-Zip IV transcription factor ATML1 via its Zn finger motif.

(A) Predicted GIR1 structure (AlphaFold;³⁹) visualized with UCSF ChimeraX⁶⁵ with labeled structural motifs including the EAR motif comprising a LxLxL pattern and four conserved cysteine residues forming a C4 zinc-finger motif with a CPxC pattern.

(B) Predicted ATML1 structure (AlphaFold;³⁹) visualized with USCf Chimera X (Chimera;⁶⁵) with respective domain configuration. HD, homeodomain; ZLZ, zipper-loop-zipper; START, steroidogenic acute regulatory (StAR) related lipid transfer domain; STAD, START adjacent domain.

(C) Yeast two-hybrid (Y2H) assays reveal that GIR1 interacts with ATML1 through the STAD segment, and that STAD residues D589, N627, and E633 are critical for interaction. Cysteine residues (C71, C87) within the predicted Zn finger of GIR1, are important for interaction with ATML1, whereas the EAR motif and Arg residues within the N-terminus of GIR1 (R23, R30) are not required. The empty prey vector served as a negative control. Growth on -Leu-Trp signifies presence of bait and prey plasmids, while growth on selective media (-Leu-Trp-His) indicates protein interaction, and four-fold dilutions of yeast are indicated left to right.

(D and E) Co-IP experiments confirmed GIR1-ATML1 interaction in planta using *Nicotiana benthamiana* transient expression. EGFP-tagged GIR1 or gir1 mutant (red arrowhead) or EGFP negative control (blue arrowhead) served as the bait while HA-tagged proteins served as the prey. Input indicates the initial protein amount for the Co-IP. Input and immunoprecipitated (IP) proteins were detected in western blots with anti-HA and anti-GFP antibodies. Predicted kilodalton (kDa) sizes are indicated.

(D) ATML1 STAD residues F578 and D589 are important for physical interaction with GIR1.

(E) The cysteine mutants (gir1^{C87A}, gir1^{C71A}) failed to interact with ATML1 while the gir1^{ΔEAR} interaction with ATML1 was unaffected, suggesting the importance of the GIR1 Zn finger for ATML1 interaction.

(F) Schematic of mutations that affect the GIR1-ATML1 interaction. For GIR1, the C71A and C87A missense mutations affect the predicted Zn finger motif. For ATML1 STAD, the F578A and D589L mutations affect aromatic and acidic residues, respectively, within an acidic patch.

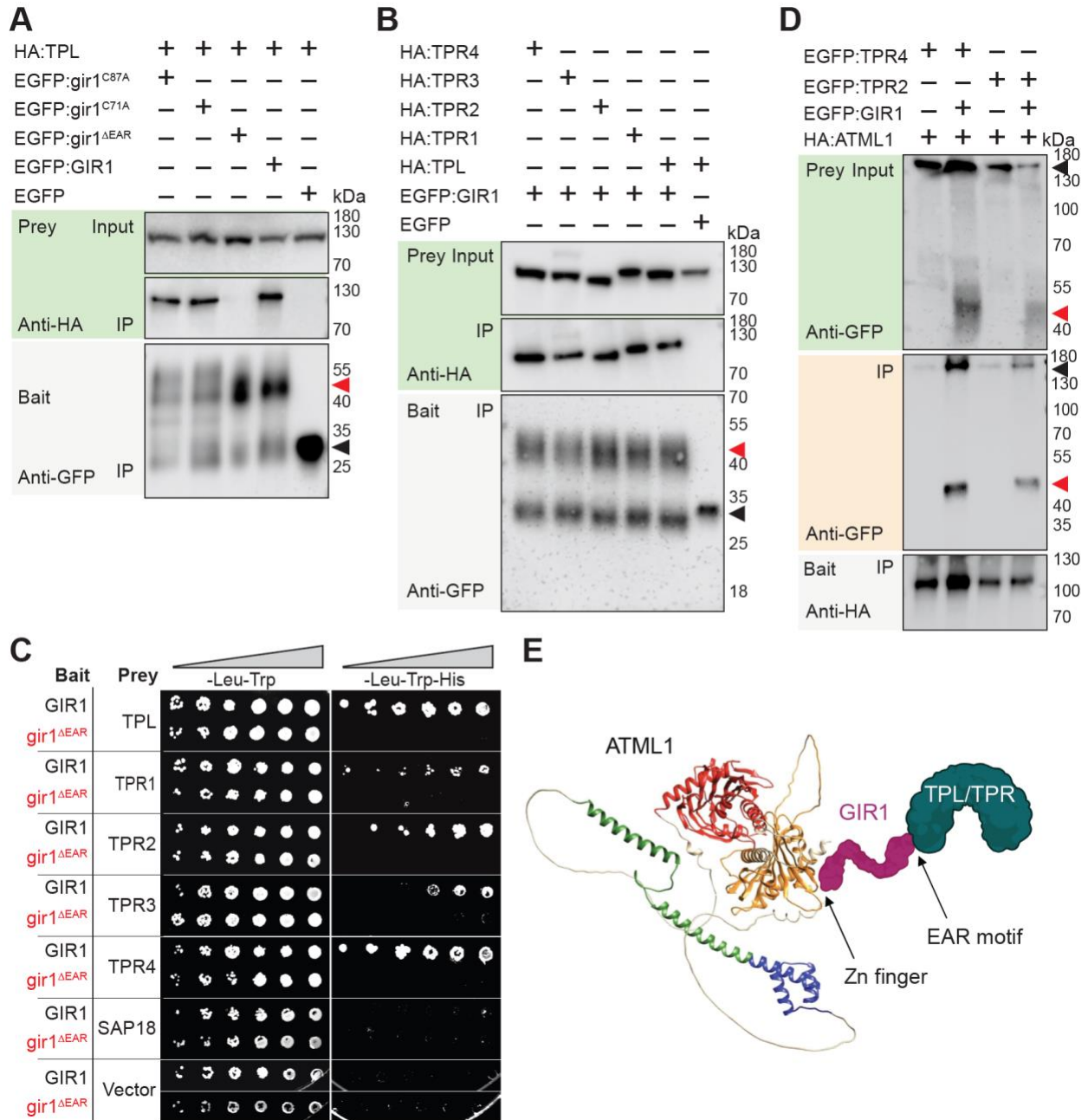


Figure 4. GIR1 interacts with TPL and TPRs via its EAR motif.

(A and B) Co-IP experiments demonstrate GIR1-TPL and GIR1-TPR interactions in *plantia* using *Nicotiana benthamiana* transient expression. EGFP- and HA-tagged proteins served as baits and preys, respectively. Input indicates the initial protein amount used for the Co-IP. Input and IP proteins were detected in western blots with anti-HA and anti-GFP antibodies. Predicted kDa sizes are indicated for the HA-tagged and the EGFP-tagged proteins.

(A) Co-IP shows that the GIR1 EAR motif, but not its Zn-finger associated Cys residues, is required for GIR1-TPL interaction. EGFP-tagged GIR1, *gir1* mutants (*gir1*^{C71A}, *gir1*^{C87A}, *gir1*^{DEAR}) (red arrowhead), or EGFP control (black arrowhead) served as baits whereas HA-tagged TPL was used as prey.

(B) Co-IP shows GIR1 interaction with TPL and TPR proteins. EGFP-tagged GIR1 (red arrowhead) or EGFP negative control (black arrowhead) served as bait whereas HA-tagged TPL or TPR proteins served as preys.

(C) Y2H assays confirm that GIR1 interacts with TPL and TPR proteins and that the GIR1 EAR motif is required for this interaction. SAP18 and empty vector served as negative controls. Growth on -Leu-Trp signifies presence of bait and prey plasmids, while growth on selective media (-Leu-Trp-His) indicates protein interaction, and four-fold dilutions of yeast are indicated left to right.

(D) Co-IP demonstrates that GIR1 serves as a bridge connecting ATML1 and TPR proteins (TPR2, TPR4). HA-ATML1 and EGFP-TPR2 (or TPR4) were coexpressed with and without EGFP:GIR1 in *N. benthamiana*, and proteins were immunoprecipitated using anti-HA magnetic beads. Proteins were detected with anti-HA or anti-GFP antibodies in western blots. Black arrowheads indicate EGFP:TPR2 and EGFP:TPR4 prey proteins in the input (top blot) and co-immunoprecipitated (middle blot) samples in the presence or absence of GIR1. Red arrowheads indicate EGFP:GIR1 in the input (top blot) and coimmunoprecipitated samples (middle blot). The presence of the HA:ATML1 bait is shown in the bottom blot. Protein sizes are shown in kDa.

(E) Model of the tripartite complex. ATML1 binds GIR1 via its C-terminal STAD segment, while GIR1 binds ATML1 via its Zn finger. The EAR motif of GIR1 is required for binding to

TPL or TPR proteins. Predicted contact sites of the Zn finger and EAR motif are indicated by arrows.

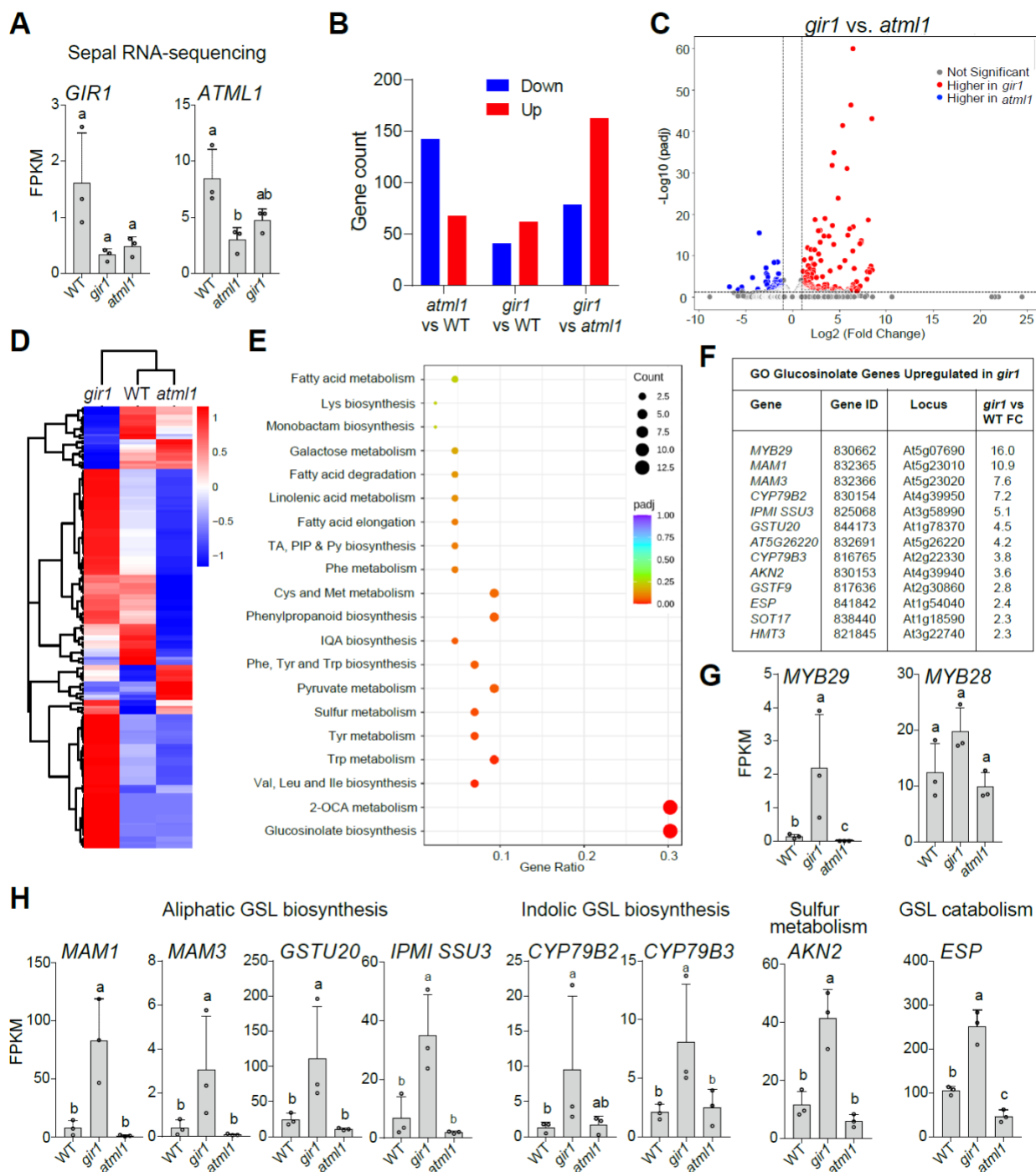


Figure 5. Differential expression analysis suggests that GIR1 acts as a repressor of ATML1 to control GSL biosynthesis genes in sepals.

(A) Sepal RNA-seq data of *GIR1* and *ATML1* expression levels in the corresponding mutants in comparison to WT. The *GIR1* and *ATML1* mRNA levels appear decreased in each mutant background. FPKM, fragments per kilobase per million mapped reads.

(B) Binary comparison of differentially expressed genes (DEGs) in *Arabidopsis* sepals for WT, *gir1* mutant, and *atml1* mutant genotypes. Blue bars correspond to downregulated genes and red bars represent upregulated genes.

(C) Volcano plot illustrates the greater number of upregulated genes in *gir1* in comparison to *atml1*. Red dots represent upregulated genes, and blue dots represent down-regulated genes. Genes with unchanged expression are indicated by grey dots.

(D) Heatmap shows overview of DEGs between *atml1*, *gir1*, and WT sepals. Averaged relative expression between n=3 biological replicates is graded from low (blue) to high (red) ($P < 0.05$). A dendrogram depicting clustering of gene expression patterns is shown on the left of the heatmap. Individual replicates are shown for each genotype in [Appendix Figure 4](#).

(E) Dot plot visualization of gene GO enrichment analysis of upregulated genes of *gir1* vs. *atml1*. The y-axis shows GO terms while the x-axis shows the proportion of total genes identified in the gene set which are also associated with the GO term. Dot color reflects the statistical significance ($-\log(P\text{-value})$) whereas the dot size corresponds to the number of genes in each GO term.

(F) List of upregulated genes in *gir1* sepals associated with the GO term GSL biosynthesis. Expression fold difference between *gir1* and WT is provided for each gene.

(G) Sepal RNA-seq data for *MYB29* and *MYB28*, two redundant transcriptional regulators of GSL biosynthesis. *MYB29* expression levels are >10-fold higher in *gir1* as compared to WT, while *MYB29* transcripts were not detected in *atml1*.

(H) Sepal RNA-seq data for other GSL biosynthesis genes upregulated in *gir1* mutants, including aliphatic (*MAM1*, *MAM3*, *GSTU20*, *IPMI SSU3*) and indolic (*CYP79B2*, *CYP79B3*)

GSL biosynthesis genes, as well as a gene involved in sulfur metabolism (*AKN2*) and GSL catabolism (*ESP*), respectively. In **(A)**, **(G)**, and **(H)**, significant differences between genotypes are marked by letters (one-way ANOVA, Tukey's test, $p < 0.05$). Additional RNA-seq data is shown in **Appendix Figures 5,6** and **Appendix Table 1**.

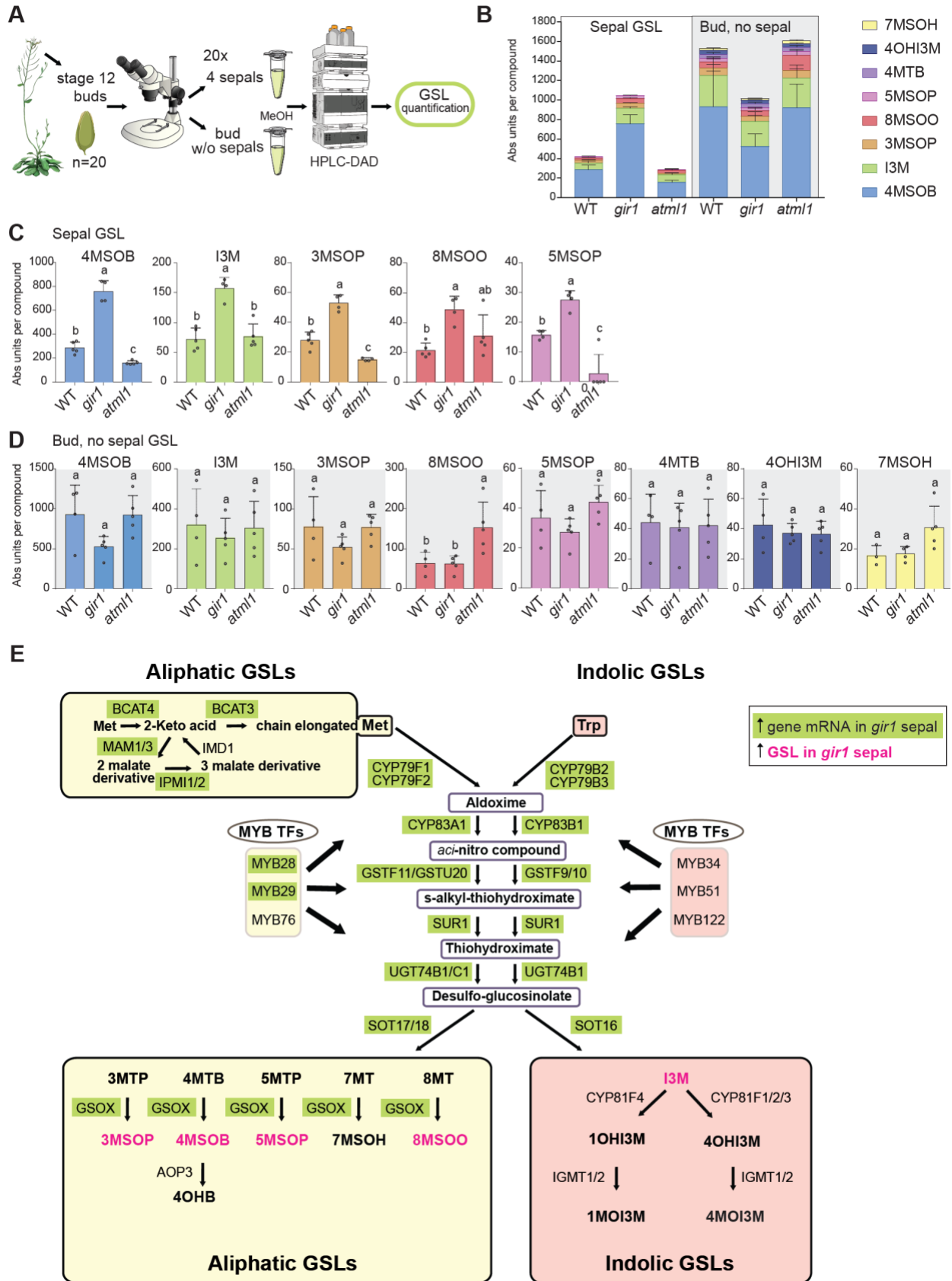


Figure 6. GSL levels are elevated in sepals from *gir1* mutants.

(A) Schematic of workflow for GSL analysis. *Arabidopsis* stage 12 buds were dissected under the stereomicroscope. Sepals were separated from the remainder of the bud tissues and transferred to 100% methanol prior to extraction, desulfation, and detection by HPLC-DAD. GSLs were quantified from absorbance peaks.

(B-D) HPLC-DAD analysis uncovered elevated levels of GSLs in *gir1* mutant sepals in comparison to WT and *atm11*. Full names of GSL compounds and information on their properties is given in [Appendix Table 2](#).

(B) Stacked bar graphs illustrate GSL composition in sepals from WT, *gir1*, and *atm11* plants (left panel) in comparison to bud tissue lacking sepals (right panel, grey shading). The key for the GSL compounds (right) is ordered from top to bottom from the least- to most-abundant compounds. The aliphatic GSL 4MSOB is most abundant in both tissues, followed by the indolic GSL I3M.

(C) Bar graphs indicate that GSLs detected in sepals are elevated in *gir1* mutants, compared to WT, for all five compounds that were detected by HPLC-DAD (4MSOB, I3M, 3MSOP, 8MSOO, 5MSOP). The *atm11* mutants exhibit decreases in 3MSOP, 4MSOP, and 8MSOO.

(D) Bar graphs (grey shading) show that buds without sepals, which were separated from the sepals and harvested concurrently, failed to show consistent differences between the genotypes. The overall GSL profile, which additionally includes the minor compounds 4MTB, 4OHI3M, and 7MSOH, appears distinct from the sepal profile.

(E) Model of the GSL biosynthesis pathway in *Arabidopsis*. Aliphatic and indolic GSLs are derived from Met and Trp, respectively. The side-chain elongation of Met is shown at the top

left. Core GSL scaffold formation is depicted in the center. The downstream enzymatic reactions specific to aliphatic and indolic GSLs are shown in the bottom left and right boxes, respectively. MYB transcription factors that control the aliphatic and indolic components are shown on the flanks of the core pathway. Genes upregulated in the *gir1* loss-of-function mutants are indicated in green boxes, and compounds detected by HPLC-DAD in sepals are shown in magenta. Modified from⁴⁶.

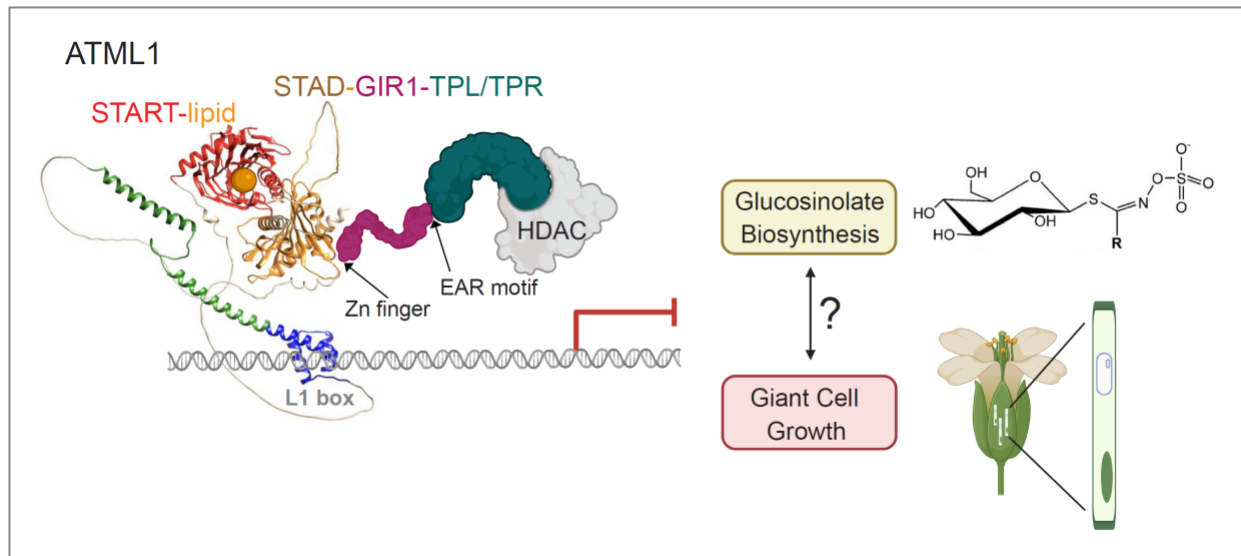


Figure 7. Model for biological function of ATML1-GIR1-TPL/TPR complex in repression of GSL biosynthesis and giant cell growth.

The HD-Zip IV transcription factor ATML1 binds to the L1 box of target genes²⁸. The START domain of ATML1 binds a lipid⁶⁶, while the C-terminal STAD segment binds to GIR1 through its Zn finger. In turn, GIR1 binds to TPL or TPR corepressor proteins through its EAR motif. The TPL and TPR proteins interact with HDAC to induce closed chromatin and transcriptional repression^{67,68}. GIR1 activity is required for both GSL biosynthesis and giant cell formation in sepals. While GIR1 appears to negatively regulate both processes, the connection between GSL biosynthesis and giant cell formation is unclear.

References

1. Fahey, J.W., Zalcmann, A.T., and Talalay, P. (2001). The chemical diversity and distribution of glucosinolates and isothiocyanates among plants. *Phytochemistry* 56, 5–51. 10.1016/s0031-9422(00)00316-2.
2. Bidart-Bouzat, M.G., and Kliebenstein, D.J. (2008). Differential levels of insect herbivory in the field associated with genotypic variation in glucosinolates in *Arabidopsis thaliana*. *J Chem Ecol* 34, 1026–1037. 10.1007/s10886-008-9498-z.
3. Bednarek, P., Pislewska-Bednarek, M., Svatos, A., Schneider, B., Doubek, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R., Sanchez-Vallet, A., et al. (2009). A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science* 323, 101–106. 10.1126/science.1163732.
4. Blazevic, I., Montaut, S., Burcul, F., Olsen, C.E., Burow, M., Rollin, P., and Agerbirk, N. (2020). Glucosinolate structural diversity, identification, chemical synthesis and metabolism in plants. *Phytochemistry* 169, 112100. 10.1016/j.phytochem.2019.112100.
5. Halkier, B.A., and Gershenzon, J. (2006). Biology and biochemistry of glucosinolates. *Annu Rev Plant Biol* 57, 303–333. 10.1146/annurev.arplant.57.032905.105228.
6. Costa-Perez, A., Nunez-Gomez, V., Baenas, N., Di Pede, G., Achour, M., Manach, C., Mena, P., Del Rio, D., Garcia-Viguera, C., Moreno, D.A., and Dominguez-Perles, R. (2023). Systematic Review on the Metabolic Interest of Glucosinolates and Their Bioactive Derivatives for Human Health. *Nutrients* 15. 10.3390/nu15061424.
7. Connolly, E.L., Sim, M., Travica, N., Marx, W., Beasy, G., Lynch, G.S., Bondonno, C.P., Lewis, J.R., Hodgson, J.M., and Blekkenhorst, L.C. (2021). Glucosinolates From Cruciferous Vegetables and Their Potential Role in Chronic Disease: Investigating the Preclinical and Clinical Evidence. *Front Pharmacol* 12, 767975. 10.3389/fphar.2021.767975.
8. Favela-Gonzalez, K.M., Hernandez-Almanza, A.Y., and De la Fuente-Salcido, N.M. (2020). The value of bioactive compounds of cruciferous vegetables (Brassica) as antimicrobials and antioxidants: A review. *J Food Biochem* 44, e13414. 10.1111/jfbc.13414.
9. Soundararajan, P., and Kim, J.S. (2018). Anti-Carcinogenic Glucosinolates in Cruciferous Vegetables and Their Antagonistic Effects on Prevention of Cancers. *Molecules* 23. 10.3390/molecules23112983.
10. Kissen, R., and Bones, A.M. (2009). Nitrile-specifier proteins involved in glucosinolate hydrolysis in *Arabidopsis thaliana*. *J Biol Chem* 284, 12057–12070. 10.1074/jbc.M807500200.
11. Barba, F.J., Nikmaram, N., Roohinejad, S., Khelifa, A., Zhu, Z., and Koubaa, M. (2016). Bioavailability of Glucosinolates and Their Breakdown Products: Impact of Processing. *Front Nutr* 3, 24. 10.3389/fnut.2016.00024.
12. Luang-In, V., Narbad, A., Nueno-Palop, C., Mithen, R., Bennett, M., and Rossiter, J.T. (2014). The metabolism of methylsulfinylalkyl- and methylthioalkyl-glucosinolates by a selection of human gut bacteria. *Mol Nutr Food Res* 58, 875–883. 10.1002/mnfr.201300377.
13. Liou, C.S., Sirk, S.J., Diaz, C.A.C., Klein, A.P., Fischer, C.R., Higginbottom, S.K., Erez, A., Donia, M.S., Sonnenburg, J.L., and Sattely, E.S. (2020). A Metabolic Pathway for

- Activation of Dietary Glucosinolates by a Human Gut Symbiont. *Cell* *180*, 717–728 e719. 10.1016/j.cell.2020.01.023.
14. Hansen, B.G., Kliebenstein, D.J., and Halkier, B.A. (2007). Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*. *Plant J* *50*, 902–910. 10.1111/j.1365-313X.2007.03101.x.
 15. Pfalz, M., Vogel, H., and Kroymann, J. (2009). The gene controlling the indole glucosinolate modifier1 quantitative trait locus alters indole glucosinolate structures and aphid resistance in *Arabidopsis*. *Plant Cell* *21*, 985–999. 10.1105/tpc.108.063115.
 16. Sonderby, I.E., Geu-Flores, F., and Halkier, B.A. (2010). Biosynthesis of glucosinolates--gene discovery and beyond. *Trends Plant Sci* *15*, 283–290. 10.1016/j.tplants.2010.02.005.
 17. Bird, K.A., Ramos, A.A., and Kliebenstein, D.J. (2025). Phylogenetic and genomic mechanisms shaping glucosinolate innovation. *Curr Opin Plant Biol* *85*, 102705. 10.1016/j.pbi.2025.102705.
 18. Frerigmann, H., and Gigolashvili, T. (2014). MYB34, MYB51, and MYB122 distinctly regulate indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Mol Plant* *7*, 814–828. 10.1093/mp/ssu004.
 19. Frerigmann, H., Glawischnig, E., and Gigolashvili, T. (2015). The role of MYB34, MYB51 and MYB122 in the regulation of camalexin biosynthesis in *Arabidopsis thaliana*. *Front Plant Sci* *6*, 654. 10.3389/fpls.2015.00654.
 20. Gigolashvili, T., Yatusovich, R., Berger, B., Muller, C., and Flugge, U.I. (2007). The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana*. *Plant J* *51*, 247–261. 10.1111/j.1365-313X.2007.03133.x.
 21. Schweizer, F., Fernandez-Calvo, P., Zander, M., Diez-Diaz, M., Fonseca, S., Glauser, G., Lewsey, M.G., Ecker, J.R., Solano, R., and Reymond, P. (2013). *Arabidopsis* basic helix-loop-helix transcription factors MYC2, MYC3, and MYC4 regulate glucosinolate biosynthesis, insect performance, and feeding behavior. *Plant Cell* *25*, 3117–3132. 10.1105/tpc.113.115139.
 22. Nintemann, S.J., Hunziker, P., Andersen, T.G., Schulz, A., Burow, M., and Halkier, B.A. (2018). Localization of the glucosinolate biosynthetic enzymes reveals distinct spatial patterns for the biosynthesis of indole and aliphatic glucosinolates. *Physiol Plant* *163*, 138–154. 10.1111/ppl.12672.
 23. Xu, D., Hunziker, P., Koroleva, O., Blennow, A., Crocoll, C., Schulz, A., Nour-Eldin, H.H., and Halkier, B.A. (2019). GTR-Mediated Radial Import Directs Accumulation of Defensive Glucosinolates to Sulfur-Rich Cells in the Phloem Cap of *Arabidopsis* Inflorescence Stem. *Mol Plant* *12*, 1474–1484. 10.1016/j.molp.2019.06.008.
 24. Koroleva, O.A., Gibson, T.M., Cramer, R., and Stain, C. (2010). Glucosinolate-accumulating S-cells in *Arabidopsis* leaves and flower stalks undergo programmed cell death at early stages of differentiation. *Plant J* *64*, 456–469. 10.1111/j.1365-313X.2010.04339.x.
 25. Shroff, R., Vergara, F., Muck, A., Svatos, A., and Gershenzon, J. (2008). Nonuniform distribution of glucosinolates in *Arabidopsis thaliana* leaves has important consequences for plant defense. *Proc Natl Acad Sci U S A* *105*, 6196–6201. 10.1073/pnas.0711730105.
 26. Schrick, K., Ahmad, B., and Nguyen, H.V. (2023). HD-Zip IV transcription factors: Drivers of epidermal cell fate integrate metabolic signals. *Curr Opin Plant Biol* *75*, 102417. 10.1016/j.pbi.2023.102417.

27. Abe, M., Katsumata, H., Komeda, Y., and Takahashi, T. (2003). Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. *Development* 130, 635–643. 10.1242/dev.00292.
28. Abe, M., Takahashi, T., and Komeda, Y. (2001). Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J.* 26, 487–494.
29. Meyer, H.M., Teles, J., Formosa-Jordan, P., Refahi, Y., San-Bento, R., Ingram, G., Jonsson, H., Locke, J.C., and Roeder, A.H. (2017). Fluctuations of the transcription factor ATML1 generate the pattern of giant cells in the Arabidopsis sepal. *Elife* 6, e191131. 10.7554/eLife.19131.
30. Roeder, A.H., Cunha, A., Ohno, C.K., and Meyerowitz, E.M. (2012). Cell cycle regulates cell type in the Arabidopsis sepal. *Development* 139, 4416–4427. 10.1242/dev.082925.
31. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. I. Interaction with the repressor GLABRA2 and regulation of root hair development. *Biochem Biophys Res Commun* 488, 547–553. 10.1016/j.bbrc.2017.05.084.
32. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. II. Interaction with the co-repressor TOPLESS and promotion of histone deacetylation of target chromatin. *Biochem Biophys Res Commun* 488, 609–613. 10.1016/j.bbrc.2017.05.085.
33. Yang, C., Li, H., Zhang, J., Luo, Z., Gong, P., Zhang, C., Li, J., Wang, T., Zhang, Y., Lu, Y., and Ye, Z. (2011). A regulatory gene induces trichome formation and embryo lethality in tomato. *Proc Natl Acad Sci U S A* 108, 11836–11841. 10.1073/pnas.1100532108.
34. Gao, S., Gao, Y., Xiong, C., Yu, G., Chang, J., Yang, Q., Yang, C., and Ye, Z. (2017). The tomato B-type cyclin gene, SlCycB2, plays key roles in reproductive organ development, trichome initiation, terpenoids biosynthesis and *Prodenia litura* defense. *Plant Sci* 262, 103–114. 10.1016/j.plantsci.2017.05.006.
35. Wu, M.L., Cui, Y.C., Ge, L., Cui, L.P., Xu, Z.C., Zhang, H.Y., Wang, Z.J., Zhou, D., Wu, S., Chen, L., and Cui, H. (2020). NbCycB2 represses Nbwo activity via a negative feedback loop in tobacco trichome development. *J Exp Bot* 71, 1815–1827. 10.1093/jxb/erz542.
36. Liu, X., Moncuquet, P., Zhu, Q.H., Stiller, W., Zhang, Z., and Wilson, I. (2020). Genetic Identification and Transcriptome Analysis of Lintless and Fuzzless Traits in *Gossypium arboreum* L. *Int J Mol Sci* 21. 10.3390/ijms21051675.
37. Feng, X., Cheng, H., Zuo, D., Zhang, Y., Wang, Q., Lv, L., Li, S., Yu, J.Z., and Song, G. (2021). Genome-wide identification and expression analysis of GL2-interacting-repressor (GIR) genes during cotton fiber and fuzz development. *Planta* 255, 23. 10.1007/s00425-021-03737-7.
38. Roeder, A.H., Chickarmane, V., Cunha, A., Obara, B., Manjunath, B.S., and Meyerowitz, E.M. (2010). Variability in the control of cell division underlies sepal epidermal patterning in Arabidopsis thaliana. *PLoS Biol* 8, e1000367. 10.1371/journal.pbio.1000367.
39. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Zidek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. 10.1038/s41586-021-03819-2.

40. Kagale, S., and Rozwadowski, K. (2011). EAR motif-mediated transcriptional repression in plants: an underlying mechanism for epigenetic regulation of gene expression. *Epigenetics* 6, 141–146. 10.4161/epi.6.2.13627.
41. Yang, J., Liu, Y., Yan, H., Tian, T., You, Q., Zhang, L., Xu, W., and Su, Z. (2018). PlantEAR: Functional Analysis Platform for Plant EAR Motif-Containing Proteins. *Front Genet* 9, 590. 10.3389/fgene.2018.00590.
42. Sonderby, I.E., Hansen, B.G., Bjarnholt, N., Ticconi, C., Halkier, B.A., and Kliebenstein, D.J. (2007). A systems biology approach identifies a R2R3 MYB gene subfamily with distinct and overlapping functions in regulation of aliphatic glucosinolates. *PLoS One* 2, e1322. 10.1371/journal.pone.0001322.
43. Gigolashvili, T., Yatusovich, R., Rollwitz, I., Humphry, M., Gershenzon, J., and Flugge, U.I. (2009). The plastidic bile acid transporter 5 is required for the biosynthesis of methionine-derived glucosinolates in *Arabidopsis thaliana*. *Plant Cell* 21, 1813–1829. 10.1105/tpc.109.066399.
44. Katz, E., Bagchi, R., Jeschke, V., Rasmussen, A.R.M., Hopper, A., Burow, M., Estelle, M., and Kliebenstein, D.J. (2020). Diverse Allyl Glucosinolate Catabolites Independently Influence Root Growth and Development. *Plant Physiol* 183, 1376–1390. 10.1104/pp.20.00170.
45. Wilkens, A., Czerniawski, P., Bednarek, P., Libik-Konieczny, M., and Yamada, K. (2024). ATML1 Regulates the Differentiation of ER Body-Containing Large Pavement Cells in Rosette Leaves of Brassicaceae Plants. *Plant Cell Physiol* 65, 1160–1172. 10.1093/pcp/pcae039.
46. Choi, D., Kim, S.H., Choi, D.M., Moon, H., Kim, J.I., Huq, E., and Kim, D.H. (2024). ELONGATED HYPOCOTYL 5 interacts with HISTONE DEACETYLASE 9 to suppress glucosinolate biosynthesis in *Arabidopsis*. *Plant Physiol* 196, 1340–1355. 10.1093/plphys/kiae284.
47. Helfer, A., Nusinow, D.A., Chow, B.Y., Gehrke, A.R., Bulyk, M.L., and Kay, S.A. (2011). LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the *Arabidopsis* core clock. *Curr Biol* 21, 126–133. 10.1016/j.cub.2010.12.021.
48. Rombauts, S., Dehais, P., Van Montagu, M., and Rouze, P. (1999). PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res* 27, 295–296. 10.1093/nar/27.1.295.
49. Michael, T.P., Mockler, T.C., Breton, G., McEntee, C., Byer, A., Trout, J.D., Hazen, S.P., Shen, R., Priest, H.D., Sullivan, C.M., et al. (2008). Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet* 4, e14. 10.1371/journal.pgen.0040014.
50. Li, L., Shi, Z.Y., Li, L., Shen, G.Z., Wang, X.Q., An, L.S., and Zhang, J.L. (2010). Overexpression of ACL1 (abaxially curled leaf 1) increased Bulliform cells and induced Abaxial curling of leaf blades in rice. *Mol Plant* 3, 807–817. 10.1093/mp/ssq022.
51. Tao, Z., Zhu, L., Li, H., Sun, B., Liu, X., Li, D., Hu, W., Wang, S., Miao, X., and Shi, Z. (2024). ACL1-ROC4/5 complex reveals a common mechanism in rice response to brown planthopper infestation and drought. *Nat Commun* 15, 8107. 10.1038/s41467-024-52436-w.
52. Paula, E.M., da Silva, L.G., Brandao, V.L.N., Dai, X., and Faciola, A.P. (2019). Feeding Canola, Camelina, and Carinata Meals to Ruminants. *Animals (Basel)* 9. 10.3390/ani9100704.

53. Riaz, R., Ahmed, I., Sizmaz, O., and Ahsan, U. (2022). Use of *Camelina sativa* and By-Products in Diets for Dairy Cows: A Review. *Animals (Basel)* *12*, 10.3390/ani12091082.
54. 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. 10.1046/j.1365-313x.1998.00343.x.
55. Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci* *7*, 193–195. 10.1016/s1360-1385(02)02251-3.
56. Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juehne, T., Song, K., and Pikaard, C.S. (2006). Gateway-compatible vectors for plant functional genomics and proteomics. *Plant J* *45*, 616–629. 10.1111/j.1365-313X.2005.02617.x.
57. Geisbrecht, B.V., Bouyain, S., and Pop, M. (2006). An optimized system for expression and purification of secreted bacterial proteins. *Protein Expr Purif* *46*, 23–32. 10.1016/j.pep.2005.09.003.
58. Mukherjee, T., Subedi, B., Khosla, A., Begler, E.M., Stephens, P.M., Warner, A.L., Lerma-Reyes, R., Thompson, K.A., Gunewardena, S., and Schrick, K. (2022). The START domain mediates *Arabidopsis* GLABRA2 dimerization and turnover independently of homeodomain DNA binding. *Plant Physiol* *190*, 2315–2334. 10.1093/plphys/kiac383.
59. Silhavy, D., Molnar, A., Lucioli, A., Szitty, G., Hornyik, C., Tavazza, M., and Burgyan, J. (2002). A viral protein suppresses RNA silencing and binds silencing-generated, 21- to 25-nucleotide double-stranded RNAs. *Embo J* *21*, 3070–3080. 10.1093/emboj/cdf312.
60. Kliebenstein, D.J., Kroymann, J., Brown, P., Figuth, A., Pedersen, D., Gershenzon, J., and Mitchell-Olds, T. (2001). Genetic control of natural variation in *Arabidopsis* glucosinolate accumulation. *Plant Physiol* *126*, 811–825. 10.1104/pp.126.2.811.
61. Reichelt, M., Brown, P.D., Schneider, B., Oldham, N.J., Stauber, E., Tokuhisa, J., Kliebenstein, D.J., Mitchell-Olds, T., and Gershenzon, J. (2002). Benzoic acid glucosinolate esters and other glucosinolates from *Arabidopsis thaliana*. *Phytochemistry* *59*, 663–671. 10.1016/s0031-9422(02)00014-6.
62. Robichaud, G., Garrard, K.P., Barry, J.A., and Muddiman, D.C. (2013). MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *J Am Soc Mass Spectrom* *24*, 718–721. 10.1007/s13361-013-0607-z.
63. Palmer, A., Phapale, P., Chernyavsky, I., Lavigne, R., Fay, D., Tarasov, A., Kovalev, V., Fuchser, J., Nikolenko, S., Pineau, C., et al. (2017). FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat Methods* *14*, 57–60. 10.1038/nmeth.4072.
64. Suwabe, K., Nagasaka, K., Windari, E.A., Hoshiai, C., Ota, T., Takada, M., Kitazumi, A., Masuko-Suzuki, H., Kagaya, Y., Yano, K., et al. (2020). Double-Locking Mechanism of Self-Compatibility in *Arabidopsis thaliana*: The Synergistic Effect of Transcriptional Depression and Disruption of Coding Region in the Male Specificity Gene. *Front Plant Sci* *11*, 576140. 10.3389/fpls.2020.576140.
65. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* *30*, 70–82. 10.1002/pro.3943.
66. Wojciechowska, I., Mukherjee, T., Knox-Brown, P., Hu, X., Khosla, A., Subedi, B., Ahmad, B., Mathews, G.L., Panagakis, A.A., Thompson, K.A., et al. (2024). *Arabidopsis*

- PROTODERMAL FACTOR2 binds lysophosphatidylcholines and transcriptionally regulates phospholipid metabolism. *New Phytol* 244, 1498–1518. 10.1111/nph.19917.
67. Long, J.A., Ohno, C., Smith, Z.R., and Meyerowitz, E.M. (2006). TOPLESS regulates apical embryonic fate in *Arabidopsis*. *Science* 312, 1520–1523. 10.1126/science.1123841.
68. Liu, Z., and Karmarkar, V. (2008). Groucho/Tup1 family co-repressors in plant development. *Trends Plant Sci* 13, 137–144. 10.1016/j.tplants.2007.12.005.

Chapter 3 - Characterization of the direct interaction between the transactivation domain ATML1 STAD and coregulator GIR1

Abstract

A novel transcriptional repression complex that controls both giant cell formation and glucosinolate production in sepals was recently described in *Arabidopsis*. GIR1 is thought to act as a bridge to connect the HD-Zip IV TF ATML1 and TPL/TPR corepressors to decrease transcription of ATML1 target genes. To better understand the biological nature of the GIR1-ATML1 interaction, we sought to elucidate the minimum requirements of this interaction and to calculate the associated relative affinity. Transcriptional activation domains commonly contain an acidic patch surrounded by bulky hydrophobic residues. Identification of this motif in ATML1 STAD prompted us to ask whether it serves as the activation domain for ATML1. Yeast two-hybrid assays established that STAD is responsible for ATML1 autoactivation while site-directed mutagenesis revealed residues within the acidic patch that were critical for autoactivation. Although previous work demonstrated the ATML1-GIR1 protein-protein interaction in yeast and plant systems, we addressed whether this interaction could be detected in a biophysical assay in vitro. We used surface plasmon resonance (SPR) and found that GIR1 binds STAD with a low micromolar affinity. In contrast, a gir1 cysteine mutant affecting the predicted zinc finger abolished this binding affinity. A STAD acidic patch mutant had a weaker effect on the interaction, although it significantly decreased affinity to GIR1 by about two-fold. Improved understanding of this repression complex will advance the knowledge of gene expression regulation underlying critical plant processes, including cell elongation and glucosinolate biosynthesis.

Introduction

The interaction between GIR1 and ATML1 was detected in vivo with both yeast and plant systems, as described in [Chapter 2](#). We sought to identify the components necessary for the interaction in vitro, and thereby also assess the strength of the interaction by calculating the binding affinity. To answer these questions, we turned to Surface Plasmon Resonance (SPR), which provides means to measure interaction affinity, binding specificity, and real time kinetics of biomolecular interactions¹. Although SPR has become increasingly prevalent in biophysical assays, especially with human proteins, limited SPR analyses have been used to study plant protein-protein interactions, with prior SPR work focusing on protein-sugar or protein-small molecule interactions¹.

Transactivation domains in TFs function by recruiting chromatin remodelers or by interacting with a mediator complex to aid in assembling the transcriptional machinery, which subsequently activates transcription of target genes². Due to the vast diversity of genes a single TF can target, specificity in activation of gene expression can be accomplished by binding various cofactors². The relatively weak nature of the interactions between activation domains and cofactors has contributed to the difficulty in quantitatively characterizing these interactions³. Various efforts have been made to fill this gap. DelRosso et al. took a high throughput approach to measure the affinities of >300 human activation domains and coactivators from isotherms generated from concentration-dependent binding curves³. The authors developed a microfluidic system to detect preferential binding of coactivators to different activation domains, with most interactions tracking to the low-mid micromolar range³. Importantly, point mutations in known interacting regions resulted in a 2- to 20-fold decrease in affinity³. In a similar approach, SPR was employed to detect the interaction with an *Arabidopsis* TF, Dreb2, and the human and plant

Med25 (a domain in the mediator complex) with a low micromolar affinity⁴. Blomberg et al showed that the Dre2 binding and activation domains were required for binding with Med25⁴. This interaction results in the recruitment of general TFs, thereby impacting the assembly of the transcription pre-initiation complex⁴.

Transcriptional activation domains are characterized by bulky hydrophobic residues surrounded by acidic patches and are commonly structurally disordered^{5,6,3}. Thus, interactions between transactivation domains and cofactors can occur through hydrophobic and electrostatic interactions^{3,6}. Further investigation of the Dre2-Med25 interaction led to the discovery of an acidic amino acid motif (DFDLMLGD) required for the interaction⁷. The same nine amino acid charged motif was found residing near the carboxyl terminus in transactivation domains of yeast, mammalian, and viral TFs, suggesting that it mediates interaction with cofactors^{8,9}.

We recently proposed a novel transcriptional repression complex comprised of ATML1, GIR1, and TPL/TPR based on the protein-protein interaction assays in yeast and plant systems (**Chapter 2**). Here, we expressed recombinant ATML1 STAD, atml1 STAD^{D589L}, GIR1, and gir1^{C71A} proteins in an *E. coli* bacterial system to use for biophysical interaction studies. In an SPR experiment, with ATML1 STAD or mutant STAD immobilized, GIR1 or gir1 mutant was flowed over the chip to generate binding sensorgrams. The equilibrium dissociation constant (K_D) between wild type STAD and GIR1 was calculated to be 4.9 μ M by a steady state analysis, a result consistent with previous activation domain-cofactor affinities. The mutant atml1 STAD negatively affected the interaction by a two-fold difference. In stark contrast, no resulting K_D could be calculated from combining STAD with the gir1 cysteine mutant, suggesting no or very weak association. The results imply that the functional interaction between GIR1 and ATML1

occurs through the previously uncharacterized STAD domain of ATML1 and the highly conserved cysteine residues in the predicted zinc finger of GIR1.

Results

ATML1 STAD contains an acidic patch important for transcriptional activation.

Upon sequence analysis of ATML1 STAD, we identified a 13 amino acid acidic patch between amino acids 577 and 589 with the sequence, DFLRDENSRSSEWD, that contains five acidic amino acids and two bulky hydrophobic residues (**Fig 1a**). Motifs with these chemical properties are key characteristics of acidic patches found in many transcriptional activation domains⁶. To test if STAD possesses any transactivation activity, we employed a yeast two-hybrid autoactivation assay. We found that the ATML1 full-length protein and STAD alone were able to autoactivate while START and GIR1 were not (**Fig 1b**). Intriguingly, the START plus STAD construct autoactivated much weaker than STAD alone indicating START may work to restrict STAD activity. To determine whether the acidic patch is required for autoactivation, STAD mutants were also subject to the autoactivation assay. Results showed two mutations in the acidic patch, F578A and D589L, abolished or decreased autoactivation of STAD along with two mutations residing outside of the acidic patch, E633L and K547E (**Fig 1b**). Thus, residues within in the acid patch contribute to STAD's transactivation but are not solely responsible for this function.

ATML1 STAD interacts with GIR1 directly with a low micromolar affinity.

In order to assess the minimum requirements of the ATML1-GIR1 interaction, we sought to detect the direct interaction in an in vitro setting and subsequently obtain an affinity of the interaction, which would provide information of the biological nature of the interaction. To

prepare purified protein samples for SPR analysis, we expressed GIR1, gir1^{C71A}, ATML1 STAD, and atml1 STAD^{D589L} proteins in an *E. coli* bacterial system (**Fig 2a-b**). Trace minerals present in purified wild-type protein samples were detected with Inductively Coupled Plasmon Mass Spectrometry (ICP-MS). While sodium, potassium, calcium, zinc, and selenium were present in all samples including the buffer, magnesium appeared in both protein samples and cobalt was only detected in GIR1, indicating the GIR1 C4 cysteine motif may additionally bind cobalt (**Table 1**).

In yeast and plant systems, C71 of GIR1 was found to be critical for the interaction with ATML1 (**Chapter 2**). In addition, **Fig 1** shows that D589 of ATML1 STAD negatively affected transactivation and resides within the acidic patch. Using SPR, we were able to detect direct binding of GIR1 to ATML1 STAD with a K_D of 4.9 μ M (**Table 2**). Conversely, no K_D could be calculated from gir1^{C71A} over STAD. GIR1 and gir1^{C71A} over purified MBP alone also had no K_D calculated indicating this interaction occurs through STAD (**Fig 3**). GIR1 over atml1 STAD^{D589L} had approximately a two-fold decrease in affinity (8.7 μ M) (**Table 2**). Although residue D589 of STAD is important for the transcriptional activity of ATML1, its removal does not completely abolish the interaction interface.

Discussion

Based on the SPR data, ATML1 STAD binds to GIR1 in the low micromolar range, a similar affinity as previously reported transcriptional activation domain-cofactor interactions³. While atml1 STAD^{D589L} severely affected autoactivation of STAD in the yeast two-hybrid system, the affinity to GIR1 was only decreased by a two-fold difference. Therefore, additional residues are likely to be important for the interaction with GIR1. This principle is in line with

human activation domain studies, where these domains usually contain more than one region involved in activation and interaction³. Measuring the affinity of STAD with the entire acidic patch deleted may more drastically affect the interaction. It has been suggested that affinity can provide information on the relative activation extent of the cofactor, with tighter binding correlating with more activation or repression³. Our data suggests that GIR1 binds directly to ATML1, but due to a more transient nature of the interaction, GIR1 may promptly respond to stimuli and dissociate from ATML1, allowing quicker transitions in ATML1 target gene expression.

GIR1 conserved cysteine motifs are predicted to fold into a zinc finger that can bind a zinc ion. ICP-MS analysis of purified GIR1 detected both zinc and cobalt, suggesting that the C4 cysteine motif may support binding not only to zinc but also cobalt. The SPR data further demonstrate that mutation of even a single cysteine within this motif negatively affects the interaction with ATML1 STAD. One possibility is the ATML1 STAD and GIR1 contact site occurs in part through an electrostatic interaction between the positively charged zinc and acidic patch. If this idea is correct, mutating the cysteines would result in the inability to bind zinc with negative consequences on the integrity of the interaction. Another possibility for the mechanism of this interaction is a conformational change of GIR1 upon STAD binding. Some activation domains are intrinsically disordered but can transition to form structural features upon cofactor binding in an induced fit model⁴. Although STAD is predicted to be mostly structured, a conformational change of GIR1 could allow increased accessibility for additional binding partners, such as TPL/TPR corepressors. Further structural studies are needed to test this model. For example, a nuclear magnetic resonance (NMR) comparison between GIR1 and gir1^{C71A} could reveal chemical shift changes consistent with changes in metal coordination or local

structure. During the course of this work, we encountered challenges with the purification of GIR1, possibly because GIR1 is predicted to be structurally disordered.

Together, these findings identify GIR1 as a direct binding partner of ATML1 STAD and suggests that this interaction is mediated by a specific interface that depends in part on the zinc finger cysteine residues of GIR1 and acidic patch of STAD. The relatively transient nature of binding advances our understanding how ATML1 rapidly fine-tunes target gene expression in response to developmental or environmental cues.

Materials and Methods

Autoactivation yeast two-hybrid assay

The bait vector, pDEST32 was transformed in the yeast strain Y2HGold as described previously¹⁰ using standard LiAc/PEG transformation. To assay bait autoactivation, cells were grown in liquid -Leu media to an OD600 of 1.0. From the normalized cultures, four serial dilutions were prepared and placed on permissive (-Leu) and selective (-Leu-His) media using a 48-pin multiplex plating tool (“Frogger,” Dankar, Inc.) to assess reporter gene expression indicating bait autoactivation. Bait vectors lacking inserts were used as negative controls. Confirmation of bait protein expression involved the extraction of total yeast proteins using a standard NaOH/trichloroacetic acid method.

Recombinant protein production

The cDNA sequences for *GIR1* and *gir1*^{C71A} flanked by *Bam*HI and *Not*I were amplified and subsequently assembled into pT7HMT¹¹ using the NEBuilder HiFi DNA Assembly Kit (New England Biolabs). Similarly, the *ATML1 STAD* and *atml1 STAD*^{D589L} cDNA sequences

were assembled into *SspI*-digested pET-His6-MBP-TEV-Lic (Addgene, Watertown, MA). Colonies from freshly transformed BL21 Rosetta (DE3) cells (Novagen, Millipore-Sigma, Burlington, MA) were inoculated in 10 ml LB and grown at 37°C overnight. Primary cell cultures were transferred to 1 L LB media and grown at 37°C to an OD₆₀₀ of 0.8 prior to inducing expression with 1 mM IPTG overnight at 18°C. Cells were collected by centrifugation and frozen at -80°C. Pellets were resuspended in native binding buffer (20 mM Tris [pH 8.0], 10 mM imidazole, and 500 mM NaCl). Cells were lysed with a LM10 high-pressure homogenizer (Microfluidics, Westwood, MA) followed by high-speed centrifugation for 30 min. Clarified lysate was then purified using a nickel affinity column and eluted using native elution buffer (20 mM Tris [pH 8.0], 500 mM imidazole, and 500 mM NaCl). The proteins were further purified with a Superdex 75pg 26/600 size exclusion column in HEPES [pH 7.4].

Trace mineral panel of aqueous samples by ICP-MS

Concentrated nitric acid suitable for trace metal analysis was obtained from Fisher Scientific (Waltham, Massachusetts), and certified reference standard stock solutions were purchased from Inorganic Ventures (Christiansburg, VA). Ultrapure 18 MΩ water was generated using a Synergi UV-R system (Millipore, Burlington, MA). Each aqueous sample (4.9 mL) was combined with 0.1 mL of 2% aqueous nitric acid in a 15-mL polypropylene centrifuge tube. A method blank was prepared using 4.9 mL of 18 MΩ water and 0.1 mL of 2% aqueous nitric acid. A check standard was used to verify the accuracy of the measurement. Samples were analyzed using an Agilent 7850 ICP-MS (Agilent Technologies, Santa Clara, CA). Calibration curves were established over the following concentration ranges: 0.0097–15 ppm for Ca; 0.0064–10 ppm for Na, Mg, P, and Fe; and 0.00064–1 ppm for Cr, Mn, Co, Cu, and Zn. An internal standard solution containing Ge and Y at 1 ppm in 2% nitric acid was used to correct for signal variation and ensure

analytical accuracy. Calibration curves were generated using linear regression, with an acceptance criterion of $R^2 > 0.99$.

Surface Plasmon Resonance

ATML1 STAD, atml1 STAD^{D589L}, and MBP alone were immobilized on a CMD200 sensor chip (XanTec bioanalytics) by amine coupling using a concentration of 20 µg/ml in 10 mM sodium acetate [pH 4.0]. A final immobilization density of 1138.2, 1066.5, and 1049.6 resonance units (RU) was obtained, respectively. Assays were performed in HBS-T running buffer (10 mM HEPES [pH 7.3], 140 mM NaCl, 0.005% [v/v] Tween- 20) at a 30 µl/min flow rate. Analytes were diluted into running buffer before injection. Interactions were assessed in multicycle experiments with an analyte injection series (0.08, 0.16, 0.31, 0.63, 1.25, 2.5, 5, and 10 µM) applied over an association time of 120 s, followed by a dissociation time of 180 s. Surfaces were regenerated with a single 60 s injection of 2 M NaCl and 3 mM EDTA. Each injection series was fit to a steady-state 1:1 binding model (Langmuir) to calculate an equilibrium dissociation constant (K_D) using Biacore T200 Evaluation Software (v 3.2, Cytiva). Standard deviations were calculated based on a triplicate injection series ($n = 3$).

Figures and Tables

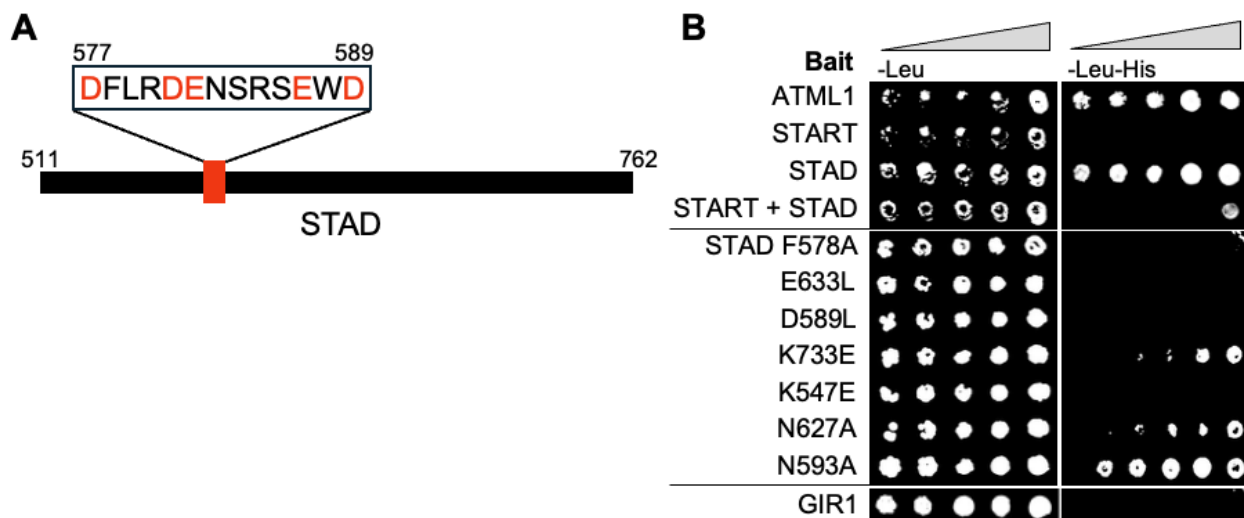


Figure 1. ATML1 STAD functions as a transactivation domain.

(A) A schematic of ATML1 STAD is illustrated highlighting the 13 amino acid acidic patch.

Acidic amino acids are colored in red.

(B) Yeast two-hybrid bait autoactivation assay reveals ATML1 STAD is an auto-activator and residues in STAD acidic patch are critical for this function. Haploid yeast expressing as bait in combination with full-length ATML1, START, STAD, START+STAD, STAD mutants, or GIR1 alone was tested for transcriptional activation without prey construct. Growth on permissive media (-Leu) indicated presence of bait plasmid while growth on selective media (-Leu -His) signified autoactivation, or transcriptional activation without requiring the prey plasmid. 32-fold dilutions of yeast are indicated left to right.

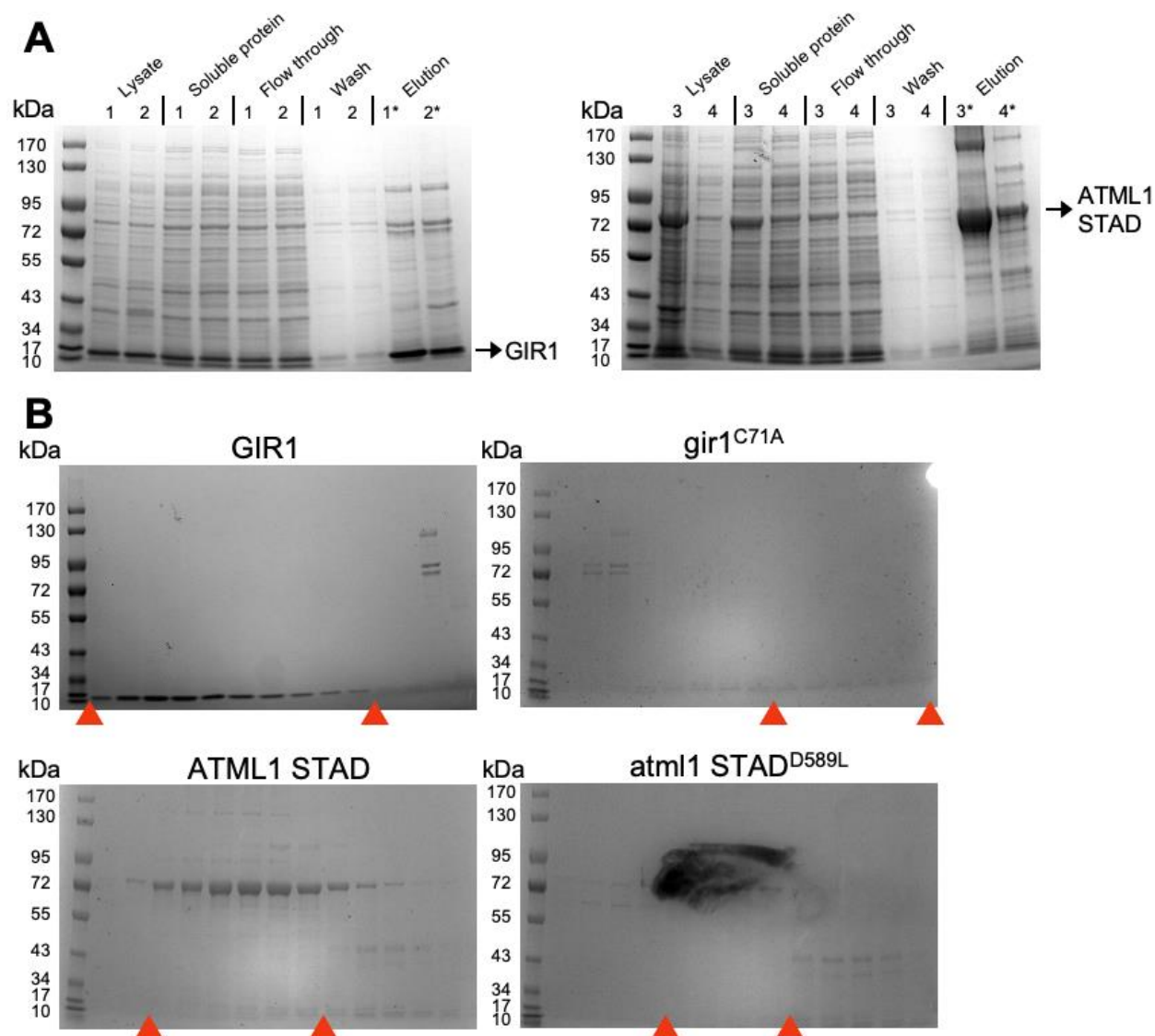


Figure 2. Purification of recombinantly expressed wild-type and mutant GIR1 and ATML1 STAD.

(A) Coomassie blue-stained gel shows protein purification of samples maltose binding protein (MBP = His6-MBP-TEV) tagged (1) wild-type ATML1 STAD, (2) mutant atml1 STAD^{D589L}, histidine tagged (His6-TEV) (3) GIR1, and (4) gir1^{C71A}. For each sample the lysate, soluble supernatant, flow through from nickel affinity column, wash from column, and 500mM imidazole elution are shown. Asterisks indicated samples taken for further purification by size exclusion chromatography. Arrows denote protein of interest band.

(B) SDS-PAGE of protein sample fractions purified by size exclusion chromatography. Fractions between red arrowheads indicate fractions that were pooled for final purified protein sample.

Table 1. Metals detected in purified protein samples from ICP-MS.

Concentration of trace metals detected in purified protein samples in HEPES buffer. Metals not detected (n.d.) were below the limit of quantitation (LOQ).

	Concentration (µg/ml)			
Sample Name	Buffer	GIR1	ATML1 STAD	LOQ
Na	10,903	805	233	0.161
Mg	n.d.	0.581	0.815	0.161
P	n.d.	n.d.	n.d.	0.161
K	1.987	2.173	1.889	0.161
Ca	2.740	4.576	4.013	0.161
Cr	n.d.	n.d.	n.d.	0.016
Mn	n.d.	n.d.	n.d.	0.016
Fe	n.d.	n.d.	n.d.	0.161
Co	n.d.	0.115	n.d.	0.016
Cu	n.d.	n.d.	n.d.	0.016
Zn	0.103	0.088	0.150	0.016
Se	0.044	0.036	0.034	0.016
Mo	n.d.	n.d.	n.d.	0.016

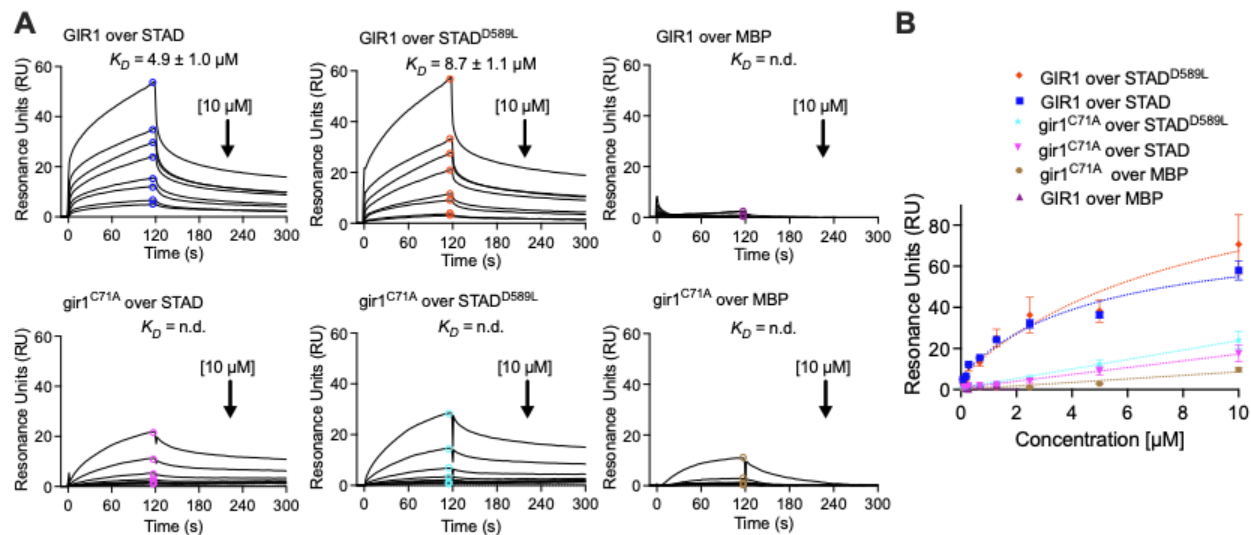


Figure 3. SPR analysis confirmed direct in vitro interaction between ATML1 STAD and GIR1.

(A-B) Purified MBP-tagged ATML1 STAD, atml1 STAD^{D589L}, or MBP alone was immobilized on a SPR sensor chip. Purified GIR1 or gir1^{C71A} was injected in a series of 2-fold dilutions ranging from 0.08 – 10 μM .

(A) Reference and background-corrected sensorgrams from a representative injection series are shown as black lines. Calculated steady-state K_D are shown as mean values from triplicate injections, and $\pm$ indicates SD. The data were fit to a steady-state model of binding. No fit or resulting K_D could be determined from the data obtained for the gir1^{C71A} mutant. See [Table 1](#).

(B) Steady-state affinity responses (circles) in **(A)** along with corresponding fits for each series are shown as concentration versus response plots (colored lines). Error bars represent the standard deviation of three independent injections.

Table 2. Steady-state binding affinities of GIR1 to ATML1 STAD and atml1 STAD^{D589L}.

K_D derived from SPR data presented in [Figures 3A-B](#). K_D values are reported alongside the mean and standard deviation (SD) for three injection series replicates (n = 3). The data were fit to a steady-state model of binding. No fit or resulting K_D could be determined for the gir1^{C71A} mutant. Asterisk indicates that the K_D for the GIR1 over mutant STAD^{D589L} binding is significantly higher than the K_D for the GIR1 over wild-type STAD binding (t-test, two-tailed distribution, equal variance, P-value = 0.011).

K _D (μM)	GIR1 over STAD	GIR1 over STAD ^{D589L}
Replicate 1	4.08	7.95
Replicate 2	4.49	8.24
Replicate 3	6.04	9.93
Mean	4.87	8.71 *
SD	1.03	1.07

References

1. Prachi Jain, D.A., and Satish C Bhatla (2016). Surface Plasmon Resonance Based Recent Advances in Understanding Plant Development and Related Processes. *Biochemistry and Analytical Biochemistry* 5. 10.4172/2161-1009.1000300.
2. Cosma, M.P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. *Cell* 97, 299–311. 10.1016/s0092-8674(00)80740-0.
3. DelRosso, N., Suzuki, P.H., Griffith, D., Lotthammer, J.M., Novak, B., Kocalar, S., Sheth, S.K., Holehouse, A.S., Bintu, L., and Fordyce, P. (2024). High-throughput affinity measurements of direct interactions between activation domains and co-activators. *bioRxiv* 2024.08.19.608698. 10.1101/2024.08.19.608698.
4. Blomberg, J., Aguilar, X., Brännström, K., Rautio, L., Olofsson, A., Wittung-Stafshede, P., and Björklund, S. (2012). Interactions between DNA, transcriptional regulator Dreb2a and the Med25 mediator subunit from involve conformational changes. *Nucleic Acids Res* 40, 5938–5950. 10.1093/nar/gks265.
5. Lee, M.S., Lim, K., Lee, M.K., and Chi, S.W. (2018). Structural Basis for the Interaction between p53 Transactivation Domain and the Mediator Subunit MED25. *Molecules* 23. 10.3390/molecules23102726.
6. Cress, W.D., and Triezenberg, S.J. (1991). Critical structural elements of the VP16 transcriptional activation domain. *Science* 251, 87–90. 10.1126/science.1846049.
7. Aguilar, X., Blomberg, J., Brannstrom, K., Olofsson, A., Schleucher, J., and Bjorklund, S. (2014). Interaction studies of the human and Arabidopsis thaliana Med25-ACID proteins with the herpes simplex virus VP16- and plant-specific Dreb2a transcription factors. *PLoS One* 9, e98575. 10.1371/journal.pone.0098575.
8. Sandholzer, J., Hoeth, M., Piskacek, M., Mayer, H., and de Martin, R. (2007). A novel 9-amino-acid transactivation domain in the C-terminal part of Sox18. *Biochem Biophys Res Commun* 360, 370–374. 10.1016/j.bbrc.2007.06.095.
9. Piskacek, S., Gregor, M., Nemethova, M., Grabner, M., Kovarik, P., and Piskacek, M. (2007). Nine-amino-acid transactivation domain: establishment and prediction utilities. *Genomics* 89, 756–768. 10.1016/j.ygeno.2007.02.003.
10. Mukherjee, T., Subedi, B., Khosla, A., Begler, E.M., Stephens, P.M., Warner, A.L., Lerma-Reyes, R., Thompson, K.A., Gunewardena, S., and Schrick, K. (2022). The START domain mediates Arabidopsis GLABRA2 dimerization and turnover independently of homeodomain DNA binding. *Plant Physiol* 190, 2315–2334. 10.1093/plphys/kiac383.
11. Geisbrecht, B.V., Bouyain, S., and Pop, M. (2006). An optimized system for expression and purification of secreted bacterial proteins. *Protein Expr Purif* 46, 23–32. 10.1016/j.pep.2005.09.003.

Chapter 4 - Conclusions and future directions

Mechanisms of ATML1-mediated gene expression regulation

This work sought to define the molecular mechanism governing regulation of target gene expression and identify downstream developmental effects of the HD-Zip IV TF ATML1, a key regulator of the plant epidermis in *Arabidopsis*. A major focus of the present study was to elucidate the function of GIR1, a poorly characterized protein that was initially reported to interact with GL2¹. We found that GIR1 also interacts with other HD-Zip IV members, including ATML1, through a conserved cysteine motif predicted to form a zinc finger, while ATML1 promotes this interaction through the STAD domain. Although GIR1 was previously reported to bind the corepressor TPL², we discovered interactions with TPR1, TPR2, TPR3, and TPR4, with the putative N-terminal EAR motif of GIR1 critical for these associations. The moderate expression of TPL and TPR1, TPR2, and TPR4 in sepals, together with the weaker expression of TPR3³, suggests that GIR1 may preferentially bind corepressors in sepals according to their coexpression patterns ([Appendix Figure 8](#)). In this way, GIR1 functions as a bridge to connect ATML1 with TPL/TPR corepressors to assemble a transcriptional repression complex. We identified that GIR1 helps govern two developmental processes: (i) giant cell formation and (ii) GSL biosynthesis in sepals.

Another layer of complexity of ATML1-mediated regulation may exist via RNA interactions. TF and non-coding RNA interactions have recently been discovered and are proposed to aid in TF-chromatin stability⁴. If ATML1 binds non-coding regulatory RNA, identification of the interacting sequences may improve our understanding of ATML1-mediated gene regulation in a gene- and time-specific manner. By utilizing the individual-nucleotide

resolution UV-crosslinking immunoprecipitation (iCLIP)⁵, we could compare wild type, *atml1* mutant, and overexpressing *ATML1* sepal lines to identify these interacting sequences. Since GIR1 functions to regulate ATML1 activity at the post-translational level, one possibility is a DNA-TF-RNA complex may interfere with GIR1 binding, leading to increased activation of ATML1 target genes. Conversely, RNA may act to direct ATML1 to specific target genes based on environmental cues or act as a scaffold to either favor activator or repressor binding to fine tune expression regulation.

Cell- and tissue-specific GSL biosynthesis regulation through adaptor proteins

ATML1 was identified as a novel regulator of GSL biosynthesis, where GIR1 acts to repress ATML1 activity and subsequent GSL production. Aliphatic and indolic GSLs showed a two-fold increase in *gir1* sepals compared to wild type while several aliphatic GSL compounds in *atml1* were significantly reduced (4MSOB, 3MSOP, and 5MSOP). Subjecting sepals overexpressing *ATML1* for GSL quantification would more clearly define the ATML1 role in GSL biosynthesis⁶. Increased GSL compounds would be expected due to increased ATML1 activity. We've defined ATML1 as a regulator of GSL accumulation in sepals, and upon promoter analysis, several GSL biosynthesis genes were identified having L1 box elements in their promoters ([Appendix Table 4](#)). HD-Zip IV TFs bind L1 box elements⁷, indicating ATML1 may directly regulate these genes. To address this possibility, an electrophoretic mobility shift assay (EMSA) could be employed to detect the interaction between ATML1 and GSL biosynthesis gene promoters.

Our GSL quantification was exclusively tested in sepals and other flower tissues due to the RNA sequencing data displaying upregulation of GSL biosynthesis genes in sepals. However, GIR1 expression goes beyond sepals¹ and thus, assessing GSL levels in other tissues would

define GIR1 GSL regulation in a wider extent. Since several GSL biosynthesis genes have previously been established from RNA sequencing to be upregulated in *gir1* sepals ([Chapter 2](#)), several candidate gene expressions could be compared between *gir1* and wild type by RT-qPCR in seeds, siliques, leaves, stem, and roots. Not only does GIR1 natively express in these tissues¹, but they represent a large part of edible portions for commercial Brassicaceae vegetables. For example, flower buds for broccoli and cauliflower, seeds for cattle feed in canola and camelina, leaves for kale and arugula, and roots for radishes and turnips.

GIR1 is a member of a small adaptor protein family in *Arabidopsis*, which also includes *GIR2* and *SEID1*. Specifically, *GIR1* and *GIR2* share a 71% sequence similarity and were previously characterized as being functionally redundant¹. Preliminary data suggests *GIR2* can interact with ATML1 STAD [Schrack lab, unpublished]. Importantly, *gir1* sepals exhibit increased giant cells, but loss-of-function *gir2* and *sied1* mutants appear to exhibit normal giant cell production [Schrack lab, unpublished]. Perhaps *GIR2* is genetically redundant with *GIR1* in the regulation of GSL biosynthesis. By using HPLC-DAD to quantify GSL levels in *gir2*, *sied1*, *gir1;gir2*, *gir1;sied1*, *gir2;sied1*, *gir1;gir2;sied1* sepals and other tissues, potential additive effects and GSL regulation by other adaptor proteins may be revealed.

Due to *GIR1* regulating both giant cell formation and GSL biosynthesis in sepals, an intriguing possibility is that these processes are interconnected. Little is understood about the cell-specific regulation of GSLs. Since *gir1* is involved in regulating cell elongation and growth, *gir1* mutants and *GIR1* overexpression lines may serve as valuable tools to investigate potential crosstalk between the regulation of these pathways. Our MALDI mass spectrometry imaging did not allow for single cell resolution of GSL localization ([Chapter 2](#)), so utilizing a method with a

higher resolution such as static secondary ion mass spectrometry (SSIMS) imaging⁸ or a single-cell transcriptomics approach may be needed to address this question.

Structural effects of GIR1 zinc finger

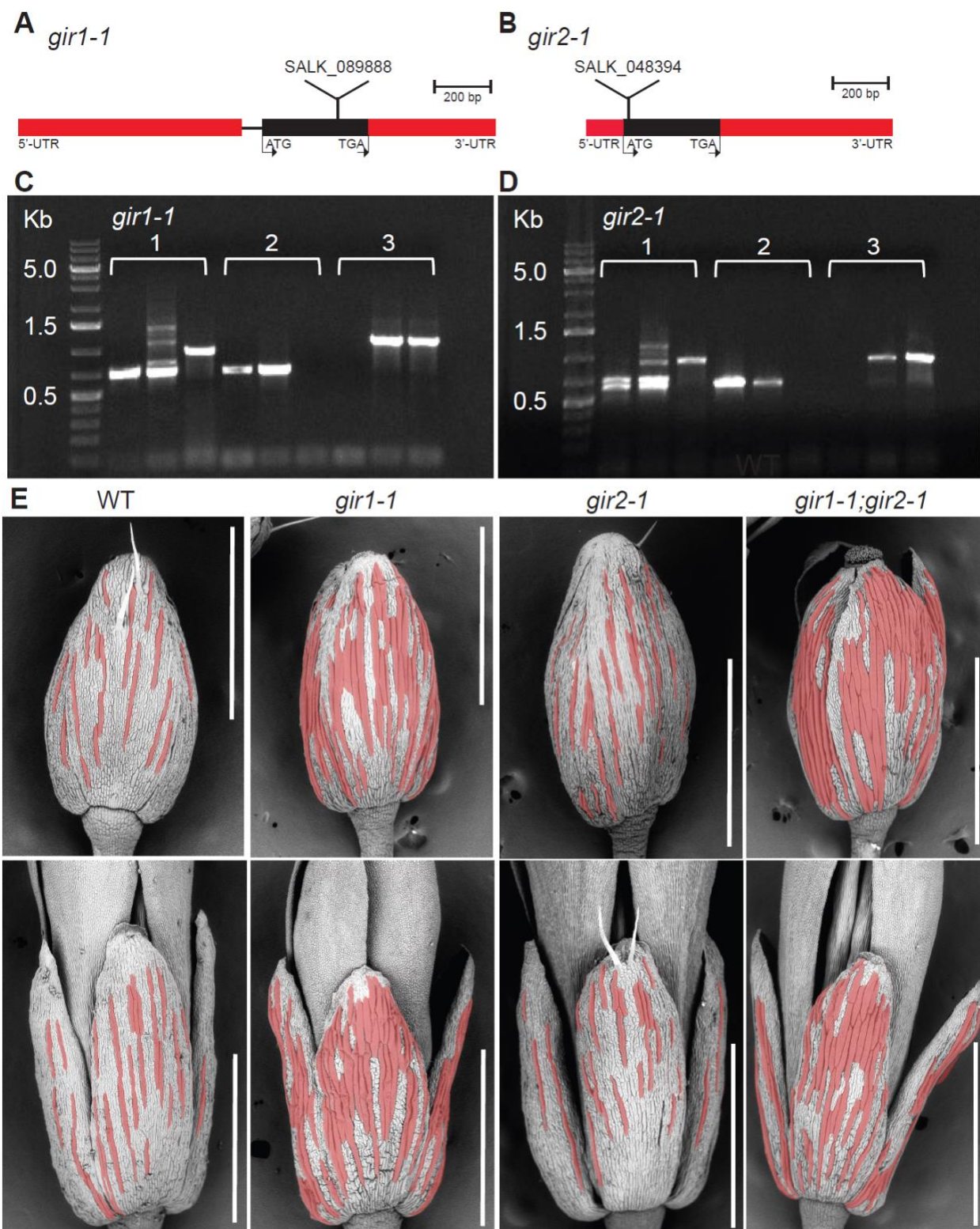
In vitro biophysical assays via SPR confirmed the direct interaction between GIR1 and ATML1 STAD ([Chapter 3](#)). The SPR analysis additionally indicated that cysteine residues in GIR1 that are predicted to form a zinc finger are critical for this interaction. The AlphaFold structure predicts GIR1 forms mostly intrinsically disorder regions resulting in a low confident model. However, structural approaches like nuclear magnetic resonance (NMR) could be employed to compare GIR1 and gir1^{C71A} to detect structural changes and effects on zinc binding. Although our previous attempts found GIR1 protein expression yields to be very low in NMR labeling media, further optimization or increased scaling could overcome this limitation. Additionally, recombinant co-expression of GIR1 with ATML1 STAD may aid in yield increase and provide information in uncovering a more complete model of residues required for interaction.

Overall, this thesis has examined the role of the GIR1-ATML1 interaction in plant development of the sepal epidermis of the plant model *Arabidopsis*. We identified a novel transcriptional repression complex involving GIR1, ATML1, and TPL/TPR proteins that control giant cell formation and GSL biosynthesis in sepals. This work improves our understanding of transcriptional regulation of GSL biosynthesis, which can be harnessed for future Brassicaceae crop development with enhanced micronutrient profiles and increased insect resistance. Moreover, this work lays the foundation for future experiments that focus on elucidating the cell- and tissue- specific adaptor protein-mediated GSL biosynthesis regulation and alternative mechanisms of ATML1 regulation of target gene expression.

References

1. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. I. Interaction with the repressor GLABRA2 and regulation of root hair development. *Biochem Biophys Res Commun* 488, 547–553. 10.1016/j.bbrc.2017.05.084.
2. Wu, R., and Citovsky, V. (2017). Adaptor proteins GIR1 and GIR2. II. Interaction with the co-repressor TOPLESS and promotion of histone deacetylation of target chromatin. *Biochem Biophys Res Commun* 488, 609–613. 10.1016/j.bbrc.2017.05.085.
3. Klepikova, A.V., Kasianov, A.S., Gerasimov, E.S., Logacheva, M.D., and Penin, A.A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *Plant J* 88, 1058–1070. 10.1111/tpj.13312.
4. Oksuz, O., Henninger, J.E., Warneford-Thomson, R., Zheng, M.M., Erb, H., Vancura, A., Overholt, K.J., Hawken, S.W., Banani, S.F., Lauman, R., et al. (2023). Transcription factors interact with RNA to regulate genes. *Mol Cell* 83, 2449–2463 e2413. 10.1016/j.molcel.2023.06.012.
5. Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* 17, 909–915. 10.1038/nsmb.1838.
6. Meyer, H.M., Teles, J., Formosa-Jordan, P., Refahi, Y., San-Bento, R., Ingram, G., Jonsson, H., Locke, J.C., and Roeder, A.H. (2017). Fluctuations of the transcription factor ATML1 generate the pattern of giant cells in the Arabidopsis sepal. *Elife* 6, e191131. 10.7554/eLife.19131.
7. Abe, M., Takahashi, T., and Komeda, Y. (2001). Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. *Plant J.* 26, 487–494.
8. Benninghoven, A. (1994). Chemical-Analysis of Inorganic and Organic-Surfaces and Thin-Films by Static Time-of-Flight Secondary-Ion Mass-Spectrometry (ToF-Sims). *Angew Chem Int Edit* 33, 1023–1043. DOI 10.1002/anie.199410231.

Appendix A - Supplemental Figures and Tables

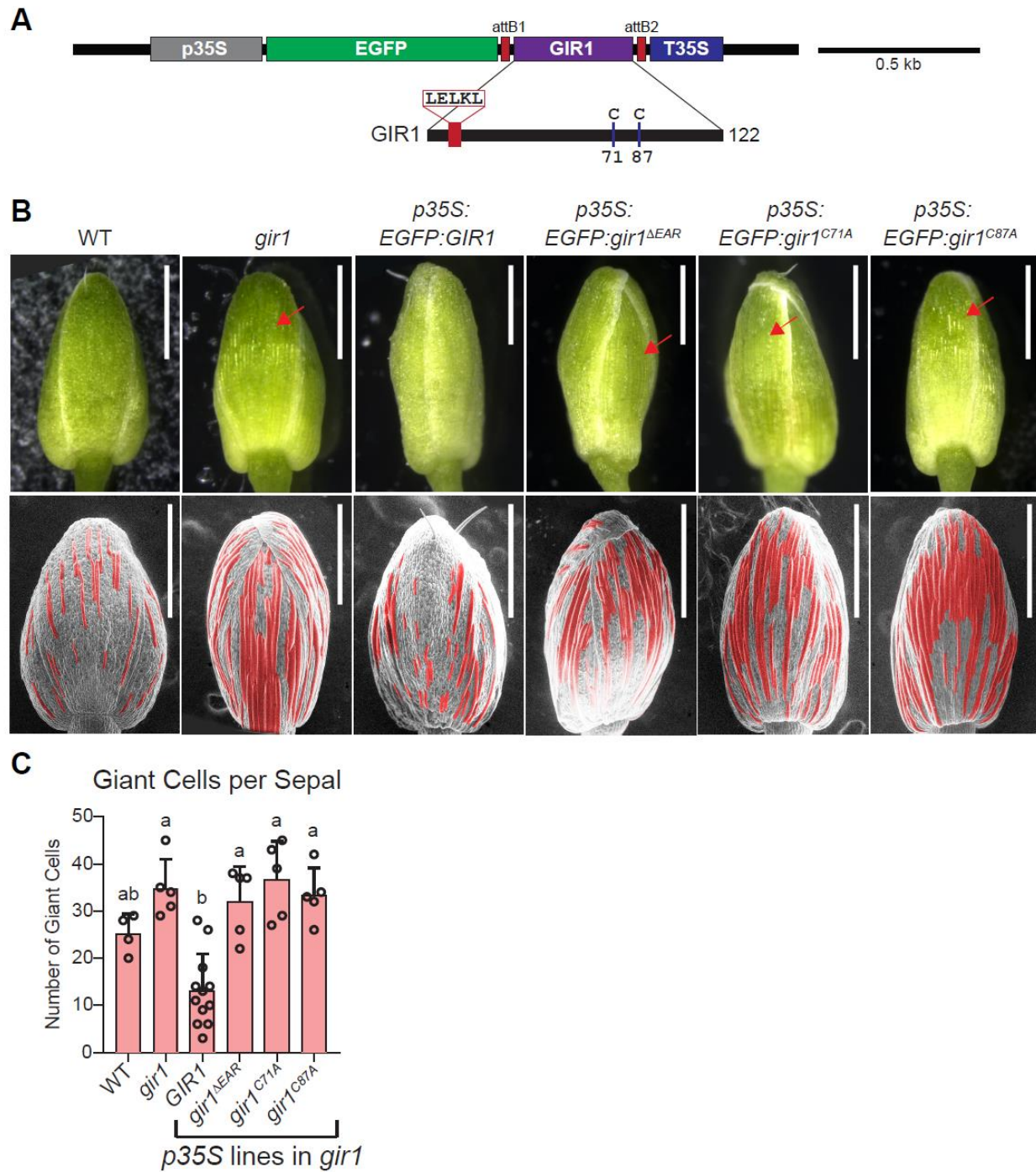


Appendix Figure 1. The *gir1* and *gir1;gir2* mutants exhibit similarly enhanced giant cell formation.

(A and B) Schematic diagrams illustrating genomic regions encoding GIR1/At5g06270 (left) and GIR2/At3g11600 (right). The positions of the respective T-DNA insertions are indicated for *gir1-1* and *gir2-1* alleles. Untranslated regions (UTRs) are depicted in red, exons and introns are shown by black thick and thin lines, respectively.

(C and D) Genotyping of *gir1-1* (SALK_089888) **(A)** and *gir2-1* (SALK_048394) **(B)** T-DNA insertion mutants given in **(A)** and **(B)**. **Panel 1:** Amplicons from three-primer PCR with corresponding gene-specific (RP and LP) and left-border (LB) primers; **Panel 2:** Amplicons from two-primer PCR with LP and the RP gene specific primers; **Panel 3:** Amplicons from the two-primer PCR with the gene-specific primers. In each panel from left to right are shown amplicons from the mutant, heterozygote, and wild-type DNA. Oligonucleotide primers are given in [Appendix Table 3](#).

(C) Giant cell phenotypes on sepals from WT in comparison to *gir1*, *gir2*, and *gir1;gir2* mutants. Giant cell formation is similarly enhanced in *gir1* and *gir1;gir2* mutants. Scanning electron microscopy (SEM) images of stage 12 (top panel) and stage 14 (bottom panel) flowers. Giant cells are falsely colored red. Bar = 1 mm. Quantification of giant cell numbers per sepal and giant cell width is given in [Figure 1C](#).

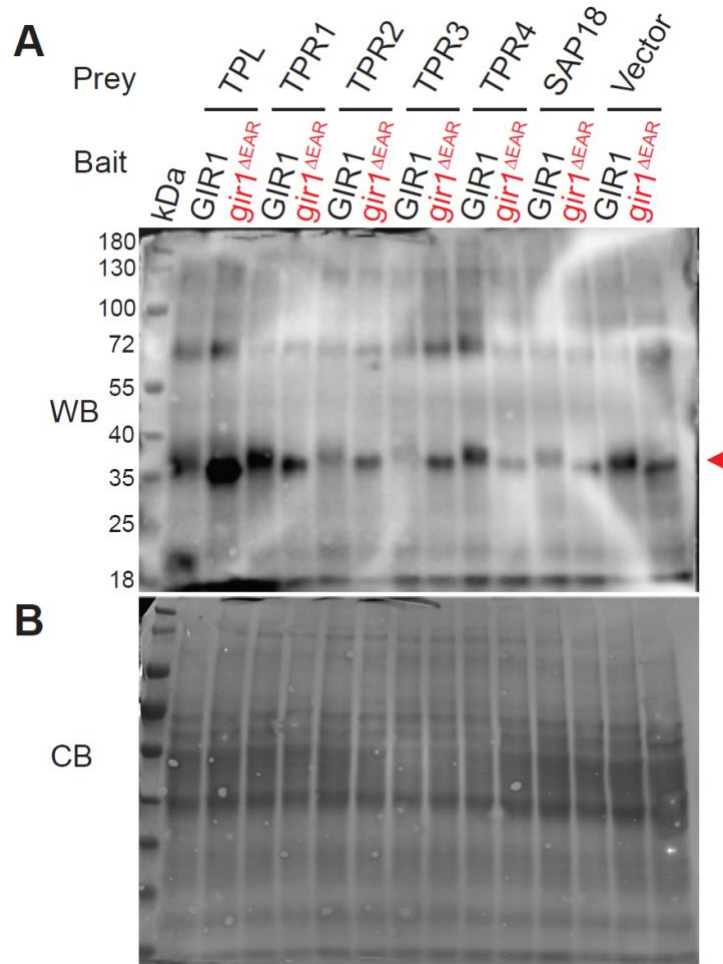


Appendix Figure 2. Overexpression of *GIR1* results in rescue of giant cell phenotype of *gir1* mutants.

(A) Schematic of the transgenic CaMV 35S (*pro35S*) promoter constructs for *GIR1* wild-type and mutants affecting the function of the EAR motif (*gir1^{ΔEAR}*) and predicted Zn finger (*gir1^{C71A}* and *gir1^{C87A}*).

(B) Stereo microscopy images of flower buds (top) and matching scanning electron microscopy images (bottom) with constructs indicated in **(A)**, in comparison to wild type and *gir1* mutant controls. Giant cells are false-colored red. Images were taken at stage 12 (closed flower). The wild-type *GIR1* construct rescued *gir1* phenotype while *gir1* mutants (*gir1*^{Δ*EAR*}, *gir1*^{C71A}, and *gir1*^{C87A}) failed to rescue. Bar = 500 μm.

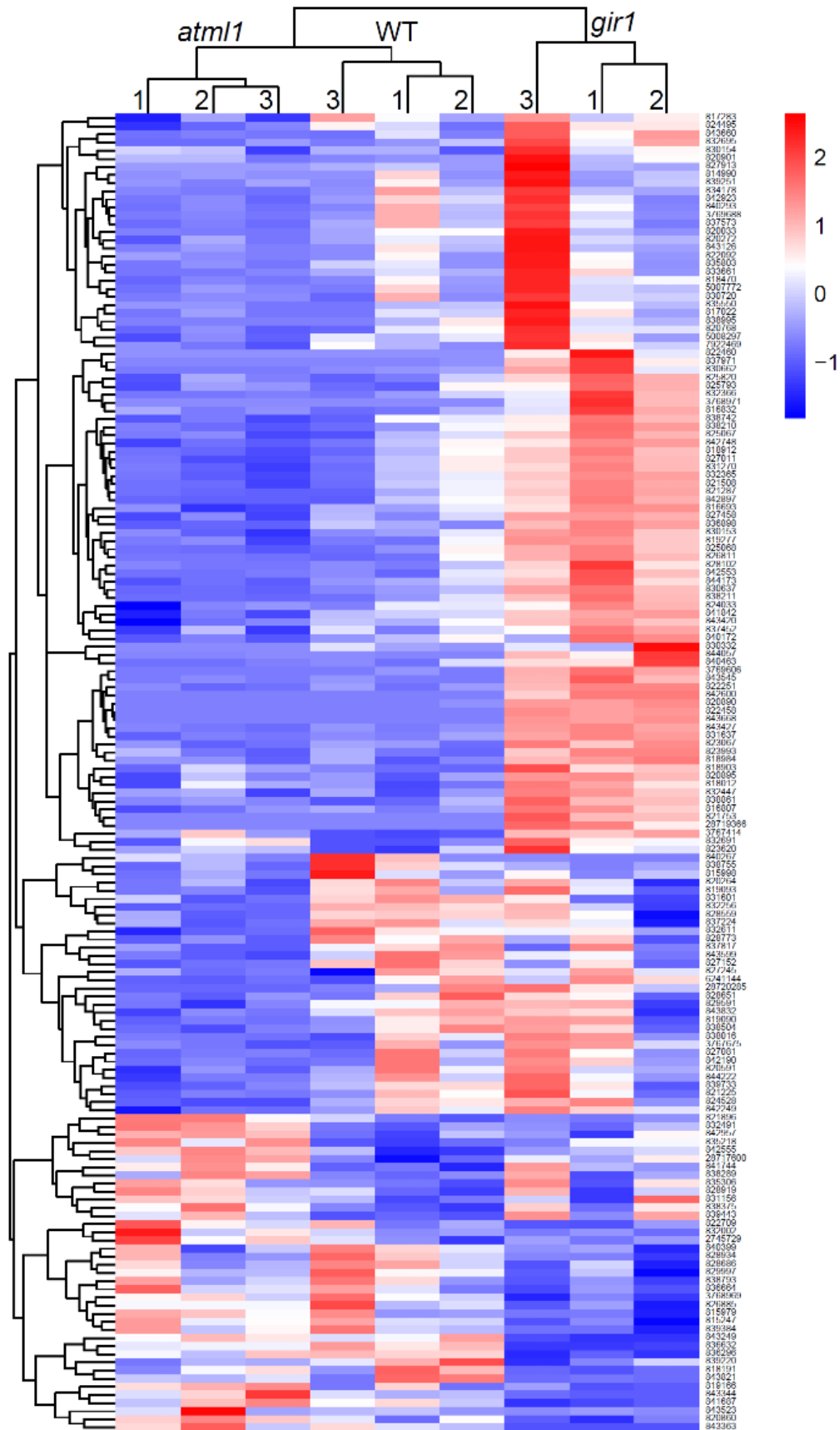
(C) Number of giant cells per sepal were quantitated for the genotypes shown in **(B)**. Standard deviations are indicated for n=4-12 biological replicates. Significant differences were assessed by one-way ANOVA, Tukey's test, and are indicated by letters (P value < 0.05).



Appendix Figure 3. Expression of GIR1 proteins in Yeast 2-Hybrid assays.

(A) Western blot (WB) confirmed the Yeast 2-Hybrid bait protein expression for the assays shown in [Figure 4C](#). Wild-type GIR1 and mutant *gir1*^{DEAR} fused to the GAL4 DNA Binding Domain (DBD) were detected using an anti-GAL4 antibody. The red arrowhead indicates migration of GIR1 and *gir1*^{DEAR} proteins, having an expected size of ~30 kDa for the GAL4 DBD fusion protein.

(B) The grayscale image (bottom) shows total protein loading after Coomassie Blue staining of the WB.

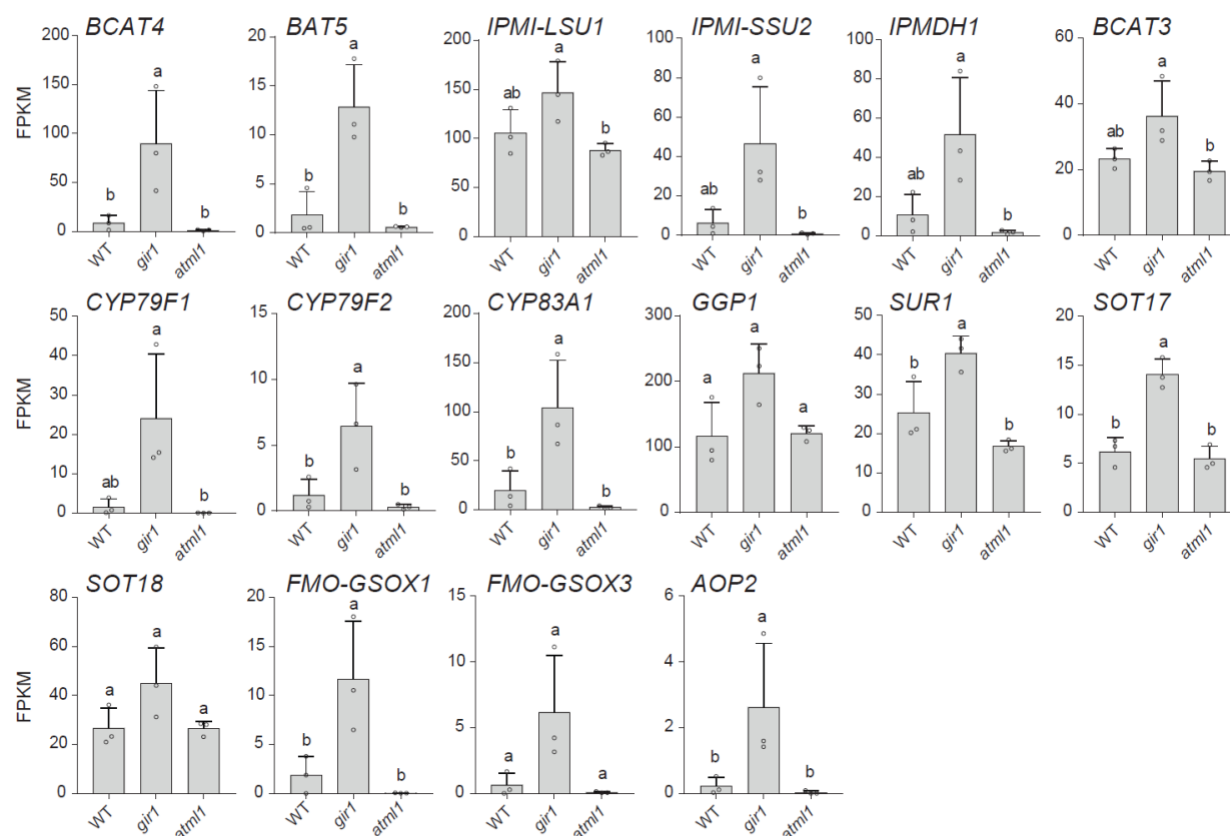


Appendix Figure 4. Heatmap from RNA-seq data showing three biological replicates for each genotype.

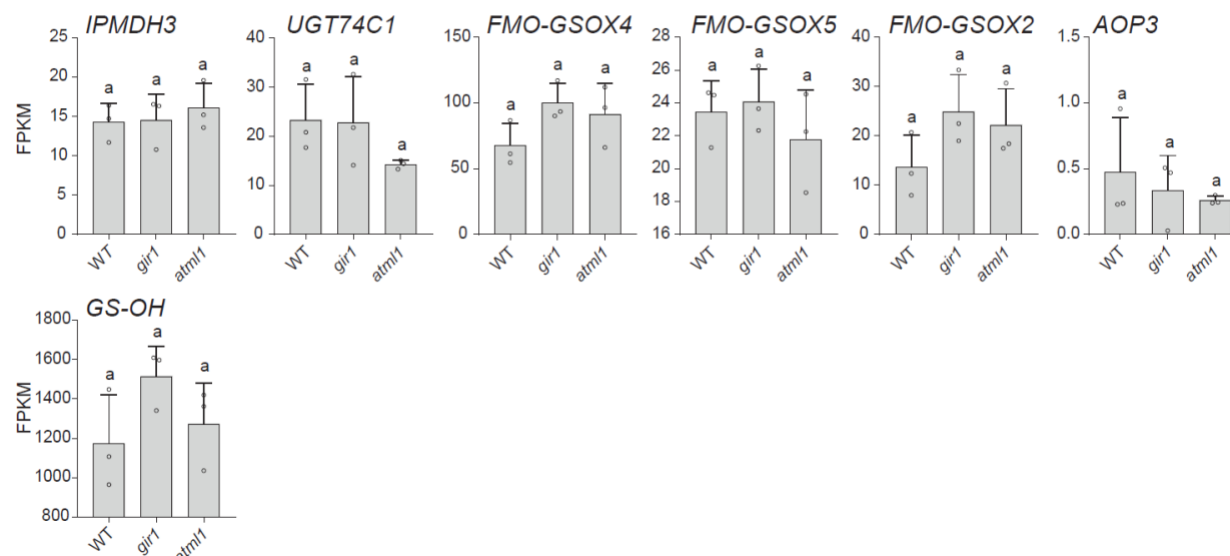
Overview of DEGs in sepals from WT, *gir1*, and *atm11* plants. Gene expression levels are shown for n=3 biological replicates for each genotype, graded from low (blue) to high (red) ($P < 0.05$).

A dendrogram depicting clustering of gene expression patterns is shown on the left of the heatmap. Gene IDs for the corresponding *Arabidopsis thaliana* genes are indicated on the right.

Aliphatic GSL biosynthesis - genes upregulated in *gir1*



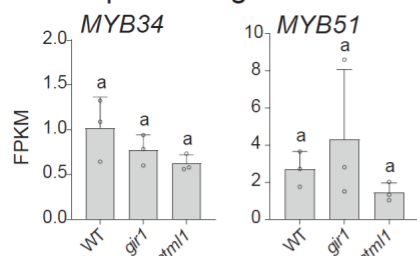
Aliphatic GSL biosynthesis - other



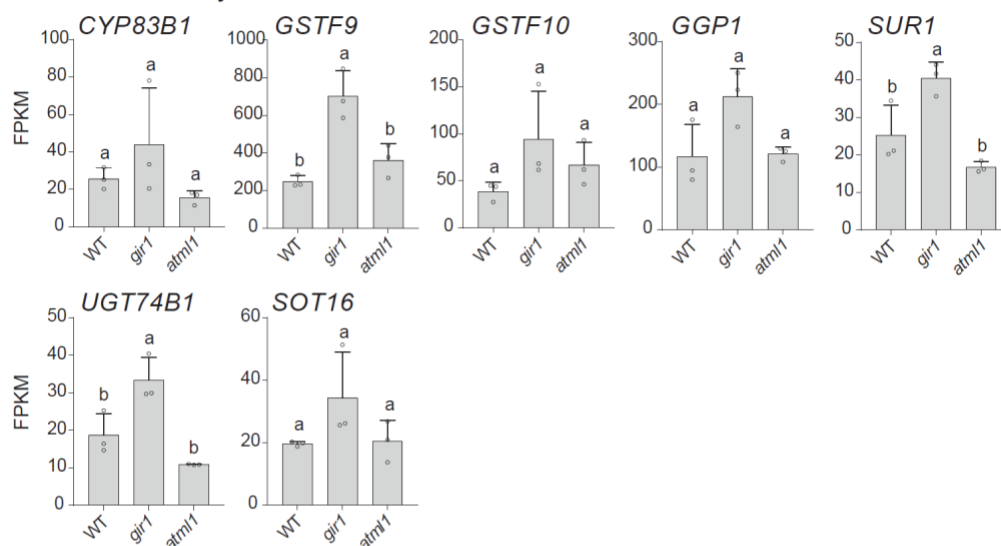
Appendix Figure 5. RNA-seq data for aliphatic glucosinolate biosynthesis genes. Related to Figure 5.

The top group of graphs displays sepal RNA-seq data for aliphatic GSL biosynthesis genes upregulated in *gir1* mutant sepals, in comparison to WT and *atml1* (*BCAT4*, *BAT5*, *IPMI-LSU1*, *IPMI-SSU2*, *IPMDH1*, *BCAT3*, *CYP79F1*, *CYP79F2*, *CYP83A1*, *GGP1*, *SUR1*, *SOT17*, *SOT18*, *FMOGSOX1*, *FMOGSOX3*, *AOP2*). The bottom group of graphs shows sepal RNA-seq data for aliphatic glucosinolate biosynthesis genes that are not regulated differentially in *gir1* mutants, in comparison to WT and *atml1*. (*IPMDH3*, *UGT74C1*, *FMO-GSOX4*, *FMOGSOX5*, *FMOGSOX2*, *AOP3*, *GS-OH*). FPKM, fragments per kilobase per million mapped reads. Genes are ordered within each group according to their approximate position in the glucosinolate biosynthesis pathway, from left to right, top to bottom. Significant differences between genotypes are marked by letters (one-way ANOVA, Tukey's test, $p < 0.05$).

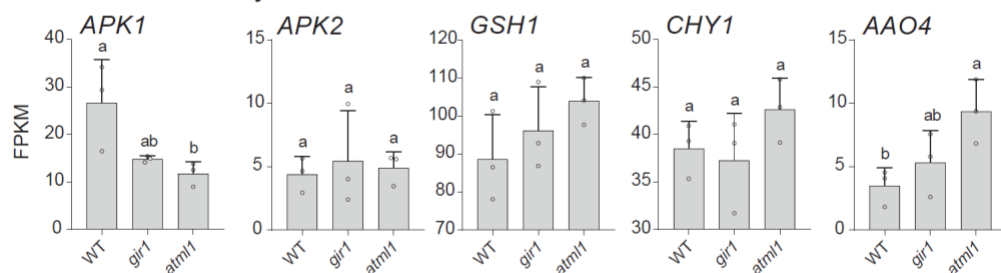
Transcriptional regulation of indolic GSL biosynthesis



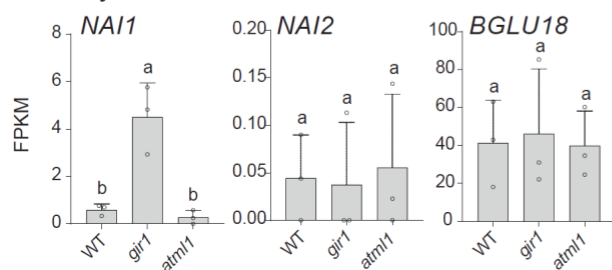
Indolic GSL biosynthesis



Benzenic GSL biosynthesis

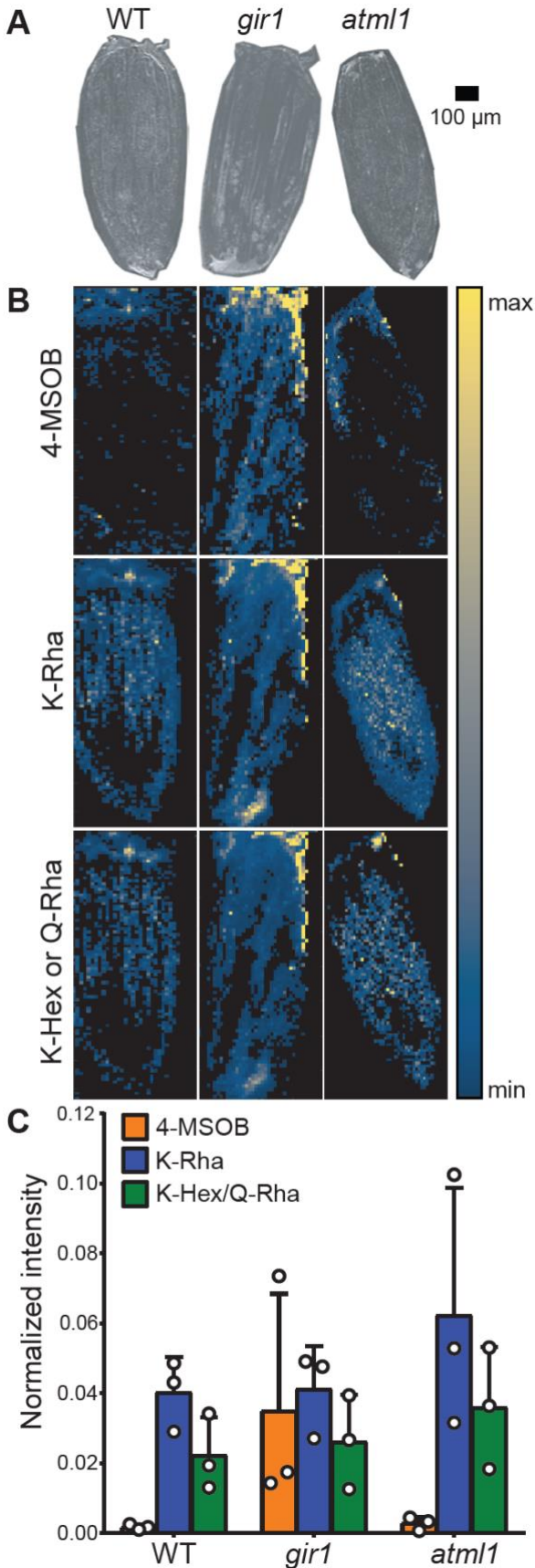


ER body - GSL-related



Appendix Figure 6. Graphs of RNA-seq data for indolic glucosinolate biosynthesis genes and other glucosinolate pathway genes. Related to Figure 5.

The top group of graphs displays sepal RNA-seq data for transcriptional regulators of the indolic GSL pathway (*MYB34*, *MYB51*). Transcripts of *MYB122* were not detected in 11/12 RNA-seq samples. The second group of graphs shows RNA-seq data for indolic glucosinolate genes, all of which appear upregulated in *gir1* mutant sepals (*CYP83B1*, *GSTF9*, *GSTF10*, *GGP1*, *SUR1*, *UGT74B1*, *SOT16*). Note that *GGP1* and *SUR1* act in biosynthesis of both aliphatic and indolic glucosinolates. The third group of graphs show sepal RNA-seq data for benzenic glucosinolate biosynthesis genes that are not upregulated in *gir1* mutants (*APK1*, *APK2*, *GSH1*, *CHY1*, *AAO4*). The bottom group of graphs show data for ER body and glucosinolate-related genes (*NAI1*, *NAI2*, *BGLU18*). FPKM, fragments per kilobase per million mapped reads. Genes are ordered within each group according to their approximate position in the glucosinolate biosynthesis pathway, from left to right, top to bottom. Error bars indicate SD. Significant differences between genotypes are marked by letters (one-way ANOVA, Tukey's test, $p < 0.05$).

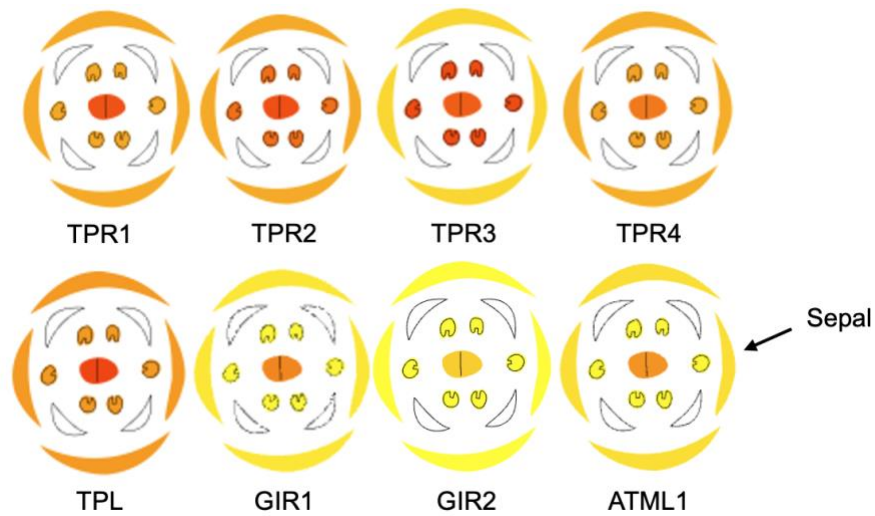


Appendix Figure 7. MALDI-MS imaging of the sepal epidermis indicates high 4-MSOB glucosinolate levels in the *gir1* mutants.

(A) Optical images of *Arabidopsis thaliana* sepals in WT in comparison to *gir1* and *atml1* mutants.

(B) MALDI-MS ion images of 4-methylsulfinylbutyl glucosinolate (4-MSOB; top), kaempferol rhamnoside (K-Rha) (middle), and kaempferol hexoside (K-Hex) or quercetin rhamnoside (Q-Rha) (bottom), shown to facilitate comparison with the glucosinolate distribution. Ion intensities were normalized to the base peak (m/z 425.0129, $[9\text{-AA+Au+Cl-H}]^-$), and the maximum color scale was set to 1.0 across metabolites and genotypes. The color scale bar represents normalized ion intensity, and all ions are displayed as $[M-H]^-$.

(C) Average normalized intensity of 4-MSOB, K-Rha, and K-Hex or Q-Rha in each genotype at defined regions of interest (ROIs; $n=3$ biological replicates). Error bars indicate SD.



Appendix Figure 8. Floral-specific expression of TPL, TPRs, and GIR1 in Arabidopsis.

The mRNA expression of young flower sepal, stamen, and carpel for TPR1, TPR2, TPR3, TPR4, TPL, GIR1, GIR2, and ATML1 were retrieved from the Arabidopsis eFP browser database.

Color scale: red (maximum) to yellow (minimum).

Appendix Table 1. Summary of expression of glucosinolate biosynthesis genes in *gir1* sepals.

The fold-change (FC) difference of the average expression between the three replicates of glucosinolate biosynthesis genes¹ in *gir1* was compared to wild type. Genes are ordered based on level of FC upregulation. Related to [Figure 5](#).

Expression Differences of Glucosinolate Biosynthesis Pathway Genes between <i>gir1</i> vs. WT			
Gene	Gene ID	Locus	<i>gir1</i> vs. WT FC
Aliphatic Pathway Genes			
MYB29	830662	At5g07690	16.0
CYP79F1	838211	At1g16410	14.9
AOP2	828102	At4g03060	12.0
MAM1	832365	At5g23010	10.9
BCAT4	821508	At3g19710	10.4
FMO-GSOX3	842553	At1g62560	9.5
MAM3	832366	At5g23020	7.6
IPMI SSU2	818912	At2g43100	7.3
BAT5	826811	At4g12030	6.9
FMO-GSOX1	842897	At1g65860	6.2
CYP79F2	838210	At1g16400	5.4
CYP83A1	827011	At4g13770	5.2
IPMI SSU3	825068	At3g58990	5.1
IPMDH1	831270	At5g14200	4.8
GSTU20	844173	At1g78370	4.5
SOT17	838440	At1g18590	2.3
GGP1	829176	At4g30530	1.8
FMO-GSOX2	842551	At1g62540	1.8
SOT18	843749	At1g74090	1.7
BCAT3	824130	At3g49680	1.6
SUR1	816585	At2g20610	1.6
MYB28	836263	At5g61420	1.6
FMO-GSOX4	842554	At1g62570	1.5
IPMI LSU1	826975	At4g13430	1.4
GS-OH	817083	At2g25450	1.3
IPMDH3	840006	At1g31180	1.0
UGT74C1	817736	At2g31790	1.0
FMO-GSOX5	837766	At1g12140	1.0
AOP3	828104	At4g03050	0.7
Indolic and Benzenic Pathway Genes			
CYP79B2	830154	At4g39950	7.2
CYP79B3	816765	At2g22330	3.8
GSTF9	817636	At2g30860	2.8
GSTF10	817637	At2g30870	2.4
GGP1	829176	At4g30530	1.8
UGT74B1	839022	At1g24100	1.8
SOT16	843750	At1g74100	1.8
CYP83B1	827011	At4g13770	1.7
SUR1	816585	At2g20610	1.6
Co-substrate Pathway Genes			
AAO4	839488	At1g04580	1.5

APK2	830153	At4g39940	1.2
GSH1	828409	At4g23100	1.1
CHY1	836724	At5g65940	1.0
APK1	815963	At2g14750	0.6

Appendix Table 2. Names and chemical properties of GSLs described in this study.

The GSLs are classified as aliphatic (top) or indolic (bottom).

Abbreviation	Full Name	Common Name	Chemical Formula	Exact Mass (Da)	Molecular Weight (g/mol)
Aliphatic GSL					
3MTP	3-Methylthiopropyl	Glucoberverin	C11H21NO9S3	407.03784477	407.5
4MTB	4-Methylthiobutyl	Glucoerucin	C12H23NO9S3	421.05349483	421.5
5MTP	5-Methylthiopentyl	Glucoberteroin	C13H25NO9S3	435.0691	435.5
6MTH	6-Methylthiohexyl	Glucosquerellin	C14H27NO9S3	449.0848	449.6
7MTH	7-Methylthioheptyl		C15H29NO9S3	463.10044502	463.6
8MTO	8-Methylthiooctyl		C16H31NO9S3	477.11609509	477.6
3MSOP	3-Methylsulfinylpropyl	Gluciberin	C11H21NO10S3	423.0328	423.5
4MSOB	4-Methylsulfinylbutyl	Glucoraphanin	C12H23NO10S3	437.0484	437.5
5MSOP	5-methylsulfinylpentyl	Glucolyssin	C13H25NO10S3	451.0641	451.5
6MSOH	6-Methylsulfinylhexyl	Glucosesperin	C14H27NO10S3	465.07970958	465.6
7MSOH	7-Methylsulfinylheptyl		C15H28NO10S3	478.08753461	478.6
8MSOO	8-Methylsulfinyloctyl	Glucobirsutin	C16H31NO10S3	493.11100971	493.6
3OHP	3-Hydroxypropyl		C10H19NO10S2	377.04503815	377.4
4OHB	4-hydroxybutyl		C11H21NO10S2	391.06068821	391.4
Benzyl	Benzyl	Glucotropaeolin	C14H19NO9S2	409.05012353	409.4
3BZOP	3-Benzoyloxypropyl	Glucomalcomiin	C17H23NO11S2	481.07125290	481.5
4BZOB	4-Benzoyloxybutyl		C18H25NO10S2	479.09198834	479.5
Indolic GSL					
I3M	Indol-3-ylmethyl	Glucobrassicin	C16H20N2O9S2	448.06102257	448.5
1OHI3M	1-hydroxy-indol-3-ylmethyl		C16H20N2O10S2	464.0559	464.5
4OHI3M	4-hydroxy-3-indolylmethyl		C16H19N2O10S2	463.04811215	463.5
1MOI3M	1-methoxy-indol-3-ylmethyl		C17H22N2O10S2	478.07158725	478.5
4MOI3M	4-methoxy-3-indolylmethyl		C17H22N2O10S2	478.07158725	478.5

Appendix Table 3. Oligonucleotides used in this study.

Primers for genotyping mutants.	
gir1_SALK_089888_F	CTTTTCTGGGGATCAGGATC
gir1_SALK_089888_R	AAGGCCTAAGCTCAAGCAATC
gir2_SALK_048394_F	CAACAAAATGGAATAACAGTTAAAGC
gir2_SALK_048394_R	CAGGTCTTCTTCTTGCCTCTG
atml1-3_F	CAGGCAGAAGAAAATCGAGAT
atml1-3_R	GAAACCAAGTGTGGCTATTGTT
LbB1.3 (T-DNA for gir1, 2)	ATTTTGCCGATTTCGGAAC
LBb1 (T-DNA for atml1-3)	GCGTGGACCGCTTGCTGCAACT
Primers for TOPO directional cloning. Additional CACC sites in forward primers required for directional cloning are bolded while stop codons are in blue. Alternative bases used for mutant generation are red.	
Topo_ATML1_F	CACC ATG TAT CAT CCA AAC ATG TTC G
ATML1_C_R_new	AT CGA TTA GGC TCC GTC GCA G
ATML1_START_F	CACC ATA CCT TCT GAG GCT GAT AAG
ATML1_START_R	CTA CAT GGA ACT GGC GAG CCG
ATML1_STAD_F	CACC GCC AGC AAC ATT CCG G
GIR1_cDNA_F	CACC ATG AGT CGA AGA AGT CC
GIR1_cDNA_R	TCA TCA GTT CCT TCG AGT C
TPL_F	CACC ATG TCT TCT CTT AGT AG
TPL_R	TCA TCT CTG AGG CTG ATC AG
Topo_TPR1_F	CACC ATG TCT TCT CTG AGC AG
Topo_TPR1_R	TCA TCT CTG AGG CTG GTC AG
atml1_K547E_F	CCG GGA AGA CCT CCA GGC ATC G
atml1_K547E_R	ATC ATC CAT GCT CTC TCG GGT CAT GAC C
atml1_D589L_F	CTT ATA CTT TCC AAT GGA GGC
atml1_D589L_R	CCA CTC GCT TCT TGA GTT TTC ATC
atml1_N593A_F	GCT GGA GGC TTG GTT CAA GAA ATG
atml1_N593A_R	GGAAAG TAT ATC CCA CTC GCT TC
atml1_N627A_F	GCC ATG TTG ATC TTA CAA GAA AGT TG
atml1_N627A_R	GCT CTG CCC TGA GTT CCC
atml1_E633L_F	TTA AGT TGT ACG GAC GCA TCA GGG
atml1_E633L_R	TTG TAA GAT CAA CAT GTT GCT CTG CCC
atml1_K733E_F	GAA CTC TCT CTC GGT TCA GTT GC
atml1_K733E_R	AGC GGT AGG AAC AGA GTC AAC AAG
gir1_R23L;R30L_F	TTG ATG GTT CGA TCT CCA AGC CTC TCA GCG
gir1_R23L_R	CCT CTG GCT TGA CGT AGG CGG CG
gir1_R30E_F	GAG TCA GCG ACG ACT TCA CCG ACC
gir1_R30L_R	GCT TGG AGA TCG AAC CAT CCT CCT CTG G
Primers for DNA assembly to clone into SR54 vector. Sequence overlapping with SR54 colored in blue while restriction sites are indicated in red. Bases used for alanine mutations are bolded.	
SR54_gir1_complement_F	AACGACGGCCAGTGCC AAGCTT GATGAAATGCATGAACGTACG
SR54_gir1_complement_R	CACCGCGGTGGAGCTC GGTACC GTTTTAGGCCTGTGATTAGC
gir1_C87A_F	GCC CCT AAA TGC AAA AGC ACC G
gir1_C87A_R_Ass	GT GCT TTT GCA TTT AGG GGC TTT AGG GTC GTC
gir1_C71A_F	GCT CCT CGT TGT CTC ATG TAC GTT ATG C
gir1_C71A_R_Ass	TACATGAGACAACGAGG AGC TCC GAC AAG AAC C
gir1_Del_EAR_F	AATCTCTCGCCGCTACG
gir1_Del_EAR_R	CTT TGG ACT TCT TCG ACT CAT
gir1_Del_21_18_EAR_F	ATGAGTCGAAGAAGTCCAAA GAA TCT CTC GCC GCC TAC G
GIR1_TPL_R	C TGG ATT TGG CCT TGG GTT CCT TCG AGT CTT GCG AC
TPL_C_Term_F	CCA AGG CCA AAT CCA GAT ATC
TPL_R	TCA TCT CTG AGG CTG ATC AG
TPL_C_GIR1_3UTR_F	CAGCCTCAGAGATGATGAAGGACCAGTAAAATCAAATCCTG

Primers for cloning cDNAs into plasmids for protein production. (Restriction sites, red; cDNA sequences, blue)	
pET-His6-MBP-TEV- Lic SspI ATML1 STAD F	ACCTGTACTTCCAATCC AATATT AGTCCTGAGGGGAGAAA
pET-His6-MBP-TEV- Lic SspI ATML1 STAD R	ATCCGTTATCCACTTCC AATATT TTAGGCTCCGTCGCA
pT7HMT BamHI GIR1 F	TACTTCCAGGGGTGCGACA GGATCC ATGAGTCGAAGAAGTCC
pT7HMT NotI GIR1 R	GTGGTGGTGCTCGAGT GCGGCCGC TCAGTTCCTTCGAGTCTTG

Appendix Table 4. L1 box elements in GSL biosynthesis gene promoters

Gene name	Number of L1 box promoter motifs	Locus	L1 box sequence	L1 box distance from transcription start site (bp)
MYB29	0	At5g07690		
MAM1	2	At5g23010	TAAATGTA	481, 642
MAM3	3	At5g23020	TAAATGTA	1579, 1822, 3281
CYP79B2	0	At4g39950		
IPMI1	0	At3g58990		
GSTU20	0	At1g78370		
AT5G26220	0	At5g26220		
CYP79B3	0	At2g22330		
AKN2	1	At4g39940	TAAATGTA	640
GSTF9	0	At2g30860		
ESP	0	At1g54040		
SOT17	0	At1g18590		
HMT3	2	At3g22740	TAAATGTA	199, 803

References

1. Sonderby, I.E., Geu-Flores, F., and Halkier, B.A. (2010). Biosynthesis of glucosinolates--gene discovery and beyond. *Trends Plant Sci* 15, 283–290. 10.1016/j.tplants.2010.02.005.

