Characterization of a novel family of repressors affecting many processes during *Arabidopsis thaliana* seed development

by

Ryan Hoy

B.Sc., The University of British Columbia, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Botany)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2022

© Ryan Hoy, 2022

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Characterization of a novel family of repressors affecting many processes during *Arabidopsis thaliana* seed development

submitted by Ryan Hoy in partial fulfilment of the requirements for

the degree of Master of Science

in Botany

Examining Committee:

Dr. Liang Song, Assistant Professor, Botany, UBC

Supervisor

Dr. Xin Li, Professor, Botany, UBC

Supervisory Committee Member

Dr. Lacey Samuels, Professor, Botany, UBC

Additional Examiner

Dr. Shawn Mansfield, Professor, Head of Botany, UBC

Additional Examiner

Additional Supervisory Committee Members:

Dr. Abel Rosado, Associate Professor, Botany, UBC

Supervisory Committee Member

Abstract

The seed is a vital component of the plant life cycle and an innovative strategy that plants use to reproduce and survive through extended periods of adverse environmental conditions. The processes involved in the development of the seed are highly complex, and one novel clade of genes that was identified to play a role in controlling the genetic network of seed development is the *DYNAMIC INFLUENCER OF GENE EXPRESSION* (*DIG*) clade. Using genetic strategies including Yeast 2 Hybrid and a mutagenesis screen, I worked to elucidate the mechanisms behind the *DIG* clade's regulation of this genetic network. My proposed model shows that through protein-protein interactions, transcription factors can recruit DIG proteins to target genes, and in turn DIGs interact with VP1/ABI3-LIKE 2 (VAL2) which recruits the Polycomb Repressive Complexes 1 and 2 to regulate the target genes expression through chromatin modifications.

Lay Summary

The seed is an important evolutionary adaptation that many plants utilize to ensure survival through adverse conditions, and a timely germination when conditions are appropriate. The development of the seed is highly complex and includes many different processes, which are controlled by the differential expression of thousands of genes. Here we characterize a novel family of genes that are expressed during seed development and believed to regulate the expression of many genes controlling these processes. My study includes a genetic screen using engineered yeast to search for proteins that work together with proteins in this novel family, in order to learn how these proteins interact. My study also includes a genetic screen in plants to identify new genes in the same pathway by using random chemical mutagenesis to create a large population of mutant plants. My findings could lead to a better understanding of the genes involved in seed development.

Preface

I was able to take high quality images of one of our mutants of interest using Transmission Electron Microscopy that shows a close look into the phenotype of interest at an organellar level.

With help from Liang Song, I set up and conducted a Yeast 2 Hybrid screen to look for protein interactors of DIG proteins. I found an interesting protein-protein interaction which provided a promising starting point for our gene family of interest's mechanism of action. Emma Laqua assisted in some of the yeast transformations and interaction tests. Many helpful suggestions and advice on Yeast 2 Hybrid were given by Gillian Dean from the Haughn Lab.

I set up a genetic screen to find suppressors of the *DIG1* overexpression phenotype and identified several putative mutants. This experiment will provide many promising projects for future students in the lab. The *DIG1* inducible overexpressor used as the background mutation was generated by Liang Song. Milad Alizadeh, Bailan Lu, and Liang Song helped with transplanting half of the population of M1 seedlings. Emma Laqua helped with harvesting and screening a several hundred M2 lines. Renwei Zheng did the ABA-sensitivity assay for some of the mutant lines (Figure 13, left side). Renwei Zheng performed Western Blot on three of the confirmed strong suppressor mutants and helped with the DNA extraction for Next Generation Sequencing of the first mutant. Dongeun Go helped harvest many plants when I was away. Several helpful suggestions and advice on this project were given by Weijie Huang from the Zhang Lab.

Together with Milad Alizadeh, I co-first authored a review paper on the transcription control of seed maturation by transcription factors. The paper was written by Milad Alizadeh

and me with help from Liang Song. Figures 1, 2, and 3 were produced by Milad Alizadeh and me. Bailan Lu generated heat maps for Figure 2C and 3B. Figure 3 from this review paper is used in this thesis as Figure 1. This review has been published as Alizadeh, M., Hoy, R., Lu, B., & Song, L. (2021). Team effort: Combinatorial control of seed maturation by transcription factors. Current Opinion in Plant Biology, 63, 102091.

Co-author on research paper on AtSDR4L with Milad Alizadeh, Bailan Lu, Emma Laqua, Dongeun Go, Liang Song, and collaborators in Zhizhong Gong's lab. Figure 1F and 1G for this paper were produced by me using my results from thesis objective 1. The experiments in this paper were conceived by MA, RH, and LS from the Song lab, and TW, JH, and ZG from the Gong lab. The experiments were conducted by MA, RH, DG, and LS from the Song lab, and TW and MB from the Gong lab. The paper was written by MA, BL, RH, EL, and LS from the Song lab, and TW and ZG from the Gong lab. Data analysis was done by MA, BL, RH, EL, DG, and LS from the Song lab, and TW, JC, MB, DT, and ZG from the Gong lab. This paper has been published as Wu, T., Alizadeh, M., Lu, B., Cheng, J., Hoy, R., Bu, M., ... & Song, L. (2022). The transcriptional co-repressor SEED DORMANCY 4-LIKE (AtSDR4L) promotes embryonic-to-vegetative transition in *Arabidopsis thaliana*. Journal of Integrative Plant Biology.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AA – Amino Acid

ABA – Abscisic Acid

ABI – ABSCISIC ACID INSENSITIVE

ABRC – Arabidopsis Biological Resource Center

AD – Activation Domain

AFL – ABI3, FUS3, LEC2

AL – ALFIN-LIKE

aLBS – atypical LUX Binding Site

ATP – Adenosine Triphosphate

BC – Back crossed

BD – Binding Domain

BiFC – Bimolecular Fluorescence Complementation

BIF – BioImaging Facility (UBC)

bZIP – BASIC LEUCINE ZIPPER TRANSCRIPTION FACTOR

CLF – CURLY LEAF

AtCLO – CALEOSIN (*Arabidopsis thaliana*)

Co-IP – Co-immunoprecipitation

CRISPR – Clustered Regularly Interspaced Short Palindromic Repeats

CRU – CRUCIFERIN

CTRL – Control

DAI – Days After Imbibition

DDO – Double Dropout (Media)

DER – Dow Epoxy Resin

DEX – Dexamethasone

DIG – DYNAMIC INFLUENCER OF GENE EXPRESSION

DIL – DIG1-LIKE

DMAE - Dimethylaminoethanol

DMSO – Dimethyl sulfoxide

DOG1 – DELAY OF GERMINATION 1

DTT – Dithiothreitol

ELF – EARLY FLOWERING

EM – Electron Microscopy

EMS – Ethyl Methanesulfonate

FIE – FERTILIZATION-INDEPENDENT ENDOSPERM

FUS3 – FUSCA 3

GA – Gibberellic Acid

HAG – HISTONE ACETYLTRANSFERASE OF THE GNAT FAMILY

HDA – Histone Deacetylase

HDAC – Histone Deacetylation Complex

HDT – Histone Deacetylase

HiFi – High Fidelity

HIS – Histidine

AtHSD1 – HYDROXYSTEROID DEHYDROGENASE 1 (*Arabidopsis thaliana*)

HSP – HEAT SHOCK PROTEIN

H2AK121ub – Histone 2a Lysine 121 monoubiquitinylation

H3ac – Histone 3 acetylation

H3K27me3 – Histone 3 Lysine 27 trimethylation

LAFL – LEC1, ABI3, FUS3, LEC2

LEC – LEAFY COTYLEDON

LEA – LATE EMBRYOGENESIS ABUNDANT

LEU – Leucine

LiAc – Lithium Acetate

LDL – LYSINE-SPECIFIC HISTONE DEMETHYLASE

LS – Linsmaier & Skoog (Nutrient Media)

MEMEsuite – Multiple Em for Motif Elicitation

MSI1 – MULTICOPY SUPPRESSOR OF IRA1

NEB – New England Biolabs

NF-Y – NUCLEAR FACTOR Y

NGS – Next Generation Sequencing

NSA – Nonenyl Succinic Anhydride

OD600 – Optical Density at wavelength 600 nm

OLE – OLEOSIN

PCK1 – PHOSPHOENOLPYRUVATE (PEP) CARBOXYKINASE 1

PEG – Polyethylene Glycol

PE150 – Paired End 150bp

PHD-Like – Plant Homeodomain-like

PKL - PICKLE

PPI – Protein-Protein Interaction

PP2C – PROTEIN PHOSPHATASE 2C

PRC – Polycomb Repressive Complex

PSV – Protein Storage Vacuole

PYR/PYL – PYRABACTIN RESISTANCE / PYRABACTIN RESISTANCE 1-LIKE

RCAR – REGULATORY COMPONENT OF ABA RECEPTOR

RFO – Raffinose Family Oligosaccharide

RGA/RGL – REPRESSOR OF GA / RGA-LIKE

ROS – Reactive Oxygen Species

SAP18 – SIN3 ASSOCIATED POLYPEPTIDE P18

SDR4 – SEED DORMANCY 4 (*Oryza sativa*)

AtSDR4L – SEED DORMANCY 4-LIKE (*Arabidopsis thaliana*)

SESA – SEED STORAGE ALBUMIN

SnRK – SUCROSE NONFERMENTING 1-RELATED PROTEIN KINASE

SRT - SIRTUIN

SSP – Seed Storage Protein

SUVH5 – SU(VAR)3-9 HOMOLOGUE 5

SWN - SWINGER

TAG – Triacyl Glycerol

TDO – Triple Dropout (Media)

TEM – Transmission Electron Microscopy

TF – Transcription Factor

TPL/TPR – TOPLESS/TOPLESS-RELATED

TRP – Tryptophan

VAL – VP1/ABI3-LIKE

VRN2 – REDUCED VERNALIZATION RESPONSE 2

WB – Western Blot

WOX – WUSHEL RELATED HOMEOBOX

WRI1 – WRINKLED 1

WT – Wildtype

WUS – WUSCHEL

YPDA – Yeast Peptone Dextrose Adenine (Media)

Y2H – Yeast 2 Hybrid

3-AT – 3-Amino-1,2,4-Triazole

Acknowledgements

I would like to give a big thank you to my supervisor Dr. Liang Song for accepting and allowing me to study in her lab. She is a wonderful and supportive teacher and is always there to help her students when needed. I would also like to thank all the other members of the Song lab for creating a great work environment. A special thank you to Milad Alizadeh for providing lots of help and training throughout my degree.

I would like to thank the members of my supervisory committee, Dr. Xin Li, Dr. Abel Rosado, and Dr. Lacey Samuels, for their suggestions, guidance, and support during my Master study.

Thank you to Miki Fujita, EunKyoung Lee, and Sam Livingston at the UBC Bioimaging Facility who helped me through my microscopy project. Thank you to Sean Shang for help with AlphaFold and NGS data analysis.

Dedication

To all of my family, for their unwavering love and support through my academic career, thank you.

Chapter 1 Introduction

1.1. The seed

The seed is an important structure thought to be vital to the terrestrialization of plants (Harris & Davies, 2016). Evolution of diverse seed shapes and structures allowed plants to disperse over great distances, ultimately resulting in rapid land colonization of plants (Harris & Davies, 2016). Seed plants can be categorized into the gymnosperms, which have naked seeds typically on the scales of cones, and angiosperms, which have their ovules encased in an ovary in their flowers. We will be focused on orthodox seeds of angiosperms, which are able to remain dormant for long periods of time, in contrast to recalcitrant seeds that cannot be stored (Berjak & Pammenter, 2002). Recalcitrant seeds may last days to months, whereas orthodox seeds can be stored for several years with the current oldest known seed to be germinated being a 2,000-year-old Judean date palm seed excavated from Israel in 2005 (Sallon et al., 2008). The major difference between recalcitrant and orthodox seeds is the desiccation tolerance of orthodox seeds that can account for the increased seed longevity (Berjak & Pammenter, 2002). The seed is quite complex and there are many steps involved in producing these important structures.

1.2.1. The stages of seed development

The seed stage makes up a vital and complex part in the plant life cycle, utilizing various strategies to overcome diverse obstacles. In developing orthodox seeds, several processes must occur to prepare the seed for a long period of dormancy where the seed must withstand harsh environmental conditions while it awaits the appropriate environmental conditions to initiate

germination (Bewley et al., 2012). After a pollination event, a seed will begin to form starting with a morphogenesis phase where differentiation and cell division produce an embryo and endosperm (Russell, 1992). During early seed development, the embryo is not ready for germination which keeps the seed under primary dormancy while it accumulates seed storage reserves and acquires desiccation tolerance (Bewley, 1997). After dehydration and dispersal, the mature seed is ready to germinate, but if the conditions are not right, the seed induces secondary dormancy until germination is triggered by exposure to water for imbibition and ideal temperatures (Finch-Savage & Leubner-Metzger, 2006). If there are problems with secondary dormancy, the seeds may start sprouting as soon as the seed is mature, but before ideal environmental conditions are met (Gubler et al., 2005). In nature, this could lead to a large problem if the conditions are not suitable for a seedling to start growing such as the onset of winter. In agriculture, this is known as preharvest sprouting and can lead to major losses in yield when harvesting grains (Gubler et al., 2005).

When germination initiates, the uptake of water will allow the cells to rehydrate, expand, and activate metabolic processes leading to the start of seedling establishment, which is signalled by the emergence of the radicle (Bewley, 1997). Early seedlings are composed of the radicle, the hypocotyl, and the cotyledons, or the embryonic root, shoot, and leaves, respectively. Seedling growth is powered from the seed storage reserves until cotyledon greening and expansion allow the seedling to start producing its own energy (Penfield et al., 2004; Josse & Halliday, 2008). In nature, a seedling's quick and efficient establishment can directly contribute to its fitness, as it may have to compete with its neighbours for sunlight and nutrients (Bergelson & Perry, 1989). The following sections of this chapter will provide a more

detailed account of the processes involved during the various stages of seed development, and some of the genes that help control them.

1.2.2. The early stages of seed development include morphogenesis and the accumulation of seed storage reserves

Seed development begins with a double fertilization event. A male pollen grain attaches to the stigma and releases two sperm cells, which creates a pollen tube going down the style to the ovules (Russell, 1992). One sperm cell will fertilize the egg cell to produce the embryo, while the other will fertilize the central cell giving rise to the endosperm (Russell, 1992). During early seed development, morphogenesis occurs where the embryo cells will divide and start to differentiate, and the seed will start to accumulate storage reserves. These storage reserves consist mainly of starch, seed storage proteins (SSPs) kept in protein storage vacuoles (PSVs), and oils in the form of triacylglycerols (TAGs) that are kept in oil bodies (Shewry et al., 1995). The SSPs help with early seedling growth as a source of nitrogen, carbon, and sulfur, and can be divided into two main groups, the 12S cruciferin proteins CRU1-3, and the 2S albumin proteins SESA1-5 (Shewry et al., 1995). TAGs are lipids that serve as high-energy storage compounds that help power the early seedling growth. In *Arabidopsis thaliana*, the oil bodies containing TAGs are encased in the structural proteins oleosin OLE1-5, caleosin AtCLO1-2, and steroleosin AtHSD1 (two identical copies), which help with oil body stability (Shao et al., 2019). These storage reserves will allow the plant to perform biological processes during the stages where it is unable to generate energy through photosynthesis.

1.2.3. Acquisition of desiccation tolerance is vital for achieving the state of dormancy

The other major step in preparing for dormancy is the acquisition of desiccation tolerance, which includes producing molecules and proteins that will protect the cells as they lose the majority of their water content (Angelovici et al., 2010). Prior to dormancy, chlorophyll in the seed must be degraded to avoid the accumulation of reactive oxygen species (ROS) from photosynthesis (Nakajima et al., 2012). Although chlorophyll is not essential for a seed to mature and subsequently germinate, as observed by the germination of albino mutants, photosynthesis in seeds does help power Adenosine Triphosphate (ATP) synthesis used for lipid biosynthesis by increasing O_2 concentrations (Sundberg et al., 1997; Baud & Lepiniec, 2010). To protect the cells from collapsing during dehydration, several molecules and proteins are formed to protect the organelles and keep the structural integrity of the cells. LATE EMBRYOGENESIS ABUNDANT (LEA) proteins help stabilize proteins and maintain membrane integrity during desiccation (Chakrabortee et al., 2007; Leprince et al., 2017). The accumulation of HEAT SHOCK PROTEINs (HSPs) and Raffinose Family Oligosaccharides (RFOs) during seed maturation are thought to also be involved in desiccation tolerance, but their exact function is still largely unknown (Leprince et al., 2017). After these all these layers of protection are in place, the seed undergoes dehydration. After maturation, when the seed is capable of germination under favourable conditions, primary dormancy is released and if the conditions are not ideal, the seed will maintain secondary dormancy until it is exposed to favourable conditions (Bewley et al., 2012).

1.2.4. Breaking dormancy

The breaking of secondary dormancy and the initiation of germination occurs when an ideal temperature is reached, and the seed coat can absorb water (Bewley et al., 2012). The uptake of water allows the cells to rehydrate and the metabolic processes to resume (Bewley et al., 2012). In *Arabidopsis thaliana*, the aleurone layer is the thin layer that remains from the endosperm and consists of the stored TAGs and storage proteins (Bethke et al., 2007). The conversion of TAGs into sucrose is catalyzed by PCK1 and allows the ensuing sucrose to be mobilized for skotomorphogenesis (Penfield et al., 2004). The switch from germination to seedling establishment occurs when the radicle emerges from the seed (Bewley, 1997). Once the cotyledons and hypocotyl are exposed to light, photomorphogenesis will begin as the cotyledons undergo greening from the synthesis of chlorophyll and then expands.

1.3. Phytohormones help control seed dormancy and germination

Phytohormones play a large role in regulating the gene networks underlying seed dormancy and germination. The two main phytohormones involved are abscisic acid (ABA) and gibberellic acid (GA), which play antagonistic roles (Liu & Hou, 2018). ABA plays a critical role in allowing plants to quickly respond to environmental stresses such as high salt or drought (Ohta et al., 2003; Santiago et al., 2009). ABA does this by influencing the expression of thousands of genes (Cutler et al., 2010). In addition, ABA is involved in many plant biological processes including stomatal opening and closing, and seed maturation and germination (Nakashima et al., 2009; Tal et al., 1970). ABA accumulates during seed maturation and induces many genes required for dormancy, while also inhibiting many genes required for germination (Cutler et al.,

2010). At the molecular level, *Arabidopsis thaliana* PYR/PYL/RCAR receptor proteins are bound by ABA, which in turn bind to and inhibit group A protein phosphatase 2Cs (PP2Cs) (Ma et al., 2009; Park et al., 2009; Santiago et al., 2009). One role of the PP2Cs is to inactivate Sucrose non-fermenting 1 - Related protein Kinase 2 (SnRK2) proteins by dephosphorylating them (Nakashima et al., 2009; Umezawa et al., 2009). The SnRK2s act to phosphorylate transcription factors involved in the ABA response (Kobayashi et al., 2005; Fujii et al., 2007). This signal pathway allows the plant to carry out complex processes quickly in response to environmental stimuli.

Antagonistically to ABA, GA accumulation is involved in the breaking of dormancy. Synthesis of GA initiates upon water uptake by the seed and promotes germination through the weakening of the tissues surrounding the embryo and activating genes involved in early growth (Ogawa et al., 2003). The presence of GA is important for germination as GA-deficient mutants can have severe defects in germination (Koorneef & Van deer veen, 1980). In the absence of GA, DELLA proteins including RGA1, RGA2, RGL1, RGL2, and RGL3 act to inhibit transcription factors involved in germination (Cao et al., 2005). When GA is released by the aleurone layer after imbibition, the DELLA proteins are ubiquitinated and under proteasomal degradation (Cao et al., 2005). DELLA protein RGL2 seems to have the largest impact on germination as *rgl2* mutants are resistant to inhibitors of germination and can suppress the phenotype of some germination deficient mutants (Lee et al., 2002). The balance between ABA and GA is the major factor that can allow for the transition between dormancy and germination.

1.4.1. The *LAFL* master transcription factors are highly involved in seed development

Each of the vital processes in seed development and germination are controlled by distinct genes sets that must be tightly regulated. The *LAFL* master transcription factor (TF) group consisting of *LEAFY COTYLEDON1* (*LEC1*), *ABSCISIC ACID INSENSITIVE3* (*ABI3*), *FUSCA3* (*FUS3*), and *LEAFY COTYLEDON2* (*LEC2*) are involved in regulating several processes during seed development (Boulard et al., 2017). *LEC1* encodes a subunit of the NF-Y complex which binds to CCAAT motifs, while *ABI3*, *FUS3*, and *LEC2* are B3-domain containing TFs that bind to Sph/RY-elements (Boulard et al., 2017). In addition to *abi3*, *abi4*, and *abi5* mutants show insensitivity to ABA during germination and display reduced dormancy (Finkelstein, 1994). To promote seed storage reserve accumulation, *LAFL* members can activate *WRINKLED1* (*WRI1*), a master activator of fatty acid biosynthesis genes, while *LEC1*, *FUS3*, and *ABI3* also activate the expression of the 12S cruciferin genes (Parcy et al., 1997; Cernac & Benning, 2004). For their role in promoting seed dormancy, *LAFL* members with the help of *bZIP67* activate the master regulator of dormancy, *DELAY OF GERMINATION1* (*DOG1*) (Bryant et al., 2019).

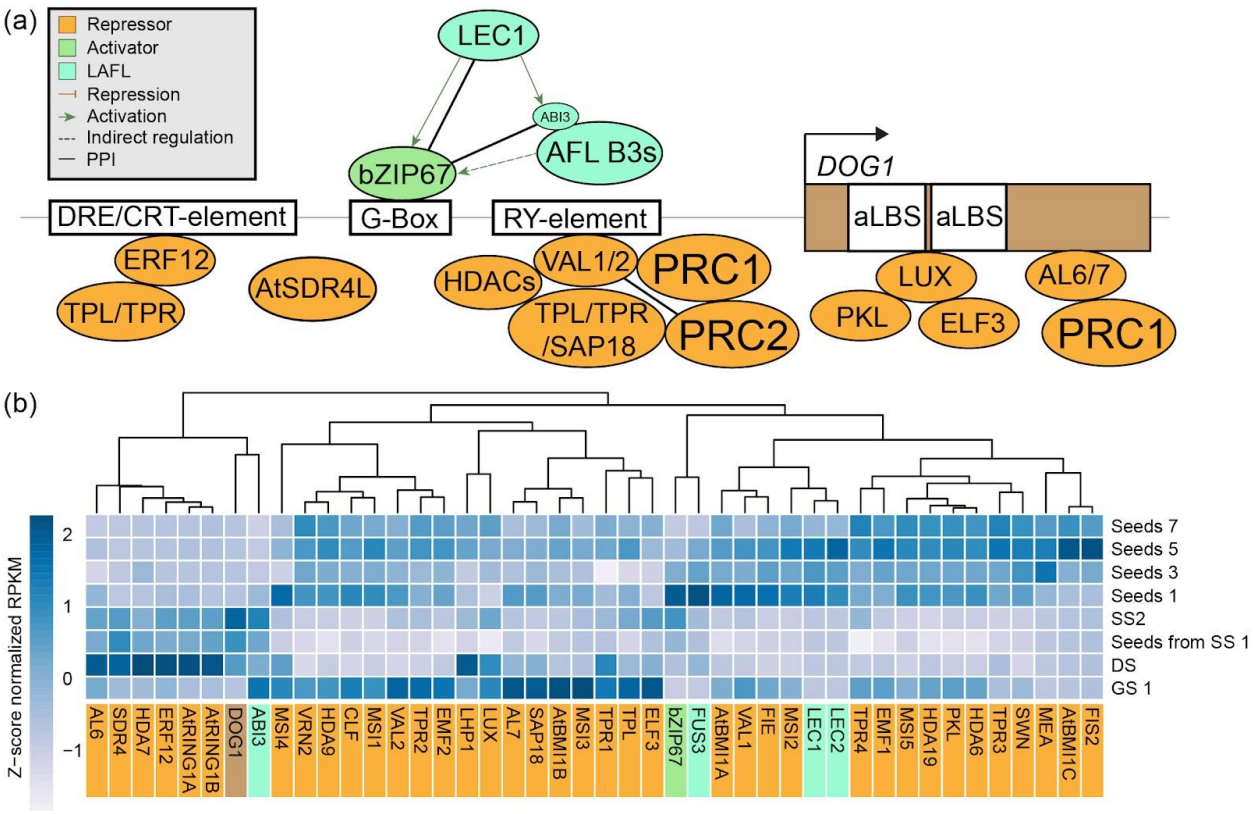

Figure 1. *DOG1* expression is controlled by many different factors.

(A) Transcriptional activation of *DOG1* controlled by bZIP67 with LAFL. Transcriptional repression controlled through PRC recruitment through VAL1/2 or AL6/7, by LUX with ELF3 and PKL, by ERF12 with TPL, or by AtSDR4L.

(B) Heatmap showing the expression patterns of activators and repressors of *DOG1*. GS1 = germinating seed 1; DS = dry seed; SS 1 = first senescent silique; SS 2 = second senescent silique.

1.4.2. Negative regulation of *DOG1* by AtSDR4L and VALs

In addition to the positive regulation from LAFL, *DOG1* is subject to negative regulation by several genes. *SEED DORMANCY4* (*SDR4*) was identified as a strong positive regulator of

dormancy in *Oryza sativa*, however, its *Arabidopsis thaliana* homolog (*AtSDR4L*) was shown to negatively regulate dormancy through the repression of *DOG1* during late seed maturation (Figure 1) (Alizadeh et al., 2021; Sugimoto et al., 2010; Cao et al., 2020; Liu et al., 2020; Winter et al., 2007). A transient trans-repression assay in *Nicotiana benthamiana* showed that *AtSDR4L* attached to a *GAL4* binding domain and was able to repress the expression of a reporter gene under the control of a promoter containing the *GAL4* upstream activation sequence showing that *AtSDR4L* functions as a repressor (Ma et al., 2021; Wu et al., 2022). In addition to the negative effects on seed dormancy, our lab has observed that *Atsdr4l* has several seedling phenotypes (Figure 2). The *Atsdr4l* seedlings have altered root and cotyledon growth (Figure 2). These phenotypes can all be observed in varying penetrance with increased expressivity in the absence of after-ripening or cold-stratification, or when germinated in the presence of exogenous sucrose (Wu et al., 2022). Staining the *Atsdr4l* seedlings with a neutral lipid stain such as Fat Red 7B results in dark staining of the hypocotyl which is indicative of the accumulation of TAGs (Figure 2B). Interestingly, there is another mutant that shows many similarities to the *Atsdr4l* phenotype. *VP1/ABI3-LIKE1* (*VAL1*) and *VP1/ABI3-LIKE2* (*VAL2*) are B3 domain-containing TFs and when both genes are knocked out, the *val1val2* double mutant shows a similar staining pattern with swollen hypocotyls and inhibited root and cotyledon growth (Tsukagoshi et al., 2007; Chen et al., 2020; Yuan et al., 2021). Furthermore, *VAL1* and *VAL2* redundantly act to negatively regulate *DOG1* as well (Chen et al., 2020; Yuan et al., 2021). An additional phenotype for the *val1val2* mutant is the formation of callus tissue at the root-hypocotyl junction (Tsukagoshi et al., 2007; Yuan et al., 2021).

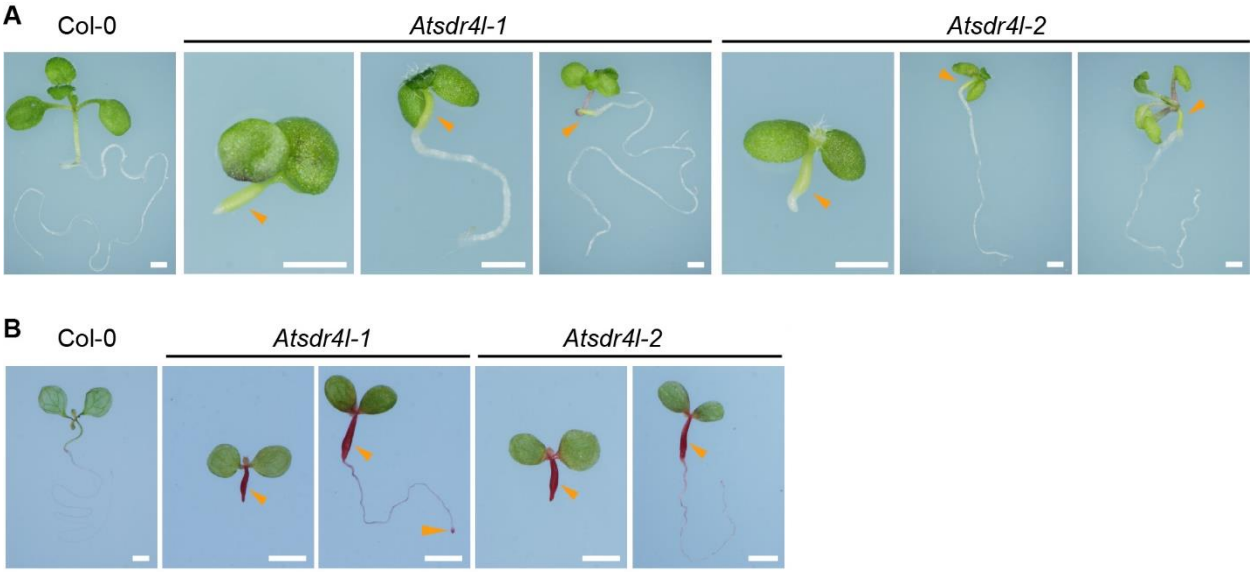

Figure 2. *Atsdr4l* mutant displays embryonic traits in seedlings.

(A) 10-day-old Col-0 and *Atsdr4l* seedlings.

(B) 10-day-old Col-0 and *Atsdr4l* seedlings stained with Fat Red 7B.

Orange arrowheads show embryonic traits in hypocotyls and root tips.

Images taken by and prepared by Milad Alizadeh.

1.4.3. VALs interact with the Polycomb Repressive Complexes to repress *DOG1*

VAL1 and VAL2 act redundantly to negatively regulate *DOG1* using the Polycomb Repressive Complexes (PRC1 and PRC2) to make histone modifications (Chen et al., 2020; Yuan et al., 2021). VAL1 and VAL2 have been found to physically interact with several components of PRC1 and PRC2, which are highly conserved in eukaryotes and involved in repressive chromatin modifications (Mozgova et al., 2015; Yang et al., 2013; Chen et al., 2020; Yuan et al., 2021). PRC1 is involved in the monoubiquitinylation of Histone 2A at Lysine 121 (H2AK121ub), while PRC2 is involved in the trimethylation of Histone 3 at Lysine 27 (H3K27me3) (Mozgova et al.,

2015). PRC1 and PRC2 can be recruited independent of each other, but PRC1 can also be involved in the recruitment of PRC2 (Baile et al., 2021). VAL1 can physically interact with PRC1 components AtBMI1A, AtBMI1B, and AtBMI1C leading to increased H2AK121ub and H3K27me3, however, VAL2's interactions were not tested (Yang et al., 2013). Both VAL1 and VAL2 can directly interact with CLF and SWN for recruitment of PRC2 for H3K27me3 (Yuan et al., 2021). Through these interactions with PRC components, VALs repress the expression of *DOG1* using histone modifications.

1.4.4. The *DIG* family of genes

AtSDR4L is paralogous to the *DYNAMIC INFLUENCER OF GENE EXPRESSION* (*DIG*) family genes that were identified in a screen for genes upregulated in response to ABA and include *DIG1*, *DIG2*, *DIG-like 1* (*DIL1*), *DIL2*, *DIL3*, and *DIL4* (Song et al., 2016). Similar to the *Atsdr4l* mutant, the *DIG* family genes are expressed during late seed maturation and plants overexpressing *DIG1* (DIG1-OE), in the presence of sucrose and ABA, results in inhibited roots and cotyledon growth, and swollen hypocotyls (Winter et al., 2007; Song et al., 2016). Similar to the *val1val2* double mutant, an additional phenotype apparent in the *DIG1*-OE is the formation of callus tissue around the root-hypocotyl junction, as we observed masses of undifferentiated tissue (Unpublished data, Song Lab). Unlike *Atsdr4l*, the hypocotyl of the *DIG1*-OE did not have intense staining when treated with Fat Red 7B, while it was apparent in callus tissue (Unpublished data, Song Lab). The DIG proteins have no known protein domains, but RNA-seq analysis showed that overexpression of *DIG1* or *DIG2* resulted in the differential expression of a few hundred genes (Song et al., 2016). DIG1 and DIG2 proteins accumulated more in the

promoters of downregulated genes, and that the DIG1 protein is localized to the nucleus and binds chromatin (Song et al., 2016). In addition, our lab has observed DIG1 does not appear to bind DNA directly (Unpublished data, Song Lab). Together, these observations suggest that DIG1 and possibly the other DIG members and AtSDR4L, act as transcriptional co-repressors involved in seed maturation. However, there are changes in some of the critical amino acid residues indicating that the DIG family may have different functions from AtSDR4L (Song et al., 2016).

1.5. Thesis objectives: uncovering the genetic network of seed development

The processes coordinating seed and early seedling development are tightly regulated by a complex genetic network, which must have strict control on the activation and repression of the genes involved. Some key players have been identified in this network, but there are many yet to be uncovered. Studying the many pathways that are involved in this network will provide crucial information about basic plant processes may allow us to better fine-tune agricultural crops.

Using genetic techniques, my research aims to identify new players in the complex network of seed and early seedling development, and to characterize how a novel family of genes function in this network at a molecular level. To study the processes involved in seed and early seedling development, the plant model organism, *Arabidopsis thaliana* will be used. *A. thaliana* is a great model organism for studying plant genetics since it is small, produces a large number of progeny, has a short generation time, a small diploid genome, and reproduces through self-pollination (Koornneef & Meinke, 2010). There are also many genetic and genomic

resources for *A. thaliana* including a fully sequenced genome, many available online community datasets, and a T-DNA mutant seed stock center (Alonso et al., 2003; Koornneef & Meinke, 2010). Additionally, it is closely related to *Brassica napus*, a crop grown for seed oils, so examining the processes involved in seed oils in *A. thaliana* could be easily applied to agriculture by identifying homologs in *B. napus* and then testing if those homologs perform the same biological function in *A. thaliana*.

Thesis objective 1: To identify which processes are affected by AtSDR4L/DIG proteins by mutant analysis of *Atsdr4l.*

Thesis objective 2: To identify the repressive mechanism used by AtSDR4L/DIG proteins by identifying novel protein interactors.

Thesis objective 3: To generate novel mutants that affect the *AtSDR4L*/*DIG* gene pathways using a mutagenesis screen with a DIG1 overexpressor.

Chapter 2 Identifying the mechanism used by AtSDR4L/DIG proteins.

2.1.1. How is the lipid distribution affected in the hypocotyls of *Atsdr4l* severe mutants?

One of the interesting phenotypes observed in the *DIG* clade was the accumulation of lipids in the hypocotyl of *Atsdr4l* seedlings that was shown by the intense staining of the neutral lipid stain, Fat Red 7B. To gain more insight into which processes may have been interrupted and to discover exactly how the lipid distribution and pattern were being affected at the cellular and organellar level, I used brightfield and transmission electron microscopy to examine sections of the hypocotyl in the severe *Atsdr4l-2* mutant.

2.1.2. Methods

2.1.2.1. Plant growth

Arabidopsis thaliana seedlings of *Atsdr4l-2* (SALK_203161) and Col-0 were grown on agar plates. To increase the severity of the mutant phenotype, freshly harvested, non-stratified seeds were used, and 1% sucrose was added to the media. The seedlings were grown in a growth chamber at 22°C on a 16h-light/8h-dark cycle. *Atsdr4l-2* mutant seedlings displaying a severe phenotype were selected to be prepared for microscopy.

2.1.2.2. Fixing the samples

Atsdrl4-2 and Col-0 seedlings were harvested 12 days after imbibition (DAI) to be fixed. The seedlings were fixed in a mixture of 4.0% formaldehyde and 2.5% glutaraldehyde in 0.1M sodium cacodylate to cross-link for 1 hour under vacuum, and then transferred to a solution of 2% osmium tetroxide in 0.1M sodium cacodylate for staining and membrane preservation for

an additional hour. Next, the samples were dehydrated with sequential application of 10%, 30%, 50%, 70%, 90%, and then 100% (x3) ethanol for 1 hour each, followed by dehydration with acetone overnight. Next, the samples were infiltrated with an epoxy resin specialized for plant tissues called Spurr's resin. The Spurr's resin was prepared 25g at a time by mixing 14.75g NSA, 10.25g ERL 4221, 3.58g DER 736, and 0.25g DMAE and placing the mixture in a vacuum for 15 minutes. The samples were incubated with acetone solutions containing 1 drop, 10%, 25%, 50%, and 75% Spurr's resin before infiltration with 100% Spurr's resin (x3). Each infiltration step was done for at least 2 hours on a rotator and then an overnight infiltration was done at 100%. After infiltration, the samples in resin were placed into moulds and polymerized in the oven at 66°C overnight (16 hours). The samples in resin blocks were then shaped into blocks with a point at the end that had a trapezoidal face containing the perpendicular hypocotyl for cross sections. The shaping of the blocks was completed using a hacksaw and razor blades as needed, and any misoriented samples were reoriented by cutting off the sample and using super glue to paste it onto an empty block of resin.

2.1.2.3. Sectioning

The Leica EM UC7 ultramicrotome was used for sectioning. Freshly prepared glass knives were employed to take 1 µm, 500 nm, and 200 nm sections for light microscopy, while a diamond knife was used to obtain 70 nm sections for TEM imaging. A cutting speed of 1.40 mm/s was used for sectioning with the glass knife, while 1.00 mm/s was used for the diamond knife. 200-mesh hexagonal copper TEM grids were used to collect the sections from the knife boat for TEM.

2.1.2.4. Staining

Toluidine blue (10 g toluidine blue and 10 g sodium borate in 1 L distilled water) was used to stain 200 nm, 500 nm, and 1 µm sections for brightfield microscopy. Using a loop, 200 nm, 500 nm, and 1 µm sections were taken from the glass knife boat and placed onto a glass slide. The glass slide was dried on a hot plate, and then one drop of toluidine blue stain was applied to the sections. The slide was placed back onto the hot plate until the border of the toluidine blue drop turned green. Excess toluidine blue was then carefully washed off with water and the glass slide was again dried on the hot plate, and then mounted for brightfield microscopy.

Uranyl acetate and lead citrate were used for positive staining of the TEM grids. 10 µL of 2% uranyl acetate was pipetted onto parafilm in a glass petri dish, and the TEM grid was placed sample side down on it. The dish was covered for 12 minutes and then the grid was washed by dipping into a beaker of distilled water 40 times and then 40 additional times in a new beaker of distilled water. The grid was dried by touching the edges with filter paper followed by air drying. Next, 10 µL of lead citrate was pipetted onto fresh parafilm between 9 pellets of sodium hydroxide in the glass petri dish. The TEM grid was placed sample side down on the lead citrate and covered for 6 minutes. Then the sample was washed in a beaker of distilled water by dipping 40 times and then another 40 times in a fresh beaker of distilled water. Water was absorbed from the edges with filter paper and then the grid was left to air dry.

2.1.2.5. Image acquisition and processing

The Olympus BX53 light/fluorescence microscope was used to take brightfield images of the 1 µm thick sections, while the Hitachi H7600 transmission electron microscope fit with an XR51 camera was used to acquire micrographs of 70 nm sections for Col-0 and *Atsdr4l*. For the Hitachi H7600, the accelerating voltage was set to 80.0 kV with magnification at 50,000x (Fig 3A, 4A, 4B) or 70,000x (Fig 3B). The camera was set to an exposure of 800 ms with gain, bin, and gamma all set to 1. All images were processed in ImageJ. Scale bars were added, and minor adjustments were made for better brightness and contrast levels. Minor cropping was used for all images to get rid of unnecessary white space.

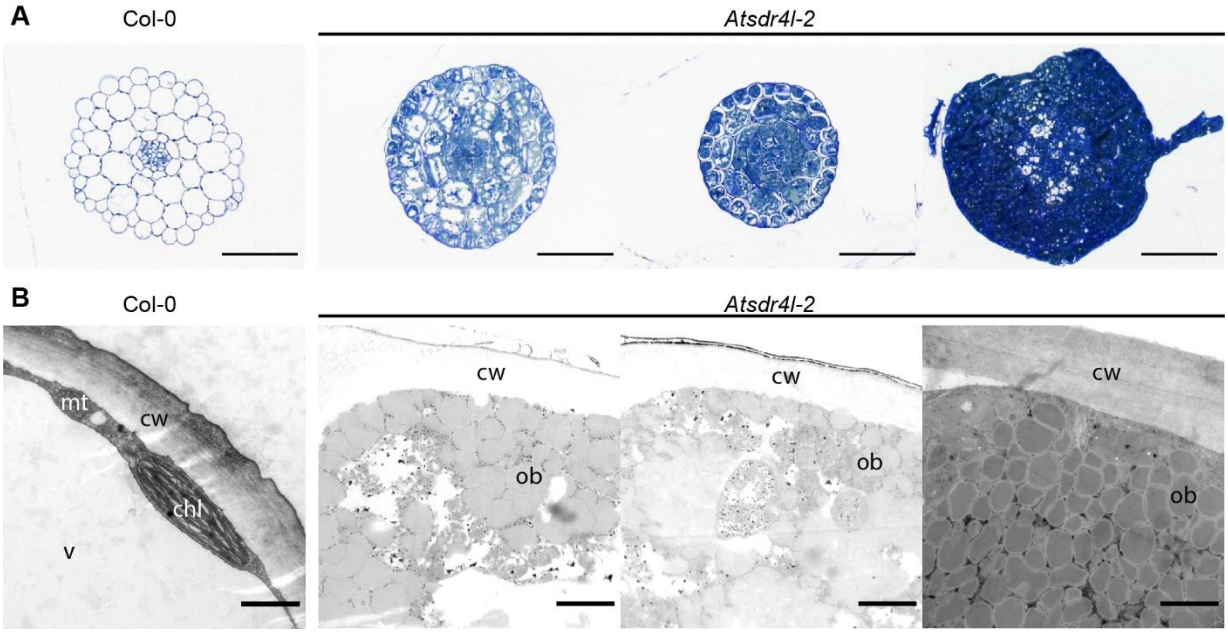

Figure 3. *Atsdr4l-2* severe mutant seedlings accumulate oil bodies in the hypocotyl.

(A) Toluidine blue stained 1 µm thick cross-sections of hypocotyls from 12-day-old seedlings of Col-0 and *Atsdr4l-2* viewed with brightfield microscopy. Scale bars = 100 µm.

(B) Transmission electron micrographs of 70 nm ultrathin cross-sections of hypocotyls from 12-day-old seedlings of Col-0 and *Atsdr4l-2* stained with uranyl acetate and lead citrate. cw = cell wall; v = vacuole; ob = oil body; ch = chloroplast; mt = mitochondrion. Scale bars = 1 µm.

2.1.3. Results

The severe *Atsdr4l-2* mutant seedlings display embryonic traits with lipids accumulating in their hypocotyls. In order to elucidate the specific processes involved that were interrupted, I used brightfield and transmission electron microscopy to view cross sections of the hypocotyl. From the brightfield images, we observed that the *Atsdr4l-2* mutant hypocotyls contained very dense cell contents compared to the Col-0 hypocotyl cells, which all had large central vacuoles (Figure 3A). A closer examination with the TEM, clearly showed that the dense cell contents in the *Atsdr4l-2* cells consisted mainly of oil bodies (Figure 3B). Together, these results show the cellular basis of the embryonic characteristics of the *Atsdr4l* seedlings, demonstrating that the hypocotyls of the severe mutant accumulate oil bodies. It is possible that the breakdown of the TAGs into sucrose is interrupted making the stored energy inaccessible to the seedling.

2.2. Identifying protein interactors for DIG proteins to illuminate their repressive mechanism

2.2.1. How do AtSDR4L and DIG family proteins repress target genes?

The combined observations of the RNA-seq of *DIG1* and *DIG2* with the trans-repression assay of *AtSDR4L* for these related proteins indicate that they are likely to be transcriptional co-repressors involved during seed maturation. However, the molecular mechanism remains unknown. Our lab has observed that these proteins do not directly bind DNA, suggesting there

must be other proteins involved that allows them to repress their targets. Identifying proteins that can physically interact with AtSDR4L and DIG proteins should help elucidate the mechanism(s) responsible for the repressive action. AtSDR4L and DIG family proteins have no known protein domains and therefore a large number of candidate proteins was necessary to identify an interaction. To identify protein-protein interactions for AtSDR4L and DIG family proteins, I used the Yeast 2 Hybrid (Y2H) system since it is a useful *in vivo* method that permits the rapid test of many different protein-protein interactions. One caveat is that false positive results are common, and therefore any identified interaction requires confirmation using another system. The similarities between AtSDR4L, the DIG family, and the VALs suggest that AtSDR4L could also use the PRC1 or PRC2 complex to repress *DOG1*, and it is possible the DIG proteins may use it to downregulate their target genes as well. For this reason, we started the Y2H assay by testing for protein-protein interactions with the PRC components, and PRC-related proteins.

2.2.2. Methods

2.2.2.1. Plasmid construction

The coding sequences of each of the DIG family proteins were cloned into a pENTR/D-TOPO vector. After confirming the correct sequence, the coding sequence was transferred using LR Recombination to the pDEST32 expression vector containing a *GAL4* DNA binding domain to be used as the bait protein for the Yeast 2 Hybrid screen. A list of candidate preys consisting of PRC components, PRC-related proteins, and histone modifiers was generated through literature search and bait plasmids that were available were ordered through the ABRC. A set of histone

modifiers was also acquired from the Qiao lab, University of Texas, Austin, and were also tested in the screen. The prey proteins were in one of pDEST22, pDEST-AD, or pEXP-AD502, all of which contain a *GAL4* activation domain.

After the initial screen, truncated proteins were created to narrow down interacting domains for novel protein-protein interactions. Initial truncations were established by dividing the protein in half, while further truncations were done based on motif scanning using MEMEsuite, and 3D protein prediction using AlphaFold. The truncated plasmids were cloned by amplifying the truncated DNA with Phusion High-Fidelity DNA Polymerase or NEB Q5 DNA Polymerase using the original bait and prey plasmids as templates. pDEST32 and pDEST22 were linearized using NEB NotI-HF and NEB AscI restriction enzymes to digest the plasmid overnight at 37°C. The new fragments were ligated into the pDEST32 and pDEST22 vectors using T4 DNA Ligase overnight at 16°C. Two domain swaps were completed using NEB HiFi DNA Assembly. The fragments were amplified using the same method as the truncations, but the primers had overlaps incorporated for NEB HiFi DNA Assembly for 3-fragment assembly. In the WUS domain swap, the swapped domain was incorporated by replacing it in the overlap region during primer design. The amount of DNA for each of the fragments and vector was calculated with NEBioCalculator to the molar ratios recommended by the NEB HiFi DNA Assembly protocol. A 20 µL reaction was incubated for 15 minutes at 50°C and then kept at -20°C until transformation into NEB 10beta chemically competent *E. coli*.

After the discovery of some protein-protein interactions and narrowing down the interacting domains, I wanted to see if the truncated proteins could complement the *Atsdr4l* mutant phenotype. Based on the Yeast 2 Hybrid results, I designed three different truncation

constructs. Truncations of *AtSDR4L* and the *AtSDR4L* native promoter were amplified from previously cloned and sequenced plasmids along with 3xHA and 3xFLAG epitope tags. The sequences were assembled and incorporated into a pCAMBIA1300 vector using NEB HiFi DNA Assembly with the same methods as previously described. The complementation constructs were then transformed into *Agrobacterium tumefaciens* C58 for transformation into *Arabidopsis thaliana Atsdr4l* CRISPR mutants. From a side-by-side comparison, Dongeun Go found the *Atsdr4l-4* (Wu et al., 2022) and *Atsdr4l-5* (generated by Liang Song) displayed more severe mutant phenotypes and were chosen for transformation with these constructs.

2.2.2.2. Yeast strains and culturing

The Y2H Gold strain was used to screen and test for protein-protein interactions. Y2H Gold yeast has several mutations that were helpful for transformation and testing protein-protein interactions. Y2H Gold has mutated *leu2* and *trp1,* which are used as transformation markers, and this strain cannot grow on media that does not contain leucine or tryptophan. pDEST32 has a functional *LEU2* gene and when expressed in the Y2H Gold, will allow the cells to grow on minimal media that does not contain leucine. Likewise, pDEST22 has a functional *TRP1* gene and when expressed in Y2H Gold, allows the cells to grow on minimal media missing tryptophan. Y2H Gold also has a mutated *his3* gene that makes it unable to synthesize histidine and the cells are unable to grow on minimal media that does not contain histidine. Y2H Gold does have a functional *HIS3* gene attached to a *GAL1* upstream activating sequence, which can be bound by a *GAL4* DNA binding domain and activated by a *GAL4* activation domain to drive expression of the *HIS3* and allow it to grow on the histidine deficient media.

2.2.2.3. Yeast transformation

Bait and prey plasmids were either co-transformed or sequentially transformed into the Y2H Gold strain. To prepare fresh competent cells for transformation, Y2H Gold cells from glycerol stocks were spread on 1x YPDA plates and incubated at 30°C for 3 days. Fresh colonies were taken from the YPDA plates, suspended in 10 mL 1x YPDA in a 250 mL Erlenmeyer flask and incubated overnight at 30°C with 200 rpm. Cells from the overnight culture were taken and resuspended in 20 mL of fresh 1x YPDA diluted to an OD600 of 0.1 – 0.15. The fresh culture was grown at 30°C with 200 rpm until the cells had doubled twice (OD600 = 0.4 – 0.6) to obtain yeast in the mid-log phase for optimal transformation efficiency. The culture was transferred to a 50 mL falcon tube and centrifuged at 3,000 rpm for 5 minutes. The cells were washed with water and then with a 0.1 M LiAc/1x TE solution using the same centrifugation settings. The cells were resuspended in the 0.1 M LiAc/1x TE solution to a concentration of ~ 10^9 cells per mL, and then 100 µL was aliquoted into 1.5 mL microcentrifuge tubes for each transformation. The tubes were centrifuged for 30 seconds at 10,000 rpm and the supernatant was removed. On ice, 400 ng of the plasmid DNA was mixed with 10 µL of prepared carrier DNA and transformation mix consisting of 240 µL PEG 3500 (50% w/v), 36 µL 1 M LiAc, 36 µL DTT, and 36 µL of water, and then gently mixed with the yeast cells. For co-transformations, 400 ng of each plasmid was added with 20 µL of carrier DNA instead. The cells were incubated at 30°C for 20 minutes, and then 10 mL of DMSO was added to the cells. The cells were gently mixed and then heat shocked at 42°C for 15 minutes. After heat shock, the cells were allowed to recover on ice for 2 minutes and then centrifuged at 10,000 rpm and the transformation mix removed. The

cells were resuspended in sterile water and a plated onto the appropriate drop out media and incubated upside-down at 30°C for 3 days.

Table 2.1. List of candidate proteins tested in Y2H

PRC1 Components	PRC2 Components	PRC-related	Histone Modifiers				
BMI1A	CLF	VAL1	HDA2	HDA14	HDT1	LDL1	HAG1
BMI1C	FIE1	VAL2	HDA5	HDA15	HDT2	LDL2	HAG2
RING1	VRN2	AL6	HDA9	HDA17	HDT3	SRT1	HAG3
	MSI1	AL7	HDA10	SUVH5	HDT4	SRT2	

2.2.2.4. Interaction screening

To screen for protein-protein interactions we employed the *HIS3* reporter gene in Y2H Gold. Activation of *HIS3* results in histidine biosynthesis and allows the yeast to grow on media that is lacking histidine, which is essential for Y2H Gold's survival. If the yeast can grow on the -His minimal media, it indicates that the two proteins interact with each other, and if the yeast cannot grow, it suggests there is no interaction. Fresh colonies from the transformations were inoculated into 4 mL of DDO/-Leu/-Trp and incubated overnight at 30°C with 225 rpm. 1 mL of the overnight culture was centrifuged and resuspended in sterile water adjusted to an OD600 of 1.0. A series of dilutions were made and 5 µL from each of undiluted, 1/10 dilution, and 1/100 dilution were plated onto TDO/-Leu/-Trp/-His plates. 3-AT, a competitive inhibitor of the product of the *HIS3* gene was added to reduce leaky expression and give a better view of the relative interaction strength. The set of plates included 0 mM, 0.1 mM, and 1.0 mM 3-AT. Negative controls were generated by transforming an empty pDEST22-AD vector with each bait protein, and an empty pDEST32-BD with each prey protein.

2.2.3. Results

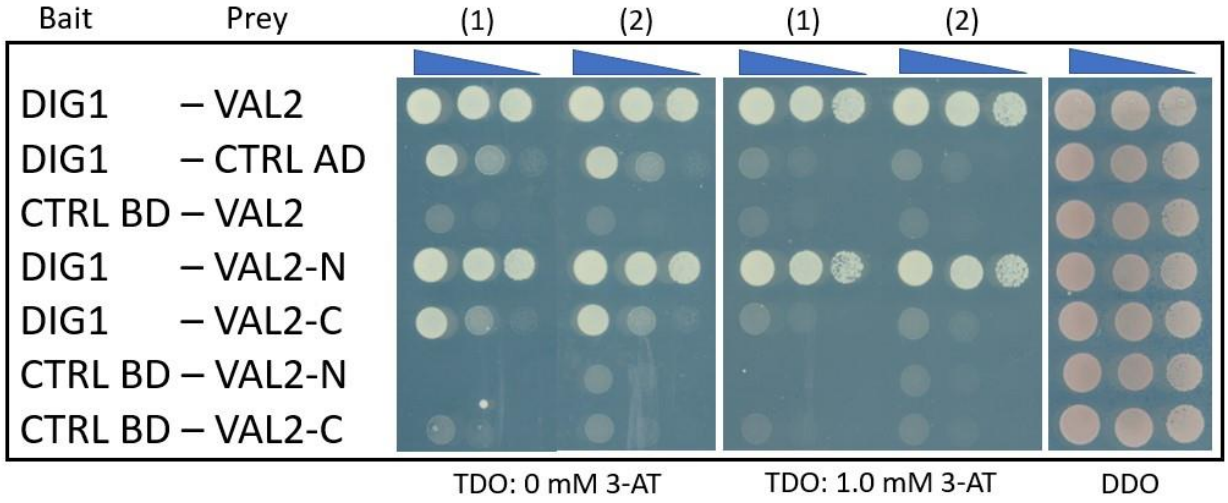

Figure 4. DIG1 physically interacts with VAL2 truncation containing PHD-like and B3 domains. 3-day-old yeast grown on TDO/-Leu/-Trp/-His with 0 and 1.0 mM 3-AT. Growth on DDO/-Leu/-Trp shown on right.

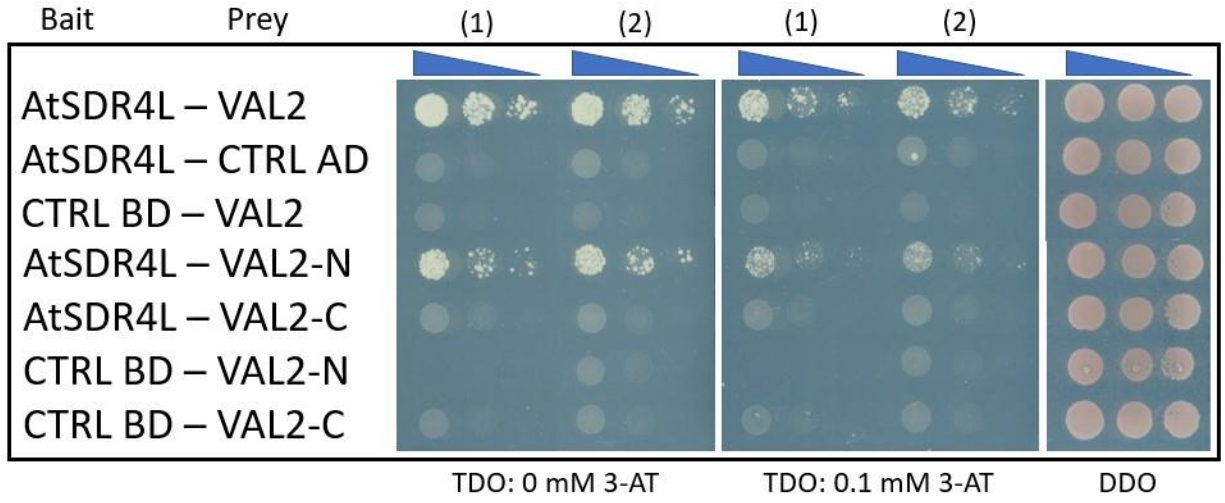

Figure 5. AtSDR4L physically interacts with VAL2 truncation containing PHD-like and B3 domains. 3-day-old yeast grown on TDO/-Leu/-Trp/-His with 0 and 0.1 mM 3-AT. Growth on DDO/-Leu/-Trp shown on right.

2.2.3.1. Results from the initial screen

The Y2H system was used to identify potential protein-protein interactions for AtSDR4L and DIG proteins to elucidate their role in seed development. From the initial Y2H screen, we found that DIG1, DIL1, and AtSDR4L are able to physically interact with VAL2 (Figures 4, 5, 11). The DIG1 and DIL1 show a very strong interaction with VAL2, while the AtSDR4L interaction is much weaker (Figures 4, 5, 11). The rest of the DIG family also showed a similar interaction strength to DIG1 and DIL1, but the negative control for each showed a positive result indicating that these proteins have autoactivation and need to be tested in another system (Figure 11). We were also able to narrow down the interacting domain of VAL2 to the N-terminal half which contains the PHD-like domain and the B3 domain (Figures 4-6). A further truncation of VAL2-N to separate the PHD-like and B3 domains showed no interaction for each, which might indicate they are both needed for the interaction. Consistently, in VAL-mediated repression of *DOG1*, Chen et al., (2020) found that the enrichment of VALs at the *DOG1* promoter require both the PHD-like and B3 domains. We also tested N- and C-terminal truncations for both DIG1 and AtSDR4L, but both pairs yielded no positive interactions.

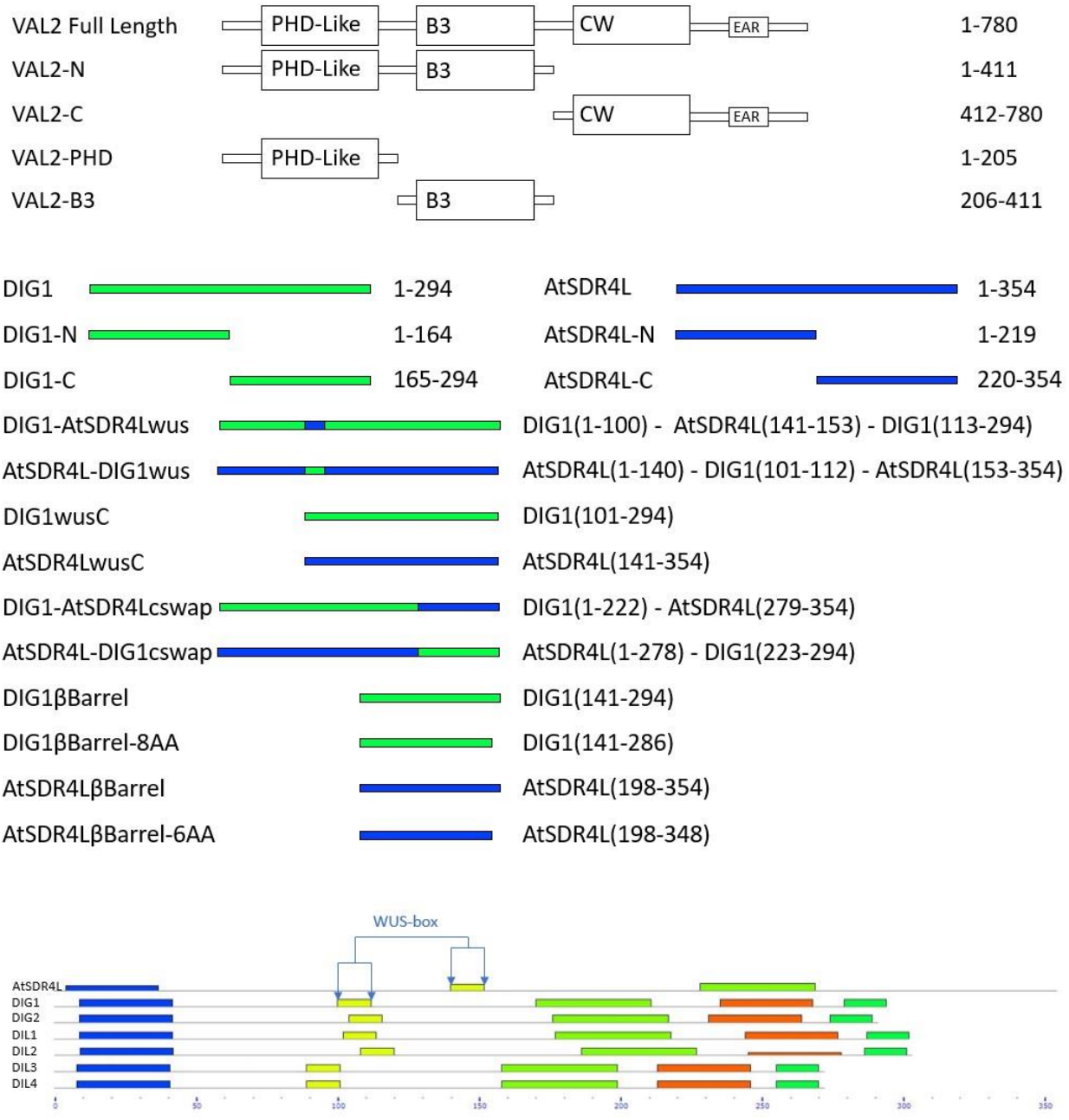

Figure 6. Schematic Diagram of the protein truncations and the domain swapped proteins tested. Top: VAL2 truncations showing the PHD-like domain, B3 domain, CW domain, and the EAR motif. Middle: DIG1 and AtSDR4L truncations, and domain swaps with amino acid numbers indicating amino acids from the original full-length proteins. Bottom: MEMEsuite reveals several conserved motifs in AtSDR4L and DIG proteins including a WUS-box like motif.

2.2.3.2. MEMEsuite analysis to test new truncations

Since the interaction strength between DIG1 and AtSDR4L with VAL2 differed greatly, we designed selective domain swaps to test for the important domains (Figure 6). To identify conserved regions in the absence of known protein domains, we used MEMEsuite to align AtSDR4L and the DIG proteins and scan for conserved motifs (Figure 6) (Bailey et al., 2009). We included a group of proteins such as WUSCHEL and WUSCHEL related homeobox (WOX) proteins after observing some similar properties and found that there is a WUS-box-like motif in AtSDR4L and the DIG proteins (Figure 6). *WUSCHEL* is a transcription factor that mainly represses genes affecting maintenance of stem cells, where the WUS-box is required for repressive activity, while the WOX proteins are involved in early embryonic patterning (Ikeda et al., 2009; Haecker et al., 2004). The MEMEsuite scan also identified two more conserved motifs in AtSDR4L and the DIGs, and an additional two motifs that were only conserved in the DIG proteins (Figure 6). The first domain swap was for the two conserved motifs found at the C-terminus of DIG1, which are found in the rest of the DIGs, but not AtSDR4L (Figure 6). This was chosen because the interaction strength appears consistent for DIG1 and DIL1, but not AtSDR4L, and so the same pattern is seen here with the absence of the conserved motifs at the end of AtSDR4L (Figure 6; Figure 11). Another domain swap was designed for the WUS-box-like motif found near the middle of the proteins (Figure 6). There were several differences between the WUS-box-like motifs of DIG1 and AtSDR4L, which is another possible candidate for the differing interaction strengths (Figure 6). The WUS-box swap did not alter the interaction strength and the C-terminus swap did not show any positive interactions. Finally, a truncation

that included the WUS-box-like motifs along with the C-terminal half of the protein, was unfortunately discarded due to design errors and new available resources (Figure 6).

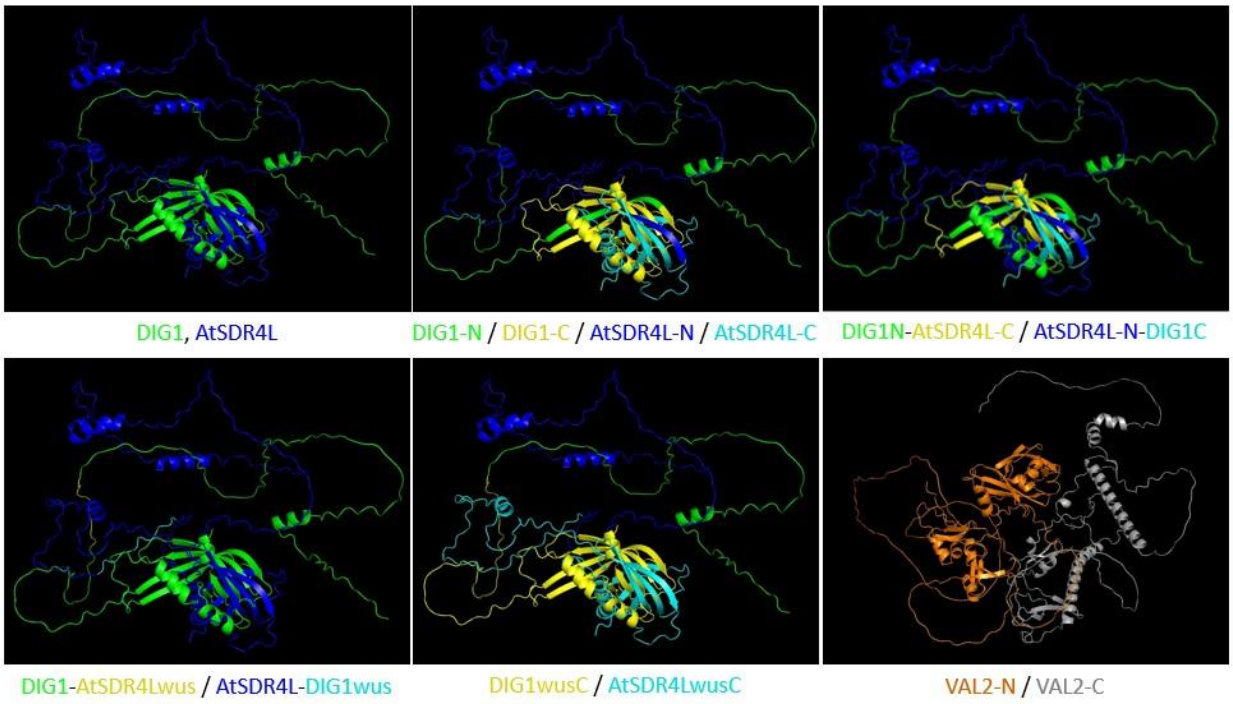

Figure 7. AlphaFold 3D prediction reveals protein structures in AtSDR4L, DIG1, and VAL2 for better structural analysis. Proteins are coloured to highlight the different truncations and domain swaps used in the Yeast 2 Hybrid assay.

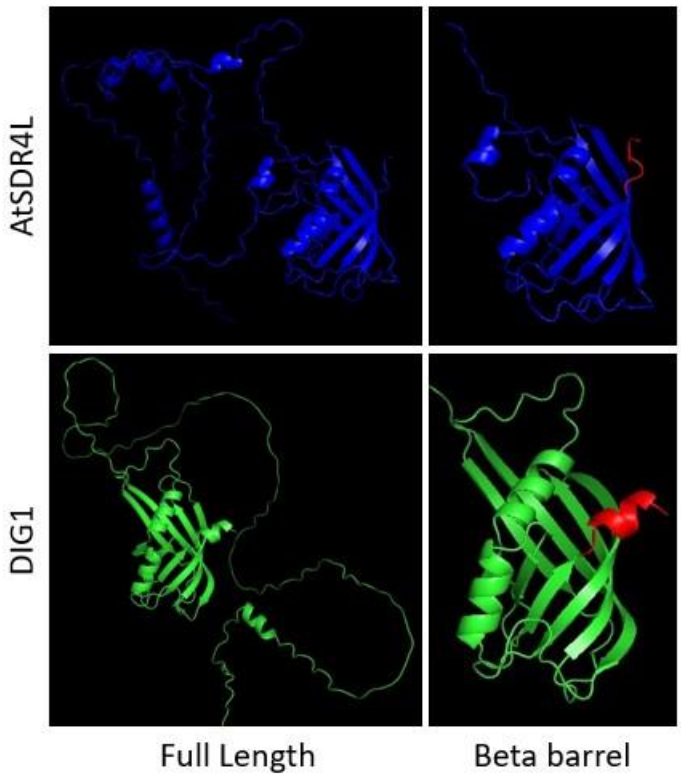

Figure 8. Alphafold 3D prediction of the new constructs designed. Truncation constructs were made for the protein with the C-terminus part indicated in red.

2.2.3.3. AlphaFold 3D protein prediction software to test important 3D structures

Recently, a new computational protein structure prediction software called AlphaFold was developed, that uses deep learning to accurately predict 3D structures even when no similar structures are known (Jumper et al., 2021). This was especially useful to our project since the DIG proteins and AtSDR4L all do not contain any known protein domains. The predicted 3D structures can help identify domains that are important to the function of the proteins. After using this software, we can see that the N and C truncations for AtSDR4L and DIG1 may have interrupted the large beta barrel-like structure, and could be the reason that no

interaction was detected for either protein truncation (Figure 7). Based on the AlphaFold predictions, we designed a few more truncations to test (Figure 8). We decided to make a truncation that only contained the beta-barrel, and another that had the beta-barrel and included the short strand that is at the C-terminus where DIG1 has an alpha helix but AtSDR4L does not, which could explain the large difference in interaction strength (Figure 8). From these new truncations, we observed that the four truncations all could physically interact with VAL2 (Figure 9; Figure 10). Interestingly, the AtSDR4LBetaBarrel was able to activate the reporter gene without having VAL2 attached to the GAL4 activation domain, suggesting that this specific truncation may have given it transcriptional activation properties (Figure 10). We can also observe that removing the 8 amino acids (AA) that includes the C-terminal alpha helix in DIG1 reduced the strength of the interaction suggesting that it may provide some function (Figure 9). Unfortunately, in the last test there were some inconsistencies in the replicates, especially evident in the DIG1BetaBarrel-8AA – VAL2 cell growth with 1.0 mM 3-AT (Figure 9). In previous tests, using the same yeast cultures, the growth between replicates was consistent. A second test shortly after, and using the same cultures showed similar results to the one shown in Figure 9. I believe that transforming fresh yeast cultures may resolve this problem.

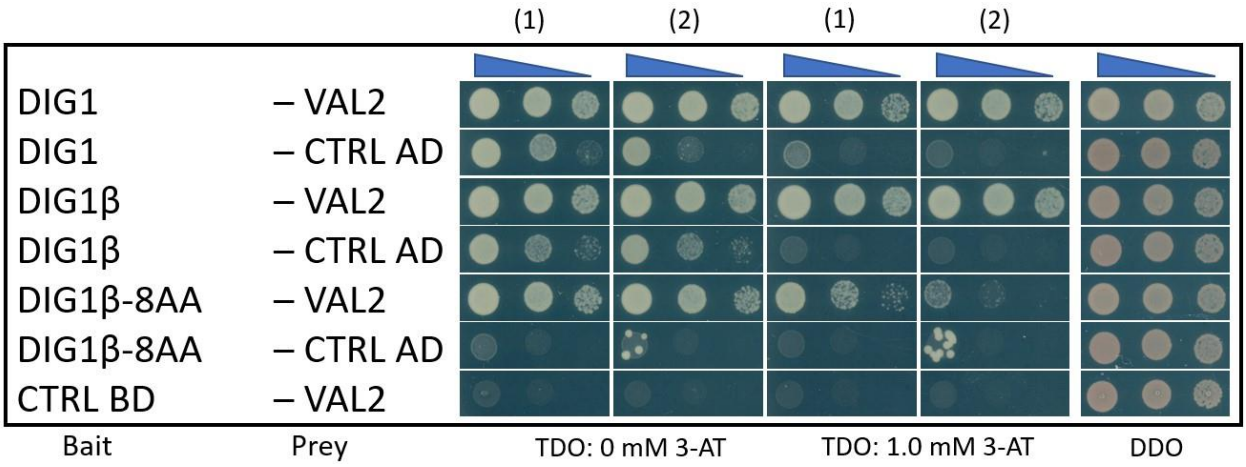

Figure 9. DIG1BetaBarrel and DIG1BetaBarrel-8AA physically interacts with VAL2 with a slightly decreased relative interaction strength for DIG1BetaBarrel-8AA. 3-day-old yeast grown on TDO/-Leu/-Trp/-His with 0 and 1.0 mM 3-AT. Growth on DDO/-Leu/-Trp shown on right.

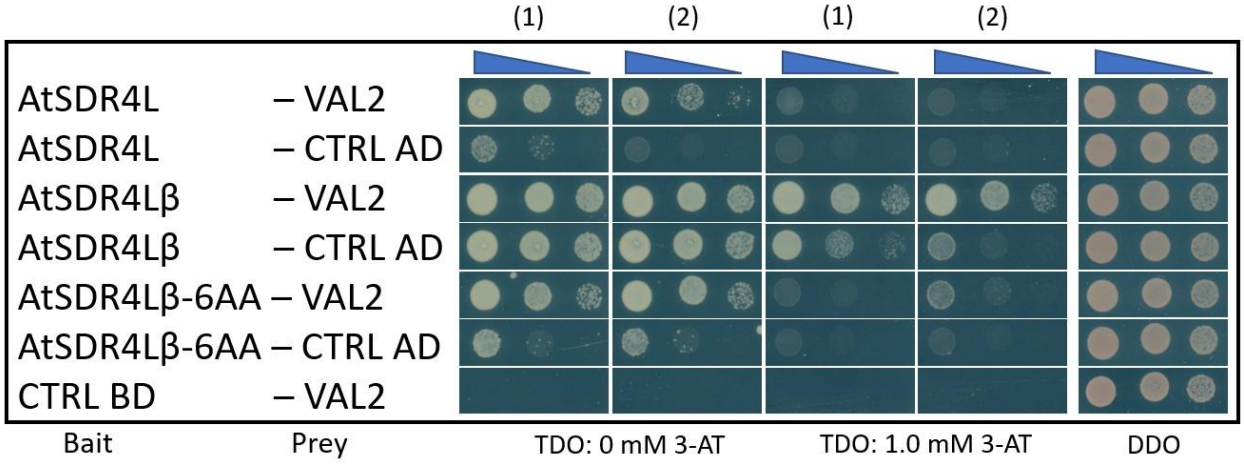

Figure 10. AtSDR4LBetaBarrel and AtSDR4LBetaBarrel-6AA physically interacts with VAL2. However, there is also autoactivation for AtSDR4LBetaBarrel which means we cannot confirm the positive interaction with VAL2. 3-day-old yeast grown on TDO/-Leu/-Trp/-His with 0, 0.1 mM, and 1.0 mM 3-AT. Growth on DDO/-Leu/-Trp shown on right.

To generate further support for the putative interacting domains, I generated constructs that would express truncated versions of AtSDR4L driven by the AtSDR4L native promoter. To examine the functionality of truncated AtSDR4L *in planta*, a set of three constructs were designed for AtSDR4L based on the Y2H data. We tested AtSDR4L that has the last 6AA missing and AtSDR4L that is missing the entire beta barrel and compared it to full length AtSDR4L. These three proteins were driven by the AtSDR4L native promoter and tagged with N-terminal 3xHA and C-terminal 3xFLAG, and were cloned into pCAMBIA1300 using HiFi DNA Assembly. The constructs were transformed into *Agrobacterium tumefaciens* C58 and then transformed into *Arabidopsis thaliana Atsdr4l* CRISPR mutants using the floral dip method. Once we have confirmed homozygous mutants, we will test to see if the truncated proteins are able to rescue the *Atsdr4l* mutant phenotype.

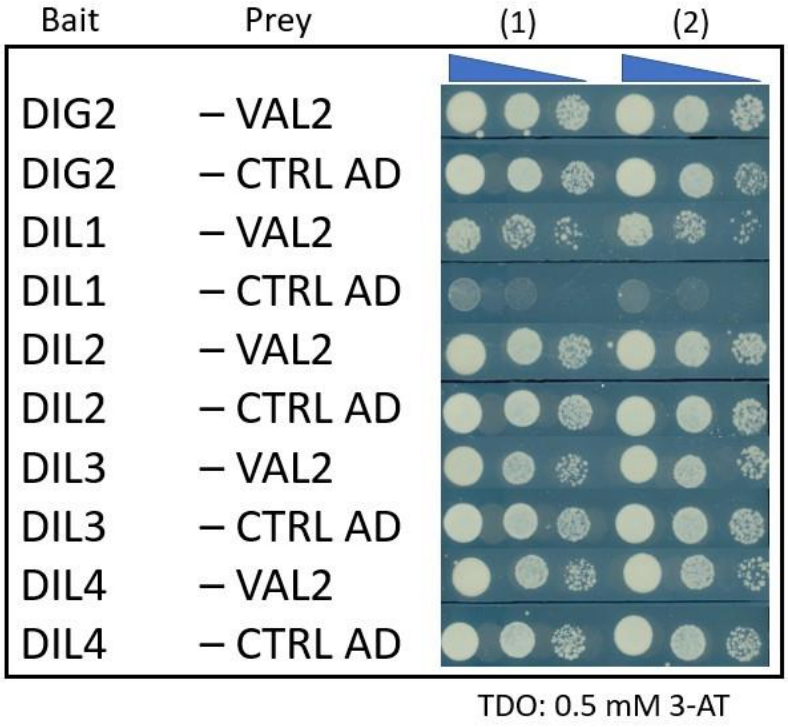

Figure 11. DIL1 shows interaction with VAL2 while DIG2, DIL2, DIL3, and DIL4 show autoactivation activity. 4-day-old yeast grown on TDO/-Leu/-Trp/-His with 0.5 mM and 1.0 mM 3-AT.

Chapter 3 EMS suppressor screen for *DIG1* OE

3.1 Introduction

To identify new potential players involved in the *DIG1* pathway, I conducted a forward genetic screen to identify mutants that could suppress the *DIG1* overexpressor phenotype. In this screen, the inhibited cotyledon growth was used as the phenotype of interest since it is the most prominent phenotype and appears earliest. We are unable to use the *dig1* knockout mutant, because the *dig1* phenotype is too subtle for a large-scale mutant screen. A population of seeds harbouring the dexamethasone inducible overexpressor of *DIG1* transgene generated by Liang Song, was treated with a chemical mutagen called Ethyl Methanesulfonate (EMS), which produces random single nucleotide mutations. The EMS converts a guanine base to a 6-ethyl guanine, which instead of pairing with a cytosine, will pair with a thymine base instead, producing random G/C to A/T mutations (Kim et al., 2006). When seeds containing the *DIG1* OE transgene were germinated on 1x LS agar plates containing 0.4 µM ABA, 0.8 µM DEX, and 1% sucrose, the resulting seedlings had severely inhibited cotyledon growth and swollen hypocotyls, and when left on the plates for several weeks, callus tissue starts to form around the root-hypocotyl junction. In the future, this screen will look for mutants that have reverted its phenotype back to wild-type-like, despite the induction of the *DIG1* OE transgene. These wild-type-like plants likely have a novel mutation that is able to suppress the effects of the *DIG1* overexpression.

3.2. Methods

3.2.1. Plants and growth conditions

The *Arabidopsis thaliana* plants were grown in growth rooms or chambers at 22°C under a 16-hour light, 8-hour dark cycle.

3.2.2. Mutagenesis and mutant screening conditions

0.1 g of seeds (~ 5,000) harbouring the dexamethasone-inducible overexpressor of *DIG1* were treated with EMS to induce random single nucleotide mutations. The seeds were put into a 50 mL falcon tube in a solution containing 116 µL 1 M Na_2HPO_4, 1.38 mL 1 M NaH_2PO_4, 0.75 mL DMSO, and 12.75 mL of water. 32.4 µL of EMS was added and the seeds were set to shake for 16 hours. The seeds were spun down and the EMS solution was removed. The seeds were washed with 13 mL 100 mM of sodium thiosulfate ($Na_2S_2O_3$) three times, shaking for 20 minutes each time. The seeds were then washed with sterile water three times. The seeds were sterilized with a 15% bleach solution and then plated on 1x LS 0.7% agar plates containing 0.5% sucrose. The seeds were stratified at 4°C for 3 days and then moved to a growth chamber.

Due to the predicted varied severity of the phenotype, the M1 plants were harvested individually and each M2 line was screened by growing approximately 50 seeds from each line on agar plates containing 1% sucrose, 0.4 µM ABA, and 0.8 µM DEX. Potential suppressors were grown, and the following generation was also subjected to the same screening conditions to identify mutants with a more consistent phenotype. PCR amplification and Western Blots were performed on several of the putative suppressors that had a consistent phenotype to confirm

the presence of the *DIG1* OE transgene. To test if any of the putative mutants were genes involved in ABA signalling, seeds from each putative mutant line were plated on agar plates with 0.4 µM ABA without inducing the *DIG1* OE transgene.

3.2.3. Next generation sequencing

In order to get rid of background mutations prior to sequencing, the putative mutants that were homozygous for both the suppressor mutation and the *DIG1* OE construct were backcrossed to the original *DIG1* OE line. The BC F1 plants were allowed to self and the BC F2 generation were then subjected to the original screening conditions to select for backcrossed mutants homozygous for the new suppressor mutation. For each of the mutant lines, leaf DNA from 80 different BC F2 plants were pooled and extracted using the Plant DNAzol extraction method. The DNA was cleaned up using a Plant DNeasy Mini Column and sent to Novogene using the Plant and Animal Whole Genome Sequencing package that uses the NovaSeq PE150 sequencing strategy. Sequencing data was analyzed on Compute Canada's server using a script provided by Sean Shang. A list of mutations generated from the NGS sequencing data will be filtered to identify the mutations most likely to be the causative mutation. From these mutations, a list of ranked candidate genes will be established based on likelihood and we will knockout these putative candidate genes in the *DIG1* OE plant using CRISPR-Cas9 to make targeted mutations. Once we have identified the suppressor, we can carry out further experiments to characterize this new gene or new interaction.

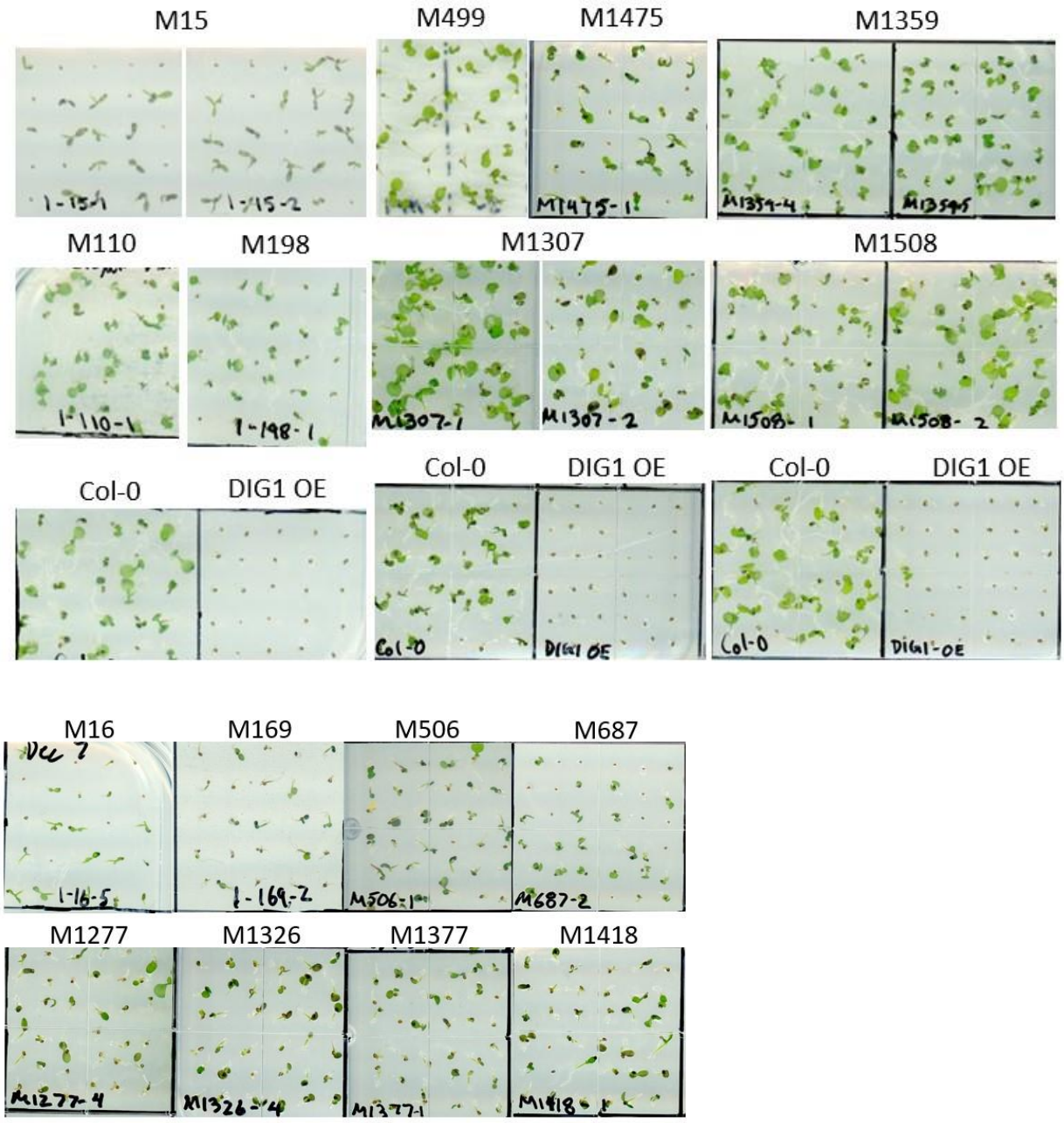

Figure 12. Phenotypes of the suppressor mutants. M3 generation of suppressor lines plated on 1x LS 0.4 µM ABA 0.8 µM DEX 1% Sucrose. Strong mutant lines shown in first three rows. Weaker lines shown in bottom two rows.

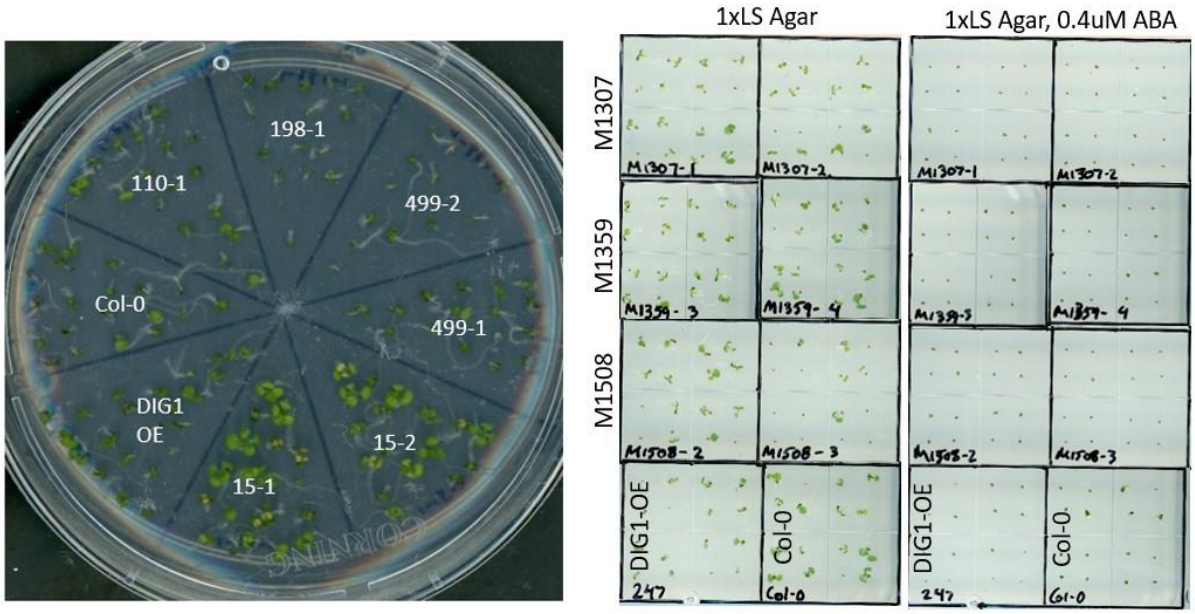

Figure 13. ABA-sensitivity assay on strong suppressor mutants. M3 generation of suppressor lines plated on 1x LS 0.4 µM ABA 1% Sucrose without the induced *DIG1* expression by DEX.

3.3. Results

We screened 1600 individual mutant lines, which was the minimum number of lines we set out to screen to cover the genome. From the 1600 lines, I found 94 potential strong or partial suppressors of the *DIG1* overexpressor phenotype. 16 of these mutant lines were found to have a more consistent phenotype, while 8 also showed a stronger level of suppression. These 8 include M15, M110, M198, M499, M1307, M1359, M1475, and M1508 (Figure 12). The M15 mutant turned out to be ABA insensitive, while M110, M198, M499, M1307, M1359, and M1508 were not (Figure 13). M1475 still needs to be tested. Expression of M15 but not M198 was detected using Western Blot by Renwei Zheng. Presence of the transgene was confirmed in lines M15 and M198 through PCR amplification. We expect that the mutants that show insensitivity to ABA are components of the ABA signalling network and will not be of significant

interest. While searching for the identities of the disrupted genes, it will be important to include genes with mutations in known binding sites in their promoter regions since we are dealing with an extensive regulatory network. Since we only found one strong mutant that was insensitive to ABA, it is fair to conclude that the screen did not reach saturation. We still have many more harvested individual lines that can be screened in the future, in attempts to identify more genes.

Table 3.1. List of suppressor lines identified in the screen.

Line	Suppressors found in M2	Suppressor lines in M3	Phenotype Consistency	Phenotype
M15	3	2	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like but with skinny cotyledons
M16	5	4	Around 50% of seedlings show phenotype in M3 for each line.	WT-like but with skinny cotyledons
M110	3	1	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M169	2	1	Around 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M198	1	1	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M499	3	2	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M506	1	1	Around 50% of seedlings show phenotype in M3 for each line.	WT-like
M687	3	1	Around 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M1277	11	10	Less than 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M1307	2	2	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M1326	4	2	Around 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M1359	8	8	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M1377	1	1	Around 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M1418	2	2	Around 50% of seedlings show phenotype in M3 for each line.	Partial suppressor
M1475	1	1	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like
M1508	4	4	Strong. Most of seedlings show phenotype in M3 for each line.	WT-like

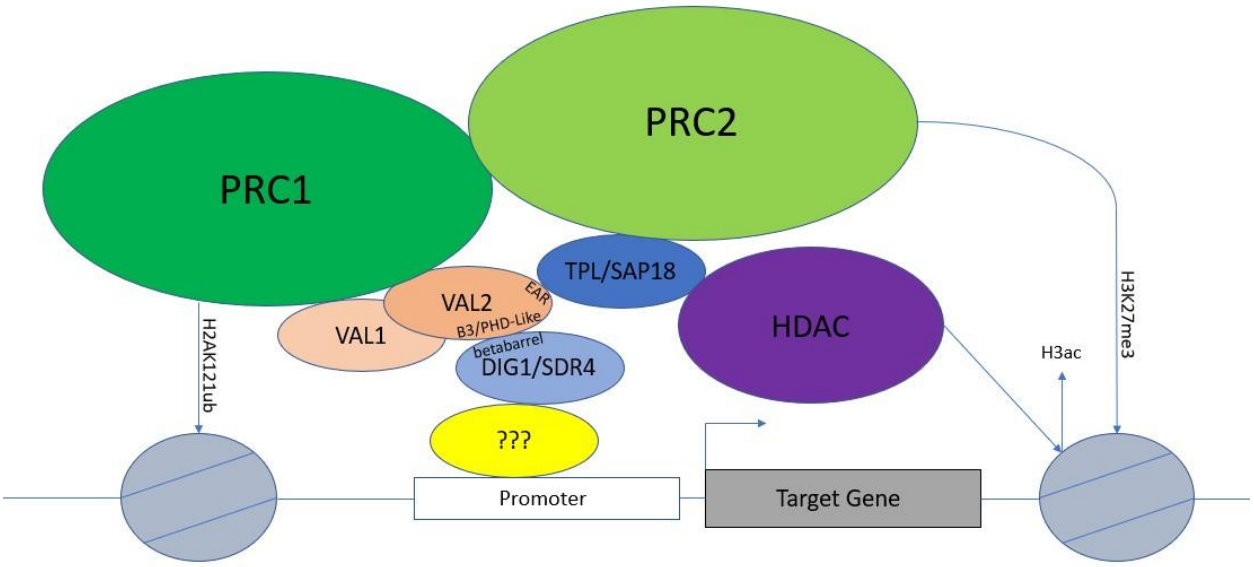

Figure 14: Proposed model of mechanism

AtSDR4L/DIG1 interacts with unknown protein to bind to target genes. AtSDR4L/DIG1 recruit VAL2 in homo- or heterodimer to these loci. VAL dimer recruits PRC1 and TPL or SAP18 to the loci. PRC1 adds repressive H2AK121ub marks to target loci to repress transcription. PRC1 along with TPL/SAP18 recruit PRC2 and HDACs to target loci. PRC2 adds repressive H3K27me3 marks to target loci to repress transcription. HDACs remove acetyl marks (H3ac) to reduce gene transcription at target loci.

Chapter 4 Conclusions and future perspectives

In chapter 2, my goal was to identify a putative mechanism involved in DIG gene repressive action by looking for protein-protein interactions. We found positive interactions for AtSDR4L, DIG1, and DIL1 with VAL2, with the other four DIG proteins being inconclusive due to autoactivation. This finding was instrumental in the development of my working model of the mechanism and led to many more experiments and hypotheses.

Since the Yeast 2 Hybrid system can result in false positive interactions, the next steps in this project will be to confirm the interactions using a different system. Using Bimolecular Fluorescence Complementation (BiFC), Renwei Zheng has preliminary data on the interaction between DIG1 and VAL2, and the results look promising. Currently, co-IP constructs, which can also be used to confirm the interactions are being cloned by Renwei Zheng. For co-IP, we will attach epitope tags to DIG1, AtSDR4L, and VAL2, which can be used to pulldown proteins using the specific antibodies. Once the interacting proteins are pulled down and if there is a positive interaction, the interacting partner will be pulled down with it. We can then detect the presence of the interacting partner using a different antibody to detect a different epitope tag. Additionally, in collaboration with Milad Alizadeh, we also generated TurboID tagged lines of AtSDR4L and DIG1 full length proteins driven by a beta-estradiol inducible promoter that can be a backup alternative for confirming the interaction. For TurboID, we fused the TurboID to AtSDR4L/DIG1 and will express it in the plant. The plant tissue of interest will then be submerged in a biotin solution and the engineered biotin ligase will use biotin to label all of the proteins that are in close proximity to our protein of interest. We will then extract those labelled proteins and analyze which proteins work in close proximity to our protein of interest.

Since evidence shows that AtSDR4L does not accumulate at RY-motifs, which can be bound by VAL1 and VAL2 through their B3 domains, it appears that AtSDR4L and possibly the DIG proteins are recruiting VAL2 to target gene loci (Wu et al., 2022). Since AtSDR4L/DIGs are not able to bind to DNA themselves, there should be other DNA binding proteins involved in this mechanism that are able to recruit AtSDR4L/DIG proteins to these loci. One potential future direction for this project would be to try and identify those specific proteins. Another future direction could be to investigate AtSDR4L/DIG-mediated recruitment of VAL2 by viewing VAL2 accumulation at shared target loci in *Atsdr4l* or *dig* mutants or looking at how the repressive chromatin marks are affected at these loci.

Since *VAL1* and *VAL2* act redundantly, I ordered T-DNA lines for *val1* and *val2* mutants and crossed them. After confirming homozygous double mutants, I was unable to recreate the phenotype from the literature. I tried an assay with different concentrations of sucrose, but I was unable to produce the embryonic seedling phenotype. From looking at the locations of the T-DNAs from the crosses I performed, the insertions appear to be near the end of the VAL1 and VAL2 proteins. Its possible that the T-DNA insertions for these lines might not have a large enough effect on the protein to fully disrupt their functions. For this reason, I generated CRISPR constructs that would result in simultaneous interruption of *VAL1* and *VAL2* with segmental deletions. I transformed these constructs into *Arabidopsis thaliana* Col-0 and harvested the T1s. After the homozygous double mutants with the CRISPR transgene segregated out are identified, they can be crossed with the *Atsdr4l* and *dig* mutants to check if the phenotype is affected in the higher order mutant.

Unfortunately, I was unable to identify the genes that were mutated in my suppressor screen at the end of my study, but I hope that future students will benefit from this project. I was able to screen 1600 individual lines, however, I did not identify the causative mutation(s) for the suppressor mutants. I was able to identify several strong, promising suppressor lines which I hope will lead to interesting results and important findings in the hands of future students. There is also an abundance of M2 lines that can still be screened since I believe the screen is far from saturation.

The field of agriculture is always in need of new technological advances to solve problems impacting the quality and yield of crops. With the recent advances in CRISPR and other gene editing technologies, many of these problems can by tackled by making precise changes to specific genes. Given the great complexity of genetic networks and interplay between genes, it is vital that we can see the whole picture when making these targeted changes. My project aimed to shed light on a vital and complex stage of the plant life cycle with respect to a family of understudied transcriptional coregulators and help provide a better understanding of the genetic network integral to seed development.

Bibliography

Alizadeh, M., Hoy, R., Lu, B., & Song, L. (2021). Team effort: Combinatorial control of seed maturation by transcription factors. Current Opinion in Plant Biology, 63, 102091.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., ... & Ecker, J. R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. Science, 301(5633), 653-657.

Angelovici, R., Galili, G., Fernie, A. R., & Fait, A. (2010). Seed desiccation: a bridge between maturation and germination. Trends in plant science, 15(4), 211-218.

Baile, F., Merini, W., Hidalgo, I., & Calonje, M. (2021). EAR domain-containing transcription factors trigger PRC2-mediated chromatin marking in Arabidopsis. The Plant Cell.

Baud, S., & Lepiniec, L. (2010). Physiological and developmental regulation of seed oil production. Progress in lipid research, 49(3), 235-249.

Bergelson, J., & Perry, R. (1989). Interspecific competition between seeds: relative planting date and density affect seedling emergence. Ecology, 70(6), 1639-1644.

Berjak, P., & Pammenter, N. W. (2002). Orthodox and recalcitrant seeds. Tropical tree seed manual, 137-147.

Bethke, P. C., Libourel, I. G., Aoyama, N., Chung, Y. Y., Still, D. W., & Jones, R. L. (2007). The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. Plant physiology, 143(3), 1173-1188.

Bewley, J. D. (1997). Seed germination and dormancy. The plant cell, 9(7), 1055.

Bewley, J. D., Bradford, K., & Hilhorst, H. (2012). Seeds: physiology of development, germination and dormancy. Springer Science & Business Media.

Boulard, C., Fatihi, A., Lepiniec, L., & Dubreucq, B. (2017). Regulation and evolution of the interaction of the seed B3 transcription factors with NF-Y subunits. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1860(10), 1069-1078.

Bryant, F. M., Hughes, D., Hassani-Pak, K., & Eastmond, P. J. (2019). Basic LEUCINE ZIPPER TRANSCRIPTION FACTOR67 transactivates DELAY OF GERMINATION1 to establish primary seed dormancy in Arabidopsis. The Plant Cell, 31(6), 1276-1288.

Cao, D., Hussain, A., Cheng, H., & Peng, J. (2005). Loss of function of four DELLA genes leads to light-and gibberellin-independent seed germination in Arabidopsis. Planta, 223(1), 105-113.

Cao, H., Han, Y., Li, J., Ding, M., Li, Y., Li, X., ... & Liu, Y. (2020). Arabidopsis thaliana SEED DORMANCY 4-LIKE regulates dormancy and germination by mediating the gibberellin pathway. Journal of Experimental Botany, 71(3), 919-933.

Cernac, A., & Benning, C. (2004). WRINKLED1 encodes an AP2/EREB domain protein involved in the control of storage compound biosynthesis in Arabidopsis. The Plant Journal, 40(4), 575-585.

Chakrabortee, S., Boschetti, C., Walton, L. J., Sarkar, S., Rubinsztein, D. C., & Tunnacliffe, A. (2007). Hydrophilic protein associated with desiccation tolerance exhibits broad protein stabilization function. Proceedings of the National Academy of Sciences, 104(46), 18073-18078.

Cutler, S. R., Rodriguez, P. L., Finkelstein, R. R., & Abrams, S. R. (2010). Abscisic acid: emergence of a core signaling network. Annual review of plant biology, 61, 651-679.

Finch-Savage, W. E., & Leubner-Metzger, G. (2006). Seed dormancy and the control of germination. New phytologist, 171(3), 501-523.

Fujii, H., Verslues, P. E., & Zhu, J. K. (2007). Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. The Plant Cell, 19(2), 485-494.

Gubler, F., Millar, A. A., & Jacobsen, J. V. (2005). Dormancy release, ABA and pre-harvest sprouting. Current opinion in plant biology, 8(2), 183-187.

Harris, L. W., & Davies, T. J. (2016). A complete fossil-calibrated phylogeny of seed plant families as a tool for comparative analyses: testing the 'time for speciation' hypothesis. PloS one, 11(10), e0162907.

Jo, L., Pelletier, J. M., & Harada, J. J. (2019). Central role of the LEAFY COTYLEDON1 transcription factor in seed development. Journal of Integrative Plant Biology, 61(5), 564-580.

Josse, E. M., & Halliday, K. J. (2008). Skotomorphogenesis: the dark side of light signalling. Current Biology, 18(24), R1144-R1146.

Kim, Y., Schumaker, K. S., & Zhu, J. K. (2006). EMS mutagenesis of Arabidopsis. In Arabidopsis protocols (pp. 101-103). Humana Press.

Kobayashi, Y., Murata, M., Minami, H., Yamamoto, S., Kagaya, Y., Hobo, T., ... & Hattori, T. (2005). Abscisic acid-activated SNRK2 protein kinases function in the gene-regulation pathway of ABA signal transduction by phosphorylating ABA response element-binding factors. The Plant Journal, 44(6), 939-949.

Koornneef, M., & Van der Veen, J. H. (1980). Induction and analysis of gibberellin sensitive mutants in Arabidopsis thaliana (L.) heynh. Theoretical and Applied genetics, 58(6), 257-263.

Koornneef, M., & Meinke, D. (2010). The development of Arabidopsis as a model plant. The Plant Journal, 61(6), 909-921.

Lee, S., Cheng, H., King, K. E., Wang, W., He, Y., Hussain, A., ... & Peng, J. (2002). Gibberellin regulates Arabidopsis seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. Genes & development, 16(5), 646-658.

Leprince, O., Pellizzaro, A., Berriri, S., & Buitink, J. (2017). Late seed maturation: drying without dying. Journal of experimental botany, 68(4), 827-841.

Liu, X., & Hou, X. (2018). Antagonistic regulation of ABA and GA in metabolism and signaling pathways. Frontiers in plant science, 9, 251.

Liu, F., Zhang, H., Ding, L., Soppe, W., & Xiang, Y. (2020). REVERSAL OF RDO5 1, a Homolog of rice Seed dormancy 4, Interacts with bHLH57 and Controls ABA Biosynthesis and Seed Dormancy in Arabidopsis. The Plant Cell.

Ma, Y., Tian, H., Lin, R., Wang, W., Zhang, N., Hussain, S., ... & Wang, S. (2021). AITRL, an evolutionarily conserved plant specific transcription repressor regulates ABA response in Arabidopsis. Scientific Reports, 11(1), 1-11.

Mozgova, I., Köhler, C., & Hennig, L. (2015). Keeping the gate closed: functions of the polycomb repressive complex PRC 2 in development. The Plant Journal, 83(1), 121-132.

Nakajima, S., Ito, H., Tanaka, R., & Tanaka, A. (2012). Chlorophyll b reductase plays an essential role in maturation and storability of Arabidopsis seeds. Plant physiology, 160(1), 261-273.

Nakashima, K., Fujita, Y., Kanamori, N., Katagiri, T., Umezawa, T., Kidokoro, S., ... & Shinozaki, K. (2009). Three Arabidopsis SnRK2 protein kinases, SRK2D/SnRK2. 2, SRK2E/SnRK2. 6/OST1 and SRK2I/SnRK2. 3, involved in ABA signaling are essential for the control of seed development and dormancy. Plant and Cell Physiology, 50(7), 1345-1363.

Ohta, M., Guo, Y., Halfter, U., & Zhu, J. K. (2003). A novel domain in the protein kinase SOS2 mediates interaction with the protein phosphatase 2C ABI2. Proceedings of the National Academy of Sciences, 100(20), 11771-11776.

Ogawa, M., Hanada, A., Yamauchi, Y., Kuwahara, A., Kamiya, Y., & Yamaguchi, S. (2003). Gibberellin biosynthesis and response during Arabidopsis seed germination. The Plant Cell, 15(7), 1591-1604.

Penfield, S., Rylott, E. L., Gilday, A. D., Graham, S., Larson, T. R., & Graham, I. A. (2004). Reserve mobilization in the Arabidopsis endosperm fuels hypocotyl elongation in the dark, is independent of abscisic acid, and requires PHOSPHOENOLPYRUVATE CARBOXYKINASE1. The Plant Cell, 16(10), 2705-2718.

Russell, S. D. (1992). Double fertilization. In International review of cytology (Vol. 140, pp. 357-388). Academic Press.

Sallon, S., Solowey, E., Cohen, Y., Korchinsky, R., Egli, M., Woodhatch, I., ... & Kislev, M. (2008). Germination, genetics, and growth of an ancient date seed. Science, 320(5882), 1464-1464.

Santiago, J., Rodrigues, A., Saez, A., Rubio, S., Antoni, R., Dupeux, F., ... & Rodriguez, P. L. (2009). Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade A PP2Cs. The Plant Journal, 60(4), 575-588.

Shao, Q., Liu, X., Su, T., Ma, C., & Wang, P. (2019). New insights into the role of seed oil body proteins in metabolism and plant development. Frontiers in Plant Science, 10, 1568.

Shewry, P. R., Napier, J. A., & Tatham, A. S. (1995). Seed storage proteins: structures and biosynthesis. The plant cell, 7(7), 945.

Song, L., Huang, S. S. C., Wise, A., Castanon, R., Nery, J. R., Chen, H., ... & Ecker, J. R. (2016). A transcription factor hierarchy defines an environmental stress response network. Science, 354(6312), aag1550.

Sugimoto, K., Takeuchi, Y., Ebana, K., Miyao, A., Hirochika, H., Hara, N., ... & Yano, M. (2010). Molecular cloning of Sdr4, a regulator involved in seed dormancy and domestication of rice. Proceedings of the National Academy of Sciences, 107(13), 5792-5797.

Sundberg, E., Slagter, J. G., Fridborg, I., Cleary, S. P., Robinson, C., & Coupland, G. (1997). ALBINO3, an Arabidopsis nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. The Plant Cell, 9(5), 717-730.

Tsukagoshi, H., Morikami, A., & Nakamura, K. (2007). Two B3 domain transcriptional repressors prevent sugar-inducible expression of seed maturation genes in Arabidopsis seedlings. Proceedings of the National Academy of Sciences, 104(7), 2543-2547.

Umezawa, T., Sugiyama, N., Mizoguchi, M., Hayashi, S., Myouga, F., Yamaguchi-Shinozaki, K., ... & Shinozaki, K. (2009). Type 2C protein phosphatases directly regulate abscisic acid-activated protein kinases in Arabidopsis. Proceedings of the National Academy of sciences, 106(41), 17588-17593.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., & Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PloS one, 2(8), e718.

Wu, T., Alizadeh, M., Lu, B., Cheng, J., Hoy, R., Bu, M., ... & Song, L. (2022). The transcriptional co-repressor SEED DORMANCY 4-LIKE (AtSDR4L) promotes embryonic-to-vegetative transition in Arabidopsis thaliana. Journal of Integrative Plant Biology.

Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F., & Calonje, M. (2013). VAL-and AtBMI1-mediated H2Aub initiate the switch from embryonic to postgerminative growth in Arabidopsis. Current Biology, 23(14), 1324-1329.

Yuan, L., Song, X., Zhang, L., Yu, Y., Liang, Z., Lei, Y., ... & Li, C. (2021). The transcriptional repressors VAL1 and VAL2 recruit PRC2 for genome-wide Polycomb silencing in Arabidopsis. Nucleic Acids Research, 49(1), 98-113.

Appendix

List of constructs:

Construct Name	Use	Primers for Cloning	Primers for Sanger
VAL2-N	Yeast 2 Hybrid	o0419, o0420	o0212, o0216
VAL2-C	Yeast 2 Hybrid	o0421, o0422	o0212, o0216
VAL2-PHD	Yeast 2 Hybrid	o0419, o0521	o0212, o0216
VAL2-B3	Yeast 2 Hybrid	o0420, o0522	o0212, o0216
DIG1-N	Yeast 2 Hybrid	o0411, o0412	0467, 0468
DIG1-C	Yeast 2 Hybrid	o0413, o0218	0467, 0468
DIG1-AtSDR4Lwus	Yeast 2 Hybrid	o0523, o0530, o0531, o0532	0467, 0468
DIG1wusC	Yeast 2 Hybrid	o0218, o0537	o0507
DIG1-AtSDR4Lcswap	Yeast 2 Hybrid	o0523, o0524, o0525, o0526	o0507
DIG1BetaBarrel	Yeast 2 Hybrid	o00218, o0800	0467, 0468
DIG1BetaBarrel-8AA	Yeast 2 Hybrid	o0800, o0816	0467, 0468
AtSDR4L-N	Yeast 2 Hybrid	o0415, o0416	0467, 0468
AtSDR4L-C	Yeast 2 Hybrid	o0417, o0418	0467, 0468
AtSDR4L-DIG1wus	Yeast 2 Hybrid	o0526, o0527, o0533, o0534	0467, 0468
AtSDR4LwusC	Yeast 2 Hybrid	o0418, o0542	o0507
AtSDR4L-DIG1cswap	Yeast 2 Hybrid	o0527, o0528, o0529, o0539	o0507
AtSDR4LBetaBarrel	Yeast 2 Hybrid	o0418, o0801	0467, 0468
AtSDR4LBetaBarrel-6AA	Yeast 2 Hybrid	o0801, o0817	0467, 0468
pER8-3xHA-DIG1-3xFLAG	Estradiol-Inducible	o0317, o0321, o0322, o0323	o0324, o0325
pER8-3xHA-DIG1-3xFLAG-TurboID	Estradiol-Inducible, TurboID	o0317, o0338, o0340, o0341	o0324, o0325
pER8-3xHA-AtSDR4L-3FLAG	Estradiol-Inducible	o0317, o0318, o0319, o0320	o0324, o0325
pCAMBIA1300-AtSDR4Lpro-3xHA-AtSDR4L-3xFLAG	Complementation Line	o0827, o0828, o0865, o0866	1160, 1440, 1441, 1442
pCAMBIA1300-AtSDR4Lpro-3xHA-AtSDR4L-6AATrunc-3xFLAG	Complementation Line	o0865, o0866, o0882, o0886, o0887, o0888, o0889	1160, 1440, 1441, 1442

pCAMBIA1300-AtSDR4Lpro-3xHA-AtSDR4L-BetaBarrelTrunc-3xFLAG	Complementation Line	o0865, o0866, o0882, o0883, o0884, o0885, o0889	1160, 1440, 1441, 1442
pHEE401E-VAL1VAL2-sg1234	CRISPR	o0941, o0942, o0943, o0944, o0945, o0946, o0947, o0948, o0949, o0950	o0080, o0081, o0951, o0952

List of primers:

Primer Name	Primer Sequence
o0201_SDR4_sg3sg4_colony_R	CGCTTGCGGGTTCTTGTAGG
o0211_pEXP_AD502_F1	CGGTCCGAACCTCATAACAACTC
o0212_pEXP_AD502_F2	TATAACGCGTTTGGAATCACT
o0216_pEXP_AD502_R	cgtaaatttctggcaaggtagac
o0217_NotI-cacc-DIG1_F	aaGCGGCCGCccccttcaccATGGACGGTAGGGGAG
o0218_AscI-DIG1stop_R	aaGGCGCGCCcaccctTCAAAGGCACAAAGCG
o0219_NotI-cacc-DIG2_F	aaGCGGCCGCccccttcaccATGAATTTCCGGGGAGGT
o0220_AscI-DIG2stop_R	aaGGCGCGCCcaccctTCACCGTCCAAGGCAA
o0253_sdr4_SALKseq_043688.2_SALK_203161_LP	CTTGTAAGCGGAGTTCACGAG
o0265_At1g27461.1_SDR4_q1F	CGGAAGTGATACAGGAGAGAGA
o0266_At1g27461.1_SDR4_q1R	TTATGGGCTGCGGTGATATG
o0267_At1g27461.1_SDR4_q2F	CTCGTGAACTCCGCTTACAA
o0268_At1g27461.1_SDR4_q2R	CACTTCTCCACAGATCCTCTTC
o0271_sdr4_CRISPR_GT_R	ACTAGAAGAACCGCATGACG
o0276_pEXP_AD502_HDA6_M1_F1	ACTTCCTCGCTTCCGTTTCG
o0277_pEXP_AD502_HAG1_M1_F1	AGAGAGGAGCAAGCAGGACG
o0278_pEXP_AD502_HAG1_M1_R1	TTTAGCACCAGATTGGAGACC
o0279_pEXP_AD502_HAG3_M1_F1	CCGTTGTGGCGGTTATGTCG
o0280_pEXP_AD502_HAM1_M1_F1	ATGACACGACACCAGAAACG
o0311_SALK059568_LP	TACAGTCCATGCAGTTGCAAG
o0317_pER8_3xHA_F	Gaagctagtcgactctagccatgtacccgtatgatgttccgga
o0318_3xHA_SDR4_R	gtatctttattcctgcgtagtctgggacgt
o0319_3xHA_SDR4_F	ctacgcaggaataaagatactcaacccccactcac
o0320_SDR4_pER8_R	cttcgaaccagggccctggcgctctgcggtcagtggttgtg
o0321_3xHA_DIG1_R	ccctaccgtctcctgcgtagtctgggacgt
o0322_3xHA_DIG1_F	ctacgcaggagacggtaggggagggtgttg

o0323_DIG1_pER8_R	cttcgaaccagggccctggcgcaaggcacaaagcggcct
o0324_pER8_Sanger_F	atgccatgtaatatgctcgactc
o0325_pER8_Sanger_R	TAGTAGGATTCTGGTGTGTG
o0336_SDR4_TurboID_R	GTATTGTCTTTTGGACCAGACCCTCCtctgcggtcagtggt tgtg
o0337_SDR4_TurboID_F	Tgaccgcagaggagggtctggtccaaaagac
o0338_TurboID_pER8_R	TTTGTAGTCTTCGAACCAGGGCCCTGGCGCcttttcggcag accgcagac
o0340_DIG1_TurboID_R	GTATTGTCTTTTGGACCAGACCCTCCaaggcacaaagcggc ct
o0341_DIG1_TurboID_F	TTTGTGCCTTggagggtctggtccaaaagac
o0344_SALKseq_033680.1_LP	GAACCCGTTTCCTAGCAGAAG
o0346_Wiscseq_DsLoxHs063_11D.2_L P	GAACCAGTCGACCAATACGAC
o0384_FIE2_AT2G35670_R	TCACGGAGACATCAACTTCC
o0385_RING1B/BMI1B_AT1G06770_R	ATCTCTCCTTCCTCTTTCCTGG
o0411_NotI-start-DIG1(1-164)_F	aaGCGGCCGCccccttcaccATGGACGGTAGGGGAG
o0412_AscI-DIG1(1-164)-stop_R	aaGGCGCGCCcaccctTCACGTTACACGCTCC
o0413_NotI-start-DIG1(165-294)_F	aaGCGGCCGCccccttcaccATGGACGCTTGGATTGACG
o0415_NotI-start-SDR4(1-219)_F	aaGCGGCCGCccccttcaccATGATAAAGATACTCAACCC
o0416_AscI-SDR4(1-219)-stop_R	aaGGCGCGCCcaccctTCAGTTTATGTAGGCTACGT
o0417_NotI-start-SDR4(220-354)_F	aaGCGGCCGCccccttcaccATGCCACTAACCAACCC
o0418_AscI-SDR4(220-354)-stop_R	aaGGCGCGCCcaccctTCTATCTGCGGTCAGTG
o0419_NotI-start-VAL2(1-411)_F	aaGCGGCCGCccccttcaccATGGAGTCAATAAAGGTT
o0420_AscI-VAL2(1-411)-stop_R	aaGGCGCGCCcaccctTCAATTCAAGCTGTTGGAA
o0421_NotI-start-VAL2(412-780)_F	aaGCGGCCGCccccttcaccATGCCGGGATGTGGTG
o0422_AscI-VAL2(412-780)-stop_R	aaGGCGCGCCcaccctTTAGTTCACAGGATCATGA
o0507_pDEST32seq_F	GACAGTTGACTGTATCGTCG
o0518_GFP_GT_R1	GGGTCAGCTTGCCGTAGG
o0519_GFP_GT_R2	CGATGCCCTTCAGCTCGATG
o0520_GFP_GT_R3	TCTCGTTGGGGTCTTTGCTC
o0521_AscI-VAL2(1-205)-stop_R	aaGGCGCGCCcaccctTCAAATCGAATTAGTCTTAC
o0522_NotI-start-VAL2(206-411)_F	aaGCGGCCGCccccttcaccATGTTCCAACTGGCCC
o0523_pD32-DIG1_F	aagcaggctccGCGGCCGCccccttcaccATGGACGGTAGG GGAG
o0524_DIG1-SDR4_R	cctcttccctctcaccattgaCGGTGCACCTTCTTCCAC
o0525_DIG1-SDR4_F	GTGGAAGAAGGTGCACCGtcaatggtgagagggaagagg
o0526_SDR4-pD32_R	AGaaagctgggtcGGCGCGCCcacccttctatctgcggtca gtggttgtg
o0527_pD32-SDR4_F	aagcaggctccGCGGCCGCccccttcaccATGATAAAGATA CTCAACCCCCAC
o0528_SDR4-DIG1_R	GAAAACTATCGCCGCTAGTTATCTCatccagccatgagcac tctg
o0529_SDR4-DIG1_F	cagagtgctcatggctggatGAGATAACTAGCGGCGATAGT TTTC

o0530_DIG1-pD32_R	AGaaagctgggtcGGCGCGCCcacccttTCAAAGGCACAAAGCGGCC
o0531_DIG1-SDR4WUS_R	agcattggagaagaggcaaagtcaccaaATTAGCACTGCCGTGAACGA
o0532_SDR4WUS-DIG1_F	tttgcctcttctccaatgctctcctcctGATCTTAAGGCGGCGGTGG
o0533_SDR4-DIG1WUS_R	TCTCCGGTAATAGAGAAAGAGTCACCGCagtagaagagacaccattctcaaaac
o0534_DIG1WUS-SDR4_F	TCTTTCTCTATTACCGGAGAAACCTATAtcctccaaatgcatggagc
o0535_VAL2_PHD_F1	GAGTGTAATTCATGTGACAAGCG
o0536_VAL2_PHD_R1	GTGTCTAAGAGCATCTAGGC
o0537_DIG1WUS_F	aaGGCGCGCCccccttcaccAGTGCTAATGCGGTGACTC
o0541_3FLAG_R	CTTATCGTCATCGTCCTTGTAATCG
o0542_SDR4_WUS_F	aaGGCGCGCCccccttcaccACTCCTACTTTGGTGACTTTGCC
o0800_NotI_DIG1-betabarrel_F	aaGCGGCCGCccccttcaccGCTTATCAGACACCGG
o0801_NotI_SDR4-betabarrel_F	aaGCGGCCGCccccttcaccagcagagtcatatcaccg
o0816_AscI_DIG1-8aaTrunc_R	aaGGCGCGCCcacccttTCAGAGCCTCCAAGCAAAAC
o0817_AscI_SDR4-6aaTrunc_R	aaGGCGCGCCcacccttTCAgtgaaacctccacgtaaag
o0826_HiFi_pCAMBIA1300_XbaI_SDR4pro_F	gcctgcaggtcgacTCTAGACAGAGGATGAATCGGCC
o0827_HiFi_SDR4pro_3xHA_R	catacgggtacatggctagaGACGTCAAGAATGATaaaaaaaaaaaaCTC
o0828_HiFi_SDR4pro_3xHA_F	tttttATCATTCTTGACGTCtctagccatgtacccgtatg
o0829_HiFi_pCAMBIA1300_KpnI_3xFLAG_R	gattacgaattcgagctcGGTACCTCACTTATCGTCATCGTCCTTG
o0865_HiFi_pCAMBIA1300_EcoRI_SDR4pro_F	CAGCTATGACCATGATTACGAATTCCAGAGGATGAATCGGCC
o0866_HiFi_pCAMBIA1300_PstI_3xFLAG_R	GATGATACGAACGAAAGCTCTGCAGTCACTTATCGTCATCGTCCTTG
o0882_HiFi_pCAMBIA1300_EcoRI_short_F	CAGCTATGACCATGATTACGAATTC
o0883_HiFi_SDR4betabarrel_3xFLAG_R	AACCAGGGCCCTGGCGCgcaaaccggccttatgg
o0884_HiFi_SDR4betabarrel_3xFLAG_F1	ccataaggccggtttgcGCGCCAGGGCCCTGGTTCGAA
o0885_HiFi_SDR4betabarrel_3xFLAG_F2	ccataaggccggtttgcGCGCCAGGGCCCT
o0886_HiFi_SDR4-6AA_3xFLAG_R	AACCAGGGCCCTGGCGCGTGAAACCTCCACGTAAAGAC
o0887_HiFi_SDR4-6AA_3xFLAG_F1	GTCTTTACGTGGAGGTTTCACGCGCCAGGGCCCTGGTTCGAA
o0888_HiFi_SDR4-6AA_3xFLAG_F2	GTCTTTACGTGGAGGTTTCACGCGCCAGGGCCCT
o0889_HiFi_pCAMBIA1300_PstI_3xFLAG_short_R	GATGATACGAACGAAAGCTCTGCAGTCA
o0081_U6-29p-F	TTAATCCAAACTACTGCAGCCTGAC
o0941_VAL1-DT1-BsF	ATATATGGTCTCGATTGCATGTAAATGGCGAAAGCGGGTT
o0942_VAL1-DT1-F0	TGCATGTAAATGGCGAAAGCGGGTTTTAGAGCTAGAAATAGC
o0943_VAL1-DT0-BsR2	ATATTATTGGTCTCAATCTCTTAGTCGACTCTACCAAT

o0944_VAL1-DT2-BsF2	ATATTATTGGTCTCAAGATTGGAAGAGTTCATGCCTCACCGGTT
o0945_VAL1-DT2-F0	TGGAAGAGTTCATGCCTCACCGGTTTTAGAGCTAGAAATAGC
o0946_VAL1-DT0-BsR3	ATATTATTGGTCTCATCACTACTTCGTCTCTAACCAT
o0947_VAL2-DT3-BsF3	ATATTATTGGTCTCAGTGATTGGTTTGCATGAACGCACTGTGGTT
o0948_VAL2-DT3-F0	TGGTTTGCATGAACGCACTGTGGTTTTAGAGCTAGAAATAGC
o0949_VAL2-DT4-R0	AACTTACCCACACTTATCGCAGACAATCACTACTTCGACTCTAGCTGTAT
o0950_VAL2-DT4-BsR	ATTATTGGTCTCTAAACTTACCCACACTTATCGCAGA
o0951_U6-1t-F	GCTAAGACAAAGTGATTGGTCCGTT
o0952_U6-1t-R	AACGGACCAATCACTTTGTCTTAGC
o0953_VAL1-sg1sg2-F	TAGACTCGGTTCAGCGTATTCG
o0954_VAL1-sg1sg2-R	CAGGAGGAGTTGAAATAGCAGG
o0955_VAL2-sg3sg4-F	GTGGATTAAAGCACCATCGTGC
o0956_VAL2-sg3sg4-R	ATGGATTGCTCGTATGCACACC
0013_M13U	TGTAAAACGACGGCCAGT
0467_GAL4DBD_F	GAACAACTGGGAGTGTCGCTAC
0468_ADH1TT_R	CAAGGTAGACAAGCCGACAAC
0517_at3g48510_CDS149_R	CCGCTGGATACGGGATTTAC
1160_pCAMBIA3300_R	CTCTAGCATTCGCCATTCAGG
1168_pMDC7_LexA35Spro_F	ACAGCAGTCGAGGTAAGATTAG
1300_At3g48510_CRISPRveri_F	GCATCATGCTTCGATTTCGTC
1355_At1g27461stop_R	CTATCTGCGGTCAGTGGTTGTGT
1421_At1g27461_gb_noStop_R	TCTGCGGTCAGTGGTTGT
1423_DIG1_At3g48510_gb_noStop_R	AAGGCACAAAGCGGCCTT
1440_At1g27461_F2	TGTTAAACGGCGAAGTCAGG
1441_At1g27461_F3	TCTCTTTGATCTCTCACCATCC
1442_At1g27461_F4	AACTACAAGGACCCATCACC
1493_At3g48510_junctionC_F	TCCTGGTTTCATATCCGACG