Universidade do Minho
Escola de Ciências

Sweta Singh

Functional analysis of a gene regulatory network involved in flower zygomorphy

julho de 2020

Functional analysis of a gene regulatory network involved in flower zygomorphy
Sweta Singh

UMinho | 2020

This work was funded by Fundação para a Ciência e Tecnologia/Ministério da Ciência, Tecnologia e Ensino Superior through national funds (Programa de Investimento e Despesas de Desenvolvimento da Administração Central) with a project grant PTDC/BIAPLA/1402/2014"Evomod-Origin and evolutionary establishment of a transcriptional module controlling flower asymmetry" and Ph.D. grant, ref. PD/BD/113645/2015. This work was supported by PEst-OE/BIA/UI4046/2014; UID/MULTI/04046/2013 center grant from FCT, Portugal to BioISI.

Universidade do Minho
Escola de Ciências

Sweta Singh

Functional analysis of a gene regulatory network involved in flower zygomorphy

Tese de Doutoramento
Doutoramento em Ciências
Especialidade em Biologia

Trabalho efetuado sob a orientação da
Professora Doutora Maria Manuela Ribeiro Costa
e do
Professor Doutor Jóse Pío Beltrán

julho de 2020

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.
Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.
Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho

Atribuição-NãoComercial-SemDerivações
CC BY-NC-ND
https://creativecommons.org/licenses/by-nc-nd/4.0/

Acknowledgments

This Journey of Ph.D. was magnificent as well as challenging for me. During all these years, many people were instrumental either directly or indirectly to shape up this project. This is an opportunity to express my gratitude to them.

First of all, I would like to express my sincere thanks to my supervisor and mentor Prof. Manuela Costa for her continuous support, patient, motivation and immense knowledge. She guided me at every single step during the research and writing this thesis. I could not have imagined having a better supervisor than her. I would like to extend my gratitude to Prof. Jóse Pío for giving me an opportunity to work in his enthusiastic and creative group. I want to express my deepest thanks to Dr. Concha Gómez, not only guided me to achieve my research objective but also supported me during my stay in Valencia. Her suggestion always helped to find the right direction and troubleshooting the problems.

I would like to thank Dr. Luis Canãs, Mari Cruz, and Edilin for not only making me familiar with Medicago but also their sincere care and support during my stay in IBMCP. Thank you, Rosa, Marisol and M. D. Gómez for unconditional help and support. A big heartfelt thanks to Rim Hamza for your friendship that made my stay so memorable.

I would also like to thank the members of CITAB and my Agrichains fellows who helped me to acclimatize in Portugal at the very beginning of my Ph.D. when I did know nothing about this country and language. I would like to thank. I would like to thank my tutor, Prof. Ana Cunha. She was always ready to help me whenever I called her. I also thank Richard, João, Weina, and Caterina who introduced me to UMinho and showed me all the equipment and facilities.

A big thanks to all the members of my lab for their care, support, and friendship. I would like to extend my heartfelt thanks to Prof. Teresa. I always learned something new about science, culture and history from her. I thank Prof. Rui Tavares for taking care of administrative issues so that we could smoothly perform our work. Thanks to Denial and Helena for all your help and friendship. I would like to appreciate the effort of Helena explaining to me very complicated Portuguese emotions and expressions in English during our café time. Thank you, Ana Teresa and Ana Maria for being my translator. I wish good luck to both of you with your Ph.D. I would like to thank, Francisca, Eunice and Nice for all their support and timely help. Thank you, Jorge, Dianna, and Arunav for helping me to access the microscope of your department. It was always delightful to talk to Mário and Sara. I wish you the very best luck for all the new ventures in your life. I could not imagine this journey without Sara and Rómulo. Thank you, Sara, for all the positivity you bring and lightening ups the lab with your vibrant spirit. Thanks, Rómulo for listening to my confused ideas and always showing me the right direction. You always had some solutions to my problems. I wish both of you all the success and love in your life.

I would like to thank all the members of the biology department who were always ready to help me. I would like to mention the name of Viviana, Henrique, Mário, and Raul, who helped me unconditionally whenever I asked them.I would like to thank my friends Rupesh, Balaji, Sushil, Praveen Sher, Mouli, Vinay, Sandra, Elena and many more for making my life very easy and joyful outside the lab. I would like to thank my parents, siblings and lovely niece Nimisha for being inspiration and motivation for me.

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration.

I further declare that I have fully acknowledged the Code of Ethical Conduct of the University of Minho.

Análise funcional de uma rede genética regulatória envolvida na zigomorfia floral

RESUMO

A arquitetura das plantas, a germinação das sementes e o tempo de floração são determinantes importantes do desempenho das plantas cultivadas. Compreender os mecanismos moleculares subjacentes a essas características pode ajudar a optimizar esses recursos e aumentar o rendimento das culturas. A zigomorfia das flores é uma característica morfológica que emergiu de forma independente durante a evolução,associada a polinizadores específicos de plantas. Em *Antirrhinum majus*, a zigomorfia da flor requer a atividade combinada de quatro fatores de transcrição: *CYCLOIDEA* (*CYC*), *DICHOTOMA* (*DICH*), *RADIALIS* (*RAD*) e *DIVARICATA* (*DIV*). O gene *CYC* é expresso no domínio dorsal do meristema floral e é um gene chave que promove as identidade das pétalas dorsais. O estudo dos promotores dos genes *CYC-like* em *Medicago truncatula*, uma espécie leguminosa com flor zigomórfica,é importante para entender a evolução dessa nova característica.Um dos objetivos desta tese foi obter conhecimento sobre os reguladores a montante do *CYC* que podem ser responsáveis pela sua ativação nos domínios dorsal do meristema das flores. Descobrimos que existem dois genes *CYC-like* em *Medicago truncatula* que são expressos apenas no domínio dorsal do meristema floral e nas pétalas dorsais de flores maduras.A sequência promotora de *CYC* em *Medicago* é similar às sequências promotoras de *CYC* em outros legumes. Este estudo revelou assim um possível papel de elementos reguladores específicos no padrão de expressão assimétrico dos homólogos de *CYC* nas espécies com flores zigomórficas.

Estudos genéticos e moleculares revelaram que RAD atua a jusante de *CYC* e que antagoniza a atividadede DIV na região dorsal de *Antirrhinum*, competindo por DIV-and-RAD-interacting factor (DRIF). Uma ação antagónica similar entre homólogos de RAD/DRIF/DIV regula diferentes processos de desenvolvimento em outras espécies. Neste trabalho, a análise funcional dos genes *DRIF* em *Arabidopsis thaliana DRIF3*, *DRIF4* e *DRIF5* revelou que estes genes estão envolvidos na germinaçãode sementes e no desenvolvimento inicial da plântula em resposta à luz. Neste estudo, descobrimos que esses homólogos de *DRIF* em *Arabidopsis* provavelmente estão envolvidos na regulação negativa da germinação de sementes e no desenvolvimento do gancho apical em plantulas crescidas no escuro, influenciando a síntese e/ ou sinalização de hormonas vegetais.Estes resultados mostram que os genes envolvidos na zigomorfia da flor foram recrutados para diferentes funções em espécies actinomórficas.

Palavras chave: *Antirrhinum majus*, fotomorfogénese, germinação da semente, regulamento de transcrição, zigomorfia floral.

Functional Analysis of a gene regulatory network involved in flower zygomorphy

ABSTRACT

The overall plant architecture, seed germination and flowering time are important determinants of the performance of crop plants. Understanding the molecular mechanisms underlying these major traits can help to optimize these features and enhance plant yield. Flower zygomorphy is a morphological trait that emerged in different independently evolutionary events, normally associated with specific plant-pollinators. In *Antirrhinum majus*, the flower zygomorphy requires the combined activity of four transcription factors: *CYCLOIDEA* (*CYC*), *DICHOTOMA* (*DICH*), *RADIALIS* (*RAD*) and *DIVARICATA* (*DIV*). *CYC* is expressed in the dorsal domain of the flower meristem and is a key gene that promotes dorsal petal identity. The study of the nature of the promoters of *CYC-like* genes in *Medicago truncatula*, a legume species with a zygomorphic flower, is important to understand the evolution of this new trait. One of the objectives of this thesis was to gain knowledge on upstream regulators of *CYC* that may be responsible for its activation in the dorsal domains of flower meristem. We found that there are two *CYC-like* genes in *Medicago truncatula* that are expressed only in the dorsal domain of the floral meristem and in dorsal petals of mature flowers. The promoter sequence of *Medicago CYC* showed similarity with the promoters of other legume *CYC-like* genes. This study proposed a role of specific regulatory elements in the asymmetric expression pattern of the *CYC* homologs in the species having zygomorphic flowers.

Genetic and molecular studies have revealed that RAD acts downstream of *CYC* and antagonizes the activity of DIV in the dorsal region of the *Antirrhinum* flower by competing for a DIV-and-RAD-interacting factor (DRIF). A similar antagonistic action between RAD/DRIF/DIV homologs regulates different developmental processes in other species. In this work, the functional analysis of the *DRIF* genes in *Arabidopsis thaliana* showed that *DRIF3*, *DRIF4* and *DRIF5* are involved in seed germination and early seedling development in response to light. In this study, we found that these *DRIF* homologs in *Arabidopsis* are likely involved in negative regulation of seed germination and development of the apical hook in dark-grown seedlings by influencing the synthesis and/or signaling of plant hormones. These results showed that the genes involved in flower zygomorphy have been co-opted for different functions in actinomorphic species.

Key words: *Antirrhinum majus,* Flower zygomorphy, photomorphogenesis, seed germination, transcription regulation.

Table of Contents

Acknowledgments iii

RESUMO v

ABSTRACT vi

List of Abbreviations x

Chapter 1 "General introduction" 1

1.1 Flower symmetry and evolution 2

1.2 Gene regulatory network controlling floral zygomorphy in *Antirrhinum majus* 5

1.3 Antagonistic interaction between RAD and DIV to maintain flower symmetry 7

1.4 DIV- DRIF-RAD interaction recruited for different functions in other species 8

1.5 Evolution of *CYC-like* genes in different species as a key regulator of flower zygomorphy 9

1.6 Independent recruitment of *CYC-like* genes in different species for floral symmetry 10

1.7 Effect of spatial-temporal change in the expression of *CYC2* genes in flower symmetry 10

1.8. Thesis outline 13

1.9 References 14

CHAPTER 2 "Characterization of *CYC2-like* genes in *Medicago* and their role in the control of flower zygomorphy" 19

2.1 Abstract 21

2.2 Introduction 21

2.3 Materials and Methods 23

2.3.1 Plant Material and growth conditions 23

2.3.2 Identification of *CYC* homologs in *Medicago truncatula*, *Glycine max*, and *Phaseolus vulgaris* and phylogenetic analysis 24

2.3.3 RNA isolation and expression analysis 24

2.3.4 Cloning and construction of pro*CYC2a*::*GFP* fusion vectors for *Medicago sativa* transformation 25

2.3.5 β-glucuronidase (GUS) staining 25

2.4 Results 26

2.4.1 Phylogenetic analysis of *Medicago CYC* homologs 26

2.4.2 Expression pattern of *MtCYC2a* and *MtCYC2b* 27

2.4.3 Analysis of *MtCYC2a* promoter elements responsible for dorsal expression in flower 29

2.4.4 Expression pattern and activities of *AtTCP1* promoter and its 5′ deletion segment in *Arabidopsis thaliana* transgenics 35

2.5 Discussion 37

2.6 References 40

2.7 Supplementary data 43

CHAPTER 3 "Co-expression and regulation of *DIV-and-RAD-INTERACTING-FACTORS* (*DRIFs*) in *Arabidopsis* - a Bioinformatics approach" 49

3.1 Abstract 51

3.2 Introduction 52

3.3 Material and Methods 53
3.3.1 *DIV*, *RAD,* and *DRIF* homologs in *Arabidopsis* 53
3.3.2 *In silico* expression analysis of *DIVs*, *RADs,* and *DRIFs*..... 53
3.3.3 Analysis of co-expression network and protein-protein interactions 54
3.3.4 *In silico* analysis of the putative promoter of *Arabidopsis DRIFs* homologs..... 54

3.4 Results 54
3.4.1 Expression pattern of *RAD*, *DIV*, and *DRIF* homologs in *Arabidopsis* 54
3.4.2 Expression of *DIV*, *RAD*and *DRIFs* in different mutant backgrounds..... 59
3.4.3 *In silico* investigation of co-expression networks and protein interactions 61
3.3.4. *In silico* analysis of promoters of *Arabidopsis DRIF* genes 64

3.5 Discussion 65

3.6 References..... 68

3.7 Supplementary data 71

CHAPTER 4 "The MYB-like encoding genes, *DRIF3*, *DRIF4*, and *DRIF5* are involved in seed germination in *Arabidopsis thaliana*" 93

4.1 Abstract 95

4.2 Introduction 96

4.3 Material and Methods 99
4.3.1 Plant Material and growth conditions..... 99
4.3.2 Seed germination and stress assays 100
4.3.3 Gene Expression Analysis 100

4.4 Results 101
4.4.1 Expression of *DRIFs* during seed maturation and germination of *Arabidopsis* seeds 101
4.4.2. *ArabidopsisDRIF3*, *DRIF4*, and *DRIF5* regulate germination in a partially redundant manner 102
4.4.3. Germination of *drif345*showed improved tolerance to salt and heat stress 103
4.4.4 Effect of ABA, GA3, and PAC on *drif345* germination..... 105
4.4.5. Effect of *drifs* mutation on ABA and GA biosynthesis and signaling genes during germination 108

4.5 Discussion 110

4.6 References..... 114

4.7 Supplementary data 119

CHAPTER 5 "Role of *ArabidopsisDIV-and-RAD-INTERACTING-FACTORS* (*DRIFs*) in skotomorphogenic andphotomorphogenic development of seedlings" 122

5.1 Abstract 124

5.2 Introduction 125

5.3 Material and Methods 127
5.3.1 Plant Material and Growth Conditions 127
5.3.2 Apical Hook and Hypocotyl Measurements 128
5.3.3 RNA Isolation and Gene Expression Analysis 128

5.4 Results 128
5.4.1 Analysis of diurnal expression of *DRIF3*, *DRIF4*, and *DRIF5* 128
5.4.2. *drif345* mutant showed hookless phenotype in dark 130
5.4.3. Expression of *HLS1* was reduced in *drif345* 132
5.4.4. Reduced hypocotyl elongation of *drif345* in light 133
5.4.5 Effect of application of GA3 on *drif345* seedlings 135
5.4.6 Expression of *PHYB* was increased in *drif345* seedling 137

5.5 Discussion 137

5.6 References 140

Concluding remarks 144

List of Abbreviations

ABA	Abscisic acid
BL	blue light
cDNA	complementary DNA
CRE	*cis*-regulatory element
CYC	CYCLOIDEA
DDR module	DIV-DRIF-RAD module
DICH	DICHOTOMA
DIV	DIVARICATA
DNase	deoxyribonuclease
DRIF	DIV-and-RAD-interacting-factor
FR	far-red light
GA	Gibberellins
gDNA	genomic DNA
GNR	gene regulatory network
LD	long day
NaCl	sodium chloride
PCR	polymerase chain reaction
RAD	RADIALIS
RL	red light
RNase	ribonuclease
TF	transcription factor
WT	wild type

Chapter 1
“General introduction”

1.1 Flower symmetry and evolution

Understanding the origin and evolution of new traits in living organisms is one of the most fascinating problems in evolutionary developmental biology. Developmental mechanisms are generally controlled by networks of transcription factors known as gene regulatory networks (GNR). The general structure of a developmental GNR comprises the regulatory regions of the genes called *cis*-regulatory elements (CREs) and the transcription factors (TFs) that bind to CREs and enhance or suppress the gene expression. Changes in the *cis*-regulatory region of genes (Doebley and Lukens, 1998; Wagner and Lynch, 2010; Wittkopp and Kalay, 2011) and changes in the activity and expression pattern of transcription factors (Lang *et al.*, 2010; de Mendoza *et al.*, 2013) are the main molecular mechanisms that underlie the arousal of structural and functional novelty during evolution. The transcription factors can influence the expression of their downstream genes by binding with their promoters either directly or by forming a complex with other proteins. Therefore, changes in protein-protein interaction and the formation of different transcriptional regulatory complexes can also lead to the development of new traits during evolution. By characterizing the GNR involved in the evolution of novel traits, we can understand how these GNRs were evolved and co-opted in different species for various developmental processes. The GNR involved in the regulation of flower symmetry in angiosperms presents an excellent opportunity to understand how the changes in the *cis*-regulatory network and the interaction of transcription factors can bring novel morphological traits.

Flower symmetry is an important structural trait in angiosperms, which is tightly related to the variation in shape and arrangement of floral parts. Generally, the symmetry of the flower is associated with the arrangement and shape of petals. A flower displaying several plans of symmetry is said to be radial symmetric (actinomorphic) (Fig.1.a) and a flower displaying only one plan of symmetry is said to be bilateral symmetric flowers (zygomorphic) (Fig.1.b). In extreme cases, a flower can possess complete asymmetry without any plan of symmetry called asymmetry (Fig.1.c). Asymmetrical flowers are present in basal angiosperms and primitive dicot family Magnoliaceae, which have a spiral arrangement of indeterminate numbers of male and female organs (Endress, 1999).

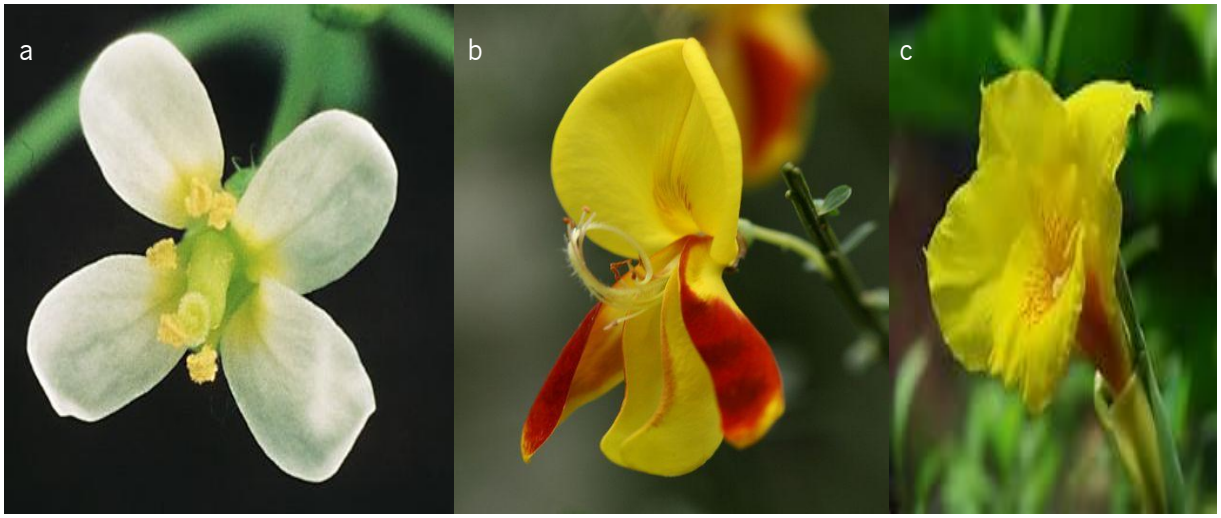


Fig.1. Types of floral symmetry. (a) **Radial symmetry** in *Arabidopsis thaliana*, (b) **Bilateral symmetry** in *Cytisus scoparius*, and (c) **Asymmetry** in *Canna indica*. Images obtained from calflora.org, phys.org, and wikipedia.org.

Flower zygomorphy is a trait that has been evolved from actinomorphy independently at least 25 times during evolution (Endress, 1999). Fossil records suggest that the origin of floral zygomorphy has coincided with the diversification of insect pollinators (Neal *et al.*, 1998). The transition of flower symmetry from actinomorphic to zygomorphic is associated with specialized plant-pollinator association as the position of petals and stamens ensures a more efficient pollen transfer to and from the pollinators (Dilcher, 2000; Ushimaru *et al.*, 2009). The structural features of zygomorphic flower that enhance the insect pollination are; the specific position of ventral petals that provide an ideal platform for pollinators, conversion of dorsal stamens into aborted staminodes and the positioning of style and fertile stamens that enhance the chances of pollen transfer from one flower to another via specific pollinators. This specific plant-pollinator interaction eventually associated with the origin of reproductive isolation and initiated extensive speciation (Neal *et al.*, 1998). The predominance of zygomorphy in species-rich families, Lamiaceae, Leguminosae, and Orchidaceae explains the evolutionary significance of this floral trait in higher plants (Endress, 2012).

Phylogenetic studies have given insight into the evolution of zygomorphy from actinomorphy in various clades (Preston *et al.*, 2011; Preston and Hileman, 2012; Zhong *et al.*, 2017). The ancestral actinomorphy of flowers dominates in basal angiosperms and is maintained in monocots and some eudicots. Zygomorphy evolved in many clades independently from the Late Cretaceous (Dilcher, 2000). Independent reversal to actinomorphy was also a common event during evolution. In some families like Orchidaceae, Fabaceae, Brassicaceae, floral zygomorphy occurred as an early event, but at later

evolutionary stages independent reversion to actinomorphy was observed (Spencer and Kim, 2018) (Fig.2).

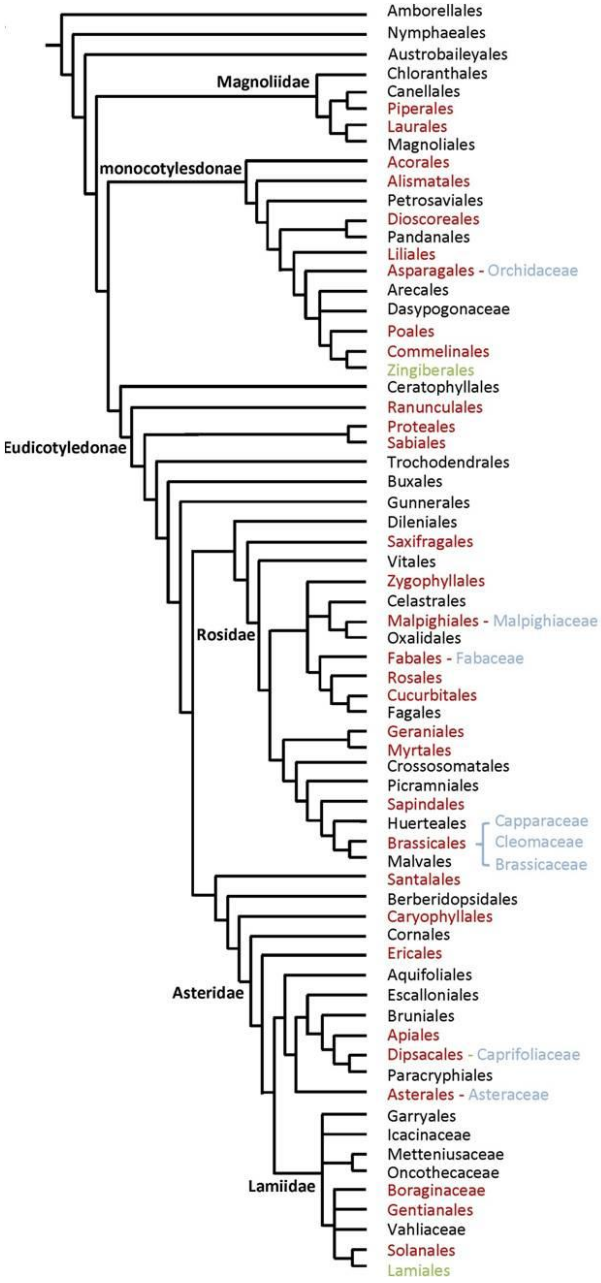


Fig.2.Phylogeny of evolution of floral symmetry in angiosperm reveals the presence of three groups 1. Presence of ancestral actinomorphy maintained in most clade members (black text), 2. Zygomorphy derived as an independent event (red text), and 3. Zygomorphy occurred as a single early event but the independent reversion to actinomorphy occurred during later stages of evolution (blue text). (Spencer and Kim, 2018).

1.2 Gene regulatory network controlling floral zygomorphy in *Antirrhinum majus*

A GNR involved in the control of flower zygomorphy has been studied in *Antirrhinum majus* using classic floral mutants such as *cycloidea* (*cyc*), *dichotoma* (*dich*), *radialis* (*rad*) and *divaricata* (*div*) (Fig.3). A wild-type *Antirrhinum* flower shows a distinct zygomorphic flower with two large dorsal petals, two lateral petals, and one ventral petal. The floral symmetry mutants of *Antirrhinum* lose the distinct identity of petals either completely or partly. Genetic studies showed that *CYC*, *DICH,* and *RAD* are responsible for the dorsal identity of a flower while *DIV* determines ventral identity (Luo *et al.*, 1996; Almeida *et al.*, 1997; Luo *et al.*, 1999; Galego and Almeida, 2002; Corley *et al.*, 2005).

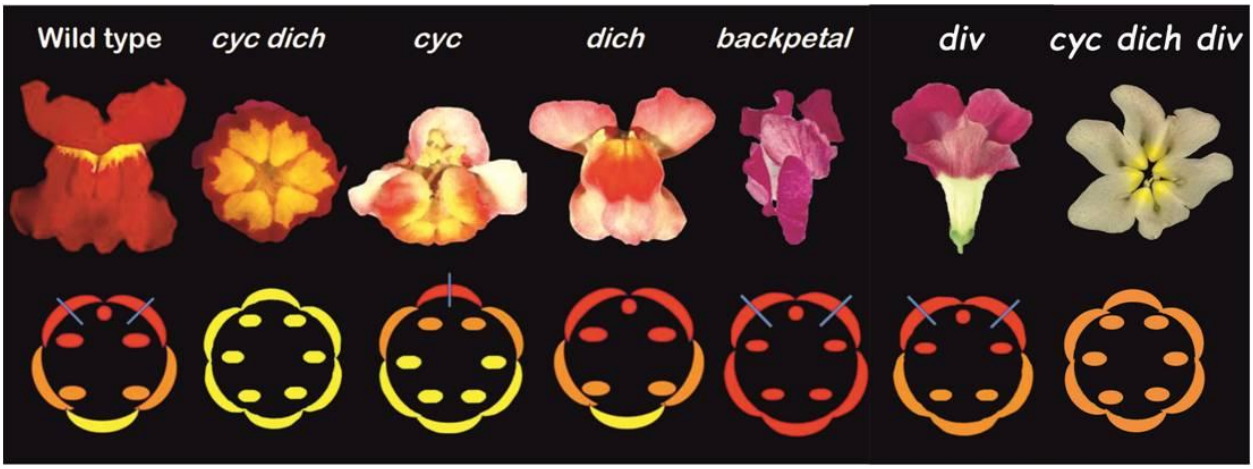


Fig.3. A flower phenotype of wild type *Antirrhinum* and its mutants. Floral diagram showing the petal and stamen identity and their distribution. The outer circle represents the petals and inner oval structures represent the stamens and a small circle represents the staminode. The color represents the identity of stamens and petals: red color (dorsal identity), yellow color (ventral identity), orange color (lateral identity), blue line (internal petal symmetry). (Raimundo, 2015)

The dorsal identity of the *Antirrhinum* flower is specified by the partially redundant functions of two closely related TCP-family transcription factors, *CYC* and *DICH* (Luo *et al.*, 1996; Luo *et al.*, 1999). Both *CYC* and *DICH* are expressed in the dorsal region of the *Antirrhinum* flower from flower meristem initiation until the early stages of petal and stamen differentiation. At the later stages of flower development, the expression of *CYC* persists throughout the dorsal domain while the expression of *DICH* becomes restricted to the outer half of the dorsal domain (Luo *et al.*, 1999). Early expression of *CYC* reduces the growth rate in the dorsal domain of the developing flower meristem by negatively regulating the expression of *D-CYCLIN*, a gene involved in cell-cycle (Gaudin *et al.*, 2000). The reduced growth rate during these early stages of development is crucial for establishing proper organ number in

the dorsal region (Luo *et al.*, 1999; Almeida and Galego, 2005). Late expression of *CYC* increases the growth rate and affects cell type specification in the petal whorl, while in the stamen whorl, late expression of *CYC* leads to abortion of the dorsal stamen (Luo *et al.*, 1996).The role of *CYC* in the maintenance of zygomorphy in the *Antirrhinum* flower was confirmed by *cyc* loss-of-function mutants, in which flowers often have one extra petal and stamen. The lateral petals of *cyc* mutant flowers adopt a ventral identity and have a fertile stamen in place of dorsal staminode (Luo *et al.*, 1996). In contrast, the *dich* mutant exhibits partial loss of internal asymmetry of the dorsal petals (Luo *et al.*, 1999). Flowers of *cyc:dich* double mutant show complete loss of dorsal identity. Both dorsal and lateral petals acquire ventral identity and two fertile stamens develop in the dorsal region (Luo *et al.*, 1996).

Flowers of *rad* mutant have a very similar phenotype to *cyc* mutant flowers. The expression of *RAD* in the dorsal part of the floral meristem occurs after the expression of *CYC* and *DICH* and the expression of *RAD* is absent in a *cyc:dich* mutant (Corley *et al.*, 2005). Further, the presence of the CYC-binding site on the promoter of *RAD* suggests the activation of *RAD* by CYC in the dorsal region of the developing flower (Costa *et al.*, 2005).

The ventral identity of the *Antirrhinum* flower is specified by *DIV* (Almeida *et al.*, 1997; Galego and Almeida, 2002). The expression of *DIV* is necessary for ventral petal shape and to some extent, the shape of lateral petals. In *Antirrhinum div* mutant, the ventral petal and the adjacent part of the lateral petals, adopt a more lateral identity, whereas the dorsal petals and the most dorsal part of the lateral petals remain as a wild type flower. In early flower development, *DIV* is expressed in both the dorsal and ventral domains of the flower meristem but its expression becomes somewhat restricted to the ventral petal at later stages of development (Almeida et al. 1997). It has been reported that in the absence of dorsal identity genes in *cyc:dich* mutant the dorsal petals acquire ventral identity (Luo *et al.*, 1996). This observation suggests that the antagonistic effect over *DIV* in the dorsal floral meristem is exerted at the post-transcriptional level by the dorsal identity genes.

The dorsal identity gene *RAD* and ventral identity gene *DIV*, both belong to MYB family proteins.*RAD* encodes a 93 amino acid protein, containing a single MYB-like domain, closely related to the N-terminal MYB-like domain of DIV. The molecular similarity between RAD and DIV proteins suggested that *RAD* may have been derived from a *DIV-like* ancestral gene through C-terminal loss and it is likely to be the dorsal factor antagonizing DIV (Corley et al. 2005). This hypothesis was supported by Cui et al. (2010), who observed that plants ectopically expressing *RAD* produce flowers without ventral identity that resemble the *div* mutant flowers.

1.3 Antagonistic interaction between RAD and DIV to maintain flower symmetry

Raimundo et al. (2013) had proposed that flower zygomorphy in *Antirrhinum majus* depends on the transcriptional regulatory module composed of three MYB-family proteins: DIV, RAD and DIV-RAD-interacting-factors (DRIFs), named as DDR-regulatory module. In that study, they determined the consensus DNA binding site of the DIV protein and showed that RAD did not bind to this site. Moreover, RAD did not displace DIV from DNA-DIV complex and RAD and DIV proteins did not interact in a yeast two-hybrid assay. Using RAD as bait in a yeast two-hybrid screen, a pair of novel MYB proteins was identified; DRIF1 and DRIF2 that also bound to DIV and form DIV-DRIF-DNA complex. The DIV-DRIF-DNA complex was disrupted by the addition of RAD, which competes with DIV not for the DNA but for the DRIF protein. Analysis of subcellular localization of these proteins revealed that RAD is present in both cytoplasm and nucleus, while DIV is found only in the nucleus. DRIF1 is nuclear localized in the presence of DIV but present in the cytoplasm in the presence of RAD.

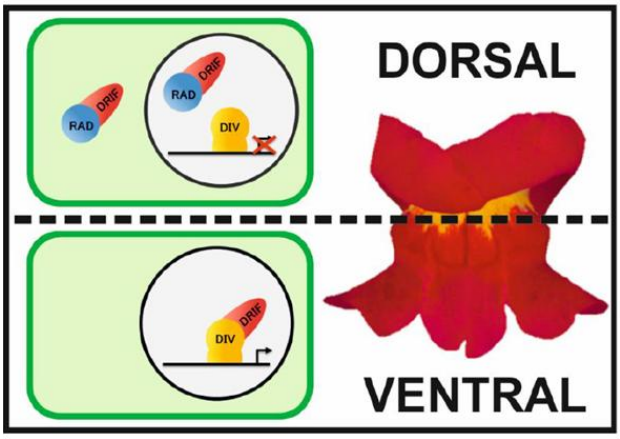


Fig.4. RAD-DIV antagonistic model proposed for *Antirrhinum* flower asymmetry. The formation of DIV-DRIF heterodimer in the ventral region promotes the expression of genes specifying the ventral identity of a flower. In the dorsal region, RAD competes with DIV for binding with DRIF. Without the formation of the DIV-DRIF complex, DIV cannot activate the ventral identity genes in the dorsal region (Raimundo *et al.*, 2013).

These results suggested that the antagonism of the function of DIV to determine ventral floral activity occurs in the dorsal part of the floral meristem where RAD is also present. This is due to the ability of RAD to sequester the DRIF proteins in the cytoplasm, preventing their interaction with DIV and consequently preventing the DIV protein from activating the transcription of genes regulated by DIV-DRIF complex (Fig.4).

1.4 DIV-DRIF-RAD interaction recruited for different functions in other species

The interaction between DIV, DRIF and RAD-like proteins occurs also in other species. For instance, in *Solanum lycopersicum* (tomato) a *RAD* homolog (*FSM1*) negatively regulates cell expansion to the tomato fruit pericarp by competing with the DIV homolog (MYB1) for the interaction with a DRIF homolog protein (FSB1). In the absence of FSM1, the MYB1 and FSB1 establish a regulatory complex that enhances cell expansion of the tomato fruit pericarp (Rose *et al.*, 1999; Machemer *et al.*, 2011). These studies suggest the DDR-regulatory network may be conserved and this module has most likely been recruited during evolution to perform different roles across the angiosperm lineage. Proteins homologous to DIV, DRIF or RAD are also involved in various functions in other species. In rice, *DIV* homologs are involved in sugar and hormone-regulated expression of the α-amylase gene (Lu *et al.*, 2002) and the *RAD* homolog in cotton regulates ovule development and fiber elongation (Zhang *et al.*, 2010a).

In *Arabidopsis thaliana*, several *RAD*, *DIV* and *DRIF* homologs are present but not well characterized because the single mutants of these genes characterized in previous studies have not shown any strong phenotypic variation from wild type (Hamaguchi *et al.*, 2008; Fang *et al.*, 2018).The overexpression lines of *Arabidopsis RAD* homolog *RSM1* (also mentioned as *RL2/ MEE3*) showed the formation of shorter hypocotyls, defect in geotropic response and prevention of apical hook formation in dark-grown plants, and shorter hypocotyl length in red light (Hamaguchi *et al.*, 2008). The *RSM1* also acts as a flowering repressor in *Arabidopsis* by binding to *FLOWERING LOCUS C* (*FLC*) promoter and activating its transcription (Li *et al.*, 2015). Recently, It has been reported that RSM1 interacts with HY5 protein and regulates the germination and seedling development in response to abscisic acid (ABA) and salt (Yang *et al.*, 2018). Some *DIV*-homologs in *Arabidopsis* have also been reported to be involved in germination and seedling development in response to sugar, ABA and salt stress (Chen *et al.*, 2017; Fang *et al.*, 2018). It has been shown that *Arabidopsis* DRIF proteins can interact with DIV and RAD proteins of *Arabidopsis* as well as to their respective homologs in tomato (Machemer et al. 2011), showing that the interaction between DIV-DRIF-RAD like proteins is conserved in *Arabidopsis* but the function of this regulatory network in this species is yet unknown.

1.5 Evolution of *CYC-like* genes in different species as a key regulator of flower zygomorphy

Since *CYC* and *DICH* are crucial for the dorsoventral asymmetry of *Antirrhinum*, an extensive phylogenetic study has been performed to understand the role of TCP family genes with special attention to *CYC-like* genes in the evolution of flower symmetry. TCP-proteins belong to a plant-specific transcription factor family that have a conserved basic helix-loop-helix (bHLH) region known as TCP-domain, involved in plant growth and development (Cubas *et al.*, 1999a). The TCP proteins are divided into two classes according to the sequence homology of TCP-domain: TCP-Class I (TCP-P) and TCP-Class II (TCP-C) (Kosugi and Ohashi, 2002; Navaud *et al.*, 2007). The difference between the TCP-domains of Class I (TCP-P) that includes *Oryza sativa* PCF1and PCF2, and Class II (TCP-C) that includes *Zea mays* TB1 and *A. majus* CYC is a four amino acid deletion in the basic region of the Class I-TCP domain (Kosugi and Ohashi, 2002).

Class II TCP is further subdivided into two clades on the basis of differences within the TCP domain: The CIN clade (includes CINCINNATA of *A. majus*) and the CYC/TB1clade, also known as ECE clade because of the presence of glutamine acid-cysteine-glutamic acid residues (Crawford *et al.*, 2004; Howarth and Donoghue, 2006).Two major *CYC-like* gene duplication events occurred in core eudicots which lead to three *CYC-like* gene groups: *CYC1*, *CYC2* and *CYC3* (Spencer and Kim, 2018) (Fig.5)

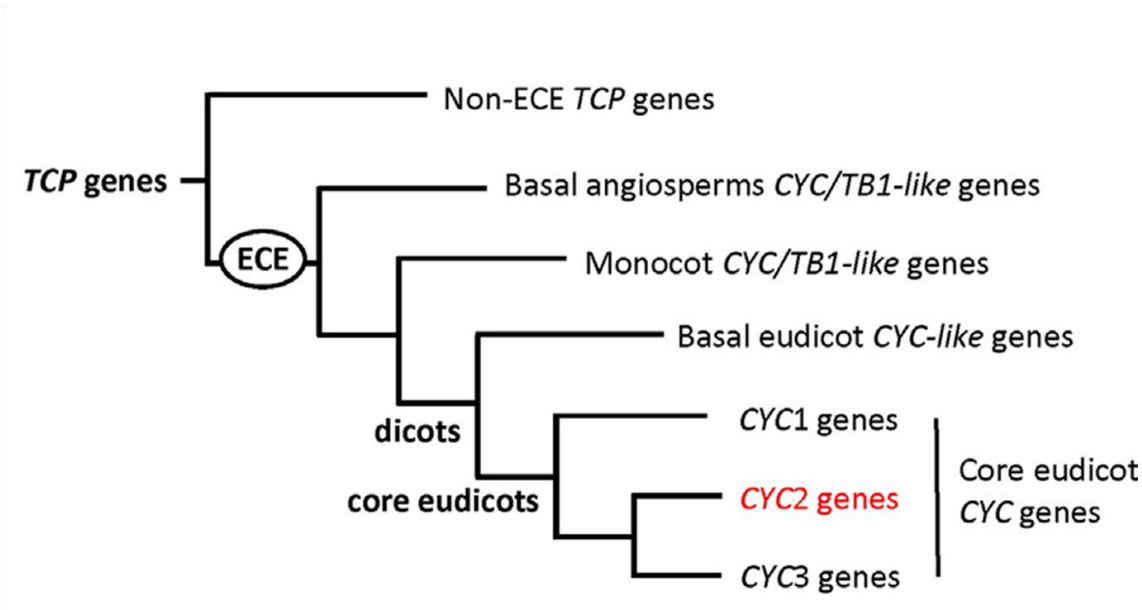


Fig.5. Gene tree of TCP-class II genes in Angiosperm. TCP-class II genes are divided into non-ECE and ECE clades. TCP proteins with ECE motifs are found in most of the angiosperms and called *CYC/TB1-like* genes. Within core eudicots, two major gene duplication events occurred, to generate *CYC1*, *CYC2*, and *CYC3-like* genes. The *CYC2-like* genes are the major regulator of flower symmetry. Adapted from (Spencer and Kim, 2018).

1.6 Independent recruitment of *CYC-like* genes in different species for floral symmetry

The role of *CYC-like* genes has been studied in different species in order to understand when these homologs have acquired their role in flower symmetry. In *Aristolochia,* a magnoliid, the expression of the *CYC-like* gene was observed after the establishment of dorsal asymmetry, which suggests that the *CYC-like* genes are unlikely to be involved in flower symmetry in *Aristolochia* (Horn *et al.*, 2015). However, the *CYC-like* genes have acquired their role in flower symmetry before the diversification of monocots and dicots as the asymmetric expression of *CYC-like* genes can be observed in some monocot species (Bartlett and Specht, 2011; Preston and Hileman, 2012). In monocots, bilaterally symmetrical flowers of *Costus*, *Heliconia* (Zingerberales), and *Commelina* (Commelinales) exhibit ventral perianth specific expression of a *CYC* homolog (Bartlett and Specht, 2011; Preston and Hileman, 2012).

In core eudicot, the *CYC2* group has been recruited independently on multiple occasions to control floral symmetry (Feng *et al.*, 2006; Wang *et al.*, 2008; Chapman *et al.*, 2012; Tahtiharju *et al.*, 2012; Garcês *et al.*, 2016). The specific role of *CYC2* genes to create zygomorphic flowers have occurred through the spatiotemporal change in *CYC2* gene expression pattern, which have been reported in at least five independent eudicot zygomorphic flower groups (Brassicales, Malpighiales, Dipsacales, Asterales, and Lamiales) (Hileman, 2014; Spencer and Kim, 2018) (Fig.3).

1.7 Effect of spatial-temporal change in the expression of *CYC2* genes in flower symmetry

Increasing evidences indicate that *CYC2-like* genes are specific for core eudicots and repeatedly recruited to control floral zygomorphy based on their strong asymmetric pattern of expression in flower meristem (Luo *et al.*, 1996; Luo *et al.*, 1999; Feng *et al.*, 2006; Busch and Zachgo, 2007; Broholm *et al.*, 2008; Preston and Hileman, 2009; Song *et al.*, 2009; Zhang *et al.*, 2010b; Howarth *et al.*, 2011). Phylogenetic analysis clearly shows that the emergence of zygomorphic flowers in Malpighiaceae is derived from the ancestral actinomorphy. In Malpighiaceae, there are two *CYC2* genes, *CYC2A* and *CYC2B*, which are originated from the duplication of common ancestor (Zhang *et al.*, 2010b). The function of *CYC2* in establishing floral symmetry in Malpighiaceae derives from expression patterns in an Old World species. In *Byrsonima crassifolia* and *Janusia guaranitica*, zygomorphic species of Malpighiaceae, the expression of *CYC2B* is absent and the expression of *CYC2A* is restricted only to the

dorsal region of flower, shows that asymmetry expression of *CYC2* genes is associated with origin and maintenance of flower zygomorphy (Zhang *et al.*, 2010b).

Leguminosae species have zygomorphic flowers with completely different shapes of petals: a single enlarge dorsal petal (standard), two lateral petals (wing), and two small ventral petals (keel) (Fig.6). In *Lotus japonicus*, *Lupinus napus* and *Pisum sativum LegCYC2* genes were identified. It has been reported that their expression is restricted to the dorsal or dorsal and lateral (*PsCYC*) petals throughout floral development (Citerne et al. 2006; Feng et al. 2006; Wang et al. 2008).In *Pisum sativum*, the two flower symmetry mutants, the *lobed standard1* (*lst1*) and *keeled wings* (*k*) are linked to the mutation in two *CYC2* genes: *PsCYC2* and *PsCYC3* (Wang et al., 2008). The study of *PsCYCs* showed the sub-fictionalization of different *CYC-like* genes, where dorsal asymmetry and internal asymmetry within dorsal petal was controlled by *PsCYC2* and *PsCYC3* (Fig.6) (Wei *et al.*, 1994), similar to two *CYC2* homologs in *Antirrhinum CYC* and *DICH* (Luo et al., 1999). In *L. japonicus,* a *CYC* homolog *LjCYC2* is expressed in the dorsal petal (Feng *et al.*, 2006). Ectopic expression of *LjCYC2* results in the dorsalization of the lateral and ventral petal, whereas reduced expression in mutant results in the ventralization of the dorsal and lateral petal (Feng *et al.*, 2006).

It has been shown that shifting spatial patterns of *CYC2* genes can also cause the conversion to radial symmetry in derived radial legumes. In *Cadia purparea*, the expression of *CYC2* was found in all petals, suggesting that the actinomorphy in this species was generated by petal dorsalization (Citerne *et al.*, 2006). This condition is similar to the *Antirrhinum backpetals* mutant where overexpression of *CYC* led to the formation of actinomorphic flowers with all five dorsal petals (Luo *et al.*, 1999).

Reversion of zygomorphy to actinomorphy is also observed in Lamiales in which the majority of species are zygomorphic. In *Plantago major*, the radial symmetry of flower is due to the loss of *CYCA* expression while *CYCB* is expressed in all petals (Preston *et al.*, 2011). Flowers of naturally occurring peloric mutant of *Linaria vulgaris* are radially symmetrical due to hyper-methylation of *LvCYC* that inhibits the expression of the *LvCYC* gene (Cubas *et al.*, 1999b).

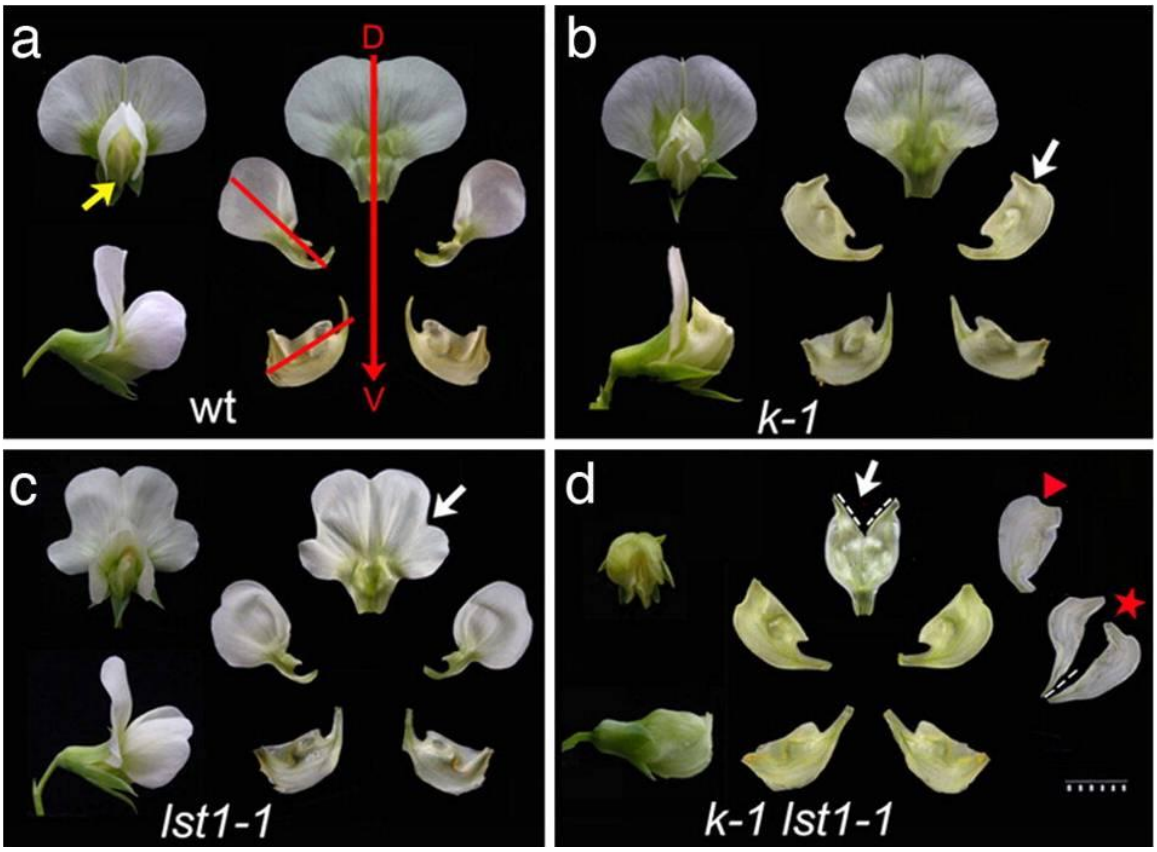


Fig.6. Mutant affecting zygomorphy in *Pisum sativum*. (a) Wild-type pea flower showing a front and side view. The yellow arrow indicates two ventral petals form a keel. Red arrow indicates the bilateral symmetry plane. The red line indicates the internal asymmetry of lateral and ventral petals (white arrow). (b) *k-1* mutant flower with ventralized lateral petals. (c) *lst1-1* mutant flower lobed dorsal petal (white arrow). (d*) k-1 lst1-1* double mutant flower. Dorsal petal is small and not fully expanded, processing a ventral petal structure (white arrow). Dashed lines represent where the petal has been cut to make it flat. Red triangle represents the dorsal petal with prominent internal asymmetry. Red star represents abnormal lateral petals occasionally found. (Wang *et al.*, 2008).

CYC orthologs also contribute to the differentiation of ray flowers (bilateral) from disc flowers (radial) across a composite (Asteraceae) inflorescence (Broholm *et al.*, 2008; Fambrini *et al.*, 2011; Chapman *et al.*, 2012; Huang *et al.*, 2016). *CYC* homologs in *Helianthus, Senecio,* and *Gerbera* are preferentially expressed around the periphery of the capitulum but at low levels, or not at all in the region of disc flower development (Broholm *et al.*, 2008; Chapman *et al.*, 2012; Tahtiharju *et al.*, 2012). The tubular-rayed mutation is due to loss-of-function mutations in a *CYC-like* gene resulting in the conversion of ray flowers to disc-like tubular flowers (Broholm et al., 2008; Fambrini et al., 2011; Chapman et al., 2012).

Like the spatial shift in the expression pattern of *CYC2* genes, the temporal shift of these genes can also change the flower symmetry. In Brassicaceae, where the occurrence of actinomorphic flowers

is more common, a *CYC* homolog *IaTCP1* also specifies dorsal identity in the zygomorphic flowers of *Iberis amara* (Busch and Zachgo, 2007). The ventral petals of *Iberis amara* are larger than the dorsal petal due to different levels of cell proliferation during flower development. At an early development stage, when petals are similar in size, there is no significant symmetric pattern of *IaTCP1* expression. However, at later stages, as petals diverge in size along the dorsoventral axis, *IaTCP1* is expressed more strongly in dorsal than in ventral petals (Busch and Zachgo, 2007). In *Arabidopsis thaliana,* the flower has radial symmetry because the dorsal expression of *AtTCP1*, a *CYC2* homolog, is transient and does not persist at later developmental stages of flower (Cubas *et al.*, 2001). It was hypothesized that the maintenance of the expression of *CYC-like* genes might have arisen through evolution by an autoregulatory loop (Yang *et al.*, 2012). The CYC-binding sites have been reported in the promoter regions of *CYC* and *RAD* homologs of *A. majus*, *Chirita heterotricha*, *Primulia heterotricha* (Costa *et al.*, 2005; Yang *et al.*, 2010; Yang *et al.*, 2012) while *AtTCP1* promoter does not contain any CYC-binding sites (Costa *et al.*, 2005).

These results suggest the presence of a positive autoregulatory feedback loop accounting for the expression maintenance of *CYC-like* genes in zygomorphic species. This feedback loop involved in the strengthening of the *CYC-like* gene expression in later developmental stages, co-opted with the dorsal-specific expression of *CYC2-like* genes, might have driven the shift of floral symmetry from actinomorphy to zygomorphy (Yang *et al.*, 2010; Yang *et al.*, 2012).

1.8. Thesis outline

In order to understand the evolution of new traits, we need to study how the gene regulatory network evolved at the molecular level. Floral zygomorphy is one of the most important morphological novelties in angiosperms, evolved from the change in gene expression pattern and interaction of TFs of a gene regulatory network. The main objective of this thesis is to understand the mechanism underlies in the specific expression and the regulatory interactions of flower symmetry genes in different species.

The establishment of an asymmetric pattern of gene expression is the key step in the generation of a zygomorphic flower. In *A. majus*, *CYC* is expressed in the dorsal domain of the floral meristem at a very early stage of development and its persistent expression is due to an autoregulation loop present in the promoter region, but what causes *CYC* to be dorsally expressed is not known yet. One of the objectives of this thesis is to gain knowledge on upstream regulators of *CYC* that may be responsible for its activation/repression in the dorsal/ventral domains of flower meristem. The other important step to bring flower zygomorphy in *A. majus* is the antagonizing effect of *RAD* over *DIV* in the dorsal floral meristem by disrupting the DIV-DRIF complex. The homologs of *DIV*, *DRIF* and *RADs* are

present in actinomorphic species also, where they may regulate different developmental processes. The other objective of this thesis is to understand the role of *DRIF* genes in *Arabidopsis*.

Chapter 2 aims to find the homologs of *A. majus CYC* in *Medicago spp.,* which is a model plant showing flower zygomorphy and ideal to investigate the role of *cis*-regulatory elements in the asymmetric expression of *CYC*. In this study, we found duplication of *CYC-like* genes in legumes in general and the presence of two *CYC* homologs in the *Medicago truncatula*. Our results show that the promoter of the *Medicago CYC2* homolog has flower specific activity.

In **Chapter 3,** we explored the expression of *Arabidopsis DIVs*, *RADs,* and *DRIFs* in *Arabidopsis thaliana* during development and in different tissues by *in silico* analysis. The aim of the study was to find the stages of plant development where all these three gene homologs are co-expressed in order to gain insight into the processes in which DIV-RAD antagonism might participate. The expression pattern of these gene homologs was analyzed in different developmental stages and anatomical parts. The results showed that the expression of some *DIVs*, *RADs,* and *DRIFs* are specific to seed, seedlings, bolting and flowering stages and give us a hint that they might be associated with several developmental stages such as germination and flowering in *Arabidopsis*.

Chapter 4 aims to characterize the role of three *Arabidopsis DRIF* homolog, *DRIF3*, *DRIF4*, and *DRIF5* during seed germination. In this study, we found that these *DRIFs* are involved in seed dormancy and germination under normal conditions, as well as under salt and heat stress. In addition, we found that these homologs may be associated with balancing the GA/ABA ratio by influencing the expression of ABA and GA biosynthesis genes.

Chapter 5 further explores the role of *DRIF3*, *DRIF4,* and *DRIF5* in seedling growth. In this chapter, we propose the functional role of these three *DRIFs* during skotomorphogenesis and photomorphogenesis in *Arabidopsis*. The results showed that these *DRIF* homologs are involved in the maintenance of the apical hook in dark-grown seedlings of *Arabidopsis*. We also found that *DRIF3*, *DRIF4* and *DRIF5* are involved in hypocotyl elongation under different light conditions.

1.9 References

Almeida, J. & Galego, L. (2005). Flower symmetry and shape in Antirrhinum. *Int J Dev Biol* 49(5-6): 527-537.

Almeida, J., Rocheta, M. & Galego, L. (1997). Genetic control of flower shape in Antirrhinum majus. *Development* 124(7): 1387-1392.

Bartlett, M. E. & Specht, C. D. (2011). Changes in expression pattern of the teosinte branched1-like genes in the Zingiberales provide a mechanism for evolutionary shifts in symmetry across the order. *Am J Bot* 98(2): 227-243.

Broholm, S. K., Tahtiharju, S., Laitinen, R. A., Albert, V. A., Teeri, T. H. & Elomaa, P. (2008). A TCP domain transcription factor controls flower type specification along the radial axis of the Gerbera (Asteraceae) inflorescence. *Proceedings of the National Academy of Sciences of the United States of America* 105(26): 9117-9122.

Busch, A. & Zachgo, S. (2007). Control of corolla monosymmetry in the Brassicaceae Iberis amara. *Proceedings of the National Academy of Sciences of the United States of America* 104(42): 16714-16719.

Chapman, M. A., Tang, S., Draeger, D., Nambeesan, S., Shaffer, H., Barb, J. G., Knapp, S. J. & Burke, J. M. (2012). Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of CYCLOIDEA genes in the Asteraceae. *PLoS Genet* 8(3): e1002628.

Chen, Y. S., Chao, Y. C., Tseng, T. W., Huang, C. K., Lo, P. C. & Lu, C. A. (2017). Two MYB-related transcription factors play opposite roles in sugar signaling in Arabidopsis. *Plant Mol Biol* 93(3): 299-311.

Citerne, H. L., Pennington, R. T. & Cronk, Q. C. (2006). An apparent reversal in floral symmetry in the legume Cadia is a homeotic transformation. *Proceedings of the National Academy of Sciences of the United States of America* 103(32): 12017-12020.

Corley, S. B., Carpenter, R., Copsey, L. & Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in Antirrhinum. *Proceedings of the National Academy of Sciences of the United States of America* 102(14): 5068-5073.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C. & Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132(22): 5093-5101.

Crawford, B. C., Nath, U., Carpenter, R. & Coen, E. S. (2004). CINCINNATA controls both cell differentiation and growth in petal lobes and leaves of Antirrhinum. *Plant Physiol* 135(1): 244-253.

Cubas, P., Coen, E. & Zapater, J. M. (2001). Ancient asymmetries in the evolution of flowers. *Curr Biol* 11(13): 1050-1052.

Cubas, P., Lauter, N., Doebley, J. & Coen, E. (1999a). The TCP domain: a motif found in proteins regulating plant growth and development. *Plant J* 18(2): 215-222.

Cubas, P., Vincent, C. & Coen, E. (1999b). An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401(6749): 157-161.

de Mendoza, A., Sebé-Pedrós, A., Šestak, M. S., Matejčić, M., Torruella, G., Domazet-Lošo, T. & Ruiz-Trillo, I. (2013). Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proceedings of the National Academy of Sciences* 110(50): E4858-E4866.

Dilcher, D. (2000). Toward a new synthesis: Major evolutionary trends in the angiosperm fossil record. *Proceedings of the National Academy of Sciences* 97(13): 7030-7036.

Doebley, J. & Lukens, L. (1998). Transcriptional Regulators and the Evolution of Plant Form. *The Plant Cell* 10(7): 1075-1082.

Endress, P. K. (1999). Symmetry in Flowers: Diversity and Evolution. *Int J Plant Sci* 160(S6): S3-s23.

Endress, P. K. (2012). The Immense Diversity of Floral Monosymmetry and Asymmetry Across Angiosperms. *Botanical Review* 78(4): 345-397.

Fambrini, M., Salvini, M. & Pugliesi, C. (2011). A transposon-mediate inactivation of a CYCLOIDEA-like gene originates polysymmetric and androgynous ray flowers in Helianthus annuus. *Genetica* 139(11-12): 1521-1529.

Fang, Q., Wang, Q., Mao, H., Xu, J., Wang, Y., Hu, H., He, S., Tu, J., Cheng, C., Tian, G., Wang, X., Liu, X., Zhang, C. & Luo, K. (2018). AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Rep* 37(11): 1499-1511.

Feng, X., Zhao, Z., Tian, Z., Xu, S., Luo, Y., Cai, Z., Wang, Y., Yang, J., Wang, Z., Weng, L., Chen, J., Zheng, L., Guo, X., Luo, J., Sato, S., Tabata, S., Ma, W., Cao, X., Hu, X., Sun, C. & Luo, D. (2006). Control of petal shape and floral zygomorphy in Lotus japonicus. *Proceedings of the National Academy of Sciences of the United States of America* 103(13): 4970-4975.

Galego, L. & Almeida, J. (2002). Role of DIVARICATA in the control of dorsoventral asymmetry in Antirrhinum flowers. *Genes Dev* 16(7): 880-891.

Garcês, H. M., Spencer, V. M. & Kim, M. (2016). Control of Floret Symmetry by RAY3, SvDIV1B, and SvRAD in the Capitulum of Senecio vulgaris. *Plant Physiol* 171(3): 2055-2068.

Gaudin, V., Lunness, P. A., Fobert, P. R., Towers, M., Riou-Khamlichi, C., Murray, J. A. H., Coen, E. & Doonan, J. H. (2000). The Expression of D-Cyclin Genes Defines Distinct Developmental Zones in Snapdragon Apical Meristems and Is Locally Regulated by the Cycloidea Gene. *Plant Physiology* 122(4): 1137-1148.

Hamaguchi, A., Yamashino, T., Koizumi, N., Kiba, T., Kojima, M., Sakakibara, H. & Mizuno, T. (2008). A small subfamily of Arabidopsis RADIALIS-LIKE SANT/MYB genes: a link to HOOKLESS1-mediated signal transduction during early morphogenesis. *Biosci Biotechnol Biochem* 72(10): 2687-2696.

Hileman, L. C. (2014). Trends in flower symmetry evolution revealed through phylogenetic and developmental genetic advances. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1648): 20130348.

Horn, S., Pabon-Mora, N., Theuss, V. S., Busch, A. & Zachgo, S. (2015). Analysis of the CYC/TB1 class of TCP transcription factors in basal angiosperms and magnoliids. *Plant J* 81(4): 559-571.

Howarth, D. G. & Donoghue, M. J. (2006). Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. *Proceedings of the National Academy of Sciences* 103(24): 9101-9106.

Howarth, D. G., Martins, T., Chimney, E. & Donoghue, M. J. (2011). Diversification of CYCLOIDEA expression in the evolution of bilateral flower symmetry in Caprifoliaceae and Lonicera (Dipsacales). *Ann Bot* 107(9): 1521-1532.

Huang, D., Li, X., Sun, M., Zhang, T., Pan, H., Cheng, T., Wang, J. & Zhang, Q. (2016). Identification and Characterization of CYC-Like Genes in Regulation of Ray Floret Development in Chrysanthemum morifolium. *Frontiers in plant science* 7: 1633.

Kosugi, S. & Ohashi, Y. (2002). DNA binding and dimerization specificity and potential targets for the TCP protein family. *Plant J* 30(3): 337-348.

Lang, D., Weiche, B., Timmerhaus, G., Richardt, S., Riano-Pachon, D. M., Correa, L. G., Reski, R., Mueller-Roeber, B. & Rensing, S. A. (2010). Genome-wide phylogenetic comparative analysis of plant transcriptional regulation: a timeline of loss, gain, expansion, and correlation with complexity. *Genome Biol Evol* 2: 488-503.

Li, C., Zhou, Y. & Fan, L.-M. (2015). A novel repressor of floral transition, MEE3, an abiotic stress regulated protein, functions as an activator of FLC by binding to its promoter in Arabidopsis. *Environmental and Experimental Botany* 113: 1-10.

Lu, C. A., Ho, T. H., Ho, S. L. & Yu, S. M. (2002). Three novel MYB proteins with one DNA binding repeat mediate sugar and hormone regulation of alpha-amylase gene expression. *Plant Cell* 14(8): 1963-1980.

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J. & Coen, E. (1999). Control of organ asymmetry in flowers of Antirrhinum. *Cell* 99(4): 367-376.

Luo, D., Carpenter, R., Vincent, C., Copsey, L. & Coen, E. (1996). Origin of floral asymmetry in Antirrhinum. *Nature* 383(6603): 794-799.

Machemer, K., Shaiman, O., Salts, Y., Shabtai, S., Sobolev, I., Belausov, E., Grotewold, E. & Barg, R. (2011). Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. *Plant J* 68(2): 337-350.

Navaud, O., Dabos, P., Carnus, E., Tremousaygue, D. & Herve, C. (2007). TCP transcription factors predate the emergence of land plants. *J Mol Evol* 65(1): 23-33.

Neal, P. R., Dafni, A. & Giurfa, M. (1998). FLORAL SYMMETRY AND ITS ROLE IN PLANT-POLLINATOR SYSTEMS: Terminology, Distribution, and Hypotheses. *Annual Review of Ecology and Systematics* 29(1): 345-373.

Preston, J. C. & Hileman, L. C. (2009). Developmental genetics of floral symmetry evolution. *Trends Plant Sci* 14(3): 147-154.

Preston, J. C. & Hileman, L. C. (2012). Parallel evolution of TCP and B-class genes in Commelinaceae flower bilateral symmetry. *EvoDevo* 3(1): 6.

Preston, J. C., Martinez, C. C. & Hileman, L. C. (2011). Gradual disintegration of the floral symmetry gene network is implicated in the evolution of a wind-pollination syndrome. *Proceedings of the National Academy of Sciences of the United States of America* 108(6): 2343-2348.

Raimundo, J. (2015). Evolution of the gene regulatory network controling flower dorsoventral asymmetry. *PhDThesis*.

Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., Coen, E. & Costa, M. M. R. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. *The Plant Journal* 75(4): 527-538.

Rose, A., Meier, I. & Wienand, U. (1999). The tomato I-box binding factor LeMYBI is a member of a novel class of myb-like proteins. *Plant J* 20(6): 641-652.

Song, C.-F., Lin, Q.-B., Liang, R.-H. & Wang, Y.-Z. (2009). Expressions of ECE-CYC2 clade genes relating to abortion of both dorsal and ventral stamens in Opithandra(Gesneriaceae). *BMC Evolutionary Biology* 9(1): 244.

Spencer, V. & Kim, M. (2018). Re"CYC"ling molecular regulators in the evolution and development of flower symmetry. *Semin Cell Dev Biol* 79: 16-26.

Tahtiharju, S., Rijpkema, A. S., Vetterli, A., Albert, V. A., Teeri, T. H. & Elomaa, P. (2012). Evolution and diversification of the CYC/TB1 gene family in Asteraceae–a comparative study in Gerbera (Mutisieae) and sunflower (Heliantheae). *Mol Biol Evol* 29(4): 1155-1166.

Ushimaru, A., Dohzono, I., Takami, Y. & Hyodo, F. (2009). Flower orientation enhances pollen transfer in bilaterally symmetrical flowers. *Oecologia* 160(4): 667-674.

Wagner, G. P. & Lynch, V. J. (2010). Evolutionary novelties. *Current Biology* 20(2): R48-R52.

Wang, Z., Luo, Y., Li, X., Wang, L., Xu, S., Yang, J., Weng, L., Sato, S., Tabata, S., Ambrose, M., Rameau, C., Feng, X., Hu, X. & Luo, D. (2008). Genetic control of floral zygomorphy in pea (Pisum sativum L.). *Proceedings of the National Academy of Sciences of the United States of America* 105(30): 10414-10419.

Wei, N., Chamovitz, D. A. & Deng, X.-W. (1994). Arabidopsis COP9 is a component of a novel signaling complex mediating light control of development. *Cell* 78(1): 117-124.

Wittkopp, P. J. & Kalay, G. (2011). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nature Reviews Genetics* 13: 59.

Yang, B., Song, Z., Li, C., Jiang, J., Zhou, Y., Wang, R., Wang, Q., Ni, C., Liang, Q., Chen, H. & Fan, L.-M. (2018). RSM1, an Arabidopsis MYB protein, interacts with HY5/HYH to modulate seed germination and seedling development in response to abscisic acid and salinity. *PLOS Genetics* 14(12): e1007839.

Yang, X., Cui, H., Yuan, Z.-L. & Wang, Y.-Z. (2010). Significance of consensus CYC-binding sites found in the promoters of both ChCYC and ChRAD genes in Chirita heterotricha (Gesneriaceae). *Journal of Systematics and Evolution* 48(4): 249-256.

Yang, X., Pang, H.-B., Liu, B.-L., Qiu, Z.-J., Gao, Q., Wei, L., Dong, Y. & Wang, Y.-Z. (2012). Evolution of Double Positive Autoregulatory Feedback Loops in CYCLOIDEA2 Clade Genes Is Associated with the Origin of Floral Zygomorphy. *The Plant Cell* 24(5): 1834-1847.

Zhang, F., Zuo, K., Zhang, J., Liu, X., Zhang, L., Sun, X. & Tang, K. (2010a). An L1 box binding protein, GbML1, interacts with GbMYB25 to control cotton fibre development. *Journal of Experimental Botany* 61(13): 3599-3613.

Zhang, W., Kramer, E. M. & Davis, C. C. (2010b). Floral symmetry genes and the origin and maintenance of zygomorphy in a plant-pollinator mutualism. *Proceedings of the National Academy of Sciences* 107(14): 6388-6393.

Zhong, J., Preston, J. C., Hileman, L. C. & Kellogg, E. A. (2017). Repeated and diverse losses of corolla bilateral symmetry in the Lamiaceae. *Ann Bot* 119(7): 1211-1223.

CHAPTER 2
"Characterization of *CYC2-like* genes in *Medicago* and their role in the control of flower zygomorphy"

Characterization of *CYC2-like* genes in *Medicago* and their role in the control of flower zygomorphy

Shweta Singh[1], Sara Laranjeira[1], Rómulo Sobral[1], Concepción Gómez-Mena[2], M. Cruz Rochina[2], José Pío Beltrán[2], M. Manuela R. Costa[1]

1. Biosystems and Integrative Sciences Institute (BioISI), Plant Functional Biology Center, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal
2. Instituto de Biología Molecular y Cellular de plantas (CSIC-UPV), Ciudad Politécnica de la Innovación, Valencia, Spain

2.1 Abstract

Flower zygomorphy is a morphological novelty of flowering plants that evolved independently several times during evolution. In *Antirrhinum majus*, the TCP family genes *CYCLOIDEA* and *DICHOTOMA* are the key players to determine the dorsal identity of the flower. Some *CYC-like* genes are also known to control flower zygomorphy in distantly related families such as Asteraceae, Leguminosae, and Orchidaceae. In legumes, *CYC-like* genes are known to control the dorsal petal identity of the flower. In this study, we show that in the *Medicago truncatula* there are two *CYC-like* genes that are expressed asymmetrically in the dorsal domain of the developing flower meristem and in the dorsal petal of the developed flower. Further, we found a putative CYC-binding site in the promoter region, at 1.1 kb upstream of the start codon of the *MtCYC2a* gene. Transgenic *Medicago* showed the flower specific activity of the *MtCYC2a* promoter. In this study, we also found that the 2.9 kb promoter of *MtCYC2a* shows the expression of GFP only in the dorsal petal of the mature flower while the 5′ truncated fragment of *MtCYC2a* promoter (1.2 kb) shows the activity in lateral and ventral petal also. This study suggests the presence of distinct regulatory sites on the promoter of *CYC-like* genes that determine the asymmetric expression of these genes. Further investigation of these *cis*-elements and transcription factors that interact with them will help to understand the persistent asymmetric pattern of *CYC-like* genes in zygomorphic species.

2.2 Introduction

Flower zygomorphy has evolved multiple times during evolution in different angiosperm lineages from actinomorphic ancestors and is recognized as a key feature of the most diverse angiosperm clades such as Fabales, Lamiales, Asterales, and Orchidales (Donoghue *et al.*, 1998; Endress, 1999; Dilcher, 2000). This shift in flower symmetry from actinomorphy to zygomorphyis supposed to have occurred due to strong selection exerted by specialized pollinators (Neal *et al.*, 1998) since zygomorphy may promote cross-pollination through increased precision in pollen placement on the insect pollinator's body (Nilsson, 1988; Lázaro and Totland, 2014).

Genes involved in the development of zygomorphic flowers were characterized in *Antirrhinum majus*. In this species, the establishment of the flower dorsal identity depends upon the early and persistent expression of *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*) genes (Luo *et al.*, 1996; Luo *et al.*, 1999). Both *CYC* and *DICH* encode proteins that belong to the plant-specific TCP transcription factor family that is generally involved in cell proliferation (Cubas *et al.*, 1999; Howarth and Donoghue, 2006). This class of proteins is characterized by a highly conserved basic helix-loop-helix (bHLH) domain at the

N-terminus called TCP domain involved in DNA-binding and protein-protein interactions (Kosugi and Ohashi, 1997; Cubas *et al.*, 1999). The members of the TCP family are further divided into two classes, TCP-class I (TCP-P) and TCP-class II (TCP-C), according to the similarity of the TCP domains (Cubas *et al.*, 1999; Navaud *et al.*, 2007). The TCP-C subfamily proteins have an additional arginine-rich domain called R-domain (Cubas *et al.*, 1999). The R-domain originated in the TCP-C family independently into two separate clades, one of which includes conserved ECE (glutamic acid-cysteine-glutamic acid) motif present between the TCP and R-domains, which include *Antirrhinum majus* CYC and *Zea mays* TB1 proteins along with *Arabidopsis thaliana* TCP1 (Howarth and Donoghue, 2006). The duplication in ECE-clade has been common in eudicot and there are three sub-class within ECE-clade, CYC1, CYC2, and CYC3 (Howarth and Donoghue, 2006). Among these three sub-classes, CYC2 contains CYC and DICH from *A. majus* (Luo *et al.*, 1996; Luo *et al.*, 1999).

The role of *CYC/TB1-like* genes in flower symmetry control has been extensively studied in eudicots. *CYC* homologs of *CYC2* sub-class are major regulators of flower symmetry, and have been recruited to establish bilateral symmetry in different families (Feng *et al.*, 2006; Wang *et al.*, 2008; Chapman *et al.*, 2012; Tahtiharju *et al.*, 2012). *CYC2* genes show the asymmetric pattern of expression, which is a key step in the generation of floral asymmetry (Luo *et al.*, 1996; Feng *et al.*, 2006; Chapman *et al.*, 2012; Tahtiharju *et al.*, 2012). Studies also showed that disruption in the promoter of the *CYC* gene by transposon insertion in *A. majus backpetals* mutant and in *Helianthus annuus,* disrupt asymmetrical expression and cause a fully dorsalized symmetric flower (Luo *et al.*, 1999; Tahtiharju *et al.*, 2012). The closest homolog of *AmCYC* in *Arabidopsis*, *AtTCP1,* is also expressed dorsally in young floral meristems but this expression is transient (Cubas *et al.*, 2001), while in another Brassica species, *Iberis amara*, expression of *IaTCP1* was observed in the dorsal petal in later stages of flower development and generates bilaterally symmetrical flowers (Busch and Zachgo, 2007). The distinct regulatory changes underlying these spatial-temporal changes in the expression of *CYC2-like* genes during evolution are still not clear.

In *A. majus*, CYC regulates the dorsal identity of a flower by directly activating the expression of *RADIALIS* (*RAD*). It has been shown that although CYC belongs to TCP-II class proteins it can bind with both TCPI binding sites (GGNCCCAC) and TCPII binding sites (GTGGNCCC) (Costa *et al.*, 2005). There are several CYC-binding sites present in the promoter and intron of *RAD* that make it a direct target of CYC (Corley *et al.*, 2005; Costa *et al.*, 2005). The CYC-binding sites were also reported in *Chirita heterotricha CYC* and *RAD* promoters and it has been thought that flower zygomorphy may have been achieved by the evolution of such autoregulatory loops for *CYC-like* genes (Yang *et al.*, 2010).

The legume family represents a group of plants with great importance to mankind (reviewed in Stagnari *et al.*, 2017). Most members of this family have zygomorphic flowers having three distinct petal types: a single dorsal petal- "standard", two lateral petals-"wings", and two lower petals-"keel" (Tucker, 2003; Roque *et al.*, 2018). In the legumes, a number of *CYC-like* genes have been identified, which have undergone repeated events of duplication (Citerne *et al.*, 2003; Ree *et al.*, 2004; Feng *et al.*, 2006; Wang *et al.*, 2008). Further, studies confirmed the role of *Lotus japonicus CYC2* in floral zygomorphy and it was also reported that the ectopic expression of *LjCYC2* alters the shape of lateral and ventral petals (Feng *et al.*, 2006).

In this study, we investigated the presence of various CYC/TCP homologs of the *Medicago truncatula* and their phylogenetic relation with other TCP family proteins. The genome of *M. truncatula* is sequenced and annotated, which provided us information about the *CYC* homologs present in the whole genome of *Medicago*. Also, we could retrieve the upstream genomic region of *MtCYC* homologs to find the *cis*-regulatory elements potentially involved in regulating the gene expression. The available efficient transformation protocol for *Medicago sativa* allowed us to study the promoter fragment and find out the endogenous regulation of *MtCYC* genes.

Here, we report two *Medicago CYC2-like* genes, *MtCYC2a* and *MtCYC2b* that are expressed dorsally throughout the development of floral meristem. Further, *in silico* analysis of the putative promoter region of *MtCYC2a* showed conserved promoter regions present in other legume *CYC-like* genes promoters eg. *Phaseolus vulgaris* and *Glycine max* and a CYC-binding site. In addition, we characterized the activity of *MtCYC2a* and *AtTCP1* promoter fragments fused to *GFP* and *GUS*, respectively, to understand the role of *cis*-regulatory elements in tissue-specific expression.

2.3 Materials and Methods

2.3.1 Plant Material and growth conditions

Medicago truncatula cv. R108, *Medicago sativa* cv. RegenSY, and *Arabidopsis thaliana* cv. Columbia (Col-0) were used in this study. Plants were grown in 16 h light/ 8 h dark photoperiod at 20 ± 2 °C. *Medicago* plants were grown in soil and sand (3:1) mixture and *Arabidopsis* plants were grown in soil and vermiculite (4:1) mixture.

2.3.2 Identification of *CYC* homologs in *Medicago truncatula*, *Glycine max*, and *Phaseolus vulgaris* and phylogenetic analysis

In order to identify *CYC* homologs in various legume species, the full-length amino acid sequence of the *Antirrhinum* CYC protein was used as a query to perform the protein BLAST search in the Phytozome database to find the TCP-family proteins from *Medicago truncatula*, *Glycine max*, and *Phaseolus vulgaris*. These proteins were aligned along with *Lotus japonicus* CYCs, *Arabidopsis thaliana* TCP1, *Iberis amara* TCP1, *Antirrhinum majus* CYC, DICH, and CIN, *Zea mays* TB1, and *Oryza sativa* PCF1 and PCF2 using Muscle. The phylogenetic tree was inferred by using the Maximum-likelihood method. The reliability of internal nodes was assessed using bootstrap with 1000 replicates. Tree inference was conducted using MEGA version7 (Kumar *et al.*, 2016). The accession numbers of all the *TCP* genes used in this phylogenetic study are mentioned in Table.2.1.

2.3.3 RNA isolation and expression analysis

Total RNA was isolated from root, hypocotyl and the shoot apical meristem along with leaf primordial of 10-days-old seedlings, and dorsal (DP) and ventral petals (VP) of flowers from *M. truncatula* plants using RNeasy Plant mini Kit (Qiagen, Germany) according to manufacturer's protocol. For quantitative Real Time-PCR (qRT-PCR) analysis, total RNA was treated with DNase I to remove gDNA and around 1 µg total RNA was used to synthesize cDNA using PrimerScriptIst strand cDNA Synthesis Kit (Takara, Japan). Primers for qRT-PCR were designed using the gene PRIMER EXPRESS software (Applied Biosystems, USA). For qRT-PCR reaction, 1 µl aliquot of reverse transcription reaction was used with 300 nM of each primer mixed with Power SYBR Green PCR master mix according to the manufacturer's manual. The reaction was performed in 96-well optical reaction plates using an ABI PRISM 7500 sequence detection system. To normalize the variance among the samples, *Secret Agent* gene (*O-linked N-acetyl glucosamine transferase*) was used as endogenous control (Serwatowska *et al.*, 2014).

RNA *in situ* hybridization was performed on 10 µm thin paraffin sections of the *M. truncatula* inflorescence as described by Gomez-Mena and Roque (2018) using dig-oxigenin-labeled probes. Around 500 bp fragments from non-overlapping regions of *MtCYC2a* and *MtCYC2b* were amplified from the cDNA obtained from the dorsal petals of *M. truncatula* flower and cloned in the pGEM T-easy vector (Promega, USA). The sense and antisense RNA probes for *MtCYC2a* and *MtCYC2b* were synthesized using SP6 and T7 RNA polymerases. The primers used in qRT-PCR and for probe synthesis are mentioned in Table.2.2.

2.3.4 Cloning and construction of pro*CYC2a*::*GFP* fusion vectors for *Medicago sativa* transformation

Genomic DNA of the *M. truncatula* was isolated from 100 mg leaf tissues using DNAzol reagent (Invitrogen) following the manufacturer's instruction. Three *MtCYC2a* promoter fragments of 2.9 kb, 1.2 kb and 0.5 kb were amplified by PCR using specific primer sets (Table.2.2) designed from three conserved regions present in the 5′ upstream genomic region of the *MtCYC2a* gene (Fig.5). Those sequences were cloned in the pENTR/TOPO8 vector. After sequencing, those fragments were cloned via Gateway LR reaction into the vector pMDC110 (Curtis and Grossniklaus, 2003) to obtain *GFP* fusions. The three final constructs *proMtCYC* (2.9 kb)::*GFP*, *proMtCYC* (1.2 kb)::*GFP*, and *proMtCYC* (0.5 kb)::*GFP* were transformed into *Agrobacterium tumifaceans* strain AGL1 by electroporation.

Genetic transformation of *M. sativa* cv. RegenSY was performed as described by Deavours and Dixon (2005). Leaf discs of *M. sativa* were incubated with the *A. tumefaciens* suspensions for 30 min and transferred to the co-cultivation medium (A1). After 2 days of co-cultivation at 28 °C in dark, explants were transferred to selection and callus induction medium (A2). After 4 to 5 weeks, green callus induced on the selection medium were transferred to embryo development medium (A3) and after 4 weeks developed embryos were transferred on embryo germination medium (A4). Regenerated shoots were transferred on plant development medium (A5) for root development and shoot elongation and transgenic lines having *proMtCYC* (2.9 kb)::*GFP*, *proMtCYC* (1.2 kb)::*GFP,* and *proMtCYC* (0.5 kb)::*GFP* were regenerated. The media used for *A. tumefaciens* culture and *M. sativa* regeneration is mentioned in Appendix.2.1 and 2.2. Leaves, internodes, and flowers of transgenic plants were observed under the fluorescence microscope to observe the *GFP* expression.

2.3.5 β-glucuronidase (GUS) staining

Transgenic *Arabidopsis* lines harboring *proTCP1* (5 kb)::*GUS* and *proTCP1* (2.1 kb)::*GUS* were used to assess the *TCP1* promoter-driven activity of *GUS* reporter in 7 and 14-day-old *Arabidopsis* plant grown in 16 h light/8 h dark conditions. The *GUS* expression was also analyzed in leaves, internodes, floral inflorescences and siliques. The seedlings and tissues of transgenic *Arabidopsis* were vacuum-infiltrated with X-gluc solution (Appendix.2.3) for 15 min and incubated at 37 °C for 24 h. After

successive washing with ethanol (25%, 50%, 75%, and 95% (v/v)), photographs were made with a stereomicroscope (Leica Microsystems, Germany).

2.4 Results

2.4.1 Phylogenetic analysis of *Medicago CYC* homologs

In order to evaluate the evolutionary and phylogenetic relationship among the TCP family genes in *M. truncatula*, a total of 46 TCP protein sequences, including 8 MtCYC, 11 GmCYC, 17 PvCYC, 4 LjCYC along with AtTCP1, AmCYC, AmDICH, AmCIN, OsPCF1, OsPCF2, and ZmTB1 were aligned using Muscle and phylogenetic tree was constructed using MEGA7 (Kumar *et al.*, 2016) (Fig.1).

According to the level of similarity of TCP domain, the TCP genes in legumes were divided into two groups, the TCP-P group clustered with PCFs and the TCP-C group clustered with CYC/TB1. Within the TCP-C group, they are further subdivided into CIN-group and CYC/TB1group. This phylogenetic study suggested that the CYC-like proteins in legumes have a highly conserved TCP domain. On the basis of their TCP domain and R domain, they are separated between TB1 and CYC cluster within the CYC/TB1 subgroup of TCP-C clade.

This study showed that seven TCP-family proteins of *M. truncatula* are grouped as TCP-C class and one as TCP-P class proteins. Out of seven MtCYC proteins of TCP-C class, three are present in CYC/TB1 sub-group while four are grouped with CIN-sub-group. The MtCYC2a, MtCYC2b and MtCYC3 are grouped with AmCYC. Out of four *L. japonicus* CYC proteins, LjCYC1, LjCYC2 and LjCYC3 are clustered together with AmCYC protein and LjCYC5 is present together with ZmTB1. The expression of *LjCYC1* and *LjCYC2* was reported in the dorsal region of flower meristem, and *LjCYC2* was functionally characterized as a dorsal symmetry gene (Feng *et al.*, 2006). In the phylogenetic study, *MtCYC2a* and *MtCYC2b* are grouped with *LjCYC1* and *LjCYC2* that belongs to legume *CYC2* clade (Feng *et al.*, 2006). The other *MtCYC3* is close to *LjCYC3* (Fig.1) that was earlier reported as a member of *LegCYC3* group and its expression was observed in different vegetative tissues also (Feng *et al.*, 2006; Howarth and Donoghue, 2006). There are four *Medicago CYC* homologs, MtCYC4, MtCYC5, MtCYC6, and MtCYC7 grouped with AmCIN while MtCYC8 is present in TCP-P group along with OsPCF1 and OsPCF2 (Fig.1).

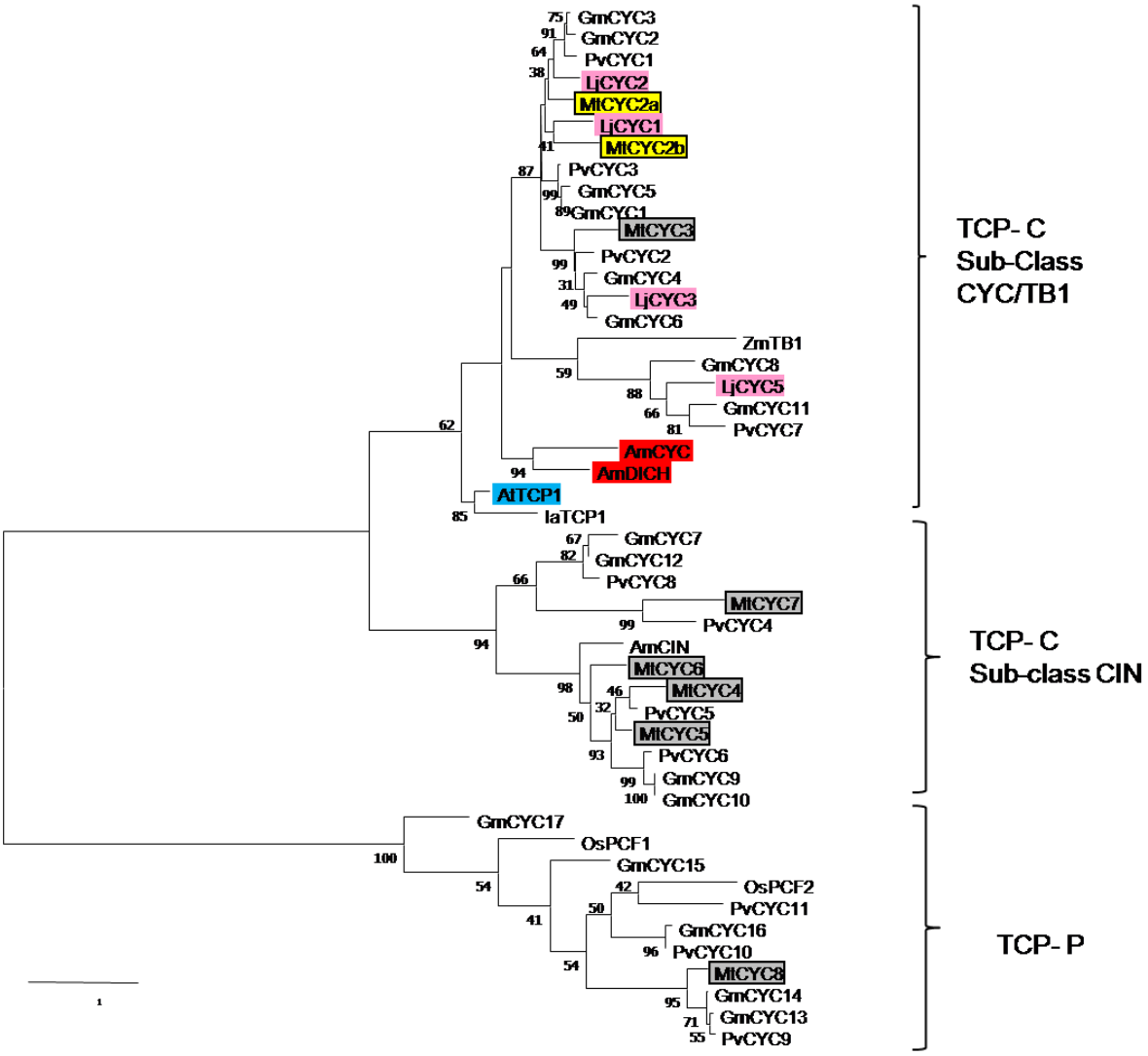


Fig.1. Phylogenetic analysis of TCP proteins of legumes along with *Zea mays* TB1, *Antirrhinum majus* CYC and *Oryza sativa* PCFs. Maximum likelihood method optimized the phylogenetic tree of the TCP gene family, based on 60 amino acids of TCP and R-domain; 1000 bootstraps were performed. Am-*Antirrhinum majus*, At-*Arabidopsis thaliana*, Gm-*Glycine max*, Lj-*Lotus japonicus*, Mt-*Medicago truncatula*, Os-*Oryza sativa*, Pv-*Phaseolus vulgaris*, Zm-*Zea mays*., AmCYC and AmDICH (red), AtTCP1 (blue), LjCYC (pink), MtCYC2a and MtCYC2b (yellow), MtCYC3 to MtCYC8 (grey) are highlighted in the tree.

We found around 80% similarity between MtCYC2a and MtCYC2b proteins, which meant that they may have arisen from a recent duplication from a common ancestral gene. For further study, we selected these two homologs of *Medicago CYC* in order to understand their role in flower symmetry.

2.4.2 Expression pattern of *MtCYC2a* and *MtCYC2b*

To understand whether *MtCYC2a* and *MtCYC2b* are involved in flower symmetry establishment in *M. truncatula*, we analyzed the expression of *MtCYC2a* and *MtCYC2b* in vegetative tissues and in

flower petals using qRT-PCR. The relative expression of *MtCYC2a* and *MtCYC2b* was almost negligible in the vegetative tissues analyzed; root, hypocotyl, and apices of seedlings (Fig.2). In contrast, *MtCYC2a* and *MtCYC2b* were highly expressed in the dorsal petals although the value was comparatively lower for *MtCYC2b*. The expression of both *MtCYC2a* and *MtCYC2b* was almost negligible in ventral petals (Fig.2).

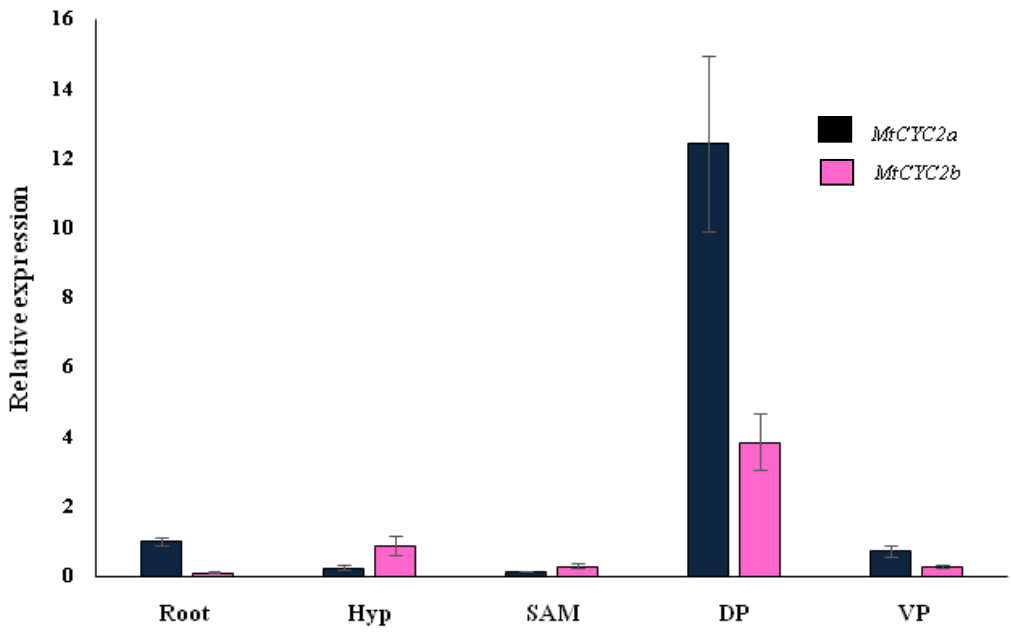


Fig.2. Expression pattern of *Medicago truncatula CYC2a* and *CYC2b* in various plant tissues. The expression level of *MtCYC2a* and *MtCYC2b* in the root, hypocotyl (HYP), shoot apical meristem along with leaf primordia (SEM), dorsal petals (DP) and ventral petals (VP) of fully developed *M. truncatula* flowers were detected by qRT-PCR. The relative expression level of the genes was normalized by using the *O-linked N-acetyl glucosamine transferase* (*Secret Agent*) gene as an internal control. Values are the mean of three technical replicates ± SD.

Further, we investigated the expression pattern of these two *MtCYC2* genes, *MtCYC2a* and *MtCYC2b* in inflorescence apices by *in situ* localization of their transcripts (Fig.3). The flower developmental stages used in the analysis were identified as described previously (Benlloch *et al.*, 2003). The development of the *Medicago truncatula* flower is divided into 8 stages. The initial common flower primordium (I1) further gives a flower primordium and a leaf primordium (I2) that arises at the ventral side of the flower. We found that *MtCYC2a* and *MtCYC2b* were expressed in only one side (dorsal) of the floral primordia away from the leaf primordia during the early developmental stage (I2) (Fig. 3.a.and b). The expression of *MtCYC2a* was higher in the early stages than *MtCYC2b* (Fig.3.a, b, c, and d).

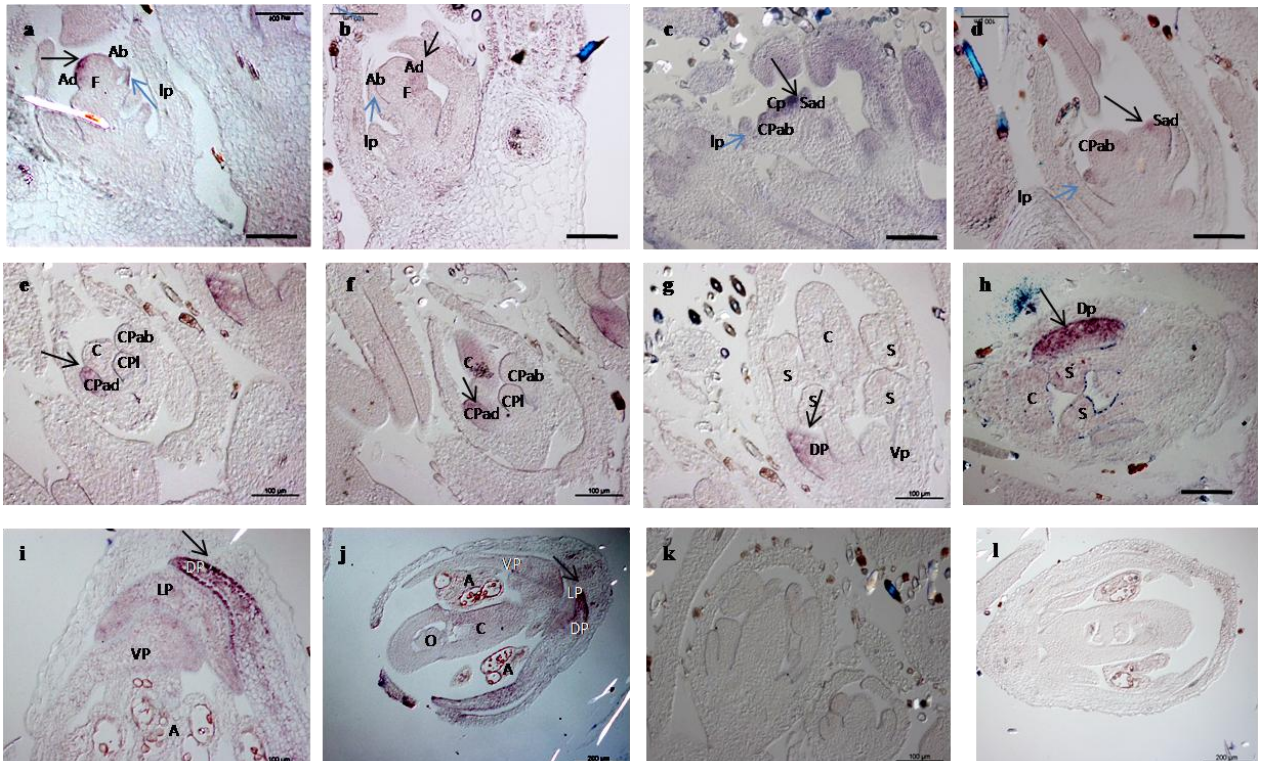


Fig.3. Expression pattern of *MtCYC2a* and *MtCY2b* genes during the development of the floral meristem of *Medicago truncatula*. *In situ* localization of *MtCYC2a* (a, c, e, g, and i) and *MtCYC2b* (b, d, f, h, and j) transcripts in developing floral meristem of *M. truncatula*. Transcripts of *MtCYC2a* were localized in the adaxial side (Ad) in the early floral meristem (a) and in later stages, they were detected in the dorsal common primordia (CPad) and in the dorsal petal (DP) (c, e, and g). The expression pattern of *MtCYC2b* was similar to *MtCYC2a* but with a lower level during the early developmental stages (b) of the flower. Black arrow indicates the transcript accumulation and blue arrow locates leaf primordial (Ip). Ab- abaxial, Ad- adaxial, C- Carpel, CP- common primodia, Cp- carpel primordial, DP- dorsal petal, F- flower meristem, LP- lateral petal, S-stamen, Sab- abaxial stamen primordial. Sense probes of *MtCYC2a* (k) and *MtCYC2* (l) was used as control Scale bar in all is 100 µm except h and l is 200 µm.

In late stages of flower development, expression of *MtCYC2a* and *MtCYC2b* was confined to the dorsal petal primordia (Fig.3.e, f, g, and h) and in the fully differentiated floral meristem, the expression of these homologs was only observed in the dorsal petal (Fig.3.i and j).

2.4.3 Analysis of *MtCYC2a* promoter elements responsible for dorsal expression in flower

The spatio-temporal expression of a gene is regulated by the *cis*-regulatory elements present in its promoter region of that gene. In order to understand the role of *cis*-regulatory elements in the asymmetric expression of *MtCYCs,* we searched the 5′upstream intergenic region of *MtCYC2a* and *MtCYC2b* along with other legume *CYC* homolog sequences available in Phytozome database. The

VISTA Webserver was used to visualize the conserved sequences present in the genomic region of *MtCYC* genes along with its homologs in *G. max*, *P. vulgaris* and also in *A. thaliana* (Fig.4 and 5). The VISTA server automatically aligned *GmCYC2*, *PvCYC1* and *AtTCP1* along with *MtCYC2a*. The previous gene present at 8 kb upstream of *MtCYC2a* is a hypothetical uncharacterized protein and we considered that 8 kb intergenic region as the putative promoter region. We found that there were three conserved regions at -0.3 kb, -2 kb and -4.5 kb of around 50 to 300 bp (Fig. 5A, B and C grey peaks in the gene plot) in the putative promoter of *MtCYC2a* genes that are similar in the promoters of *GmCYC2* and *PvCYC1*. The *CYC* homologs of *P. vulgaris* and *G. max*, which showed similarity in their putative promoter with *MtCYC2a* were *PvCYC1* and *GmCYC2* that appeared close to *MtCYC2a* and *LjCYC2* in the phylogenetic tree (Fig.1). There was also a CYC-binding site present at -1100 bp from the transcription start site of *MtCYC2a* (Fig.4). There were no similar regions present between the promoter of *AtTCP1* and Legume *CYCs*.

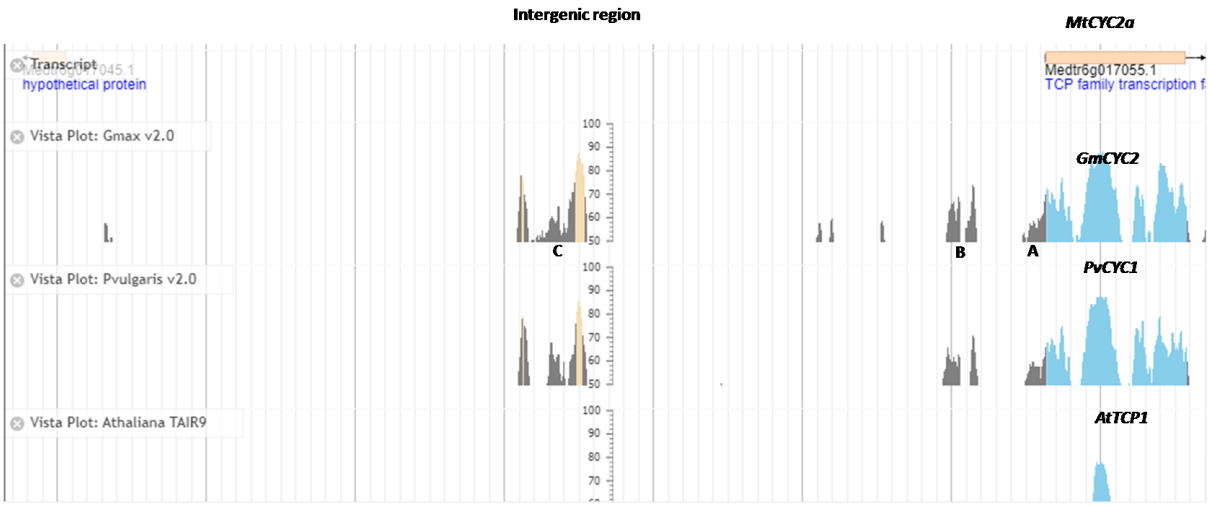


Fig.4. Alignment of *Medicago truncatula CYC2a* along with *Glycine max*, *Phaseolus vulgaris*, and *Arabidopsis thaliana*. Vista plot showing the alignment of the *MtCYC2a* gene along with *GmCYC2, PvCYC1*, and *AtTCP1.* Blue peaks represent the conserved regions present in *CYC* homologs of all legumes and *Arabidospsis* and gray peaks represent the conserved regions present in the upstream genomic region of these homologs. Legume CYC genes show more similarity with each other while *AtTCP1* shows less similarity with the three legume *CYC* homologs. *MtCYC2a* shows three conserved regions present across these legume species while there is no similarity between the *AtTCP1* promoter region and other legume *CYC* promoters.The scale that appeared vertically in the plot represents the similarity percentage between the conserved sequences.

In case of *MtCYC2b,* the length of the intergenic region was around 5 kb and there was only one common region present at around 0.7 kb upstream of start codon in *M. truncatula,* present in both *G. max*, *P. vulgaris* and there was no CYC-binding sequence in the putative promoter region of *MtCYC2b* gene (Fig.5. a). *AtTCP1* was also aligned with *MtCYCb* but the promoter region of *AtTCP1* has no similar region as *MtCYC2b.* We also checked the promoter of *AtTCP1* along with two other species of *Arabidopsis, A. halleri* and *A. lyrata*, and other Brassicaceae species *Brassica rapa*. We found that the *TCP* promoter of all the *Arabidopsis spp.*are highly conserved and also conserved in *B. rapa* (Fig.5.b). This result showed that not only the TCP domain but also the *cis*-regulatory elements of *CYC* homologs of the same family are conserved and may be associated with their functional diversity.

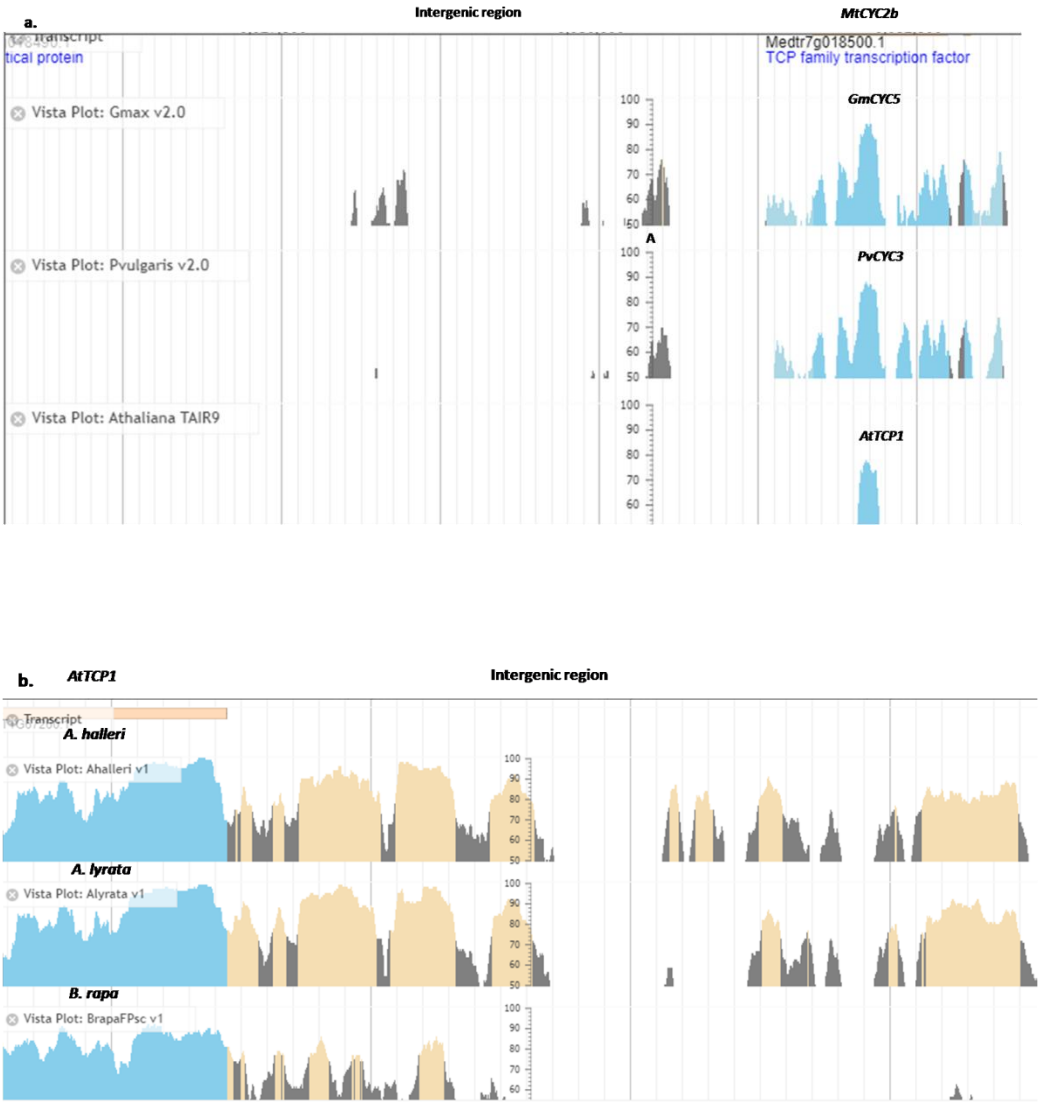


Fig. 5. Alignment of *Medicago truncatula CYC2b* and *Arabidopsis thaliana TCP1* with other legume and Brassicaceae CYC homologs. Vista plot showing the alignment of (a) *MtCYC2b* along with *GmCYC5*, *PvCYC3,* and *AtTCP1*, (b) *AtTCP1* along with *A. halleri*, *A. lyrata* and *Brassica rapa* TCP genes. Blue peaks represent the conserved gene

region, gray and brown peaks represent conserved promoter regions across the species. *MtCYC2b* is very similar to *GmCYC5* and *PvCYC3* while these legume *CYC* homologs show less similarity with *AtTCP1*. The *TCP* gene of all the four species of Brassicaceae show higher similarity with each other and their promoter regions are also very similar (b). The vertical scale appears in the plot represents the percentage similarity between the inter-species.

In order to understand the role of the conserved regions of the *MtCYC2a* promoter, we amplified three fragments of different sizes having only one (0.5 kb fragment), two (1.2 kb fragment), and all three (2.9 kb fragment) conserved regions from the gDNA of *M. truncatula* cv. R108. We used 5′ specific primers designed from the conserved region along with a 3′ primer designed just after the start codon of *MtCYC2a*. After sequencing, we found that the largest fragment containing all the conserved regions and a CYC-binding site was 2.9 kb instead of 5 kb as appeared in the Phytozome database. It could be due to the difference in the cultivars of *M. truncatula* used in this study. The cultivar used for the genome sequencing project of *M. truncatula* was Jemalong A17 (Ané *et al.*, 2008) while we used *M. truncatula* R108 in this study.

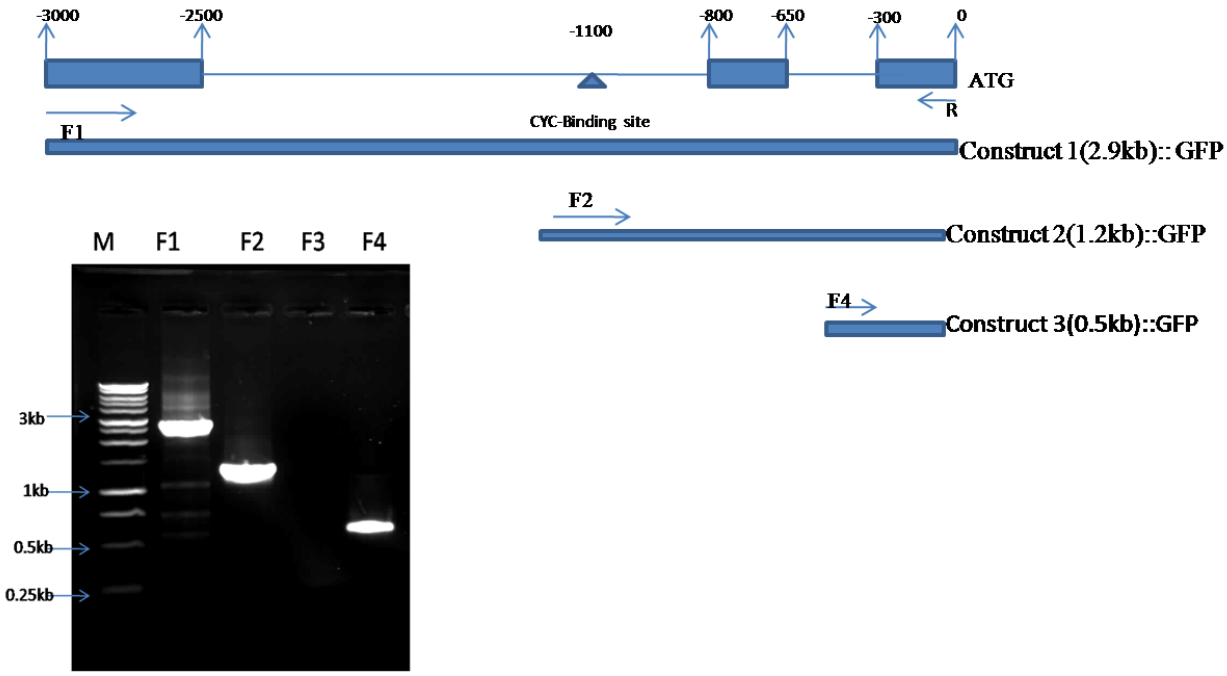


Fig.6. Amplification of *Medicago truncatula CYC2a* promoter fragments. Conserved regions present in the promoter of *MtCYC2a* were amplified by PCR. The 2.9 kb fragment contains all the conserved regions present at -0.3 kb, -1.2 kb and -2.8 kb upstream of *MtCYC2a* and a CYC-binding site present at -1.1 kb. The 1.2 kb fragment contains the conserved regions near the beginning of the transcription start site and CYC- binding site, and 0.5 kb fragment contains only one conserved region close to start codon of the *MtCYC2a* gene.

All the three fragments of the *MtCYC2a* promoter (Fig.6) were cloned upstream of the *GFP* and the constructs transformed into *M. sativa* by *Agrobacterium* mediated transformation. We obtained five transgenic *Medicago* plants harboring *proMtCYC2a* (0.5 kb)::*GFP*, four *proMtCYC2a* (1.2 kb)::GFP and five *proMtCYC2a* (2.9 kb)::*GFP* regenerated from independent callus. Transgenics with *proMtCYC2a* (0.5kb)::*GFP* did not reach the flowering stage and died during hardening, while two plants having *proMtCYC2a* (1.2 kb)::*GFP* and one plant having *proMtCYC2a* (2.9 kb)::*GFP* showed flowering.

M. sativa flower shows one big dorsal petal, two lateral petals, and two fused ventral petals (Fig.9.a). We examined three areas of each flower petals: the free end of petal, the middle part, and the part attached with the axis. *Medicago* transgenic plants harboring *proMtCYC2a* (1.2 kb)::*GFP* showed GFP signal at the margins of ventral and all parts of lateral petal along with dorsal petal (Fig. 7.b, 7.d, 7.f, and 7.j, and 9.b), anther lobe (Fig. 7.j) and pollen (Fig. 7.l). The GFP signal was not observed in the filament of stamen, style, stigma and ovary.

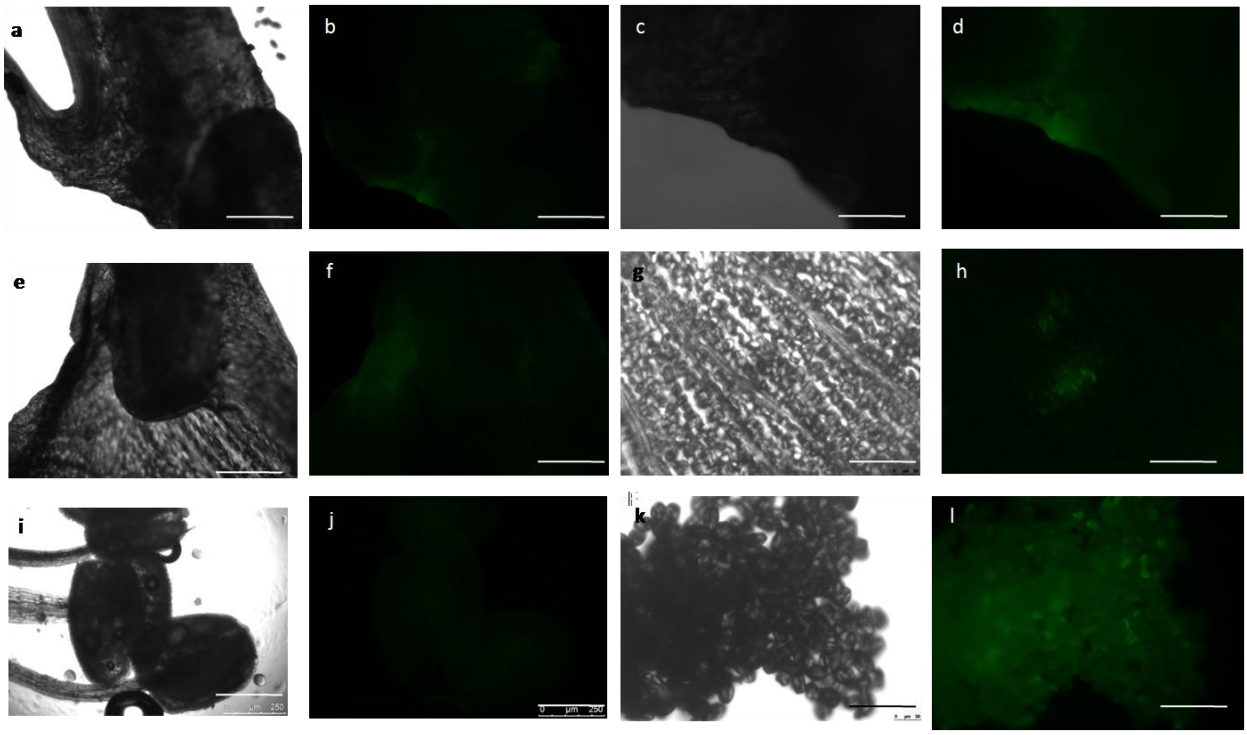


Fig.7. *MtCYC2a* promoter (1.2 kb) driven expression of *GFP* in flower parts of *Medicago sativa*. (b) Expression of GFP in ventral petal attached with the axis, (d) middle part of the ventral petal, (f) lateral petal, (h) dorsal petal, (j) stamen and (l) pollen. a, c, e, g, i, and k represent the bright field image of the corresponding fluorescent images on the right. Scale bar 250 μm.

In *Medicago* transgenic plants harboring *proMtCYC2a* (2.9 kb)::*GFP,* GFP signal was not observed in either ventral or lateral petals. A few green spots were observed at the distal end (free end) of the dorsal petal but no GFP signal was observed in the middle region and lateral sides of the dorsal petal (Fig. 8.d, 8.f, and 9.c). The GFP signaling in the anther lobes of *proMtCYC2a* (2.9 kb)::*GFP* was similar to the GFP signals observed in transgenics having *proMtCYC2a* (1.2 kb)::*GFP* (Fig.8.j and 8.l).

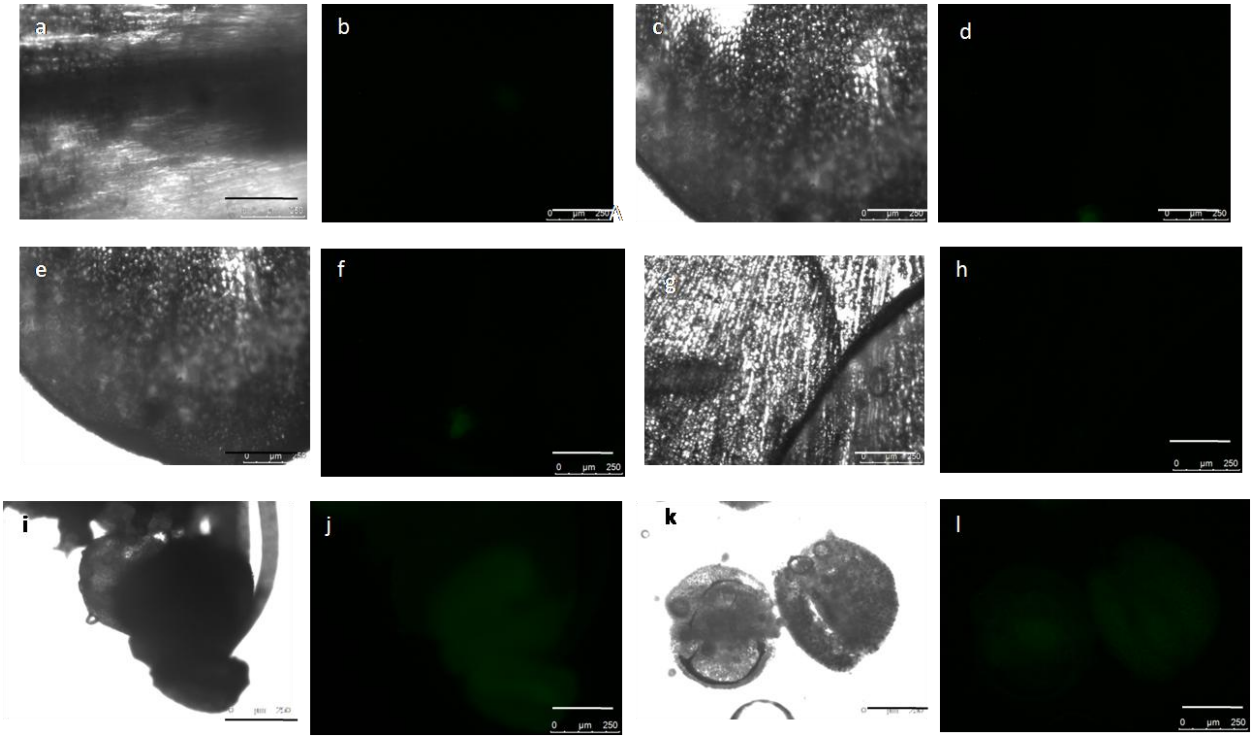


Fig.8. *MtCYC2a* promoter (2.9 kb) driven expression of *GFP* in flower parts of *Medicago sativa*. (b) Expression of *GFP* in central, (d and f) distal and (h) lateral margins of dorsal petals, (j) stamen, and (l) anther lobes. a, c, e, g, i, and k represent the bright field image of the corresponding fluorescent images on the right. Scale bar 250 µm.

Transgenic plants harboring *proMtCYC2a* (0.5 kb)::*GFP* did not flower so the expression of *GFP* in those plants floral tissues was not observed, but there was no GFP signal in vegetative tissues of those transgenic lines.The expression of *GFP* was also not detected in leaves, petioles, and internodes in transgenic plants harboring *proMtCYC2a* (1.2 kb)::*GFP*, and *proMtCYC2a* (2.9 kb)::*GFP*.

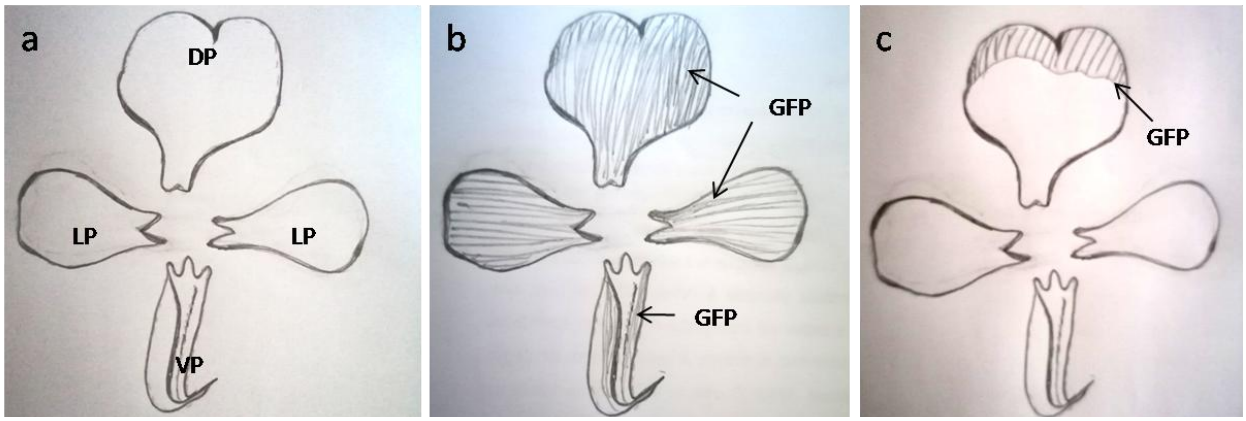


Fig.9. Diagramatic representation of *MtCYC2a* promoter activity in of *Medicago sativa* flower petals. (a) Structure of dorsal (DP), lateral (LP), and ventral petal (VP), (b) expression of *GFP* driven by *MtCYC2a* promoter fragment (1.2 kb), and (c) expression of *GFP* driven by *MtCYC2a* promoter (2.9 kb). The shaded area of petals represents the expression of *GFP*.

These results suggest that the 2.9 kb upstream region of *MtCYC2a* had flower specific promoter activity because there was no GFP expression observed in any vegetative tissue. Also the1.2 kb fragment, which has two conserved regions and a CYC-binding site, was sufficient to sustain the flower specific promoter activity of *MtCYC2a*. The distal conserved promoter region of the *MtCYC2a* promoter (2.9 kb upstream) may be responsible for the suppression of *MtCYC2a* expression in the ventral petals of the flowers.

2.4.4 Expression pattern and activities of *AtTCP1* promoter and its 5′ deletion segment in *Arabidopsis thaliana* transgenics

TCP1 is the closest homolog of *AmCYC* present in *Arabidopsis* and has been functionally correlated to the longitudinal growth of leaves and stem (Koyama *et al.*, 2010). After aligning the promoter regions of *MtCYC2a* and *TCP1*, we found that there is no conserved region between the promoters of *MtCYC2a* and *AtTCP1*. In order to check the promoter activity of the upstream region of the *TCP1* gene, two independent homozygous lines of transgenic *Arabidopsis* plants having *proTCP1* (2.1 kb)::*GUS* and *proTCP1* (5 kb)::*GUS* were used for GUS staining. We checked *GUS* expression in 7days old seedlings, 14-days-old plants, and different other organs such as internodes, flowers, and siliques. We found that plants expressing *GUS* driven by the *proTCP1* (2.1 kb) fragment showed more GUS activity than plants expressing *GUS* driven by the *proTCP1* (5 kb) fragment (Fig.10). Previously, it was shown that the expression of *GUS* controlled by the 2.98 kb fragment of the *TCP1* promoter, confined to the apical meristem of 7-days-old seedlings (Koyama *et al.*, 2010). Here, a 2.1 kb-*TCP1*

promoter fragment leads to higher *GUS* activity in all parts of 7-day-old seedlings and 14-days-old plants, except in roots (Fig.10.a and e). The expression of *GUS* was also observed in midrib and veins of true rosette leaves but absent at the border of the leaf blade (Fig.10.i). Expression of *GUS* was also observed at the internode, and junctions between pedicle and siliques, the filament of stamen and stigma, and in siliques (Fig.10.b, f, and j). The 5 kb-*TCP1* promoter fragment showed much less *GUS* activity in 7-days old seedlings and 14 days old plant (Fig.10.c and g). The expression of *GUS* was observed at internodes and petiole (10.d and k). Expression of *GUS* was lower in blades of cotyledons of 7-day old seedling and in the 14-day old plant and appeared only at the center of the rosette (Fig.10.c and g). In true leaves, the expression of *GUS* driven by 5 kb promoter was concentrated in midrib only, the expression in veins was very faint (Fig.10.k). The expression of *GUS* was present at the internodes and junctions of inflorescence and siliques (Fig.10.d, h, and i). In a previous study, expression of *GUS* did not appear in floral parts of *Arabidopsis* transgenics driven by 2.98 kb *TCP1* promoter (Koyama *et al.*, 2010). Those results confirmed the expression of *TCP1* in midrib and petioles of leaves, internodes and inter-junctions of inflorescence that explains the role of *TCP1* in longitudinal growth of stem and leaves in *Arabidopsis* (Guo *et al.*, 2010; Koyama *et al.*, 2010).

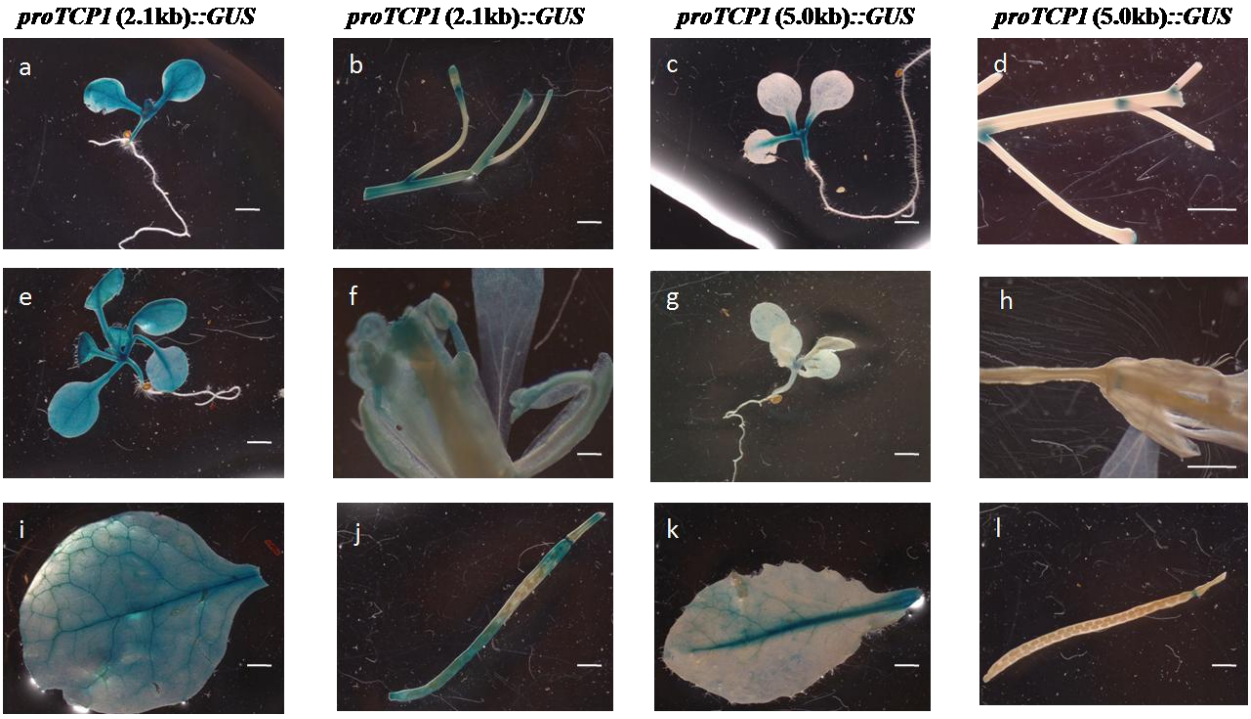


Fig.10. *AtTCP1* promoter driven expression of *GUS* in *Arabidopsis thaliana*. GUS staining was performed using 7 days old seedling (a, c), 14 days old seedling (e, g), rosette leaves (j, l), internodes (e, k) flowers (f, h), and siliques (j, l) of transgenic *A. thaliana* having 2.1 kb and 5 kb *AtTCP1* promoter with *GUS* gene. Scale bar represents 1 mm.

Our results showed that the activity of the *AtTCP1* promoter was present in vegetative tissues while the promoter activity of *MtCYC2a* was specific to floral tissues. When the 5′ extreme end of the promoter was absent in the *AtTCP1* and *MtCYC2a* promoter, the higher and non-specific expression of reporter genes was observed while the stringent expression of reporter genes was observed when they were cloned with large promoter fragments. These results suggested that there may be distinct activator and repressor sites on these promoters present near the core promoter site and its distal 5′ end.

2.5 Discussion

Zygomorphic flower symmetry in legumes is believed to have evolved separately from the Lamiales. The role of *CYC-like* genes on flower zygomorphy of legumes has been reported in *L. japonicus* and *P. sativum* (Feng *et al.*, 2006; Wang *et al.*, 2008). There were 21 *MtTCPs* reported in the *M. truncatula,* where they used *AtTCPs* and *OsTCPs* proteins as a query for blast search (Wang *et al.*, 2018). They also characterized the expression profile of these genes in different tissues and found that *MtTCP* class I genes had a relatively high level of expression in all tissues/organs (Wang *et al.*, 2018).

In our study, we obtained only eight MtCYC proteins when *A. majus* CYC protein was used as a query for blast search. The phylogenetic study showed that two *CYC-like* genes present in *M. truncatula*, *MtCYC2a,* and *MtCYC2b,* are the closest homologs of *AmCYC* and *LjCYC2*. *MtCYC2a* and *MtCYC2b* may have evolved from gene duplication as in other legume-*CYC* homologs (Citerne *et al.*, 2006), and maybe paralogs of each other with a sub-functional role in maintaining the dorsal identity of flower, like *CYC* and *DICH* in *A. majus* (Luo *et al.*, 1999). In a previous study, there were four *MtTCP* genes mentioned in CYC/TB1 clade, *MtTCP1a*, *MtTCP2a*, *MtTCP12* and *MtTCP18*. (Wang *et al.*, 2018). The expression of *MtTCP1a* and *MtTCP1b* was only observed in flower and shoot bud while *MtTCP12* is expressed in root and leaves also. A higher relative expression of *MtTCP18* was reported in vegetative tissues (Wang *et al.*, 2018). In this study, we found that the *MtCYC2a* (*MtTCP1a*) and *MtCYC2b* (*MtTCP1b*) are expressed only in the dorsal region of floral meristem and in the dorsal petal of a mature flower.

The asymmetric expression patterns of these two *MtCYCs* in dorsal floral meristem that is similar to other legume-*CYC* homologs are likely to be involved in the flower zygomorphy of *Medicago* (Citerne *et al.*, 2006; Feng *et al.*, 2006). In *M. truncatula*, both *CYC* homologs are expressed during early floral development in the adaxial side of the floral meristem. The expression of both *MtCYC2a* and *MtCYC2b* is maintained in the dorsal region of developing flower primordia. *MtCYC2a* and *MtCYC2b* are expressed only in the dorsal petal of the developed flower. This result provides a comparative framework

to examine distinct regulatory changes that occurred during evolution to maintain the asymmetrical expression of *CYC2-like* genes in zygomorphic species.

Expression modification of key developmental genes / transcription factors is a crucial step to bring phenotypic novelty in an organism (Yang *et al.*, 2015). It was reported that the radial symmetry of *Cadia* is the result of the expanded expression of *LegCYC1A* across the ventral floral meristem (Citerne *et al.*, 2006). The presence of different *cis*-regulatory regions upstream of the genes is one of the important factors that regulate the diversity in the expression pattern of these *CYC*-homologs. In the study, we found that in legume-*CYC2 like* genes, there are some conserved sequences present in the putative promoters regions of these genes that may be related to the regulation of the spatial and temporal expression of these genes. In the *MtCYC2a* promoter, there is a consensus DNA binding site GGGCCCTC, resembles that of the *A. majus* CYC binding site (GGNCCCNC) reported in previous studies (Costa *et al.*, 2005). The persistent expression of *CYC* in later stages of flower development is also important to generate a zygomorphic flower in *Antirrhinum* (Cubas, 2004). In this regard, the maintenance of *CYC* expression requires the presence of the CYC-binding site in the promoter of *CYC*, which in turn creates an autoregulatory loop (Costa *et al.*, 2005; Yang *et al.*, 2010; Yang *et al.*, 2012). In Gesneriaceae, a basal family of Lamiales, the presence of the CYC-binding site was reported in *Chirita heterotricha CYC* orthologs (Yang *et al.*, 2010).

In this study, we found the activity of the *MtCYC2a* promoter is flower specific while the activity of the *AtTCP1* promoter was observed in vegetative tissues, which suggest the difference in the *cis*-regulatory elements of two closely related *CYC* homologs in actinomorphic and zygomorphic species. Further, we report the more intense and less tissue-specific expression of reporter genes driven by smaller fragments of *MtCYC2a* and *TCP1* promoters in which 5′ upstream region was deleted while more tissue-specific and less intense expression of reporter genes driven by the larger promoter fragment of *MtCYC2a* and *TCP1.* In the case of *AtTCP1*, two independent transgenic lines showed less intense expression of *GUS* driven by 5 kb promoter compared to 2.1 kb promoter. Also, the expression of *GUS* is more precise with 5 kb *TCP1* promoter and the expression of *GUS* was restricted to the petiole, leaf midrib, and internodes. *AtTCP1* is expressed in young flower meristem, vegetative shoot, axillary meristem but not observed in old flower meristem (Cubas, 2004). In this study, we found almost negligible *GUS* expression in floral part of transgenic plants having *proTCP1* (5 kb)::*GUS* but in case of *proTCP1* (2.1 kb)::*GUS* the faint expression of *GUS* was observed in filament and in siliques. This change in the expression pattern of the transgene may be due to the length of the promoter that has different transcription factor binding sites. These results suggest the organization of functional modules

within *TCP* promoters, which regulates different aspects of the overall gene expression profile. Further investigation of the *cis*-regulatory elements present in these fragments and the corresponding TFs interacting with them, will help to understand the change in the expression pattern of *CYC* homologs in actinomorphic plants.

The expression of *GFP* driven by three different fragments of *MtCYC2a* promoter was not found in any vegetative tissues of *M. sativa* and specific for developing flowers (Fig.3). The expression of *GFP* driven by 2.9 kb *MtCYC2a* promoter was only observed in the dorsal petal of mature flower that recapitulates the expression of *MtCYC2a* in wild type (Fig.7.c). The removal of the distal end of the *MtCYC2a* promoter causes the extended-expression of *GFP* ventral petal also (Fig.7.b), which shows that the distal end of *MtCYC2a* promoter may contain the motif required for the dorsal specific expression of the gene. This 5′ truncated promoter fragment of *MtCYC2a* also contains the regulatory elements to regulate the level of expression of the gene. Further, analyzing the *cis*-regulatory elements present in these fragments will help to understand the evolution of asymmetric expression of *CYC* homologs in zygomorphic species.

The remarkable difference in the expression of *CYC* homologs in an actinomorphic species *Arabidopsis* and a zygomorphic species *Medicago is* likely to be mediated by *cis*-regulatory changes. Further investigation and functional characterization of these specific *cis*-elements and trans-regulatory factors would be important to unveil the spatial-temporal change in the expression of *CYC2-like* genes and infer the origin of floral zygomorphy during evolution.

2.6 References

Ané, J. M., Zhu, H. & Frugoli, J. (2008). Recent Advances in Medicago truncatula Genomics. *Int J Plant Genomics* 2008: 256597.

Benlloch, R., Navarro, C., Beltrán, J. & Cañas, L. A. (2003). Floral development of the model legume Medicago truncatula: ontogeny studies as a tool to better characterize homeotic mutations. *Sexual Plant Reproduction* 15(5): 231-241.

Busch, A. & Zachgo, S. (2007). Control of corolla monosymmetry in the Brassicaceae Iberis amara. *Proceedings of the National Academy of Sciences of the United States of America* 104(42): 16714-16719.

Chapman, M. A., Tang, S., Draeger, D., Nambeesan, S., Shaffer, H., Barb, J. G., Knapp, S. J. & Burke, J. M. (2012). Genetic analysis of floral symmetry in Van Gogh's sunflowers reveals independent recruitment of CYCLOIDEA genes in the Asteraceae. *PLoS Genet* 8(3): e1002628.

Citerne, H. L., Luo, D., Pennington, R. T., Coen, E. & Cronk, Q. C. (2003). A phylogenomic investigation of CYCLOIDEA-like TCP genes in the Leguminosae. *Plant Physiol* 131(3): 1042-1053.

Citerne, H. L., Pennington, R. T. & Cronk, Q. C. (2006). An apparent reversal in floral symmetry in the legume Cadia is a homeotic transformation. *Proceedings of the National Academy of Sciences of the United States of America* 103(32): 12017-12020.

Corley, S. B., Carpenter, R., Copsey, L. & Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in Antirrhinum. *Proceedings of the National Academy of Sciences of the United States of America* 102(14): 5068-5073.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C. & Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132(22): 5093-5101.

Cubas, P. (2004). Floral zygomorphy, the recurring evolution of a successful trait. *Bioessays* 26(11): 1175-1184.

Cubas, P., Coen, E. & Zapater, J. M. (2001). Ancient asymmetries in the evolution of flowers. *Curr Biol* 11(13): 1050-1052.

Cubas, P., Lauter, N., Doebley, J. & Coen, E. (1999). The TCP domain: a motif found in proteins regulating plant growth and development. *Plant J* 18(2): 215-222.

Curtis, M. D. & Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* 133(2): 462-469.

Deavours, B. E. & Dixon, R. A. (2005). Metabolic Engineering of Isoflavonoid Biosynthesis in Alfalfa. *Plant Physiology* 138(4): 2245-2259.

Dilcher, D. (2000). Toward a new synthesis: Major evolutionary trends in the angiosperm fossil record. *Proceedings of the National Academy of Sciences* 97(13): 7030-7036.

Donoghue, M. J., Ree, R. H. & Baum, D. A. (1998). Phylogeny and the evolution of flower symmetry in the Asteridae. *Trends in Plant Science* 3(8): 311-317.

Endress, P. K. (1999). Symmetry in Flowers: Diversity and Evolution. *Int J Plant Sci* 160(S6): S3-s23.

Feng, X., Zhao, Z., Tian, Z., Xu, S., Luo, Y., Cai, Z., Wang, Y., Yang, J., Wang, Z., Weng, L., Chen, J., Zheng, L., Guo, X., Luo, J., Sato, S., Tabata, S., Ma, W., Cao, X., Hu, X., Sun, C. & Luo, D. (2006). Control of petal shape and floral zygomorphy in Lotus japonicus. *Proceedings of the National Academy of Sciences of the United States of America* 103(13): 4970-4975.

Gomez-Mena, C. & Roque, E. M. (2018). Non-isotopic RNA In Situ Hybridization for Functional Analyses in Medicago truncatula. *Methods Mol Biol* 1822: 133-144.

Guo, Z., Fujioka, S., Blancaflor, E. B., Miao, S., Gou, X. & Li, J. (2010). TCP1 Modulates Brassinosteroid Biosynthesis by Regulating the Expression of the Key Biosynthetic Gene DWARF4 in Arabidopsis thaliana. *The Plant Cell* 22(4): 1161-1173.

Howarth, D. G. & Donoghue, M. J. (2006). Phylogenetic analysis of the "ECE" (CYC/TB1) clade reveals duplications predating the core eudicots. *Proceedings of the National Academy of Sciences* 103(24): 9101-9106.

Kosugi, S. & Ohashi, Y. (1997). PCF1 and PCF2 specifically bind to cis elements in the rice proliferating cell nuclear antigen gene. *The Plant Cell* 9(9): 1607-1619.

Koyama, T., Sato, F. & Ohme-Takagi, M. (2010). A role of TCP1 in the longitudinal elongation of leaves in Arabidopsis. *Biosci Biotechnol Biochem* 74(10): 2145-2147.

Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33(7): 1870-1874.

Lázaro, A. & Totland, O. (2014). The influence of floral symmetry, dependence on pollinators and pollination generalization on flower size variation. *Ann Bot* 114(1): 157-165.

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J. & Coen, E. (1999). Control of organ asymmetry in flowers of Antirrhinum. *Cell* 99(4): 367-376.

Luo, D., Carpenter, R., Vincent, C., Copsey, L. & Coen, E. (1996). Origin of floral asymmetry in Antirrhinum. *Nature* 383(6603): 794-799.

Navaud, O., Dabos, P., Carnus, E., Tremousaygue, D. & Herve, C. (2007). TCP transcription factors predate the emergence of land plants. *J Mol Evol* 65(1): 23-33.

Neal, P. R., Dafni, A. & Giurfa, M. (1998). FLORAL SYMMETRY AND ITS ROLE IN PLANT-POLLINATOR SYSTEMS: Terminology, Distribution, and Hypotheses. *Annual Review of Ecology and Systematics* 29(1): 345-373.

Nilsson, L. A. (1988). The evolution of flowers with deep corolla tubes. *Nature* 334(6178): 147-149.

Ree, R. H., Citerne, H. L., Lavin, M. & Cronk, Q. C. (2004). Heterogeneous selection on LEGCYC paralogs in relation to flower morphology and the phylogeny of Lupinus (Leguminosae). *Mol Biol Evol* 21(2): 321-331.

Roque, E., Gomez-Mena, C., Ferrandiz, C., Beltran, J. P. & Canas, L. A. (2018). Functional Genomics and Genetic Control of Flower and Fruit Development in Medicago truncatula: An Overview. *Methods Mol Biol* 1822: 273-290.

Serwatowska, J., Roque, E., Gómez-Mena, C., Constantin, G. D., Wen, J., Mysore, K. S., Lund, O. S., Johansen, E., Beltrán, J. P. & Cañas, L. A. (2014). Two euAGAMOUS Genes Control C-Function in Medicago truncatula. *PloS one* 9(8): e103770.

Stagnari, F., Maggio, A., Galieni, A. & Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: an overview. *Chemical and Biological Technologies in Agriculture* 4(1): 2.

Tahtiharju, S., Rijpkema, A. S., Vetterli, A., Albert, V. A., Teeri, T. H. & Elomaa, P. (2012). Evolution and diversification of the CYC/TB1 gene family in Asteraceae–a comparative study in Gerbera (Mutisieae) and sunflower (Heliantheae). *Mol Biol Evol* 29(4): 1155-1166.

Tucker, S. C. (2003). Floral Development in Legumes. *Plant Physiology* 131(3): 911-926.

Wang, H., Wang, H., Liu, R., Xu, Y., Lu, Z. & Zhou, C. (2018). Genome-Wide Identification of TCP Family Transcription Factors in Medicago truncatula Reveals Significant Roles of miR319-Targeted TCPs in Nodule Development. *Frontiers in plant science* 9: 774.

Wang, Z., Luo, Y., Li, X., Wang, L., Xu, S., Yang, J., Weng, L., Sato, S., Tabata, S., Ambrose, M., Rameau, C., Feng, X., Hu, X. & Luo, D. (2008). Genetic control of floral zygomorphy in pea (Pisum sativum L.). *Proceedings of the National Academy of Sciences of the United States of America* 105(30): 10414-10419.

Yang, X., Cui, H., Yuan, Z.-L. & Wang, Y.-Z. (2010). Significance of consensus CYC-binding sites found in the promoters of both ChCYC and ChRAD genes in Chirita heterotricha (Gesneriaceae). *Journal of Systematics and Evolution* 48(4): 249-256.

Yang, X., Pang, H.-B., Liu, B.-L., Qiu, Z.-J., Gao, Q., Wei, L., Dong, Y. & Wang, Y.-Z. (2012). Evolution of Double Positive Autoregulatory Feedback Loops in CYCLOIDEA2 Clade Genes Is Associated with the Origin of Floral Zygomorphy. *The Plant Cell* 24(5): 1834-1847.

Yang, X., Zhao, X. G., Li, C. Q., Liu, J., Qiu, Z. J., Dong, Y. & Wang, Y. Z. (2015). Distinct Regulatory Changes Underlying Differential Expression of TEOSINTE BRANCHED1-CYCLOIDEA-PROLIFERATING CELL FACTOR Genes Associated with Petal Variations in Zygomorphic Flowers of Petrocosmea spp. of the Family Gesneriaceae. *Plant Physiol* 169(3): 2138-2151.

2.7 Supplementary data

Table.2.1 TCP family gene codes and names

Species	Gene code	Gene name	Other name
Antirrhinum majus	Y16133	*AmCYC*	-
Antirrhinum majus	205607	*AmCIN*	-
Antirrhinum majus	AF199465	*AmDICH*	-
Arabidopsis thaliana	At1G67260	*AtTCP1*	-
Glycine max	Glyma08G256400	*GmCYC1*	-
Glycine max	Glyma19G044400	*GmCYC2*	-
Glycine max	Glyma13G047400	*GmCYC3*	-
Glycine max	Glyma20G148500	*GmCYC4*	-
Glycine max	Glyma18G280700	*GmCYC5*	-
Glycine max	Glyma10G246200	*GmCYC6*	-
Glycine max	Glyma05G142000	*GmCYC7*	-
Glycine max	Glyma17G121500	*GmCYC8*	-
Glycine max	Glyma06G284500	*GmCYC9*	-
Glycine max	Glyma12G208800	*GmCYC10*	-
Glycine max	Glyma04G152400	*GmCYC11*	-
Glycine max	Glyma15G092500	*GmCYC12*	-
Glycine max	Glyma01G045500	*GmCYC13*	-
Glycine max	Glyma02G105900	*GmCYC14*	-
Glycine max	Glyma09G284500	*GmCYC15*	-
Glycine max	Glyma07G080300	*GmCYC16*	-
Glycine max	Glyma11G196000	*GmCYC17*	-
Iberis amara	EU145779	*IaTCP1*	
Lotus japonicus	DQ202475	*LjCYC1*	
Lotus japonicus	DQ202476	*LjCYC2*	
Lotus japonicus	DQ202477	*LjCYC3*	-
Lotus japonicus	DQ202478	*LjCYC5*	-
Medicago truncatula	Medtr6G017055	*MtCYC2a*	*MtTCP1a*

Medicago truncatula	Medtr7G018500	*MtCYC2b*	*MtTCP1b*
Medicago truncatula	Medtr1G103380	*MtCYC3*	*MtTCP12*
Medicago truncatula	Medtr4G079580	*MtCYC4*	*MtTCP10a* (CIN)
Medicago truncatula	Medtr2G090960	*MtCYC5*	*MtTCP10b* (CIN)
Medicago truncatula	Medtr8G463380	*MtCYC6*	*MtTCP4* (CIN)
Medicago truncatula	Medtr3G026050	*MtCYC7*	*MtTCP5a* (CIN)
Medicago truncatula	Medtr5G039600	*MtCYC8*	*MtTCP15* (PCF)
Oryza sativa	D87260	*OsPCF1*	-
Oryza sativa	D87261	*OsPCF2*	-
Phaseolus vulgaris	Phuvl004G072100	*PvCYC1*	-
Phaseolus vulgaris	Phuvl007G060100	*PvCYC2*	-
Phaseolus vulgaris	Phuvl008G020400	*PvCYC3*	-
Phaseolus vulgaris	Phuvl003G183700	*PvCYC4*	-
Phaseolus vulgaris	Phuvl003G183700	*PvCYC5*	-
Phaseolus vulgaris	Phvul011G136115	*PvCYC6*	-
Phaseolus vulgaris	Phvul009G200600	*PvCYC7*	-
Phaseolus vulgaris	Phvul006G166600	*PvCYC8*	-
Phaseolus vulgaris	Phvul002G086600	*PvCYC9*	-
Phaseolus vulgaris	Phvul010G089100	*PvCYC10*	-
Phaseolus vulgaris	Phvul001G089100	*PvCYC11*	-
Zea mays	JQ900509	*ZmTB1*	-

Table.2.2 Primers used in this study

Primer name	Sequence (5′ -3′)	Application
qRT-MtCYC2aF	TGGTGGTGATGCTAAAAGCTT	qPCR for *MtCYC2a* (forward primer)
qRT-MtCYC2aR	CTCGGAATCTTCGCAATCTGA	qPCR for *MtCYC2a* (reverse primer)
qRT-MtCYCbF	TTCTGGCGATGACGACGAT	qPCR for *MtCYC2b* (forward primer)
qRT-MtCYCbR	CATCAGAAGACGAAGAGAAGCTA	qPCR for *MtCYC2b* (reverse primer)
qRT-SecretAgent F	TGGCTACTAGGGTTGCTGGC	qPCR for internal control (forward primer)
qRT-SecretAgent R	CCTCACCCAGTCCAGTGGAA	qPCR for internal control (reverse primer)
IS-MtCYC2a-F	ACCAGCTTGTGTTAGAGCAAG	Cloning of *MtCYC2a* fragment for *in situ*
IS-MtCYC2a-R	TAGACGTGGATTGGTGCATTCC	Cloning of *MtCYC2a* fragment for *in situ*
IS-MtCYC2b-F	AGAAGCAAGAACAACATAGCTTC	Cloning of *MtCYC2b* fragment for *in situ*
IS-MtCYC2b-R	AACATTGGAGAGTCATGATCAAAAC	Cloning of *MtCYC2b* fragment for *in situ*
Pro-MtCYC2a-F1	GTGTGATTACTAGATTTACTATATAAG GCAGG	Cloning of *MtCYC2a* promoter (2.9kb)
Pro-MtCYC2a-F2	TCCACCTTTACTTCAGGGCCCTGCCTTGG	Cloning of *MtCYC2a* promoter (1.2kb)
Pro-MtCYC2a-F3	GAGAAGCTATAAGCTATAAGCTA	Cloning of *MtCYC2a* promoter (0.5 kb)
Pro-MtCYC2a-R	TGTGAGAAGAAGAGGACCCCTTATGGTGT	Cloning of *MtCYC2a* promoter (reverse primer for all the fragments)

Appendix.2.1 *Agrobacterium* culture media

Ingredients	Per litre
Peptone	10 g
Yeast extract	10 g
Sodium chloride	5 g

Final pH-7.0. 25 mg/l Rifampcin and 50 mg/l Kanamycin was added after autoclaving.

Appendix.2.2 *Medicago sativa* culture media

1. A1 media (co-cultivation media)

Ingredients	Per litre
MS salt mix	4.30 g
Casein hydrolase	0.25 g
Sucrose	30.00 g
Gamborg's B5 vitamins (1000x stock)	1 ml
2,4-D (10 mM stock)	1 ml
BAP (10 mM stock)	0.5 ml
Phytagar	8 g

Final pH-5.8 (Adjusted with NaOH). 0.1 ml of 1 M sterile acetosyringon was added after autoclaving.

2. A2 media (selection and callus induction media)

Ingredients	Per litre
MS salt mix	4.30 g
Casein hydrolase	0.25 g
Sucrose	30.00 g
Gamborg's B5 vitamins (1000 x stock)	1 ml
2,4-D (10 mM stock)	1 ml

BAP (10 mM stock)	0.5 ml
Phytagar	8 g

Final pH-5.8 (Adjusted with NaOH). After autoclaving, remaining sterilized antibiotics solutions 250 mg/l Ticarcillin and 15 mg/l Hygromycin were added.

3. A3 media (embryo development media)

Ingredients	**Per litre**
MS salt mix	4.30 g
Casein hydrolase	0.25 g
Sucrose	30.00 g
Gamborg's B5 vitamins (1000 x stock)	1 ml
2,4-D (10 mM stock)	1 ml
Phytagar	7.5 g

Final pH-5.8 (Adjusted with NaOH). After autoclaving, sterilized solutions of 30 mM L-Proline, 250 mg/ml Ticarcillin, and 15 mg/ml Hygromycin were added.

4. A4 media (embryo germination media)

Ingredients	**Per litre**
MS salt mix	4.30 g
Casein hydrolase	0.25 g
Sucrose	30.00 g
Gamborg's B5 vitamins (1000 x stock)	1 ml
Phytagar	7.5 g

Final pH-5.8 (Adjusted with NaOH). After autoclaving, sterilized solutions of 30 mM L-Proline, 250 mg/ml Ticarcillin, and 15 mg/ml Hygromycin were added.

5. A5 media (plant development media)

Ingredients	Per litre
MS salt mix	4.30 g
Casein hydrolase	0.25 g
Sucrose	30.00 g
Gamborg's B5 vitamins (1000 x stock)	1 ml
Phytagar	7.5 g

Final pH-5.8 (Adjusted with NaOH). After autoclaving, 250 mg/ml Ticarcillin and 15 mg/ml Hygromycin were added.

Appendix.2.3.GUS-staining solution

Ingredients	Final concentration
EDTA	1 mM (pH-8)
Potassium ferricyanide	5 mM
Potassium ferrocyanide	5 mM
Sodium phosphate	100 mM (pH-7)
Triton-X-100	1 per cent (v/v)
X-Gluc	20 mM*

*The X-Gluc was pre-dissolved in a drop of *N*, *N* dimethyl formamide immediately prior to use.

CHAPTER 3
"Co-expression and regulation of *DIV-and-RAD-INTERACTING-FACTORS* (*DRIFs*) in *Arabidopsis* - a Bioinformatics approach"

Co-expression and regulation of *DIV-and-RAD-INTERACTING-FACTORS* (*DRIFs*) in *Arabidopsis*- a Bioinformatics approach

Shweta Singh, Rómulo Sobral, Sara Larangeira, João Raimundo, M. Manuela R. Costa

Biosystems and Integrative Sciences Institute (BioISI), Plant Functional Biology Center, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal

3.1 Abstract

The MYB gene family represents one of the largest transcription factor families in plants that are involved in the various developmental processes and environmental responses. In *Antirrhinum majus,* three *MYB* genes, *DIVARICATA* (*DIV*), *RADIALIS* (*RAD*), and *DIV-and-RAD-interacting-factor* (*DRIF*) are involved in the establishment of flower zygomorphy. There are several *DIV*, *RAD* and *DRIF* homologs present in *Arabidopsis thaliana*, a model plant used for molecular studies. In this study, we analyzed the expression of 12 *DIVs*, 6 *RADs* (*RL*), and 5 *DRIFs* present in *Arabidopsis* using publically available transcriptome data. This analysis showed the diversity in the expression profile of the *DIV*, *DRIF*, and *RAD* (DDR) homologs in *Arabidopsis*. We also found that out of five *DRIFs*, *DRIF3*, *DRIF4*, and *DRIF5* showed a more similar expression profile. Using the bioinformatics approach, we analyzed the co-expression gene network and predicted *cis*-regulatory elements in the promoter region of all five *DRIF* homologs. The analysis showed the co-expression of genes involved in hormonal responses and light responses and also the presence of light-responsive and hormone-responsive elements in the promoter regions. This study provides information about the DDR genes in *Arabidopsis* and paves a way for further investigation into the function of these genes.

3.2 Introduction

Transcription factors play an important role in plant growth and development by regulating the transcription of target genes. They bind to the conserved promoter elements in a precise way on the basis of both protein-DNA and protein-protein interactions. This type of combinational regulation of gene expression at the transcription level occurs in higher eukaryotes and regulates various developmental processes and morphological characters. An example of such a combinational gene regulatory network is the one that regulates flower zygomorphy in *Antirrhinum majus*. Three MYB family transcription factors; *RADIALIS* (*RAD*), *DIVARICATA* (*DIV*), and *DIV-and-RAD-interacting-factor* (*DRIF*) are involved in the control of asymmetry between the dorsal and ventral region of flower along with two TCP family transcription factors; *CYCLOIDEA* (*CYC*) and *DICHOTOMA* (*DICH*) (Luo *et al.*, 1996; Almeida *et al.*, 1997; Luo *et al.*, 1999; Galego and Almeida, 2002; Corley *et al.*, 2005; Raimundo *et al.*, 2013).

The dorsal identity of the flower of *Antirrhinum* is determined by *CYC* and *DICH* (Luo *et al.*, 1996; Luo *et al.*, 1999). Both *CYC* and *DICH* are expressed in the dorsal domain of the flower meristem in very early stages of flower development and persist in the 2nd and 3rd whorls of the flower during later stages of development (Luo *et al.*, 1996; Luo *et al.*, 1999). The expression of *RAD* is similar to the *CYC* but *RAD* expression is also extended to the dorsal part of lateral petal primordial (Corley *et al.*, 2005). CYC regulates the expression of *RAD* in the dorsal region of the flower by directly binding to its promoter region (Costa *et al.*, 2005). On the other hand, *DIV* determines the ventral identity of the *Antirrhinum* flower but unlike the dorsal identity genes, expression of *DIV* is observed in both the dorsal and ventral region of the flower meristem (Almeida *et al.*, 1997; Galego and Almeida, 2002).

Raimundo *et al.* (2013) proposed that RAD antagonizes the function of DIV in the dorsal region of *A. majus* flower by competing for a common interactor DRIF. This is due to the ability of RAD to sequester the DRIF proteins in the cytoplasm, preventing their interaction with DIV in the nucleus, and consequently preventing the DIV protein from controlling the transcription of genes regulated by the DIV-DRIF complex. The DIV-DRIF-RAD (DDR) regulatory network is conserved in other species, also. For instance, in *Solanum lycopersicum* (tomato), a *RAD* homolog (*FSM1*) negatively regulates cell expansion in the tomato fruit pericarp by competing with a DIV homolog (MYB1) for the interaction with a DRIF homolog (FSB1). In the absence of FSM1, the MYB1 and FSB1 establish a regulatory complex that enhances cell expansion of the tomato fruit pericarp (Rose *et al.*, 1999; Machemer *et al.*, 2011).

Arabidopsis thaliana, used as a model plant for molecular studies and in functional characterization of genes in flowering plants. There are several *DIV*, *DRIF*, and *RAD* homologs present in *A. thaliana*, which indicates that these homologs may adopt different functions in this species. Some *A.*

thaliana RAD and *DIV* homologs are involved in seed germination, photomorphogenesis and abiotic stress tolerance (Kwon *et al.*, 2013; Chen *et al.*, 2017; Fang *et al.*, 2018; Yang *et al.*, 2018). DRIF proteins of *A. thaliana* were shown to interact with RADs and DIVs in yeast hybrid assays (Machemer *et al.*, 2011), but the function of *DRIF* genes and their role in association with *RADs* and *DIVs* is not known yet.

The availability of gene expression and co-expression data in *Arabidopsis* provides us an opportunity to analyze the *in silico* expression profile of the *RAD*, *DIV* and *DRIFs* homologs during different developmental stages and to infer their possible functions. The *in silico* analysis performed in this study suggested that the *DRIFs* are co-expressed along with *DIV* and *RAD* during different developmental stages and organs. We also found changes in the expression level of some *DIV*, *RAD*, and *DRIF* homologs in mutants associated with photomorphogenesis such as *constitutively photomorphogenic1* (*cop1*), *phytochrome interacting factors* (*pifs- pif1* and *pif1,3,4,5-pifq*), *ethylene insensitive 3 ethylene insensitive 3-like 1* (*ein3eil1*); and with flowering such as *leafy* (*lfy*), constans (*co*) and *flowering locus-t* (*ft*). The co-expression network of genes associated with *A. thaliana DRIF* homologs and their promoter analysis suggests the functional diversity of the *DRIF* homologs in this species.

3.3 Material and Methods

3.3.1 *DIV*, *RAD,* and *DRIF* homologs in *Arabidopsis*

In this study we used twelve *DIV*, six *RAD* and five *DRIF* homologs of *Arabidopsis* mentioned in a previous study (Raimundo *et al.*, 2018). The gene accession number for each homolog and the synonym used in this study are mentioned in table.3.1.

3.3.2 *In silico* expression analysis of *DIVs*, *RADs* and *DRIFs*

The expression pattern of *A. thaliana DIVs*, *RAD* and *DRIFs* in different developmental stages and anatomical parts were obtained from microarray and RNASeq datasets using online tools (http://bar.utoronto.caand http://geneinvestigator.com/). Fluorescent pictograph of gene expression data in different tissues was obtained from the Arabidopsis eFP Browser (Winter *et al.*, 2007). Expression of all *DIVs*, *RADs* and *DRIFs* was also checked in *pif1*, *pifq*, *cop1*, *ein1eil3*, *lfy*, *co*, and *ft* mutant background in microarray data and compared with wild type (Columbia ecotype).

3.3.3 Analysis of co-expression network and protein-protein interactions

Gene co-expression network and interacting partners of the *A. thaliana DRIF* homologs were analyzed using web-based tools (Geneinvestivator, GeneMANIA, STRING, and BioGRID) (Zimmermann *et al.*, 2004; Warde-Farley *et al.*, 2010; Oughtred *et al.*, 2018; Szklarczyk *et al.*, 2019).

3.3.4 *In silico* analysis of the putative promoter of *Arabidopsis DRIFs* homologs

To analyze the promoter region of *DRIFs*, the 5′ upstream intergenic region of each *Arabidopsis DRIF* homolog was obtained from the TAIR database. We used those sequences for *in silico* prediction of *cis*-regulatory elements (CREs) using PlantCARE and PlantPAN (Lescot *et al.*, 2002; Chow *et al.*, 2016).

3.4 Results

3.4.1 Expression pattern of *RAD*, *DIV* and *DRIF* homologs in *Arabidopsis*

There are six *RAD*, twelve *DIV*, and five *DRIF* homologs reported in *Arabidopsis*. In order to investigate possible functions, in which these genes may be involved, we examined the *in silico* expression of these homologs during different developmental stages of *Arabidopsis*. There are only three *RAD* homologs present in the *Arabidopsis* microarray databases ((http://bar.utoronto.caand http://geneinvestigator.com/), *RL2*, *RL4* and *RL5*. The expression of *RL2* was observed in young flower buds and in the carpel, while the expression of *RL5* and *RL6* was observed in cauline leaves and sepals. The expression of *RL5* was also observed in sepals and the expression of *RL6* was observed during seed development (Fig.1).

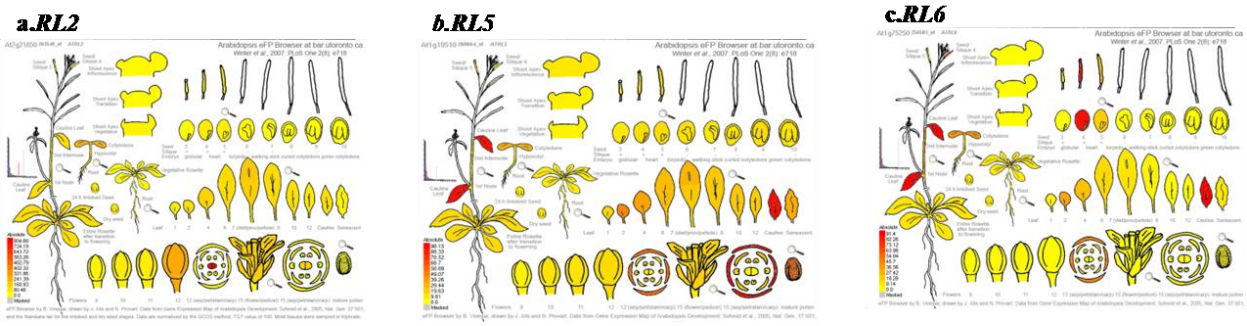


Fig.1. Fluorescent pictographs of *Arabidopsis thaliana RAD* homologs showing the expression during different developmental stages. Expression pattern of (a). *RL2*, (b). *RL5,* and (c). *RL6.* The pictograph is the representation of data obtained from the Gene Expression Map of *Arabidopsis* development (Schmid *et al.*, 2005).

Expression profiles of ten *DIV* homologs (*DIV1*, *DIV2*, *DIV4*, *DIV5*, *DIV6*, *DIV7*, *DIV8*, *DIV9*, *DIV10*, and *DIV12*) were available in the microarray database. These *DIV* homologs were expressed differently depending on the developmental stage. The expression profile of the *DIVs* showed that some homologs were expressed during all the developmental stages; *DIV6*, *DIV7,* and *DIV10* while the expression of *DIV9* and *DIV12* was only observed in dry seeds and during seed development, respectively (Fig.2). Also, we observed that the expression of *DIV1*, *DIV4,* and *DIV5* was high during the vegetative stages while the expression of *DIV8* and *DIV12* was higher during seed development (Fig.2).

The expression of *DIV1* was appeared low in the dry seeds but peak after 24 h of seed imbibitions (Winter *et al.*, 2007). Further, expression of *DIV1* was higher in vegetative tissues but very low in floral organs and during seed development stages (Fig.2.a). Significant expression of *DIV2* was observed in the developing flower, carpel and early developmental stages of seeds (Fig.2.b). The expression of *DIV4* appeared in seedlings and plant roots (Fig.2.c). Expression of *DIV5* was observed in seedlings and leaves (Fig.2.d) while expression of *DIV6* was observed in all developmental stages (Fig.2.e) and expression of *DIV7* was also observed in all stages, except in dry and imbibed seeds (Fig.2.f). The expression of *DIV8* was appeared high in the fresh seed (Fig.2.g) while the expression of *DIV9* was high in dry seeds and pollen (Fig.2.h). The expression of *DIV10* was higher in the young flower, pollen and in the mature embryo (Fig.2.i) and *DIV12* was almost negligible in all stages. It was increased during the early stage of seed development and again decreased in seeds (Fig.2.j).

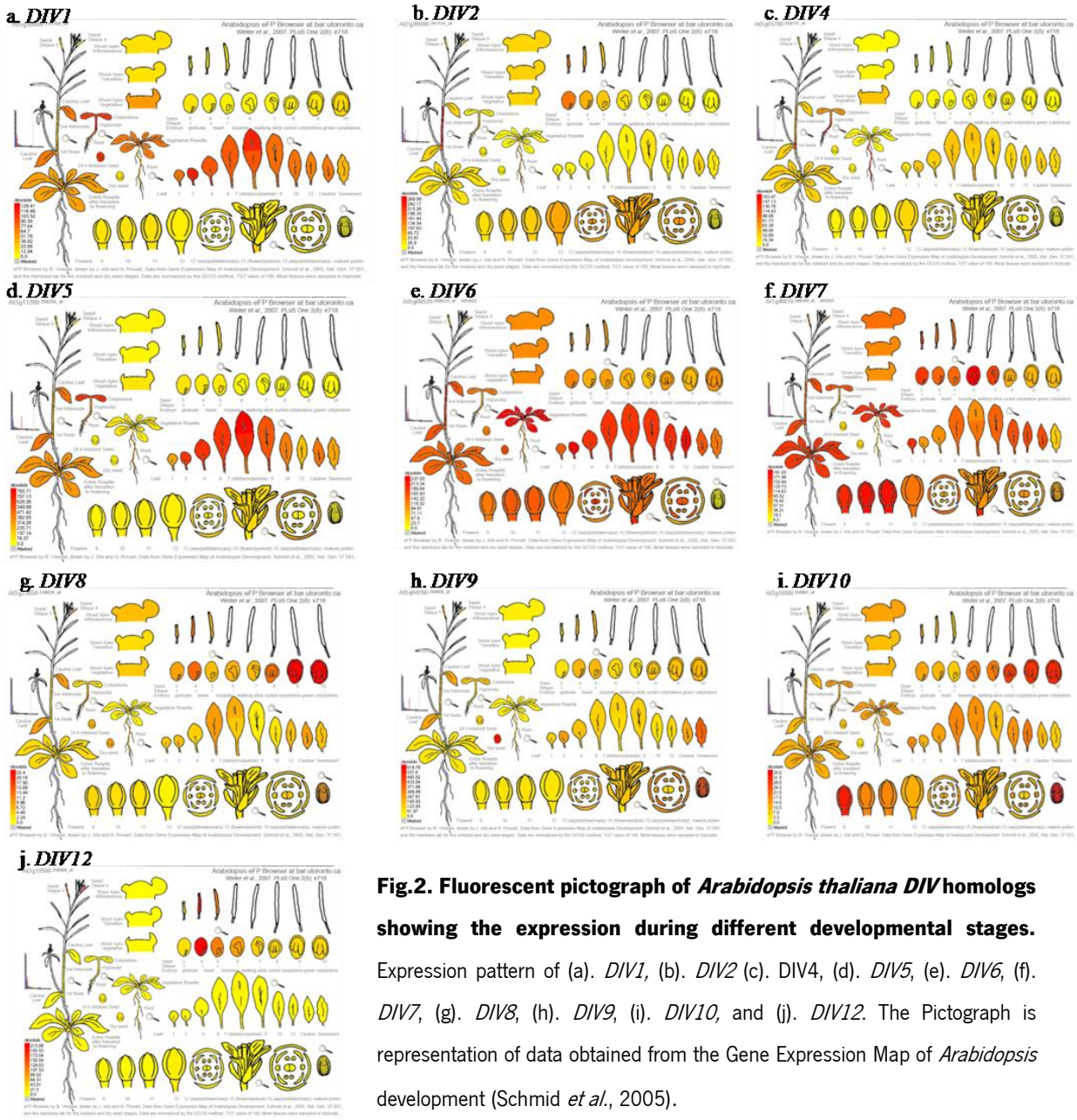


Fig.2. Fluorescent pictograph of *Arabidopsis thaliana DIV* homologs showing the expression during different developmental stages. Expression pattern of (a). *DIV1,* (b). *DIV2* (c). DIV4, (d). *DIV5*, (e). *DIV6*, (f). *DIV7*, (g). *DIV8*, (h). *DIV9*, (i). *DIV10,* and (j). *DIV12*. The Pictograph is representation of data obtained from the Gene Expression Map of *Arabidopsis* development (Schmid *et al.*, 2005).

The expression pattern of three *DRIF* homologs, *DRIF1*, *DRIF4,* and *DRIF5,* were obtained from microarray data. The expression of *DRIF1* was more specific to seeds (Fig.3.a) while the expression of *DRIF4* and *DRIF5* was observed during all the developmental stages (Fig.3.b and c). The expression of *DRIF4* was observed in all stages and organs except pollen. The expression of *DRIF4* was high in the dry seed and reduced in imbibed seed (Fig.3.b). The expression of *DRIF5* was observed in all stages and it was also expressed in mature pollen (Fig.3.c). There was no change in the expression level of *DRIF5* in the dry and imbibed seed.

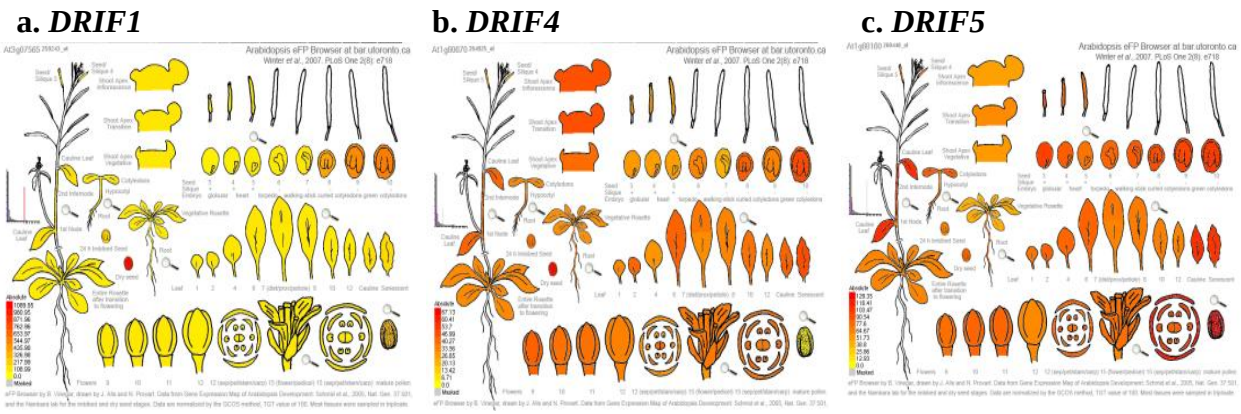


Fig.3. Fluorescent pictograph of *Arabidopsis DRIF* homologs showing expression patterns of genes during different developmental stages. Expression pattern of (a). *DRIF1*, (b). *DRIF4*, and (c). *DRIF5*. The Pictograph is a representation of data obtained from the Gene Expression Map of *Arabidopsis* development (Schmid *et al.*, 2005).

The expression of *DIV*, *RAD* and *DRIF* homologs were also investigated in RNASeq data in 9 developmental stages : germinated seed, seedling, young rosette, developed rosette, bolting, young flower, developed flower, flower and siliques, and mature siliques (Fig.4). We found that the expression of *RL1* was present in seedling and young rosette. The expression of *RL1* was lower in the developed rosette, flowers, and siliques (Fig.4). The expression of *RL3* was present in all the stages except in mature siliques, while the expression of *RL4* starts increasing in the developed rosette and a maximum level of *RL4* was observed in mature flower (Fig.4). The expression profile of *RL2*, *RL4,* and *RL6*, obtained from the RNAseq dataset, was similar to their expression profile found in microarray dataset. Expression of *DIV3* was observed during all developmental stages while expression of *DIV11* was negligible (Fig.4). The expression of *DIV11* was observed only in flowers and siliques. The expression profile of all the *DIVs* with available microarray data was similar to the RNAseq data (Fig.4).

Expression of *DRIF2* and *DRIF3* was observed in RNAseq data obtained from the germinated seed, seedling, young rosette, developed rosette, bolting, young flower, developed flower, flower and siliques, and mature siliques. The expression level of *DRIF2* was lower in comparison to the other four *DRIFs* and it was almost absent in mature siliques. The expression of *DRIF3* was observed in all the stages and the level of expression was almost similar to *DRIF4* and *DRIF5*. The expression level of *DRIF3* was higher in developed flower compared to other stages. RNAseq data did not include the dry seed stage, in which the expression of *DRIF1* appeared the maximum in microarray data (Fig.3.a). The results showed a divergent expression pattern in *DIV*, *DRIF,* and *RAD* (DDR) homologs in *Arabidopsis*. The more specific expressions of some *RL*, *DIV*, and *DRIFs* during different developmental stages may be related to their functional role at that stage of development.

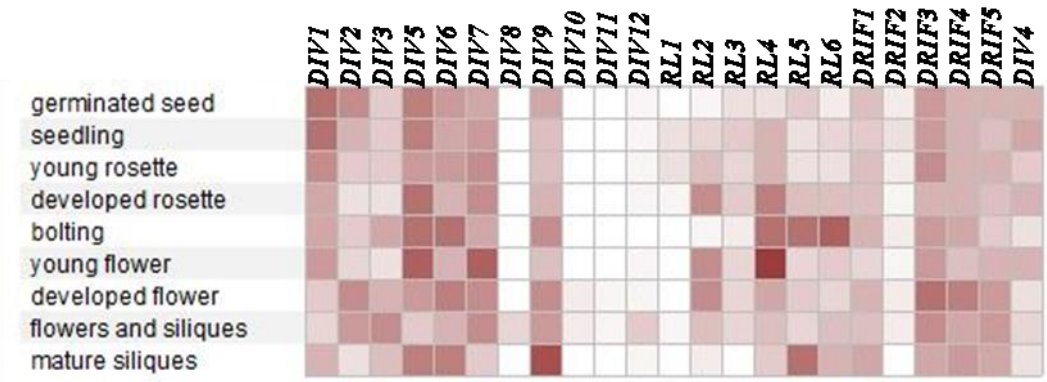


Fig.4. Expression of *Arabidopsis DIV, RAD*, and *DRIF* homologs in different developmental stages. The expression of *DIV*, *RAD* and *DRIF* homologs in germinated seeds, seedling, young rosette, developed rosette, bolting, young flower, developed flower, flower and siliques, and mature siliques. A heat map was generated using the RNAseq dataset by Geneinvestigator.

Further, we performed hierarchical clustering of this *DIV*, *DRIF*, and *RAD* (DDR) homologs on the basis of their expression at the different stages of development of the plant (Fig.5). This study showed that these DDR homologs are divided into two main clusters. *DIV1*, *DIV2*, *DIV3*, *DIV5*, *DIV6*, *DIV7*, and *DIV9* appeared in the same cluster along with *RL2*, *RL4*, *DRIF1*, *DRIF3*, *DRIF4*, and *DRIF5*. The expression level of these genes was higher during all the developmental stages compared to the genes in the second cluster. The second cluster of genes comprises *DIV4*, *DIV8*, *DIV10*, *DIV11*, and *DIV12* along with *RL1*, *RL3*, *RL5*, *RL6* and *DRIF2* (Fig.5).

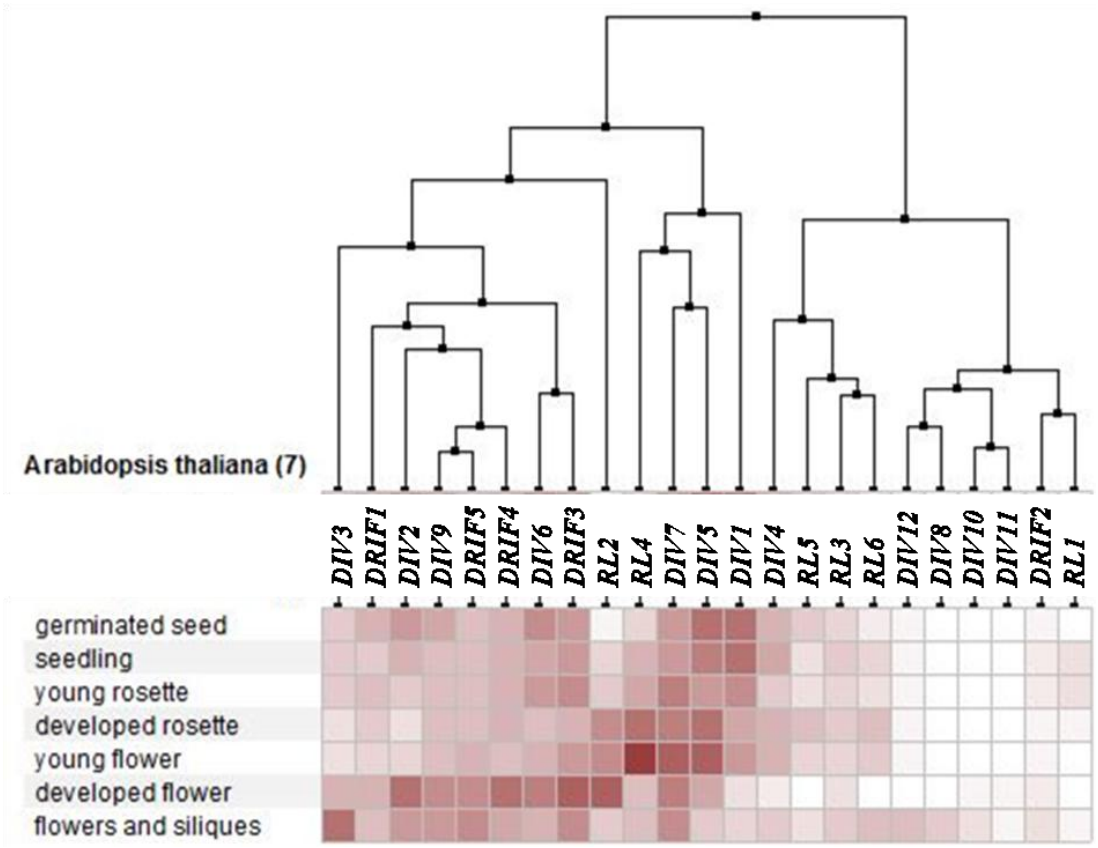


Fig.5. Hierarchical clustering of *Arabidopsis DIV, RAD,* and *DRIF*. Clustering of 12 *DIV,* 6 *RAD* and 5 *DRIF* homologs of *Arabidopsis* on the basis of their expression level in germinated seed, seedling, young rosette, developed rosette, young flower, developed flower and flower and siliques. Heatmap obtained from Geneinvestigator using the RNASeq dataset.

3.4.2 Expression of *DIV, RAD* and *DRIFs* in different mutant backgrounds

The expression profile of the DDR homologs in *Arabidopsis* suggested that the expression of some genes are more specific to seeds and seedlings while others are expressed during different stages of flower and seed maturation. In order to investigated whether these genes are involved in the early seedling development and flowering, we checked the expression of these DDR homologs in different mutant backgrounds associated with seedling development (*pif1*, *pifq*, *cop1,* and *ein3eil1*) and flowering (*lfy*,*co*, and *ft*) mutants.

Microarray data showed that the expression level of *RL2*, *RL5,* and *RL6* in *pif1* and *pifq* mutants were similar to wild-type (WT) in germinated seeds (Fig.6). Expression of *DIV3*, *DIV4*, *DIV8*, *DIV9,* and *DIV12* was also unchanged. The expression of *DIV1* was reduced in *pif1* but increased in *pifq* mutant. Expression of *DIV2* was increased in *pif1* but in *pifq*, the expression level was unchanged. Expression of *DIV5* was unchanged in *pif1* and it was slightly higher in *pifq* while the expression of *DIV6* was higher in

pif1 and unchanged in *pifq*. The expression of *DIV7* was very high in *pifq*. The expression of *DIV10* was slightly increased in *pif1* but unchanged in *pifq* (Fig.6).

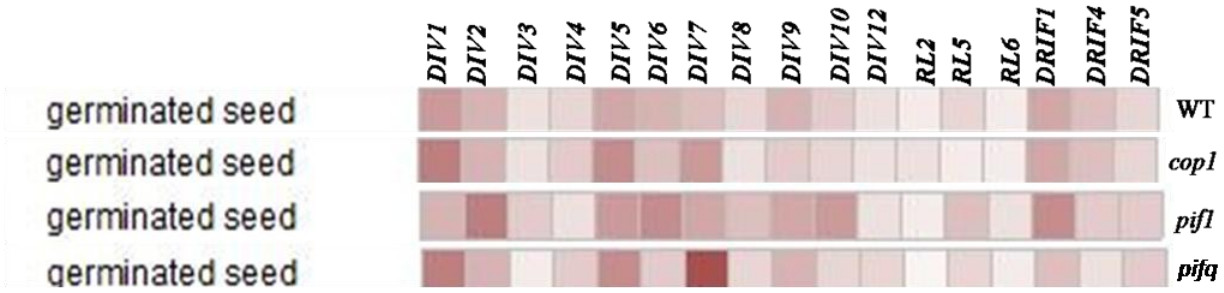


Fig.6. Expression of *Arabidopsis DIVs*, *RLs,* and *DRIFs* in *pif1*, *pifq*, and *cop1* mutants compared with WT in the germinated seed. Heatmap obtained from Geneinvestigator using a microarray dataset.

The expression of *DRIF1* was increased in *pif1* and slightly reduced in *pifq* mutant. Expression of *DRIF4* and *DRIF5* was not much changed in *pif1* but in *pifq*, the expression level of *DRIF4* was slightly reduced and *DRIF5* was slightly increased (Fig.6).The pattern of change in the expression of *DIVs* in the *cop1* mutant was similar to the *pifq* mutant. The expression of *DRIF1*, *DRIF4,* and *DRIF5* was not altered in the *cop1* mutant. Expressions of *RL2* and *RL6* were not changed in *cop1* but the expression of *RL5* was slightly reduced in this mutant (Fig.6).

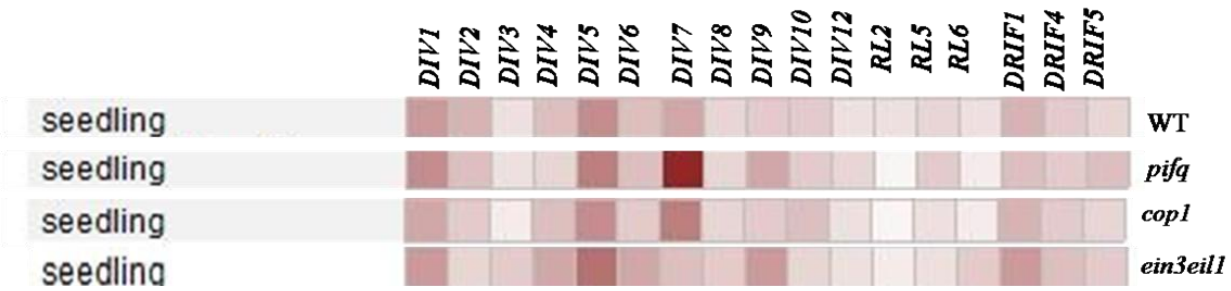


Fig.7. Expression of *Arabidopsis DIVs*, *RLs,* and *DRIFs* in *pifq*, *cop1*, and *ein3eil1* compared with WT in the seedling. Heatmap obtained from Geneinvestigator using microarray dataset.

At the seedling stage, the pattern of change in the expression levels of *DIV*, *RAD,* and *DRIFs* was similar to the seed germination stage of *pifq* and *cop1* mutants. Expression data of these genes at the seedling stage of the *pif1* mutant was not available. We also checked the expression of DDR homologs in the *ein3eil1* mutant, associated with the early development of seedlings like *cop1* and *pifq*. We observed that the expression of *DIV7* was increased in *pifq.* In *ein3eil1* mutant the expression of *DIV7* was slightly decreased (Fig.7).

Further, we checked the expression data of these DDR homologs in flowering time mutants. Microarray data available for *lfy* at bolting stage and *co* and *ft* at the rosette stage showed variation at the level of expression in *DIV*, *RL* and *DRIFs* from WT. Expression of *DIV1*, *DIV5* and *DIV9* were reduced in *lfy* mutant while expression of *DIV7* and *DIV10* were increased in *co* and *ft* (Fig.8). Expression of *RL5* and *RL6* was highly reduced in *lfy* while in *co* and *ft*, the expression of *RLs* was similar to WT. The expression of *DRIF4* was slightly increased in *lfy* while the expression of *DRIF5* was reduced in *co* and *ft* (Fig.8).

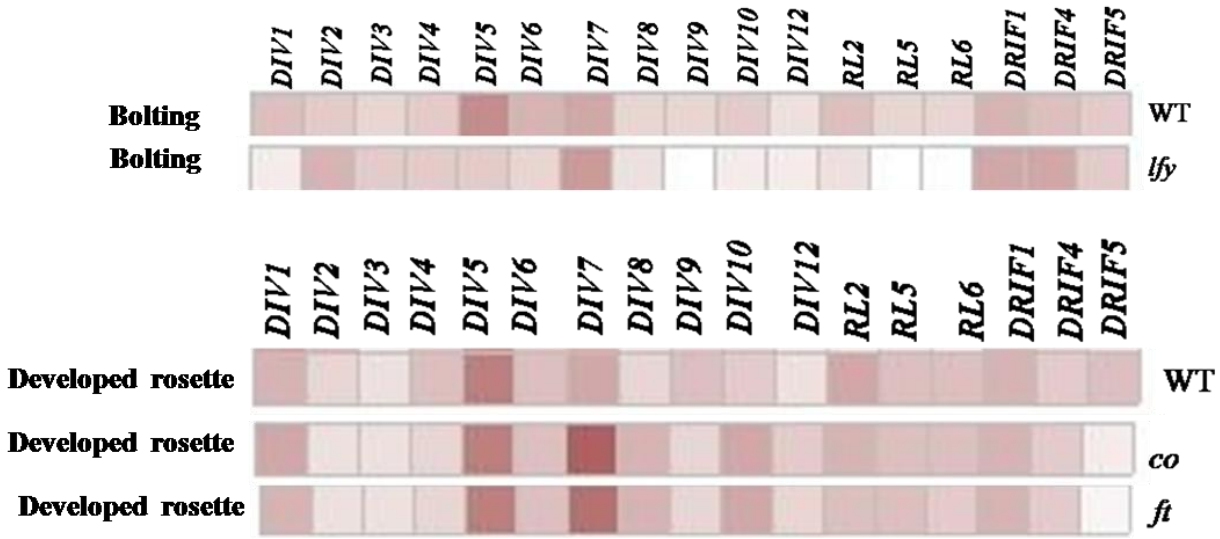


Fig.8.Expression of *Arabidopsis thaliana DIV*, *RL* and *DRIF* homologs in *lfy* at bolting stage and in *co* and *ft* in developed rosette compared with WT. Heatmap obtained from Geneinvestigator using microarray dataset.

These results indicated the association of *DIVs*, *RLs,* and *DRIFs* with gene regulatory networks involved in seedling development and flowering. Amongst the 11 *DIVs* that appeared in the microarray datasets, the expression of *DIV7* was up-regulated in *pifq*, *cop1*, *co*, and *ft* mutants. The expression of *RL2* was reduced in *cop1* and *pifq* mutant seedling while the expression of *RL5* and *RL6* was reduced in *lfy*. Expression of *DRIF4* and *DRIF5* were more affected in the flowering mutants *co* and *ft*. In *pifq* also, the expression of *DRIF4* and *DRIF5* were altered. These results showed the possible role of some of the DDR genes during early seedling development and flowering in *Arabidopsis*.

3.4.3 *In silico* investigation of co-expression networks and protein interactions

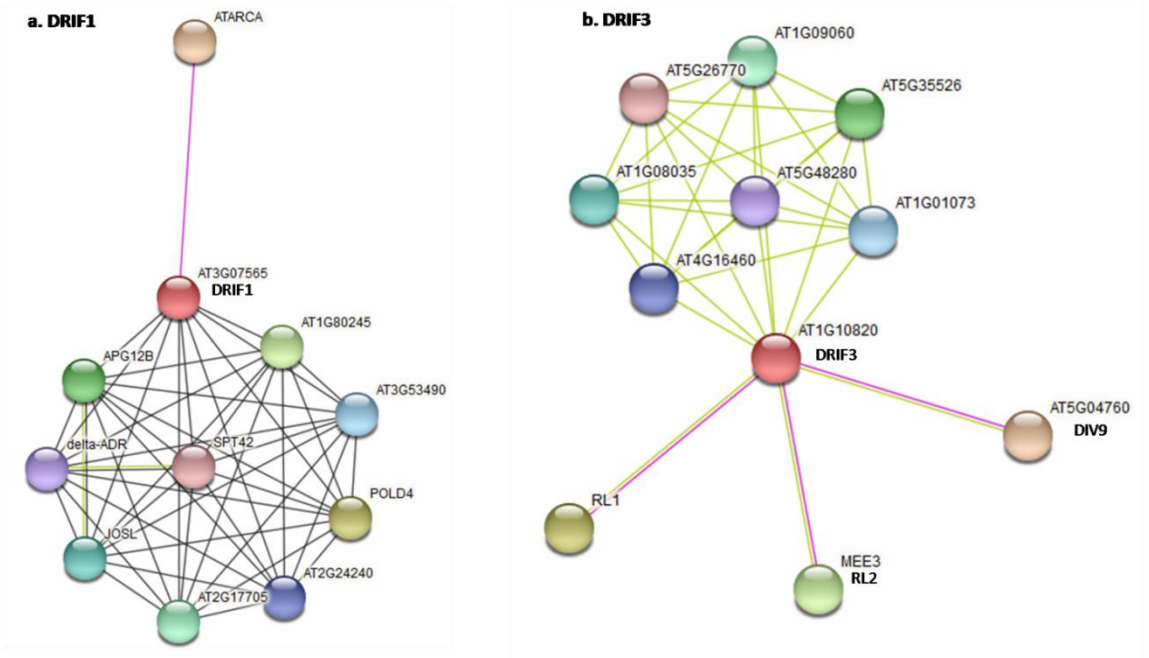
According to the idea of 'guilt-by-association', co-expressed genes along with the gene of interest across many samples, help to generate hypotheses about the function of that gene (Wu *et al.*,

2002; Lee *et al.*, 2004). To decipher the possible role of the *Arabidopsis DRIF* homologs, the set of co-expressed genes and protein partners that possibly interact with *DRIFs* were investigated using a co-expression tool in Geneinvestigator, along with GeneMANIA and STRING. We checked the genes correlated with the *Arabidopsis DRIF* homologs across 691 perturbations from the RNAseq dataset (Table.3.2). The genes that appeared in this channel of selection showed very low Pearson's correlation coefficient value (>0.8). Further, we searched the co-expression network only in seed and siliques as the expression of all *DRIFs* was observed at these stages (Fig.5). Across 16 perturbations from seed and silique RNAseq datasets, we found the top 25 genes with Pearson's correlation coefficient value higher than 0.8 co-expressing along with *DRIFs*. There were no *DIV* or *RAD* homologs found in that dataset

The *DRIF1* co-expression data contained genes involved in protein modification, fatty acid biosynthesis, seed dormancy and germination along with cell differentiation and root growth. Genes involved in circadian regulation, photoperiod response to flowering (*VIL2*) were also present in this dataset with a 0.98 score (Table.3.3.1). Genes are co-expressed with *DRIF2* included membrane proteins synthesizing genes, transposable elements, and some uncharacterized proteins. Genes related to RNA and protein modifications were also obtained in this dataset (Table.3.3.2.). The set of genes co-expressing with *DRIF3* included genes involved in ABA, GA, and brassinosteroid signaling pathways but also genes involved in RNA and protein processing and modification, meristem maintenance and floral organ formation (Table.3.3.3). Gene co-expressing with *DRIF4* included the genes coding for membrane proteins, genes involved in ion and protein transport, kinases and chaperons involved in stress responses. Genes involved in hormonal responses were also present (Table.3.3.4). The genes that appeared in the *DRIF5* co-expression dataset represented the protein biosynthesizing genes and transposons elements. Genes involved in stress-mediated responses were also found in the co-expression network of *DRIF5* along with some uncharacterized genes/proteins (Fig.3.3.5).

To find the interacting partners of DRIFs, we searched the STRING and the BioGRID databases. There was no information available for DRIF2 in STRING database. For DRIF1, we lowered the stringency to obtain the interaction network. The interacting partners for DRIF1 proteins appeared from three channels of association: 1. The experimental channel based on laboratory experiments (includes biophysical, biochemical and genetic experiments), 2. Textmining that shows proteins mentioned along with query protein and 3. Co-expression, based on gene expression data.

The proteins that interact experimentally with DRIF1 are ATARCA that is a major component of RACK1 proteins (Olejnik *et al.*, 2011). The RACK1 regulatory proteins are involved in multiple signal transduction pathways regulating multiple hormone responses and developmental processes (Chen *et al.*, 2006; Guo *et al.*, 2009). Other proteins that co-expressed with DRIF1 are ubiquitin family proteins, membrane proteins, and DNA replication (Table.3.4.1). Interaction of DRIF3, DRIF4, and DRIF5 with RADs and DIVs obtained from experimental data (Machemer *et al.*, 2011) while the other partners appeared by textmining. Three proteins, RL1, RL2, and DIV9 are common interactors of DRIF3, DRIF4, and DRIF5. The other interactors of DRIF3 are ubiquitin-like protein and zinc-finger proteins (Table.3.4.2). DRIF4 interacts with all the RLs except RL4 and only one DIV (DIV9) while DRIF5 interacts with six DIVs (DIV1, DIV2, DIV3, DIV4, DIV5, and DIV9) and only two RLs (RL1 and RL2) (Table.3.4.3 and 3.4.4).



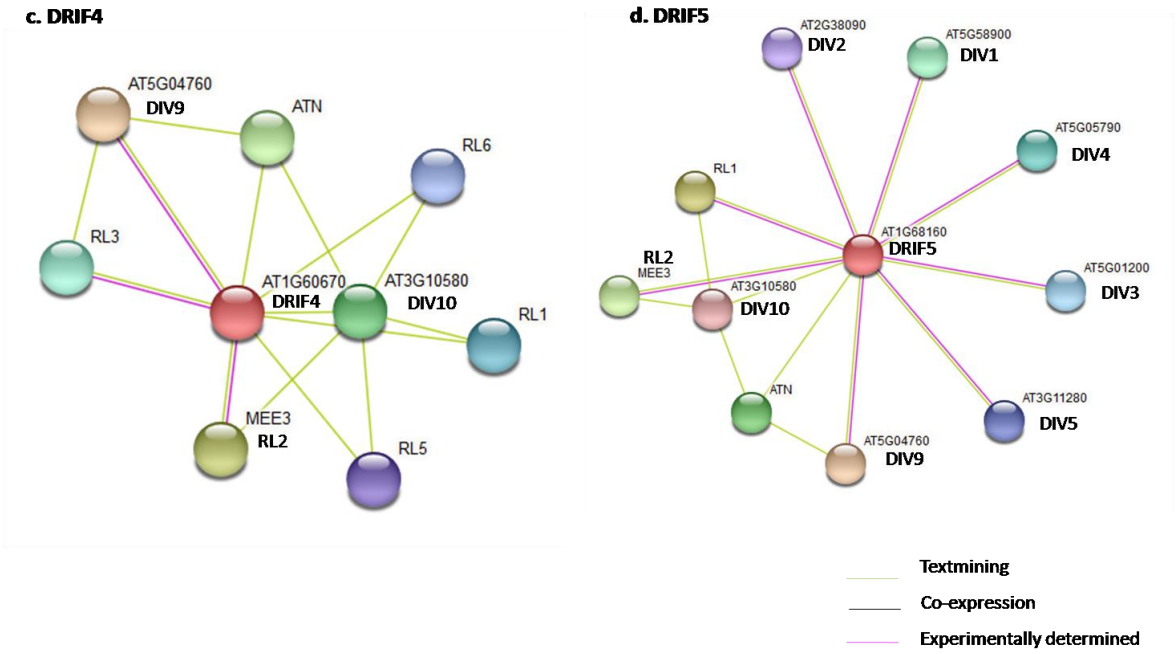


Fig.9. Protein network associated with *Arabidopsis* DRIF proteins. Predicted functional partners of (a). DRIF1, (b). DRIF3, (c). DRIF4, and (d). DRIF5. Nodes represent the proteins and edges represent protein-protein interactions obtained through different channels. Nodes connected with green lines represent the proteins that appeared in literature along with query protein. Nodes connected with black lines represent the proteins that appeared in different co-expression datasets. Nodes connected with pink lines represent the proteins interact with query proteins in laboratory experiments.

These results indicate that all five *DRIFs* may be involved in various developmental processes. Although *RAD* and *DIV* do not co-express along with *DRIFs* in datasets that were analyzed in this study, DRIF proteins have the ability to bind with RADs and DIVs. On the basis of expression and interacting partners we speculated that out of the five *DRIFs*, *DRIF3*, *DRIF4,* and *DRIF5* are more similar to each other and may have overlapping functions.

3.3.4. *In silico* analysis of promoters of *Arabidopsis DRIF* genes

The differential expression of a gene depends on the *cis*-regulatory elements present in its promoter region. From the expression data of *Arabidopsis DRIFs*, we found that the expression pattern of *DRIFs* were different and on the basis of this they were divided into two clusters (Fig.5). In order to understand the role of *cis*-regulatory elements in their expression pattern, we analyze the 5′ upstream intergenic sequences of *DRIF1*, *DRIF2*, *DRIF3*, *DRIF4*, and *DRIF5* using PlantCARE and PlantPAN database to identify the TF-binding motifs in the sequences.

DRIF1 has the smallest putative promoter region of 347 base pairs (bp), containing two motifs, GATAGGA, and TCTCCT, that are part of light-responsive elements. The *cis*-acting elements involved in

cold response and methyl jasmonate (MeJA) responses, were also found in this promoter along with various transcription promoters and enhancer regions (Appendix 1.A).The intergenic region of *DRIF2* used for promoter analysis was 1,783 bp long. This putative *DRIF2* promoter had several *cis*-elements involved in light responsiveness. The abscisic acid-responsive element (ABRE) with high matrix score (GCAACGTGTC) and auxin-responsive elements were also present in that region (Appendix.1.B). The promoter region of *DRIF2* showed various MYB binding sites.

Putative promoter region for *DRIF3* (921 bp) was analyzed and we found that two MeJA responsive *cis*-binding elements along with MYB binding sites and light-responsive elements are present in this region (Appendix.1.C). The putative promoter region of *DRIF4* was small having 444 bp. This region contained *cis*-regulatory elements responsible for ABA and MeJA response along with light-responsive elements. The promoter of *DRIF4* also has MYB binding sites (Appendix.1.D). *DRIF5* had the longest intergenic region of 3,025 bp and we found that this region has many MYB binding sites and light-responsive elements along with ABA, MeJA and GA responsive elements. This promoter also has an RY-element (CATGCATG) that regulates seed-specific expression (Appendix.1.E)

3.5 Discussion

Predicting gene function by using available global transcriptomic data is one of the most important roles of bioinformatics. *In silico* gene expression profile and characterization of co-expression networks in different conditions provides an opportunity to infer on the possible function of uncharacterized genes. In the present study, we attempt to understand the function of the DDR homologs by using web-tools for gene expression analysis. The expression of *RAD*, *DIV,* and *DRIF* homologs during different developmental stages shows the functional diversity of these homologs. Some of these genes are expressed throughout all the stages while others are specific to particular developmental stages. The specific expression of these genes was found in seeds, the transition from vegetative to reproductive phase, specific floral organs and during the maturation of seed and siliques that suggests their involvement in different developmental processes and also the functional diversity among these homologs.

Transcriptome data showed the expression of all *DIVs, RLs*, and *DRIFs* either in germinating seeds or in siliques. This result indicated that DDR genes in *Arabidopsis* may be involved in seed germination. *PIFs* and *COP1* are genes involved in the regulation of germination and photomorphogenesis (Oh *et al.*, 2006; Xu *et al.*, 2014). Microarray data obtained from *pif1*, *pifq,* and *cop1* mutants at two stages, germinating seeds and seedling, are available, which allowed us to

compare the expression of DDR genes in these mutants. The expression of *DIVs* was altered in *pif1*, *pifq*, and *cop1* seeds during germination. Expression of *DIV2*, *DIV6*, and *DIV10* was increased in *pif1* mutant while expression of *DIV1* and *DIV7* were increased in *pifq* mutant and the expression of *DIV5* and *DIV7* were increased in *cop1*. This result shows that the *DIV* homologs may regulate the germination of seeds by acting as a downstream target of *PIFs* and *COP1*. This result supports our hypothesis about the role of DDR genes in the early development of seedling. Up-regulation or down-regulation DDR genes in absence of *pifs* and *cop1* show the possible role of these genes in germination and photomorphogenesis regulated by *PIFs* and *COP1*.

PIFs are known to be involved in the formation of the apical hook along with *EIN3* and *EIL1.* Dark grown seedlings show hookless phenotype in the absence of these genes (Zhang *et al.*, 2018). It has been reported that *RL2* overexpression causes disruption in the apical hook of dark-grown seedling in *Arabidopsis* (Hamaguchi *et al.*, 2008) but the expression of *RL2* was slightly down-regulated in *pifq* and unchanged in the *ein3eil1* mutant. On the other hand, the expression of *DIV7* is up-regulated in *pifq* and *cop1* mutant and expression of *DIV9* is up-regulated in the *ein3eil1* mutant background (Fig.7). It may be possible that these *DIV* homologs are involved in photomorphogenesis and suppressed by *PIFs* and *COP1* in dark. In the absence of these negative regulators of *DIVs*, expression of *DIVs* increases and causes photomorphogenic phenotype in the dark-grown seedling. DRIF proteins interacts with their partners RADs and DIVs. The interaction of DRIFs with RAD or DIV may be dependent on the availability of these proteins at different developmental stages and in different plant organs. Change in the level of expression of *DIVs* and *RADs* in these mutants might alter the interaction of DIV and RAD with DRIFs and the downstream processes regulated by DIV-DRIF/RAD-DRIF complexes. Transition from the vegetative to the reproductive phase and flowering time of *Arabidopsis* in response to photoperiod is controlled by *LFY*, *CO* and *FT* (Blazquez *et al.*, 1997; Kim *et al.*, 2008). Mutation in these genes causes delayed flowering and distorted flowers. Microarray data of these mutants showed that the expression of DDR genes was also altered in these mutants. Expression of *DIV7* was up-regulated while the expression of *RL5* and *RL6* was down-regulated in *lfy* that indicates that these *DIV* and *RAD* homologs, may associated with flowering.

In order to know more specifically about the function of the *DRIF* homologs, we explored the co-expression network and promoter sequence of all five *DRIF* homologs. The *DRIFs* co-expression network included genes involved in transcriptional regulation, modification of nucleic acid and proteins alongside with genes involved in hormone signaling, stress responses, circadian regulation, and photoperiod response. The presence of binding sites responsive to hormones, stress, and light supports the co-

expression network. We could not find *DIVs* and *RADs* co-expressing with *DRIFs* with a higher score of confidence but experimentally DRIF3, DRIF4 and DRIF5 interact with RAD and DIV. *In silico* expression data of *DIVs* and *RADs* are tissue and stage-specific but the RNAseq data for the stages that highly express some DDR homologs are not available. For instance, expression of *DIV1* and *DRIF1* appeared highest in dry seed but the co-expression dataset is not available for dry seeds. So, it may be possible that the samples used for co-expression analysis do not represent all the *DIV* and *RAD* homologs.

In previous studies, the role of *DIV* and *RAD* homologs in *Arabidopsis* was studied with the help of loss-of-function mutants but because of the presence of several homologs and functional redundancy among them, their functions were not completely understood (Hamaguchi*et al.*, 2008; Fang *et al.*, 2018). The availability of transcriptome data for *Arabidopsis* provided us an exceptional opportunity to explore the expression profile of all the twenty-three DDR genes. The transcriptome data shows the expression of *DRIF* homologs in seed, seedling, and flower and similar expression profile of *DRIF3*, *DRIF4*, and *DRIF5* indicates that these three *DRIF* homologs may be involved in similar functions. This initial analysis of *DRIF* genes in *Arabidopsis* conducted in this study helped us to build a hypothesis regarding the possible role of *DRIFs* in seed germination and seedling development. The presence of hormone-responsive elements in the promoter region of *DRIFs* and co-expression of genes involved in hormone signaling shows a possibility of regulation of *DRIF* genes in response to ABA and GA. The presence of light-responsive elements in the promoter of *DRIFs* and also the co-expression of light-responsive genes along with *DRIFs* gives a clue about a putative role of *DRIFs* in response to light. In the future, this information will be useful to reveal the function of*DRIF* genes and it will also help to understand the regulatory mechanism by which they are involved in plant growth and developmental processes.

3.6 References

Almeida, J., Rocheta, M. & Galego, L. (1997). Genetic control of flower shape in Antirrhinum majus. *Development* 124(7): 1387-1392.

Blazquez, M. A., Soowal, L. N., Lee, I. & Weigel, D. (1997). LEAFY expression and flower initiation in Arabidopsis. *Development* 124(19): 3835-3844.

Chen, J.-G., Ullah, H., Temple, B., Liang, J., Guo, J., Alonso, J. M., Ecker, J. R. & Jones, A. M. (2006). RACK1 mediates multiple hormone responsiveness and developmental processes in Arabidopsis. *Journal of Experimental Botany* 57(11): 2697-2708.

Chen, Y. S., Chao, Y. C., Tseng, T. W., Huang, C. K., Lo, P. C. & Lu, C. A. (2017). Two MYB-related transcription factors play opposite roles in sugar signaling in Arabidopsis. *Plant Mol Biol* 93(3): 299-311.

Chow, C. N., Zheng, H. Q., Wu, N. Y., Chien, C. H., Huang, H. D., Lee, T. Y., Chiang-Hsieh, Y. F., Hou, P. F., Yang, T. Y. & Chang, W. C. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res* 44(D1): D1154-1160.

Corley, S. B., Carpenter, R., Copsey, L. & Coen, E. (2005). Floral asymmetry involves an interplay between TCP and MYB transcription factors in Antirrhinum. *Proceedings of the National Academy of Sciences of the United States of America* 102(14): 5068-5073.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C. & Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132(22): 5093-5101.

Fang, Q., Wang, Q., Mao, H., Xu, J., Wang, Y., Hu, H., He, S., Tu, J., Cheng, C., Tian, G., Wang, X., Liu, X., Zhang, C. & Luo, K. (2018). AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Rep* 37(11): 1499-1511.

Galego, L. & Almeida, J. (2002). Role of DIVARICATA in the control of dorsoventral asymmetry in Antirrhinum flowers. *Genes Dev* 16(7): 880-891.

Guo, J., Wang, J., Xi, L., Huang, W. D., Liang, J. & Chen, J. G. (2009). RACK1 is a negative regulator of ABA responses in Arabidopsis. *J Exp Bot* 60(13): 3819-3833.

Hamaguchi, A., Yamashino, T., Koizumi, N., Kiba, T., Kojima, M., Sakakibara, H. & Mizuno, T. (2008). A small subfamily of Arabidopsis RADIALIS-LIKE SANT/MYB genes: a link to HOOKLESS1-mediated signal transduction during early morphogenesis. *Biosci Biotechnol Biochem* 72(10): 2687-2696.

Kim, S. Y., Yu, X. & Michaels, S. D. (2008). Regulation of CONSTANS and FLOWERING LOCUS T expression in response to changing light quality. *Plant Physiol* 148(1): 269-279.

Kwon, Y., Kim, J. H., Nguyen, H. N., Jikumaru, Y., Kamiya, Y., Hong, S. W. & Lee, H. (2013). A novel Arabidopsis MYB-like transcription factor, MYBH, regulates hypocotyl elongation by enhancing auxin accumulation. *J Exp Bot* 64(12): 3911-3922.

Lee, H. K., Hsu, A. K., Sajdak, J., Qin, J. & Pavlidis, P. (2004). Coexpression analysis of human genes across many microarray data sets. *Genome Res* 14(6): 1085-1094.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. & Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 30(1): 325-327.

Luo, D., Carpenter, R., Copsey, L., Vincent, C., Clark, J. & Coen, E. (1999). Control of organ asymmetry in flowers of Antirrhinum. *Cell* 99(4): 367-376.

Luo, D., Carpenter, R., Vincent, C., Copsey, L. & Coen, E. (1996). Origin of floral asymmetry in Antirrhinum. *Nature* 383(6603): 794-799.

Machemer, K., Shaiman, O., Salts, Y., Shabtai, S., Sobolev, I., Belausov, E., Grotewold, E. & Barg, R. (2011). Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. *Plant J* 68(2): 337-350.

Oh, E., Yamaguchi, S., Kamiya, Y., Bae, G., Chung, W.-I. & Choi, G. (2006). Light activates the degradation of PIL5 protein to promote seed germination through gibberellin in Arabidopsis. *The Plant Journal* 47(1): 124-139.

Olejnik, K., Bucholc, M., Anielska-Mazur, A., Lipko, A., Kujawa, M., Modzelan, M., Augustyn, A. & Kraszewska, E. (2011). Arabidopsis thaliana Nudix hydrolase AtNUDT7 forms complexes with the regulatory RACK1A protein and Ggamma subunits of the signal transducing heterotrimeric G protein. *Acta Biochim Pol* 58(4): 609-616.

Oughtred, R., Stark, C., Breitkreutz, B.-J., Rust, J., Boucher, L., Chang, C., Kolas, N., O'Donnell, L., Leung, G., McAdam, R., Zhang, F., Dolma, S., Willems, A., Coulombe-Huntington, J., Chatr-aryamontri, A., Dolinski, K. & Tyers, M. (2018). The BioGRID interaction database: 2019 update. *Nucleic Acids Res* 47(D1): D529-D541.

Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., Coen, E. & Costa, M. M. R. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. *The Plant Journal* 75(4): 527-538.

Raimundo, J., Sobral, R., Laranjeira, S. & Costa, M. M. R. (2018). Successive Domain Rearrangements Underlie the Evolution of a Regulatory Module Controlled by a Small Interfering Peptide. *Mol Biol Evol* 35(12): 2873-2885.

Rose, A., Meier, I. & Wienand, U. (1999). The tomato I-box binding factor LeMYBI is a member of a novel class of myb-like proteins. *Plant J* 20(6): 641-652.

Schmid, M., Davison, T. S., Henz, S. R., Pape, U. J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D. & Lohmann, J. U. (2005). A gene expression map of Arabidopsis thaliana development. *Nat Genet* 37(5): 501-506.

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J. & Mering, C. V. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47(D1): D607-d613.

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D. & Morris, Q. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 38(Web Server issue): W214-220.

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-e718.

Wu, L. F., Hughes, T. R., Davierwala, A. P., Robinson, M. D., Stoughton, R. & Altschuler, S. J. (2002). Large-scale prediction of Saccharomyces cerevisiae gene function using overlapping transcriptional clusters. *Nat Genet* 31(3): 255-265.

Xu, X., Paik, I., Zhu, L., Bu, Q., Huang, X., Deng, X. W. & Huq, E. (2014). PHYTOCHROME INTERACTING FACTOR1 Enhances the E3 Ligase Activity of CONSTITUTIVE PHOTOMORPHOGENIC1 to Synergistically Repress Photomorphogenesis in Arabidopsis. *Plant Cell* 26(5): 1992-2006.

Yang, B., Song, Z., Li, C., Jiang, J., Zhou, Y., Wang, R., Wang, Q., Ni, C., Liang, Q., Chen, H. & Fan, L.-M. (2018). RSM1, an Arabidopsis MYB protein, interacts with HY5/HYH to modulate seed germination and seedling development in response to abscisic acid and salinity. *PLOS Genetics* 14(12): e1007839.

Zhang, X., Ji, Y., Xue, C., Ma, H., Xi, Y., Huang, P., Wang, H., An, F., Li, B., Wang, Y. & Guo, H. (2018). Integrated Regulation of Apical Hook Development by Transcriptional Coupling of EIN3/EIL1 and PIFs in Arabidopsis. *The Plant Cell* 30(9): 1971-1988.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. & Gruissem, W. (2004). GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol* 136(1): 2621-2632.

3.7 Supplementary data

Table.3.1 Accession number of *Arabidopsis DIV*, *DRIF* and *RAD* homologs

S.N.	Accession number	Name given in this study
1.	At5g58900	*DIV1*
2.	At2g38090	*DIV2*
3.	At5g01200	*DIV3*
4.	At5g05790	*DIV4*
5.	At3g11280	*DIV5*
6.	At5g08520	*DIV6*
7.	At1g49010	*DIV7*
8.	At5g23650	*DIV8*
9.	At5g04760	*DIV9*
10.	At3g10580	*DIV10*
11.	At4g09450	*DIV11*
12.	At3g10590	*DIV12*
13.	At4g39250	*RL1*
14.	At2g21650	*RL2*
15.	At4g36570	*RL3*
16.	At2g18328	*RL4*
17.	At1g19510	*RL5*
18.	At1g75250	*RL6*
19.	At3g07565	*DRIF1*
20.	At2g43470	*DRIF2*
21.	At1g10820	*DRIF3*
22.	At1g60670	*DRIF4*
23.	At1g68160	*DRIF5*

Table.3.2.1 Genes co-expressed along with *DRIF1* in 691 perturbations using RNAseq dataset

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.75	At3g07560	PEX13	Peroxisome biogenesis, Protein transport, Translocation, Transport
2.	0.72	At2g36900	MEMB11	Protein transport
3.	0.72	At4g05460	SKIP19	Protein ubiquitination
4.	0.71	At3g25910	MPE11	Kinase, Transferase
5.	0.69	At3g06190	BPM2	Protein ubiquitination
6.	0.68	At5g05100	MUG13.4	Single-stranded nucleic acid binding R3H protein
7.	0.68	At3g62770	ATG18A	Autophagy, Protein transport, Transport
8.	0.67	At1g14740	OBE3	Transcription, transcription regulation
9.	0.67	At3g57340	DnaJ-like protein	Stress response
10.	0.67	At5g18110	NCBP	Protein biosynthesis, Translation regulation
11.	0.67	At5g53300	UBC10	Protein ubiquitination, Transferase
12.	0.67	At3g01310	VIP1	Kinase, Transferase, Jasmonic acid signaling pathway
13.	0.67	At3g13200	EMP2769	RNA binding, mRNA splicing via spliceosome
14.	0.67	At2g17520	IRE1A	Serine/Threonine kinase, RNA splicing and processing
15.	0.67	At4g39910	UBP3	Hydrolase, Protease, Protein deubiquitination
16.	0.67	At1g34370	STOP1	Transcription regulation
17.	0.67	At1g16240	SYP51	Protein transport
18.	0.67	At2g19270		Uncharacterized
19.	0.66	At4g25180		DNA-directed RNA polymerase activity
20.	0.66	At3g54850	PUB14	Protein ubiquitination, Protein modification
21.	0.66	At1g11480		Protein biosynthesis
22.	0.66	At1g67580		Kinase activity, regulation of long-day photoperiodism and flowering
23.	0.66	At1g08630	THA1	Threonine catabolism, glycine biosynthesis
24.	0.66	At2g01100		
25.	0.66	At4g14030	SBP1	Cellular response to Selenium and Cadmium

Table.3.2.2 Genes co-expressed along with *DRIF2* in 691 perturbations using RNAseq dataset

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.54	At3g32920		Putative DNA-repair protein
2.	0.47	At3g29470		Unknown
3.	0.46	At3g51644		Unknown
4.	0.46	At1g31772		Defensin-like protein, involved in plant defense
5.	0.46	At1g68090	ANN5	Response to cold, heat, and R/FR light, pollen development
6.	0.46	At4g03630		SCF-dependent proteasomal ubiquitin-dependent protein catabolic process
7.	0.44	At5g27660	DEGP14	Serine-protease, hydrolase activity
8.	0.44	At5g38790		S-acetyltrasferase
9.	0.44	At3g45070	SOT5	Involved in flavonoid metabolic process
10.	0.44	At5g11440	CID5	Increased polyploidy level in dark and hypocotyl elongation
11.	0.43	At4g29305	LCR25	Defensin-like protein
12.	0.43	At2g14700		Unknown
13.	0.43	At3g24670		Involve in pectin degradation pathway
14.	0.43	At3g22570		Unknown
15.	0.41	At3g12955	SAUR74	Small auxin upregulated RNA74
16.	0.41	At5g10572	SNOR77	Unknown
17.	0.40	At1g10680	ABCB10	ATPase-coupled transmembrane transport activity
18.	0.40	At1g32950	SBT3.4	Serine-type endopeptidase activity
19.	0.40	At4g01930		Cysteine/histidine-rich C1 domain family protein
20.	0.40	At2g40920		F-box/LRR-repeat protein
21.	0.40	At3g5930		Natural antisense of GLP8 involved in germination
22.	0.40	At4g23590		Tyrosine transaminase family protein
23.	0.40	At5g03960	IQD12	Member of CaM domain-containing family
24.	0.40	At2g17064		Pseudogene of *At2g17080* involved in tetrahydrofolate biosynthesis
25.	0.39	At1g72870		Involved in methyltransferase

Table.3.2.3 Genes co-expressed along with *DRIF3* in 691 perturbations using RNAseq dataset

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.71	At5g06440	MHF15	lipid transport superfamily protein
2.	0.71	At5g45760	MRA19	
3.	0.70	At2g28070	ABCG3	Transmembrane transport
4.	0.69		CPuORF38	Conserved protein upstream ORF, known to regulate transcription
5.	0.69		CPuORF39	Conserved protein upstream ORF, known to regulate transcription
6.	0.69		CPuORF40	Conserved protein upstream ORF, known to regulate transcription
7.	0.69	At5g64340	SAC51	Transcription regulation, Stem elongation
8.	0.69	At1g70100		Protein kinase domain-containing protein of unknown function
9.	0.69	At1g80260	emb1427	Microtubule organization and spindle formation
10.	0.69	At3g20820		Leucine-rich repeat family protein of unknown function
11.	0.68	At1g56230	Endolase(DUF1399)	
12.	0.68	At1g54730		Sugar transport
13.	0.68	At5g04510	PDPK1	Intracellular signal transduction, Ser/The kinase activity
14.	0.68	At5g53020	MNB8.8	Ribonucelase P protein subunitP38-like protein
15.	0.66	At1g15215	SHH1	Involved in RNA-directed DNA methylation, gene silencing
16.	0.66	At1g12930	ARM repeat superfamily protein	Protein import into the nucleus
17.	0.66	At2g18850	AXX17	Protein kinase domain-containing protein
18.	0.66	At3g06440	B3GALT16	Involve in protein glycosylation and protein modification
19.	0.65	At1g62330	O-fucosyltransferase 15	Glycosyl transferase activity and glycan metabolism
20.	0.65	At5g06180	MBL20.6	
21.	0.65	At3g16260	TRZ4	3'-tRNA processing endoribonuclease activity
22.	0.65	At2g02070	IDD5	TF acting as a positive regulator of the starch synthase SS4
23.	0.65	At1g33400	TPR9	Tetra tricopeptide repeat (TPR)-like superfamily protein
24.	0.65	At1g20540		Involve in protein ubiquitination
25.	0.65	At1g19835	FPP4	Trichome morphogenesis and microtubule binding

Table.3.2.4 Genes co-expressed along with *DRIF4* in 691 perturbations using RNAseq dataset

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.70	At5g41350	MYC6.6	Metal-binding protein
2.	0.69	At3g14830	AXX17	RBR-type E3 ubiquitin transferase
3.	0.69	At1g49590	Zinc finger protein (ZOP1)	Involve in mRNA processing, splicing and RNA induced gene silencing
4.	0.69	At2g32700	Transcriptional co-repressor LEUNIG-HOMOLOG	Transcriptional corepressor of abiotic response genes as epigenetic regulator, involve in various developmental process
5.	0.68	At3g19770	Vacuolar protein sorting-associated protein 9A(VPS9A)	Involve in cell wall biogenesis, embryo development ending in seed dormancy and post-embryonic root development
6.	0.68	At5g61500	Autophagy-related protein 3(ATG3)	Involve in autophagy, protein transport and Ubi-conjugation pathway
7.	0.68	At5g41150	RAD1/UVH1	DNA repair endonuclease
8.	0.67	At2g21270	Ubiquitin fusion degradation1(UFD1)	Involve in ubiquitin dependent protein catabolic processes
9.	0.67	At2g16405		
10.	0.67	At1g75850	Vacuolar protein sorting-associated protein 35B(VPS35B)	Protein transport, Involve in efficient sorting of seed storage proteins
11.	0.67	At3g61710	Beclin1-like protein(ATG6)	Required for normal plant development and pollen germination, may involve in vacuolar protein sorting
12.	0.67	At5g13020	EML3	Defense response to fungus
13.	0.65	At2g01460		Hydrolase activity
14.	0.65	At1g55520	TBP2	TATA-box-binding protein
15.	0.65	At1g39340	SAC3A	Involve in RNA splicing
16.	0.65	At3g27260	GTE3	Bromodomain containing protein of unknown function
17.	0.65	At1g79020		Enhancer of the polycomb-like gene of unknown
18.	0.65	At1g09730	ULP2B	Involve in protein sumoylation
19.	0.65	At3g03120	ARFB1C	ADP-ribosylation factor
20.	0.65	At3g11960		Proteosome-mediatedUbi-dependent protein catabolism
21.	0.65	At3g09180	MED27	Mediator of RNA polymerase II transcription subunit
22.	0.64	At4g38360	LAZ1	Involved in negative regulation of brassinosteroid mediated signaling pathway
23.	0.64	At5g67320	HOS15	Promotes deacetylation of histone H4, acts as a repressor of cold stress-regulated gene expression
24.	0.64	At4g31410		Involved in serine-threonine kinase activity and pollen recognition
25.	0.64	At5g43560	TRAF1A	Required for SINAT1- and SINAT2-mediated ubiquitination

Table.3.2.5 Genes co-expressed along with *DRIF5* in 691 perturbations using RNAseq dataset

S. no.	Score	Gene id	Gene name	Functions
1.	0.68	At5g53530	VPS26A	Involved in intracellular protein transport
2.	0.66	At4g14147	ARPC4	Functions as actin-binding component
3.	0.66	At1g69980		Unknown
4.	0.66	At4g21160	AGD12	GTP activation and protein transport
5.	0.65	At1g71820	SEC6	involved in the docking of exocytic vesicles, polarized cell growth and organ morphogenesis
6.	0.65	At1g14230	APY4	Catalyzes the hydrolysis of phospho-anhydride bonds of nucleoside tri- and di-phosphates
7.	0.64	At5g26990	DI19-6	Dehydration induced protein homolog-19
8.	0.64	At1g09645		Unknown
9.	0.64	At1g78870	UBC35	Involve in protein ubiquitination and modification
10.	0.64	At5g02040	PRA1A1	PRA1 family protein may involve in secretory and endocytic intracellular trafficking
11.	0.64	At1g77140	VPS45	Vascular protein sorting-associated protein
12.	0.64	At3g53670		Uncharacterized protein
13.	0.64	At4g17270		Signal transduction by protein phosphorylation
14.	0.64	At5g22950	VPS24-1	Protein transport
15.	0.63	At1g09920		Membrane protein
16.	0.63	At2g01650	PUX2	Involve in UBI conjugation pathway and metal-binding
17.	0.63	At1g10720		BDS-domain containing protein
18.	0.63	At5g61450	MCI2.1	Involve in hydrolase activity
19.	0.63	At5g51290	CERK	Ceramide kinase, ceramide is involved in programmed cell death
20.	0.63	At1g65820		Glutathione peroxidase activity
21.	0.63	At1g16540	ABA3	ABA-biosyhtesis and abiotic stress response
22.	0.63	At1g16800		Acyl-CoA N-acyl transferases superfamily protein, involved in histone acetylation
23.	0.63	At3g03560		Uncharacterized protein
24.	0.63	At2g33120	VAMP722	Vesicle mediated protein transport
25.	0.63	At2g44280		Carbohydrate and organic substance transport

Table.3.3.1 Genes co-expressed along with *DRIF1* in 16 perturbations using RNAseq dataset from seed and siliques

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.98	At4g30200	VIL2	Circadian regulation of gene expression, flower development, photoperiod response
2.	0.98	At1g12780	UGE1	Plays a role in D-galactose detoxification
3.	0.98	At3g17850	IREH1	May be involved in root hair elongation
4.	0.98	At4g36650	PBRP	Transcriptional regulation
5.	0.98	At5g63610	CDKE-1	Involve in cell differentiation, stamen and carpel identity
6.	0.97	At1g61040	VIP5	Flowering and transcriptional regulation
7.	0.97	At1g61040	TTM1	Pyrophasphatase activity
8.	0.97	At4g15530	PPDK	May be involved in regulating the flux of carbon into starch and fatty acids in seeds
9.	0.97	At3g47610	F1P2	Transcription regulation
10.	0.97	At5g03720	HSFA3	Heat stress transcription factor
11.	0.97	At1g11480		Protein biosynthesis
12.	0.97	At2g27200	LSG1-1	Hydrolase activity
13.	0.97	At4g34370	ARI1	May act as E3-ubiquitin-protein ligase
14.	0.97	At1g03140		mRNA splicing
15.	0.97	At1g03200		Serine /threonine-protein kinase
16.	0.97	At3g12570	FYD	
17.	0.97	At1g09430	ACLA-3	Lipid biosynthesis and metabolism
18.	0.97	At3g06380	TULP9	ABA signaling pathway and UBI conjugation pathway
19.	0.97	At4g26450	WIP1	Nuclear organization, protein heterodimerization
20.	0.97	At1g55510	BCDH BETA1	Branched chain amino acid catabolism
21.	0.97	At4g31430	KAKU4	Nuclear membrane organization
22.	0.97	At5g13010	EMB3011	Seed dormancy, RNA splicing, root development and root hair elongation
23.	0.97	At5g49820	RUS6	Protein root UVB sensitive 6 protein
24.	0.97	At5g05100	MUG13.4	Nucleic acid binding
25.	0.97	At1g08720	EDR1	Serine/threonine protein kinase

Table.3.3.2 Genes co-expressed along with *DRIF2* in 16 perturbations using RNAseq dataset from seed and siliques

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.89	At2g01029		Transposable element gene
2.	0.83	At3g13855	U6-26	Involved in RNA splicing
3.	0.83	At2g27150	AAO3	ABA and auxin biosynthesis pathway
4.	0.82	At3g10510		Putative F-box/kelch-repeat protein
5.	0.82	At2g21470	SAE2	Protein sumoylation and protein modification
6.	0.81	At1g75170		unknown
7.	0.81	At4g00525		unknown
8.	0.81	At3g58347		Pseudogene of phospholipid-like protein
9.	0.81	At1g09980		Putative serine esterase family protein
10.	0.79	At5g64820		transmembrane protein
11.	0.79	At5g53930		transcriptional regulator ATRX-like protein
12.	0.78	At3g27600		RNA-binding
13.	0.78	At2g19610		RING/U-box superfamily protein
14.	0.77	At1g69880	TRX8	Electron transport
15.	0.77	At3g51090	F24M12.130	unknown
16.	0.77	At5g05570	MOP10-11	Regulation of exocytosis and vesicle-mediated transport
17.	0.77	At5g58003	CPL4	Dephosphorylation of RNA polymerase II, and salt response
18.	0.77	At5g61190	MAF19.19	Regulation of gene expression
19.	0.77	At2g47150		NAD(P) – binding superfamily protein
20.	0.77	At3g03740	BPM4	Protein ubiquitination
21.	0.77	At2g23150	NRAMP3	Ion transport
22.	0.77	At1g32670		unknown
23.	0.76	At1g40118		Transposable element gene
24.	0.76	At1g74458		Transmembrane protein
25.	0.76	At2g30700		GPI-anchored protein

Table.3.3.3 Genes co-expressed along with *DRIF3* in 16 perturbations using RNAseq dataset from seed and siliques

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.93	At3g62970		Zn-finger protein involved in protein ubiquitination
2.	0.92	At4g37180		Floral organ formation and gene regulation
3.	0.92	At5g65260	PABN2	RNA-binding
4.	0.91	At5g24460		Hydrolase activity
5.	0.91	At1g54830	NFYC3	ABA and GA mediated signaling pathways, seed germination, photomorphogenesis and flowering
6.	0.91	At1g13690	ATE1	Regulation of RNA metabolic processes
7.	0.90	At1g20890		Membrane protein
8.	0.90	At5g08335	ICMTB	Negative ABA signaling, post-translational methylation
9.	0.90	At1g20560	AAE1	Acyl activating enzyme involved in AMP binding
10.	0.90	At5g06950	TGA2	Transcription regulation, plant defense and hypersensitivity
11.	0.90	At4g10110	F28M11.30	RNA-binding
12.	0.90	At5g02130	NDP1	Meristem maintenance
13.	0.90	At5g54170		Lipid binding and response to wound
14.	0.90	At3g47120		Nucleic acid and metal binding
15.	0.89	At4g02020	EZA1	Polycomb group protein
16.	0.89	At3g19290	ABF4	ABA signaling pathway and transcriptional regulation
17.	0.89	At2g38465		Unknown
18.	0.89	At1g18335		Acetylation and regulation of chromatin silencing
19.	0.89	At1g75340	CG1	Negative regulation of brassinosteroid signaling pathway, protein and mRNA transport
20.	0.89	At1g72900		Signal transduction
21.	0.89	At5g64730		Unknown
22.	0.89	At3g51530		F-box/FBD/LRR-repeat protein
23.	0.89	At1g30480	DRT111	Regulation of alternative mRNA splicing
24.	0.89	At4g27540	PRA1H	Vesicle-mediated transport
25.	0.88	At1g78590	NADK3	NADP biosynthesis

Table.3.3.4 Genes co-expressed along with *DRIF4* in 16 perturbations using RNAseq dataset from seed and siliques

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.92	At4g23930		Membrane protein
2.	0.92	At3g60400		Transcription termination
3.	0.90	At5g39360	EDL2	F-box protein
4.	0.89	At1g22020	SHM6	Interconversion of serine to glycine
5.	0.88	At3g46920		Protein kinase
6.	0.87	At3g22180	PAT20	Protein acyltransferase
7.	0.86	At1g05940	CAT9	Amino acid transport
8.	0.86	At4g38360	LAZ1	Brassinosteroid signaling pathway, hypersensitivity response
9.	0.86	At1g79970	F19K16.7	Unknown
10.	0.86	At3g27390		Uncharacterized membrane protein
11.	0.86	At4g31540	ATEXO70G1	Exocyte sub-unit family protein involved in protein transport
12.	0.85	At3g47110		Receptor protein involve in serine/threonine kinase activity
13.	0.85	At1g22770	GI	Circadian regulation and photoperiodic flowering
14.	0.84	At2g35360		Serine/threonine-protein kinase
15.	0.84	At5g35980	YAK1	Dual specific protein kinase may involve in the regulation of ABA responsive genes
16.	0.84	At2g37570	SLT1	Chaperon protein involve in osmotic regulation
17.	0.84	At4g34310		Hydrolase
18.	0.84	At5g06710	HAT14	Homeodomain leucine-zipper involve in transcriptional regulation
19.	0.83	At2g29970	SMXL7	Component of co-repressor complex, controls branching and cotyledon expansion, Promotes auxin transport and PIN1 accumulation
20.	0.83	At4g11970		RNA binding
21.	0.83	At1g01490	HIPP39	Probable heavy metal binding protein
22.	0.82	At1g73130	F3N23.33	Heat acclimation
23.	0.82	At3g54350	emb1967	Histone acetylation
24.	0.82	At4g11845		Interleukin receptor associated protein kinase
25.	0.82	At3g17740		Involve in RNA processing

Table.3.3.5 Genes co-expressed along with *DRIF5* in 16 perturbations using RNAseq dataset from seed and siliques

S. no.	Score	Gene id	Gene name	Description/ Functions
1.	0.90	At5g18015		Pre-tRNA, tRNA-try, Involved in translational elongation
2.	0.90	At3g03852		Pre-tRNA, tRNA-try, Involved in translational elongation
3.	0.90	At3g03845		Pre-tRNA, tRNA-try, Involved in translational elongation
4.	0.90	At1g80250		Pre-tRNA, tRNA-try, Involved in translational elongation
5.	0.90	At1g61880		Pre-tRNA, tRNA-try, Involved in translational elongation
6.	0.90	At1g15450		Pre-tRNA, tRNA-try, Involved in translational elongation
7.	0.90	At1g11640		Pre-tRNA, tRNA-try, Involved in translational elongation
8.	0.90	At4g17612		Pre-tRNA, tRNA-try, Involved in translational elongation
9.	0.90	At4g17765		Pre-tRNA, tRNA-try, Involved in translational elongation
10.	0.90	At4g39615		Pre-tRNA, tRNA-try, Involved in translational elongation
11.	0.90	At5g18005		Pre-tRNA, tRNA-try, Involved in translational elongation
12.	0.88	At1g35405	ATRE2-1	Retrotransposon family gene
13.	0.87	At4g39290		Galactose oxidase
14.	0.87	At1g15045		Uncharacterized
15.	0.87	At1g47690		Uncharacterized
16.	0.87	At1g47700		F-box associated domain containing protein
17.	0.86	At5g39635	MIR169C	microRNA involved in phosphate starvation
18.	0.86	At1g11760	MED32	Required for expression of CBF1 controlled gene expression in response to cold
19.	0.85	At2g04980		Transposable element
20.	0.85	At3g46230	HSP17.4A	Involved in stress responses
21.	0.85	At3g24280	SMAP2	Involved in auxin activated signaling pathway
22.	0.85	At1g80960		F-box family protein
23.	0.85	At5g35425		Transposable element
24.	0.85	At5g07430	Portable pectinsterase50	Involved in synthesis of Pectin
25.	0.85	At2g06025		Aceyl transferase subfamily protein

Table.3.4.1 Interacting partners of DRIF1

Partner name	Description/ Function	Interactions	Score
ATTRCA	Transducin/WD40 repeat-like superfamily protein; Major component of the RACK1 regulatory proteins that play a role in multiple signal transduction pathways. Involved in multiple hormone responses and developmental processes. MAPK cascade scaffolding protein involved in the protease IV and ArgC signaling pathway but not the flg22 pathway; Belongs to the WD repeat G protein beta family. Ribosomal protein RACK1 subfamily	Experimentally determined	0.216
POLD4	Polymerase delta 4 (POLD4); Its function is described as DNA-directed DNA polymerase activity; Involved in DNA replication; Located in the nucleus; Expressed in 24 plant structures; Expressed during 15 growth stages; Contains the following InterPro domains- DNA polymerase delta, subunit 4	Co-expression	0.211
At1g80245	Spc97 / Spc98 family of spindle pole body (SBP) component; Involved in microtubule cytoskeleton organization; Located in spindle pole, microtubule organizing center; Expressed in 20 plant structures; Expressed during 12 growth stages; Contains the following InterPro domains- Spc97/Spc98	Co-expression	0.209
APG12B	Ubiquitin-like superfamily protein; Ubiquitin-like protein involved in the cytoplasm to vacuole transport (Cvt) and autophagy vesicles formation. Conjugation with ATG5 through a ubiquitin-like conjugating system involving also ATG7 as an E1-like activating enzyme and ATG10 as an E2-like conjugating enzyme, is essential for its function. ATG12/ATG5 conjugate has an essential role in plant nutrient recycling; Belongs to the ATG12 family	Co-expression	0.208
At2g17705	Methionine-S-oxide reductase; Unknown protein; Its function is described as molecular function unknown; Involved in biological process unknown; Located in plasma membrane; Expressed in cultured cell	Co-expression	0.202
JOSL	JOSEPHIN-like protein; May act as a deubiquitinating enzyme	Co-expression	0.195
At3g53490	Uncharacterized protein F4P12_190; Unknown protein; Located in endomembrane system	Co-expression	0.187
At2g24240	BTB/POZ domain with WD40/YVTN repeat-like protein; May act as a substrate-specific adapter of an E3 ubiquitin-protein ligase complex (CUL3-RBX1-BTB) which mediates the ubiquitination and subsequent proteasomal degradation of target proteins	Co-expression	0.187
Delta-ADR	AP-3 complex subunit delta; Part of the AP-3 complex, an adaptor-related complex. The complex is associated with the Golgi region as well as more peripheral structures. It facilitates the budding of vesicles from the Golgi membrane and may be directly involved in trafficking to the vacuole. It also functions in maintaining the identity of lytic vacuoles and in regulating the transition between storage and lytic vacuoles; Belongs to the adaptor complexes large subunit family	Co-expression	0.181
SPT42	Transcription elongation factor SPT4 homolog 2; May regulate transcription elongation by RNA polymerase II. May enhance transcriptional pausing at sites proximal to the promoter, which may, in turn, facilitate the assembly of an elongation competent RNA polymerase II complex (By similarity)	Co-expression	0.177

Table.3.4.2 Interacting partners of DRIF3

Partner name	Function	Interactions	Score
At5g04760/DIV9	Duplicated homeodomain-like superfamily protein; Contains SANT, DNA-binding domain	Experimentally and Textmining	0.806
RL1	Protein RADIALIS-like 1; Probable transcription factor	Experimentally and Textmining	0.770
MEE3/RL2	Homeodomain-like superfamily protein; RSM1 is a member of a small sub-family of single MYB transcription factors. Analysis of overexpressing lines indicate its involvement during early morphogenesis	Experimentally and Textmining	0.770
At5g35526	Unknown protein	Textmining	0.696
At1g09060	Zinc finger, RING-type;Transcription factor jumonji/aspartyl beta-hydroxylase; Its function is described as sequence-specific DNA binding transcription factor activity, zinc ion binding; Expressed in 24 plant structures; Expressed during 15 growth stages; Contains the following InterPro domains- Transcription factor jumonji/aspartyl beta-hydroxylase (InterPro-IPR003347), Zinc finger, RING-type (InterPro-IPR001841), Transcription factor jumonji (InterPro-IPR013129), WRC (InterPro-IPR014977), Cell division cycle-associated protein	Textmining	0.696
At1g08035	Cyclin-dependent protein kinase inhibitor SMR16; Probable cyclin-dependent protein kinase (CDK) inhibitor that functions as a repressor of mitosis in the endoreduplication cell cycle	Textmining	0.696
At1g01073	Unknown protein	Textmining	0.696
At4g16460	Zinc finger CCCH domain protein; Unknown protein; Expressed in sperm cell, flower; Expressed during 4 anthesis	Textmining	0.667
At5g48280	Unknown protein; Located in endomembrane system	Textmining	0.666
At5g26770	Nuclear envelope-associated protein 2	Textmining	0.666

Table.3.4.3 Interacting partners of DRIF4

Partner name	Function	Interactions	Score
At5g04760/DIV9	Duplicated homeodomain-like superfamily protein; Contains SANT, DNA-binding domain	Experimentally and Textmining	0.744
MEE3/RL2	Homeodomain-like superfamily protein; RSM1 is a member of a small sub-family of single MYB transcription factors. Analysis of overexpressing lines indicate its involvement during early morphogenesis	Experimentally and Textmining	0.693
ATN	Encodes a protein with moderate sequence similarity to the maize microtubule-binding protein TANGLED1. A single base-pair deletion (-A) at position Chr3-1519176 in Columbia relative to the Landsberg erecta and Achkarren-2 ecotype (see ESTs DR378436 and CB26450) introduces a frame-shift and premature termination codon. The protein encoded from the Columbia gene is truncated by 29 amino acids relative to the Landsberg erecta and Achkarren-2 encoded proteins. Involved in the identification of the division plane during mitosis and cytokinesis	Textmining	0.637
At3g10580/DIV10	Duplicated homeodomain-like superfamily protein; Contains SANT, DNA-binding domain	Textmining	0.573
RL3	Protein RADIALIS-like 3; Probable transcription factor	Experimentally and Textmining	0.526
RL1	Protein RADIALIS-like 1; Probable transcription factor	Textmining	0.491
RL6	Protein RADIALIS-like 6; Probable transcription factor	Textmining	0.491
RL5	Protein RADIALIS-like 5; Probable transcription factor	Textmining	0.491

Table.3.4.4 Interacting partners of DRIF5

Partner name	Function	Interactions	Score
At5g04760/DIV9	Duplicated homeodomain-like superfamily protein; Contains SANT, DNA-binding domain	Experimentally and Textmining	0.806
RL1	Protein RADIALIS-like 1; Probable transcription factor	Experimentally and Textmining	0.770
MEE3/RL2	Protein RADIALIS-like 2; Probable transcription factor	Experimentally and Textmining	0.770
ATN	Encodes a protein with moderate sequence similarity to the maize microtubule-binding protein TANGLED1.	Textmining	0.637
At5g58900/DIV1	Similar to Myb-related transcription factor; Homeodomain-like transcriptional regulator; Its function is described as DNA binding, sequence-specific DNA binding transcription factor activity; Involved in regulation of transcription	Experimentally	0.580
At5g05790/DIV4	Duplicated homeodomain-like superfamily protein; Contains the following InterPro domains- Molecular chaperone, heat shock protein, Hsp40, DnaJ	Experimentally	0.580
At5g01200/DIV3	Duplicated homeodomain-like superfamily protein; Contains the following InterPro domains- Molecular chaperone, heat shock protein, Hsp40, DnaJ	Experimentally	0.580
At3g11280/DIV5	Duplicated homeodomain-like superfamily protein; Putative transcription factors interacting with the gene product of VHA-B1	Experimentally	0.580
At2g38090/DIV2	Duplicated homeodomain-like superfamily protein; Contains the following InterPro domains- Molecular chaperone, heat shock protein, Hsp40, DnaJ	Experimentally	0.580
At3g10580/DIV10	Homeodomain-like superfamily protein; Contains the following InterPro domains- SANT, DNA-binding	Textmining	0.573

Appendix.3.1.1 Promoter sequence of *DRIF1* showing *cis*-acting obtained from PlantCARE.Color code for motifs is presented at the end.

```
+ TAGGTTATTA TGTATCAGCC CCTATAGTTA TAATCCTTAC CCAAAAAGCC TCTTAATTAC TAAAATTGAC
- ATCCAATAAT ACATAGTCGG GGATATCAAT ATTAGGAATG GGTTTTTCGG AGAATTAATG ATTTTAACTG

+ GTATTTGACT CTCAAAAATA GAAAACCAAA ATCCGAAAAA GGATAGAGAT GAAATTATAT GTTTATATTA
- CATAAACTGA GAGTTTTTAT CTTTTGGTTT TAGGCTTTTT CCTATCTCTA CTTTAATATA CAAATATAAT

+ TTATAAGCAA AGATTCCATA TTTCGAGGGA AGCAGAAGAA GAAGCTCAGT CGGCGGCCGG ATTAATCTCC
- AATATTCGTT TCTAAGGTAT AAAGCTCCCT TCGTCTTCTT CTTCGAGTCA GCCGCCGGCC TAATTAGAGG

+ CCATCATCTC TCTCTCTCCC TCTCTCTGTT CCTATCCCGG CGAATTCATA TCTGAATTTC CCAAAATTCA
- GGTAGTAGAG AGAGAGAGGG AGAGAGACAA GGATAGGGCC GCTTAAGTAT AGACTTAAAG GGTTTTAAGT

+ AACTTTTCCG GTGGAATTAT TTTAGGAAAC CAAGTGACTT CTTAAGCTCC AAAAATTGGG TATTT
- TTGAAAAGGC CACCTTAATA AAATCCTTTG GTTCACTGAA GAATTCGAGG TTTTTAACCC ATAAA
```

+	
+	ARE
+	Box 4
+	CAAT-box
+	CGTCA-motif
+	DRE core
+	GATA-motif
+	LTR
+	O2-site
+	STRE
+	TATA-box
+	TCCC-motif
+	TGACG-motif
+	Unnamed__4
+	as-1

Appendix.3.1.2 Promoter sequence of *DRIF2* showing *cis*-acting obtained from PlantCARE. Color code for motifs is presented at the end.

```
+ GAAGCTGTTG ACTGTTGAGT GTTGAATAGA GAAGCCAAAG AGGTGTGTTT TTGAGTACTA CTAGTGAGGG
- CTTCGACAAC TGACAACTCA CAACTTATCT CTTCGGTTTC TCCACACAAA AACTCATGAT GATCACTCCC

+ TTTATATACG TAGAGATTGA TAGATTGTGA GGGCACTTAT GGGTTTTTGT CCACGTGGCA TATCTGACTA
- AAATATATGC ATCTCTAACT ATCTAACACT CCCGTGAATA CCCAAAAACA GGTGCACCGT ATAGACTGAT

+ GGATTAATTC TTAAGTATAT AGTAAGGTCC TTAAAATTAA ATTGAAATAC AAAGTTGTGC AAGCATAATT
- CCTAATTAAG AATTCATATA TCATTCCAGG AATTTTAATT TAACTTTATG TTTCAACACG TTCGTATTAA

+ ATCGGATCTT GTTGAGTTTC GTTTATGTGC CGTGCAAGCT AAGTGGCCAG CCCAATGTCG TCTAGGCATT
- TAGCCTAGAA CAACTCAAAG CAAATACACG GCACGTTCGA TTCACCGGTC GGGTTACAGC AGATCCGTAA

+ TAATTGCTGT GTCTACTTTT TGATTTTTTT TGAACTACTA ATTTGCATGA AAGTAAAAAT AGGTTTTAGA
- ATTAACGACA CAGATGAAAA ACTAAAAAAA ACTTGATGAT TAAACGTACT TTCATTTTTA TCCAAAATCT

+ TAAAGACGAT GATTTGATAA TTGAAATACA TAGAATCCTT TTTAATAGCA TTTTAGTGAA AATTTGAAAT
- ATTTCTGCTA CTAAACTATT AACTTTATGT ATCTTAGGAA AAATTATCGT AAAATCACTT TTAAACTTTA

+ GGAACTAAAT TCAGATTGGG GTTTATAGAT AAAGGTCTTA CATTTTTATT TTTTTTTGGG CAAAGGAGAT
- CCTTGATTTA AGTCTAACCC CAAATATCTA TTTCCAGAAT GTAAAAATAA AAAAAAACCC GTTTCCTCTA

+ CTTACATATG TGACATACTG ACATGTAAGA AGCAAAGGTC AAAAGAGAAA ATAATTGTTA TCACGATTGA
- GAATGTATAC ACTGTATGAC TGTACATTCT TCGTTTCCAG TTTTCTCTTT TATTAACAAT AGTGCTAACT

+ AGTTTGAGAT ATAGAAATAC TAAATTTTTT TAAAACTAAT ATTCTTCGTG TAATTTATTG AAATATCAAC
- TCAAACTCTA TATCTTTATG ATTTAAAAAA ATTTTGATTA TAAGAAGCAC ATTAAATAAC TTTATAGTTG

+ TAATTAAGCA AATGTACAGT ATATCGTTTA CGTTTTGATC GTTTAGATTA GTGAAGCAAC GTGTGAACAT
- ATTAATTCGT TTACATGTCA TATAGCAAAT GCAAAACTAG CAAATCTAAT CACTTCGTTG CACACTTGTA

+ GTGATCGTAA CTCGGATTCT GCGCGTCGTT TGATTTCCCA TGCAAAATTA TGAGACCCCA TTTCTCTTTT
- CACTAGCATT GAGCCTAAGA CGCGCAGCAA ACTAAAGGGT ACGTTTTAAT ACTCTGGGGT AAAGAGAAAA

+ TGCTTTCTTT TTTTATATAG TTTTTGTTCT TATTATAAAT TATTATAGTG TATAAGACTG TTTCTAATAT
- ACGAAAGAAA AAAATATATC AAAAACAAGA ATAATATTTA ATAATATCAC ATATTCTGAC AAAGATTATA

+ CTAAATATAA AGGATGATTT TGATTGAAAA TTCCCACAAA ATACTGATCG TGATTGTAGC ACAATAACTA
- GATTTATATT TCCTACTAAA ACTAACTTTT AAGGGTGTTT TATGACTAGC ACTAACATCG TGTTATTGAT

+ CGGATTTTTA TTTCACTGGT CCAATAATAA TTTGAACGAC GACGATAGGA AATAATGGGC AAATGATCGA
- GCCTAAAAAT AAAGTGACCA GGTTATTATT AAACTTGCTG CTGCTATCCT TTATTACCCG TTTACTAGCT

+ ATAATGATAA ATTAGACAAA TCCGTTTTTC ATTTCACATT AAATTTTCCA CAGTTAGATC AACCTCAATT
- TATTACTATT TAATCTGTTT AGGCAAAAAG TAAAGTGTAA TTTAAAAGGT GTCAATCTAG TTGGAGTTAA

+ TTCGATAGAA GTAATGAAGC ATCAGTACTG AAATACACAA AAGAATTAGA CATTCACCAG TGTTGACATC
- AAGCTATCTT CATTACTTCG TAGTCATGAC TTTATGTGTT TTCTTAATCT GTAAGTGGTC ACAACTGTAG

+ AAGGATATCG CAATAGGTAG GATCACGGAA GATCCTCACT ACTTTATAGC TACCCTTAAC ATTGTCTCTA
- TTCCTATAGC GTTATCCATC CTAGTGCCTT CTAGGAGTGA TGAAATATCG ATGGGAATTG TAACAGAGAT

+ CCGAATCCCA TCGCACAATA TCCCGGGAAA TACGTCGTCC ATGATTCTAA ATATATATTT CAGATCATAT
- GGCTTAGGGT AGCGTGTTAT AGGGCCCTTT ATGCAGCAGG TACTAAGATT TATATATAAA GTCTAGTATA

+ ACAGTATGCA TGATCCTACA AAAAAAAGGA CGACTCAATA AGAAAATTAA TTATCCCCAA CAAATCAAAT
- TGTCATACGT ACTAGGATGT TTTTTTTCCT GCTGAGTTAT TCTTTTAATT AATAGGGGTT GTTTAGTTTA

+ TATATTGATA ATTAAGGGAG CATATGCGTT ATTATAAATT ACAACCTAAC CGACTTTACC GACCCAAAGC
- ATATAACTAT TAATTCCCTC GTATACGCAA TAATATTTAA TGTTGGATTG GATGAAATGG CTGGGTTTCG

+ CTAAAACAAA GCTGACATTC TTTTGTTTTC CCTCTCTCTA GAAGTTCCTG TTTTCATCTT CTTCTCTCTC
- GATTTTGTTT CGACTGTAAG AAAACAAAAG GGAGAGAGAT CTTCAAGGAC AAAAGTAGAA GAAGAGAGAG

+ TAGCAATTCA ATCTCGACTT CGCTTTTCTG GCTTTTCTGG GAATTTTGAT TTATTTGCCC CCATCGTTCG
- ATCGTTAAGT TAGAGCTGAA GCGAAAAGAC CGAAAAGACC CTTAAAACTA AATAAACGGG GGTAGCAAGC

+ CTTGAATTAC GAATCAGATA TCGCCGATCG TCGTAGATCT GCGTTGATAT TGTTCGCTTG ATTTATTTTT
- GAACTTAATG CTTAGTCTAT AGCGGCTAGC AGCATCTAGA CGCAACTATA ACAAGCGAAC TAAATAAAAA

+ GTTGTTGTTC TTTTTAATAT TTCCTTACAC CTATTCACCC AAAAAAAATA TATTTCCGTA CACCTCTTCT
- CAACAACAAG AAAAATTATA AAGGAATGTG GATAAGTGGG TTTTTTTTAT ATAAAGGCAT GTGGAGAAGA

+ AGAAGGATTT CAAGATTACC TGAATATTTA CGTATATTTA TCCACATCCG TACACATTGT ACGTAGTTAT
- TCTTCCTAAA GTTCTAATGG ACTTATAAAT GCATATAAAT AGGTGTAGGC ATGTGTAACA TGCATCAATA

+ ACGATAAACT AGGGCTTTAG ACAAACGAAG A
- TGCTATTTGA TCCCGAAATC TGTTTGCTTC T
```

+		
+		ABRE
+		ABRE2
+		AT~TATA-box
+		Box 4
+		Box II
+		CAAT-box
+		CAT-box
+		G-Box
+		G-box
+		GA-motif
+		GATA-motif
+		MRE
+		MYB
+		MYC
+		Myb
+		Myb-binding site
+		O2-site
+		TATA-box
+		TCA
+		TCT-motif
+		TGA-element
+		Unnamed__1
+		Unnamed__2
+		Unnamed__4
+		W box
+		WUN-motif
+		circadian

Appendix.3.1.3 Promoter sequence of *DRIF3* showing *cis*-acting obtained from PlantCARE. Color code for motifs is presented at the end.

```
+ ACACTCTGTA ACCACTATGT GTGAGTTCGA TCCAACAAGT GGGTTTAGAT TAGTATTAAC TAGGAATGTT
- TGTGAGACAT TGGTGATACA CACTCAAGCT AGGTTGTTCA CCCAAATCTA ATCATAATTG ATCCTTACAA

+ TAATAGTACT GCTAAATGAA TATTTAGGTA AAACTATTTT GTATATAAAT TCCCACTCAG ATGTCATGGG
- ATTATCATGA CGATTTACTT ATAAATCCAT TTTGATAAAA CATATATTTA AGGGTGAGTC TACAGTACCC

+ TCACTCTGAC GCCTACCACT ATACTAATAG AAGTAAATTT TTCTTTTTTC GATGTTTAGT TACAATCGGA
- AGTGAGACTG CGGATGGTGA TATGATTATC TTCATTTAAA AAGAAAAAAG CTACAAATCA ATGTTAGCCT

+ AAATGAATCT AAATTTTAAC TAATATTAAT AATATGCCAA GATCTCCATC CAAATCTAGA TTGATATCTA
- TTTACTTAGA TTTAAAATTG ATTATAATTA TTATACGGTT CTAGAGGTAG GTTTAGATCT AACTATAGAT

+ CATATTGATG CCTTGATATC TCTAAACTCT AAAGATAAAC TTAAAGGCAA GTGATTAATT GTTAACTAAA
- GTATAACTAC GGAACTATAG AGATTTGAGA TTTCTATTTG AATTTCCGTT CACTAATTAA CAATTGATTT

+ ATTAAAAATT AACAATACAA TTTTAATTCG CTAGTTATAA AAATTTTATT ACCAAGAAGA ATATGGAAAA
- TAATTTTTAA TTGTTATGTT AAAATTAAGC GATCAATATT TTTAAAATAA TGGTTCTTCT TATACCTTTT

+ AAAGGAATTG GAGGCAAAAA AAAAAAAAAA AAAAAAAAAA GAGGCCAAAA GAAAAGAAAG CAAAGGAGTA
- TTTCCTTAAC CTCCGTTTTT TTTTTTTTTT TTTTTTTTTT CTCCGGTTTT CTTTTCTTTC GTTTCCTCAT

+ TGAATACAAC AAAGAGGTTC TGCAAGAAAC AAACAACACA GCCAATTACT ACTAAACAAA CACACTCTCT
- ACTTATGTTG TTTCTCCAAG ACGTTCTTTG TTTGTTGTGT CGGTTAATGA TGATTTGTTT GTGTGAGAGA

+ CTTTCATTCT TTCTCTCTCT CTCTCTCTCG ATTTTCTCCC ATTGGAGTTC GCAATTAGCT GCCACTCTCT
- GAAAGTAAGA AAGAGAGAGA GAGAGAGAGC TAAAAGAGGG TAACCTCAAG CGTTAATCGA CGGTGAGAGA

+ GATCGATCTT CTTCTCACAT CACATCCTTT ATAATCTTCA GTTCTACTGC TCGGAGGTGT TTGCTTGCGT
- CTAGCTAGAA GAAGAGTGTA GTGTAGGAAA TATTAGAAGT CAAGATGACG AGCCTCCACA AACGAACGCA

+ TCTCTCTTGT CGGTCACATC CCTAATTACC ATCACTGCCC CTGATTAAAG TTTTACTCTA TCTTCTTTGT
- AGAGAGAACA GCCAGTGTAG GGATTAATGG TAGTGACGGG GACTAATTTC AAAATGAGAT AGAAGAAACA

+ TGATGCCAAA TCTGAGGTGG GTTTCTCATC TTTTTATTCT TTCAGCTGTG TTTTGTTTCA TTTATTTGTT
- ACTACGGTTT AGACTCCACC CAAAGAGTAG AAAAATAAGA AAGTCGACAC AAAACAAAGT AAATAAACAA

+ TCTGGATTTG CTAAAACCCT AAATTTGATA TCTTTGGGGG CATCATCGTA GCTCAAATCA GCTTCAATTA
- AGACCTAAAC GATTTTGGGA TTTAAACTAT AGAAACCCCC GTAGTAGCAT CGAGTTTAGT CGAAGTTAAT

+ TCAAATCTCC ACG
- AGTTTAGAGG TGC
```

+		+	MYB
+	AAGAA-motif	+	MYB-like sequence
+	AE-box	+	O2-site
+	ATCT-motif	+	STRE
+	AT~TATA-box	+	TATA-box
+	Box 4	+	TCA-element
+	Box II	+	TGACG-motif
+	CAAT-box	+	Unnamed__1
+	CAT-box	+	Unnamed__4
+	CGTCA-motif	+	WRE3

Appendix.3.1.4 Promoter sequence of *DRIF4* showing *cis*-acting obtained from PlantCARE. Color code for motifs is presented at the end.

```
+ AAAACCATTA CAAAACAAAA CCACCATTTT CTCTCTCTCT CTCCTCGCCA GCTTCTGCAA TTCGAAAGTT
- TTTTGGTAAT GTTTTGTTTT GGTGGTAAAA GAGAGAGAGA GAGGAGCGGT CGAAGACGTT AAGCTTTCAA

+ CTTCGCATCA TCTTGCTGTC ACCACTCTCT CGCGAATCGA ATCACGATTC TTCAATCTTC TTCTTCTCTA
- GAAGCGTAGT AGAACGACAG TGGTGAGAGA GCGCTTAGCT TAGTGCTAAG AAGTTAGAAG AAGAAGAGAT

+ TTGCGATCTC TCTGAATTTT GTAGCTTCAC GATTTCAGTA TCTCTACTTT CATCACCAAT CGTTGGGGGT
- AACGCTAGAG AGACTTAAAA CATCGAAGTG CTAAAGTCAT AGAGATGAAA GTAGTGGTTA GCAACCCCCA

+ TTAGGGTTTT TGCATCGACG CCATCTATCG AGCTACGGAG AAGCATACTC ATCTTCTATT TTGGCTAAAT
- AATCCCAAAA ACGTAGCTGC GGTAGATAGC TCGATGCCTC TTCGTATGAG TAGAAGATAA AACCGATTTA

+ CCGAGGTGGG TTTTGCTCAG GTTTTGCTCT TTTAGGTTTC TGTTGGGTTT TAATCGGATT TTTTTCCGTA
- GGCTCCACCC AAAACGAGTC CAAAACGAGA AAATCCAAAG ACAACCCAAA ATTAGCCTAA AAAAAGGCAT

+ TGATTCTGAA ATTGCTAAAA CCCTAAATTG AACATGTTTT TGGGCGTGCT TCGTTGTAGG TCAACTTAAT
- ACTAAGACTT TAACGATTTT GGGATTTAAC TTGTACAAAA ACCCGCACGA AGCAACATCC AGTTGAATTA

+ CTCTACGAAA TAGCTTTTGA GT
- GAGATGCTTT ATCGAAAACT CA
```

+		
+		ARE
+		CAAT-box
+		GARE-motif
+		MRE
+		MYB
+		Myb-binding site
+		TATA-box
+		Unnamed__4
+		W box
+		WRE3

Appendix.3.1.5 Promoter sequence of *DRIF5* showing *cis*-acting obtained from PlantCARE. Color code for motifs is presented at the end.

```
+ CTCCACATTT CACATATTCT CTCCATAATT TCTTATTCCT TTCTAGTTCC TTAATTACCT CACACTCATT
- GAGGTGTAAA GTGTATAAGA GAGGTATTAA AGAATAAGGA AAGATCAAGG AATTAATGGA GTGTGAGTAA

+ TATTTATCGA TTTATATATA GGGAGAGACA GAGAGAGAGA AAGAGAGAGA TAGGCTAGCT AAAATGAGTT
- ATAAATAGCT AAATATATAT CCCTCTCTGT CTCTCTCTCT TTCTCTCTCT ATCCGATCGA TTTTACTCAA

+ AAGGAATCCT GACTTATCCT GAGCCTCTGA GCTTTATTTA TCACACCAAT ACTTCTCTTT TACTAGCTGT
- TTCCTTAGGA CTGAATAGGA CTCGGAGACT CGAAATAAAT AGTGTGGTTA TGAAGAGAAA ATGATCGACA

+ TTTCCACTAA ATTTCAGTAC CATAAATGAG TTATGCGATA TAAGTAGATG CTAAATTCAA ATTCAGGAAT
- AAAGGTGATT TAAAGTCATG GTATTTACTC AATACGCTAT ATTCATCTAC GATTTAAGTT TAAGTCCTTA

+ ATTATTATAG ATATGTCATA TATAGAGTCT AATGAAAAGT GCTTCTAAAT TTTTTAATAC TCTAAAACAC
- TAATAATATC TATACAGTAT ATATCTCAGA TTACTTTTCA CGAAGATTTA AAAAATTATG AGATTTTGTG

+ TAAACTTATC CAAAACACTT TAGAAATCTA CTCTCTGGTC ACAAACGGGT GGTGACAGAT AATAGATTCT
- ATTTGAATAG GTTTTGTGAA ATCTTTAGAT GAGAGACCAG TGTTTGCCCA CCACTGTCTA TTATCTAAGA

+ CAAACATTTC GTATAATAAG GTTTTAAACT TTTAATAGAT ATATAAGTTT CATGAAAAAA AAAAAAAGTT
- GTTTGTAAAG CATATTATTC CAAAATTTGA AAATTATCTA TATATTCAAA GTACTTTTTT TTTTTTTCAA

+ TCATGAATAT TAACCAATGA TGGTAGTTGC ATTTATTTAT TTTTGGCTAG TGCATTCACA TAACATAACG
- AGTACTTATA ATTGGTTACT ACCATCAACG TAAATAAATA AAAACCGATC ACGTAAGTGT ATTGTATTGC

+ AATGCATTTT AAGAATAATG TGATAAGAAC ATTCTAAATT TCCTGACCAA AAAAATAAAT TTAAGAACTT
- TTACGTAAAA TTCTTATTAC ACTATTCTTG TAAGATTTAA AGGACTGGTT TTTTTATTTA AATTCTTGAA

+ CTTGAAATAA ATTACTAGAC CTCCACGGAT ATGAACTGTT TTGCACTCCA CTTTACGTAG ATAGTGTACG
- GAACTTTATT TAATGATCTG GAGGTGCCTA TACTTGACAA AACGTGAGGT GAAATGCATC TATCACATGC

+ TGTAAGTGGG TGATTTTGTT ACAAAACTAG TTTAAAGATA GAGAAAAAAA AAGTATAAAT TGTCCAAAAA
- ACATTCACCC ACTAAAACAA TGTTTTGATC AAATTTCTAT CTCTTTTTTT TTCATATTTA ACAGGTTTTT

+ GTAGTTCACC GTTCACGTTT ATATACAAAT CGGAATGCAT AAGAACAAAA AGAGAATCCA GCAACGTGTT
- CATCAAGTGG CAAGTGCAAA TATATGTTTA GCCTTACGTA TTCTTGTTTT TCTCTTAGGT CGTTGCACAA

+ CAGATGCTGT TGACACGTGC GATGCATGCT TTCTGCTTGT AGCCCATCAA TAGTTTTTAT TTTTCCAAAA
- GTCTACGACA ACTGTGCACG CTACGTACGA AAGACGAACA TCGGGTAGTT ATCAAAAATA AAAAGGTTTT

+ TACGAAGTGT TTCGAATTGT TTAGTGTCTG TCGCAAGTAG TTTCATGTAG GCATGCATGG TTTATGAAAA
- ATGCTTCACA AAGCTTAACA AATCACAGAC AGCGTTCATC AAAGTACATC CGTACGTACC AAATACTTTT

+ TATGCTTCAC CTTGGCCGTT GATTTTTTAT CTAGCTTTAA GGAATCTCAT CTTTATTTCT ATTCACGAAC
- ATACGAAGTG GAACCGGCAA CTAAAAAATA GATCGAAATT CCTTAGAGTA GAAATAAAGA TAAGTGCTTG
```

```
+ ATAGTTCTAG TAACGTACGT CAGTTGTGTG TTATAACTCA GTATACACCG TTACTTTAAA GTTATACAGA
- TATCAAGATC ATTGCATGCA GTCAACACAC AATATTGAGT CATATGTGGC AATGAAATTT CAATATGTCT

+ TGAATTATCT CTGACAATTT TACACTTTTG AGATTTTCTT TAGTGGAACA AACTTTAATT ATATATAGCT
- ACTTAATAGA GACTGTTAAA ATGTGAAAAC TCTAAAAGAA ATCACCTTGT TTGAAATTAA TATATATCGA

+ TCTAATTTTA TTTTATTTTC TAATTTCTAA ACATATTACG AAAAATTAAT ATGAGAGAGT GGTTTCAGCA
- AGATTAAAAT AAAATAAAAG ATTAAAGATT TGTATAATGC TTTTTAATTA TACTCTCTCA CCAAAGTCGT

+ AAATTAGCTA GGACTTACTT ACATGATGAG GACTATATAT ATGAGAAAAA AACTGTGAAT TAAGGTTGGG
- TTTAATCGAT CCTGAATGAA TGTACTACTC CTGATATATA TACTCTTTTT TTGACACTTA ATTCCAACCC

+ ATTAAAGTCT TAAAGAAACC AAAAGCAGGG ATAATGTAGA AATTTGGAGA TGCTAAGGCC ACAAGTTAAT
- TAATTTCAGA ATTTCTTTGG TTTTCGTCCC TATTACATCT TTAAACCTCT ACGATTCCGG TGTTCAATTA

+ AGTTGAGACT TGATAGATAT GATAACTTAT TTGTTTTGTT TTTTGTTTTT TTTGTAGTTA ATAACATTGT
- TCAACTCTGA ACTATCTATA CTATTGAATA AACAAAACAA AAAACAAAAA AAACATCAAT TATTGTAACA

+ AAATCAGTAA AACCAAAATC TAGTTCCTGC CAACAATTTT GAAATTAAAA TGTATGGATC ACCTTAATGA
- TTTAGTCATT TTGGTTTTAG ATCAAGGACG GTTGTTAAAA CTTTAATTTT ACATACCTAG TGGAATTACT

+ GCGAAAGAGT GCATCAAGTC CAACAACAAA GTCTTTCTCT TCTTCATAAC ATAATGTTCG GAGTTAAATT
- CGCTTTCTCA CGTAGTTCAG GTTGTTGTTT CAGAAAGAGA AGAAGTATTG TATTACAAGC CTCAATTTAA

+ AACTTGAATG AAACAAATCT TTAATTAATT ATATTGATTT GCATTATTAG GAAACTGTTC AGATTAGCCT
- TTGAACTTAC TTTGTTTAGA AATTAATTAA TATAACTAAA CGTAATAATC CTTTGACAAG TCTAATCGGA

+ ACCAAAAAAA AGGAAACTGT ATTACTGTTA GTGGAGCAAG TATAATGTGC AGTTTATTGG GCTCATAGCA
- TGGTTTTTTT TCCTTTGACA TAATGACAAT CACCTCGTTC ATATTACACG TCAAATAACC CGAGTATCGT

+ TTTTAGAGCC CAAATTTATT CGTTTCTTAA GCAAAATCTT AAATGCGTAT AAAATTAGTC TTTTACAAAT
- AAAATCTCGG GTTTAAATAA GCAAAGAATT CGTTTTAGAA TTTACGCATA TTTTAATCAG AAAATGTTTA

+ ATAGCATGTG GCCCAATATA AACGTCAGGC CTGATAATAT GTTTAAGTGT CCATGACTTT TAACATGTTT
- TATCGTACAC CGGGTTATAT TTGCAGTCCG GACTATTATA CAAATTCACA GGTACTGAAA ATTGTACAAA

+ CATTTGATAA GCATTTTTTA CCAGAAGCTT ATGTTAATTC ACTTATACGG GTTGGGGATA ATGAAAATTG
- GTAAACTATT CGTAAAAAAT GGTCTTCGAA TACAATTAAG TGAATATGCC CAACCCCTAT TACTTTTAAC

+ AATTTGGACT AATAGAGCAT ACATCAATCG AGGAAAATTG CAACGACAAA GAAACGACGG TACATCAATT
- TTAAACCTGA TTATCTCGTA TGTAGTTAGC TCCTTTTAAC GTTGCTGTTT CTTTGCTGCC ATGTAGTTAA

+ TTACCTTAGT AAAATCTTGA TAAGATTTGA AGTGTGAGTT AATAACAAAT TTTAAAAAAC CTAACTTATC
- AATGGAATCA TTTTAGAACT ATTCTAAACT TCACACTCAA TTATTGTTTA AAATTTTTTG GATTGAATAG

+ TAGTGAAAAA AAATGAAAAT TATAAAAGCT GATGGCATTT AGATAGTAAA CAGAAATCAA AATTGTTAAG
- ATCACTTTTT TTTACTTTTA ATATTTTCGA CTACCGTAAA TCTATCATTT GTCTTTAGTT TTAACAATTC

+ TGATAGTAGT CTTGTTGCAA ATGATAAGTC CACTGAGAAC ATGACATCGA GGATCGAGTC TCGCTCCTTA
- ACTATCATCA GAACAACGTT TACTATTCAG GTGACTCTTG TACTGTAGCT CCTAGCTCAG AGCGAGGAAT

+ TAAATGTTGA CAAAAAAAAA CAGAAATCAA AACATCATTT TGGTGTATGA ATATATTTCT TGATCTATTT
- ATTTACAACT GTTTTTTTTT GTCTTTAGTT TTGTAGTAAA ACCACATACT TATATAAAGA ACTAGATAAA

+ AGTGTATCAA TAATTTCTTA AATATCAAAA AAAAAAAAAA AAACTCTTCA TCGATTTTTA TATGCTTGCT
- TCACATAGTT ATTAAAGAAT TTATAGTTTT TTTTTTTTTT TTTGAGAAGT AGCTAAAAAT ATACGAACGA

+ TCTGTTATTG GTTCATAACT ATGAGTCCAA TTATTTTCAT AGTCGCATAT AATTGTAAAA GTATTGTACG
- AGACAATAAC CAAGTATTGA TACTCAGGTT AATAAAAGTA TCAGCGTATA TTAACATTTT CATAACATGC
```

```
+ AACGGATCAG TTTATCCCAC AATTATTGAG ATGCTAAGAT TAATTTTTGT ATTAACATAA AGAAAAAATA
- TTGCCTAGTC AAATAGGGTG TTAATAACTC TACGATTCTA ATTAAAAACA TAATTGTATT TCTTTTTTAT

+ TGAATCAAAT TCATATATCT TGAGATTTTG ACAACGAGGA TGTCACACAT TGACATTATT AGAAAACAAA
- ACTTAGTTTA AGTATATAGA ACTCTAAAAC TGTTGCTCCT ACAGTGTGTA ACTGTAATAA TCTTTTGTTT

+ AAAAAAAAAA AAAAATCTTG TCTACTTGAT AAAACTTTTT GGAATAAAAA AAGTGAGGTA GAAGAAAGAA
- TTTTTTTTTT TTTTTAGAAC AGATGAACTA TTTTGAAAAA CCTTATTTTT TTCACTCCAT CTTCTTTCTT

+ AGGCATCAAA TGCTTTATTA TTGGGATTCG CACTATTTAC TAAATTCCAC GATCTGCCAA CCATCTCTTT
- TCCGTAGTTT ACGAAATAAT AACCCTAAGC GTGATAAATG ATTTAAGGTG CTAGACGGTT GGTAGAGAAA

+ TATTTATTTA CCTATTTTCT TATTTTTTTC TACTAATTTT GGTAGTAAAA AAAAAAAAAT GAATATGTTC
- ATAAATAAAT GGATAAAAGA ATAAAAAAAG ATGATTAAAA CCATCATTTT TTTTTTTTTA CTTATACAAG

+ TTAGTCATAT CTCTCCCTTC ATTTAGGAAG AGAATCTTTT CTAAAATTCC CCATTCCTAC TTACTTTTAT
- AATCAGTATA GAGAGGGAAG TAAATCCTTC TCTTAGAAAA GATTTTAAGG GGTAAGGATG AATGAAAATA

+ AGTCAAAATC CATTTTTTTC CATTTCAACT AATTGTGTTG GAAAATAATT AATATTAGAT TTCAATTAAG
- TCAGTTTTAG GTAAAAAAAG GTAAAGTTGA TTAACACAAC CTTTTATTAA TTATAATCTA AAGTTAATTC

+ TTTATCATCA TATGGTGAAA AGTCATAAAG TAATGTATTA ATAATTTTTT CATTATTTGT CTAAAAATCC
- AAATAGTAGT ATACCACTTT TCAGTATTTC ATTACATAAT TATTAAAAAA GTAATAAACA GATTTTTAGG

+ TTCATTGGCT TAT
- AAGTAACCGA ATAT
```

+		+	Myb-binding site		
+	AAGAA-motif	+	NON		
+	ABRE	+	RY-element		
+	ABRE3a	+	TATA		
+	ABRE4	+	TATA-box		
+	ARE	+	TATC-box		
+	AT-rich element	+	TC-rich repeats	+	GATA-motif
+	AT~TATA-box	+	TCA	+	GT1-motif
+	Box 4	+	TCA-element	+	HD-Zip 1
+	CAAT-box	+	TCCC-motif	+	MBS
+	CCAAT-box	+	TGA-element	+	MRE
+	CGTCA-motif	+	TGACG-motif	+	MYB
+	CTAG-motif	+	Unnamed__1	+	MYB recognition site
+	ERE	+	Unnamed__2	+	MYB-like sequence
+	G-Box	+	Unnamed__4	+	MYC
+	G-box	+	WUN-motif	+	Myb
		+	as-1		

CHAPTER 4
"The MYB-like encoding genes, *DRIF3*, *DRIF4*, and *DRIF5* are involved in seed germination in *Arabidopsis thaliana*"

The MYB-like encoding genes, *DRIF3*, *DRIF4*, and *DRIF5* are involved in seed germination in *Arabidopsis thaliana*

Shweta Singh, Rómulo Sobral, Sara Larangeira, João Raimundo, M. Manuela R. Costa

Biosystems and Integrative Sciences Institute (BioISI), Plant Functional Biology Center, University of Minho, Campus de Gualtar, 4710-057, Braga, Portugal

4.1 Abstract

Seed dormancy is an adaptive trait of plants that ultimately affects crop production and its yield. Due to the involvement of a complex network of transcription factors and cross-talk with several phytohormones, the mechanism of dormancy acquirement in fresh seeds and its release with time is not completely understood. The timing of germination is coordinated by two main phytohormones, Abscisic acid (ABA) and Gibberellic acid (GA). The biosynthesis and signaling of these hormones are regulated by various transcription factors. Members of the MYB family of transcription factors play an important role in seed germination and abiotic stress tolerance by modulating ABA signaling. Here, we report the involvement of three *DIV-and-RAD-interacting-factors* (*DRIF*), *DRIF3*, *DRIF4*, and *DRIF5* in the maintenance of seed dormancy and delay in germination. The *drif345* triple mutant displayed early germination in normal as well as in salt and heat stress conditions. The analysis of the expression of ABA and GA-related genes in the tripled mutant revealed increased expression of *GA3ox1* and lower expression of *ABA1* and *ABA3* genes. Taken together, our results suggest that these three MYB-like transcription factors, *DRIF3*, *DRIF4*, and *DRIF5* are involved in the maintenance of seed dormancy and delay of germination by a possible regulation of the ABA/GA ratio.

4.2 Introduction

Seed dormancy is considered an adaptive trait that synchronizes the onset of the vegetative and reproductive phases of plants with seasonal changes in the environment. It is more prominent in wild plants enabling them to survive under environmental stress, while most domesticated crops show decrease levels of dormancy, which ensures high levels of germination. Induction of dormancy during seed maturation and its release in an appropriate environment requires a fine-tuning of a complex gene network and involvement of various phytohormones (reviewed in Penfield and King, 2009; reviewed in Shu *et al.*, 2018). The inappropriate loss of seed dormancy causes germination of fresh seeds and sometimes pre-harvest sprouting leads to substantial loss of yield and quality of crops(Fang and Chu, 2008; Shu *et al.*, 2016). As seed dormancy and its timely release decide the fate of plant survival and productivity, it is important to investigate this mechanism more deeply.

Freshly harvested seeds show a high level of dormancy that can be broken by stratification or 'after-ripening' of seeds (Vleeshouwers *et al.*, 1995; Bewley, 1997; Lee *et al.*, 2015). After the release of seed dormancy, the germination takes place in response to various environmental factors, such as light, moisture and temperature integrated into internal developmental programs (Penfield *et al.*, 2005; Penfield and Hall, 2009b; Chen *et al.*, 2014). The phytohormones, abscisic acid (ABA) and gibberellins (GAs) are two important hormones, involved in this internal developmental program of seed germination, playing crucial roles in the transition from seed dormancy to germination by antagonizing the effect of each other. ABA positively regulates the induction and maintenance of dormancy, while GA is known for releasing dormancy and promoting seed germination (Penfield and King, 2009; Penfield and Hall, 2009a; Ge and Steber, 2018; reviewed in Shu *et al.*, 2018). Therefore, changes in the ratio of ABA/GA and components of their signaling pathways constitute a central regulatory mechanism underlying the transition from seed dormancy to germination.

The pathways that regulate ABA synthesis and signaling are important in the control of germination (Gubler *et al.*, 2005; Finkelstein *et al.*, 2008; Shu *et al.*, 2013). ABA is a C-15 organic acid synthesized from the cleavage of a precursor C-40, xanthophyll carotenoid. The first committed step of ABA biosynthesis is the oxidative cleavage of epoxycarotenoid 9-*cis* - neoxanthin by the *9 -cis* - epoxycarotenoid deoxygenase (NCED), yields a C-15 intermediate (Schwartz *et al.*, 1997). There are five *NCED* genes reported in *Arabidopsis thaliana*, *NCED2*, *NCED3*, *NCED5*, *NCED6*, and *NCED9,* have been shown to be involved in ABA biosynthesis and its regulation in response to various abiotic stresses (Iuchi *et al.*, 2001; Lefebvre *et al.*, 2006; Frey *et al.*, 2012). Further conversion of xanthoxin to ABA aldehyde is carried out by short-chain alcohol dehydrogenase/reductase, encoded by *ABA2* (González-

Guzmán et al., 2002). Another gene that is important for ABA biosynthesis is *ABA3* that encodes a Molybdenum cofactor (MoCo) synthase, required by aldehyde oxidase and xanthine dehydrogenase activity (Bittner *et al.*, 2001). In plants, ABA content is determined by the precise balance between ABA biosynthesis and its catabolism. The member of *CYP707A* family genes (*CYP707A1-CYP707A4*) synthesizing ABA 8′-hydroxylases are known to be involved in the inactivation of ABA (Saito *et al.*, 2004).

Mature dry seeds of *A. thaliana* contain high ABA content, which slowly decreases after seed ripening and imbibition (Ali-Rachedi *et al.*, 2004). ABA regulation during seed germination requires five ABA-responsive transcription factors controlling the plant sensitivity to ABA; *ABA INSENSITIVE 1* (*ABI1*), *ABI2*, *ABI3, ABI4,*and *ABI5,* out of which *ABI1* and *ABI2* are negative regulators while *ABI3*, *ABI4,* and *ABI5* are positive regulators of the ABA signaling pathway (Shu *et al.*, 2013). The *abi3* mutant shows reduced dormancy (Koornneef *et al.*, 1989) and it can rescue the non-germinating phenotype of *GA-REQUIRING 1* (*ga1*) mutant showing GA-deficiency (Nambara *et al.*, 1992). *ABI4* is responsible for primary seed dormancy and negatively regulates the expression of *CYP707A1* and *CYP707A2* by directly binding to their promoter in freshly harvested seeds of *Arabidopsis* (Penfield *et al.*, 2006b; Shu *et al.*, 2013). The increase in the expression of *CYP707A1* and *CYP707A2* genes in *abi4* mutant ultimately decreases the level of ABA and causes reduced dormancy. The expression of *ABI5* have been reported in various vegetative tissues and is regulated by endogenous ABA levels (Lopez-Molina *et al.*, 2001) and various environmental stresses (Brocard *et al.*, 2002).

In contrast to ABA, GA inhibits seed dormancy and promotes seed germination. Biosynthesis of GA is divided into three steps : 1. biosynthesis of *ent*-kaurene, 2. conversion of *ent*- kaurene into GA12 (non-bioactive GA), 3. formation of bioactive GA in the cytoplasm. Synthesis of bioactive GA in cytoplasm is the most important step that includes a series of oxidation. Two important enzymes encoded by *GA20 oxidase* (*GA20x*) and *GA3ox* involved in the biosynthesis of active GA while *GA2ox* encodes an enzyme involved in the conversion bioactive GA into inactive form (reviewed in Sun, 2008).

Mutants defective in GA biosynthesis, cannot germinate without exogenous application of GA (Koornneef *et al.*, 1982). GA regulates seed germination by activating GA responsive genes and destabilizing the growth repressors DELLA proteins (Peng *et al.*, 1997; Penfield *et al.*, 2006a; Achard and Genschik, 2009). There are five gene homologs that code for DELLA proteins in *Arabidopsis*; *GA INSENSITIVE* (*GAI*), *REPRESSOR OF GA* (*RGA*), *REPRESSOR OF GA LIKE-1* (*RGL1*), *RGL2* and *RGL3*. Among these five *DELLAs*, *RGL2* is involved in seed germination by repressing the GA signaling cascade

(Lee *et al.*, 2002), along with *GAI* and *RGA* (Silverstone *et al.*, 1998). The integration of the GA and ABA biosynthesis and signaling pathways is a key to balance the GA/ABA ratio in seeds (Seo *et al.*, 2006).

During stress conditions, seed shows delayed germination. It has been reported that during the stress the expression of *GA20ox* and *GA3ox* are down-regulated, while the expression of *GA2ox* is up-regulated and reduces the overall synthesis of the GA (reviewed in Colebrook *et al.*, 2014). On the other hand, expression of ABA synthesis and signaling genes increases during environmental stress (Iuchi *et al.*, 2001; Brocard *et al.*, 2002).

Various transcription factors are also involved in the release of seed dormancy and initiation of seed germination. The transcriptional regulation of these gene networks by the external stimulus (light, temperature, moisture) and internal signals (phytohormones) is required to break the seed dormancy and initiate germination (Nelson *et al.*, 2017). Phytochrome-interacting factors (PIFs), a small group of bHLH transcription factor family, are known to be involved in regulating various developmental processes including seed germination, in response to light (Leivar and Monte, 2014; Pham *et al.*, 2018). The mutant of *pif1* shows germination in far-red light followed by dark (Oh *et al.*, 2004; Penfield *et al.*, 2005). PIFs compete with GA receptors for DELLA binding sites and hence stabilize DELLAs in the presence of GA (de Lucas *et al.*, 2008). In this way PIFs can stabilize DELLA proteins in the presence of GA and modify the GA signaling in seeds (Oh *et al.*, 2007). PIFs are also required for the normal induction of ABA and GA related genes expression associated with germination in response to light (Penfield *et al.*, 2005; Oh *et al.*, 2006).

Other group of transcription factors that is known to be involved in seed germination in response to ABA is the MYB family transcription factors. The MYB family is one of the largest in plants and these proteins are involved in various developmental processes and environmental responses (reviewed in Dubos *et al.*, 2010). In *Arabidopsis*, two MYB genes, *MYBS1* and *MYBS2* are involved in germination in response to sugar (Chen *et al.*, 2017). There are several other MYB homologs in *Arabidopsis*, involved in ABA synthesis or signaling and mediates various abiotic stresses (Lee *et al.*, 2015; Wang *et al.*, 2015; Yu *et al.*, 2016; Yang *et al.*, 2018). These MYB genes showed involvement in ABA biosynthesis or signaling to mediate plant tolerance to salt stress. *DIV*-and-*RAD-INTERACTION FACTORS* (*DRIFs*) are MYB-family transcription factors, which were identified as a common interactor for DIVARICATA (DIV) and RADIALIS (RAD) in *Antirrhinum majus* (Raimundo *et al.*, 2013). *DIV* and *RAD* are also MYB-family transcription factors and are involved in the establishment of dorsoventral floral asymmetry of *A. majus* (Galego and Almeida, 2002; Corley et al. 2005). RAD antagonizes the action of DIV by binding to the DRIFs in the dorsal region of flower and disrupting the formation of DIV-DRIF

complex (Raimundo et al., 2013). The homologs of *RAD*, *DIV* and *DRIFs* are also found in tomato where they regulate pericarp expansion that shows the co-option of these genes for different functions in other species (Machemer *et al.*, 2011). In *Arabidopsis*, there are several *DIV*, *RAD* and *DRIF* homologs (Raimundo *et al.*, 2018). According to available microarray and RNASeq databases, some of these genes are highly expressed in seeds and seedlings (Winter *et al.*, 2007; Klepikova *et al.*, 2016). Recently, it was also reported that in *Arabidopsis, div2* mutant shows reduced germination while *RAD Like-2* (RL2; also named as *RSM1*) affects germination by directly binding to the promoter of *ABI5* enhancing its expression (Fang et al., 2018; Yang et al., 2018). In those studies, it was also reported that the DIV and RAD homologs regulate salt stress tolerance by modulating ABA signaling.

There are five *DRIF* homologs present in *A. thaliana*, *DRIF1*, *DRIF2*, *DRIF3*, *DRIF4*, and *DRIF5*. Phylogenetic study showed that *DRIF3*, *DRIF4*, and *DRIF5* appeared in same clade along with tomato *DRIF* homolog *FSB1* while *DRIF1* and *DRIF2* appeared in different clade along with *AmDRIFs* (Raimundo *et al.*, 2013). The transcriptome data showed the expression of *DRIF3*, *DRIF4*, and *DRIF5* during all stages of plant development. In this study, we characterized the role of these three *DRIFs* in the germination of seeds. We found that *drif345* mutant germinates earlier than WT in normal conditions. We also report that these *DRIFs* are negative regulators of germination in salt and heat stress and in the absence of these genes the germination rate and total percentage of germination of seeds of *drif345* were less affected. Reduced expression of ABA synthesizing genes, *ABA1* and *ABA3* and higher expression of GA synthesizing gene, *GA3ox1* in *drif345* mutant seeds suggests the faster germination of *drif345* mutant in normal as well as in stress conditions may be because of the low level of the ABA/GA ratio in the seeds of mutant. Therefore, we propose that *DRIF3, DRIF4,* and *DRIF5* in *Arabidopsis* are likely involved in the maintenance of ABA/GA ratio during the transition from seed dormancy to germination.

4.3 Material and Methods

4.3.1 Plant Material and growth conditions

Wild type *Arabidopsis thaliana* ecotype Columbia (Col-0) and *drif3* (SAIL_636_A08), *drif4* (SALK_204381C) and *drif5* (SALK_121145C) mutants were used in this study. The T-DNA insertion lines were obtained from the Nottingham Arabidopsis Stock Center (NASC, UK) and double and triple mutants for *drifs* were generated by multiple crossings. Homozygous mutants were selected by PCR and primers used for screening the insertions are mentioned in table.4.1. For germination studies, seeds of the *drif345* and WT grew in soil under a long-day (LD) photoperiod with a regime of 16 h light

at 22 °C and 8 h dark at 18 °C temperature. Synchronized seeds from WT and *drif* mutants were harvested and stored under dark for two to four months (after-ripening period) before seed germination assays.

4.3.2 Seed germination and stress assays

For the seed germination assay, around 50-100 seeds of *drif* mutants and WT were surface sterilized and plated on half-strength Murashige and Skoog (1/2 MS) salt medium containing 0.8% (w/v) agar (pH=5.7). After 4 days stratification in dark at 4°C, the plates were transferred into a growth chamber with controlled light (80 µmoles /m^2/s) and temperature (20 °C ± 2) to examine the rate of germination. To examine the effect of salt and various growth hormones on germination, different concentrations of NaCl (0 to 150 mM), GA3 (0 to 10 µM), ABA (0 to 5 µM), and Paclobutrazol (PAC) (0 to 5 µM) were used. The surface-sterilized seeds were directly sown in the medium supplemented with different concentrations of salt and hormones.

In order to examine the effect of heat stress on germination, seeds were imbibed in water after surface sterilization and placed in the dark at 4 °C for 4 days. After stratification, they were subjected to different temperatures (38, 45, and 50 °C) for an hour in a water bath. After that treatment, seeds of *drif345* and WT were plated on germination medium and transferred to similar growing conditions as mentioned above. The seed germinated on medium without stress treatment was used as a control for this assays and protruded radicle through the seed coat was considered as the germination point.

All the experiments were repeated three to four times using four technical replicates. The germination rates and standard errors (SEs) were calculated on the basis of four biological replicates of each experiment. The numerical data were subjected to multiple t-tests (Kim *et al.*, 2003) to find statistical significance.

4.3.3 Gene Expression Analysis

Total RNA was isolated from green siliques, dry seeds, imbibed seeds after 24 h, and germinated seeds at different time points using a two-step TRIzol-based extraction method (Meng and Feldman, 2010). Around 500 ng of total RNA was digested with DNAseI to remove the gDNA contamination and first-strand cDNA was synthesized using Superscript IV enzyme (Invitrogen) according to manufacturer's manual. Resulting cDNA was diluted in 1:25 ratio and 1 µl aliquot was used for semi-quantitative and quantitative real time-PCR to check the expression of *DRIFs* during germination of WT and alteration in the expression of genes involved in ABA and GA biosynthesis and

signaling pathways in WT and *drif345*. *ACTIN2* (*ACT2*) and *Protein Phosphatase 2A subunit A3* (*PP2AA3*) were used as an internal control. The list of primers used in qRT-PCR and RT-PCR are mentioned in table.4.1.

4.4 Results

4.4.1 Expression of *DRIFs* during seed maturation and germination of *Arabidopsis* seeds

The *Arabidopsis* genome contains five *DRIF* homologs designated as *DRIF1*, *DRIF2*, *DRIF3*, *DRIF4*, and *DRIF5*. Phylogenetic study showed that *DRIF3*, *DRIF4*, and *DRIF5* are present in the same clade (Machemer *et al.*, 2011; Raimundo *et al.*, 2013) and transcriptome analysis (Winter *et al.*, 2007; Klepikova *et al.*, 2016) showed they are expressed in all *Arabidopsis* tissues. To investigate the role of these *DRIF* genes in germination, we first assessed the transcription profile of these genes in green siliques, three-month-old after-ripened seeds, 24 h imbibed seeds at 4 °C, and during different hours after exposure to light (Fig.1).

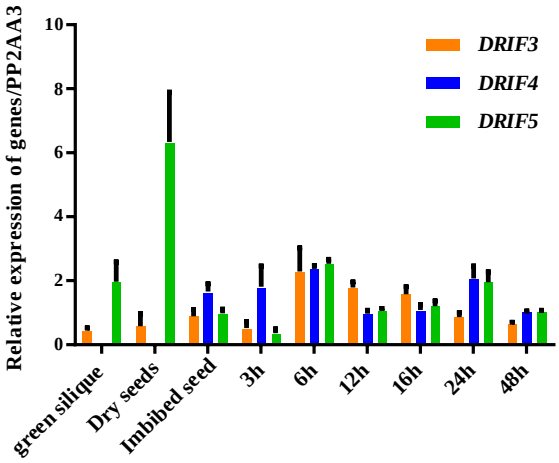


Fig.1. Expression profile of *DRIF3*, *DRIF4,* and *DRIF5* during different stages of germination and development in *Arabidopsis*. The relative expression of *DRIF3*, *DRIF4,* and *DRIF5* in green siliques, dried after ripened seeds, imbibed seeds at 4 °C for 24 h and after 3, 6, 12, 16, 24, and 48 h in long-day (LD) photoperiod after 4 days of stratification at 4 °C in dark. *Arabidopsis Protein Phosphatase 2A subunit A3* (*PP2AA3*) gene was used as an internal control. Data represent the mean of three tecnhical replicates ± SE.

In this study, we found that the relative expression of *DRIF4* was negligible in green siliques and in 3-months old dried seeds while the expression of *DRIF5* was higher in dried seeds than in any other stage of germination. The expression level of *DRIF3* was almost similar in green siliques and dried seeds (Fig.1).

The expression level of *DRIF5* was very high in the dried seed that drastically reduced after imbibition and further after exposure of seeds in light (Fig.1). The expression level of *DRIF3* was almost similar in green siliques, dry seeds, and imbibed seeds.

Differential expression of these *DRIF3*, *DRIF4*, and *DRIF5* in green siliques, dry and imbibed seeds suggested the involvedment of these genes of in seed dormancy and germination. The higher expression of *DRIF5* in dried seed, while the increase in expression of *DRIF4* after imbibition showed that these homologs might be involved in regulating different aspects of seed dormancy and germination.

4.4.2. *Arabidopsis DRIF3*, *DRIF4*, and *DRIF5* regulate germination in a partially redundant manner

In order to examine the role of *DRIF3*, *DRIF4*, and *DRIF5* in germination, we used *drif3*, *drif4* and *drif5* single mutants, *drif34*, *drif35,* and *drif45* double mutants and *drif345* triple mutant to scored germination in LD conditions. In this study, we found that at 16 h after light exposure, the germination percentage of *drif345* was significantly higher ($P \leq 0.0001$) than WT, while the germination percentage of single mutants and double mutants was similar to WT (Fig.2). The difference between the germination percentage of *drif345* and WT was reduced after 24 h. After 48 h of light exposure, all the genotypes had achieved maximum (around 100%) germination (Fig.2).

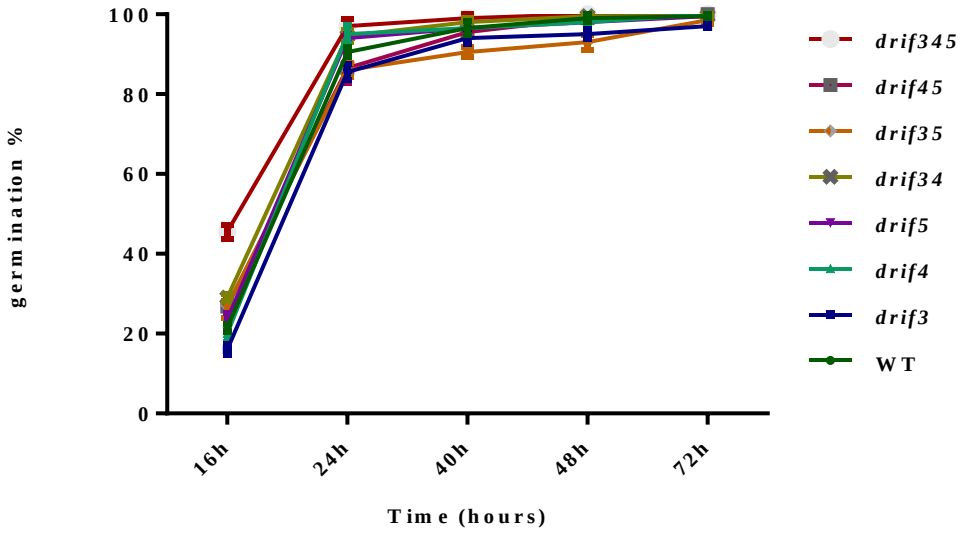


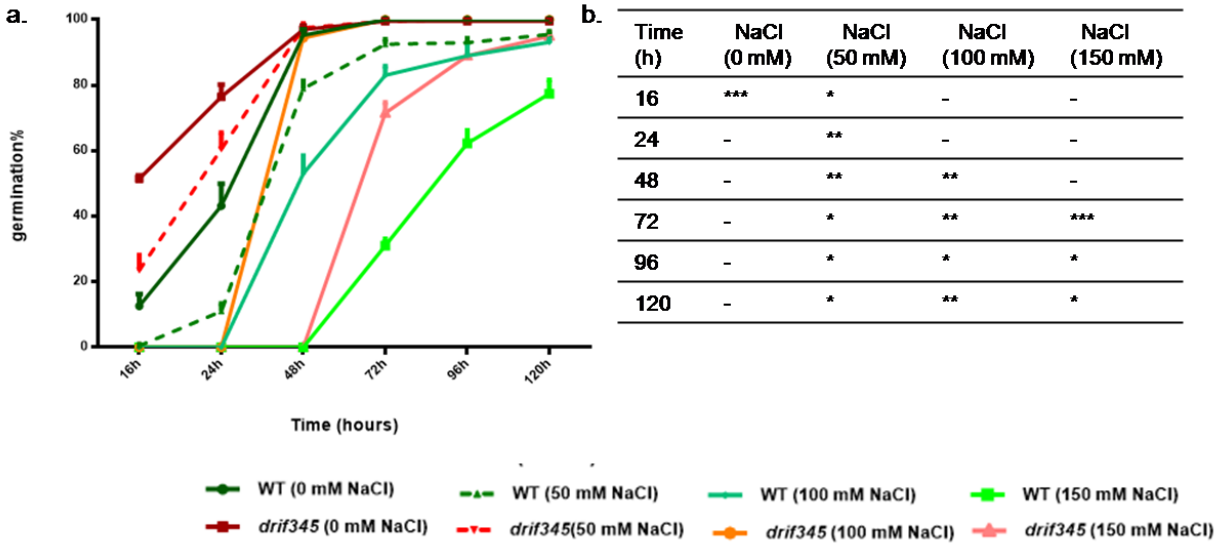
Fig.2. Germination of single, double and triple *drif* mutants in long-day (LD) photoperiod. Around 100 seeds of *drif3*, *drif4*, *drif5*, *drif34*, *drif35*, *drif45,* and *drif345* were germinated on ½ MS media. Germination assay was performed in LD after 4 days of stratification at 4 °C in dark. The germination rate of WT and *drif* mutants represents the mean of four biological replicates ± SE.

This result suggested that the *DRIF3*, *DRIF4*, and *DRIF5* may be involved in the early germination of *Arabidopsis*. Single and double *drif* mutants did not show any difference in the rate of germination while triple mutant, *drif345* showed early germination in LD (Fig.2). This result suggested that *DRIF* homologs may act as negative regulators of germination in a functionally redundant manner. In this study, only triple mutant *drif345* showed significantly higher rate of germination, so for further assays, we used only the *drif345* triple mutant.

4.4.3. Germination of *drif345* showed improved tolerance to salt and heat stress

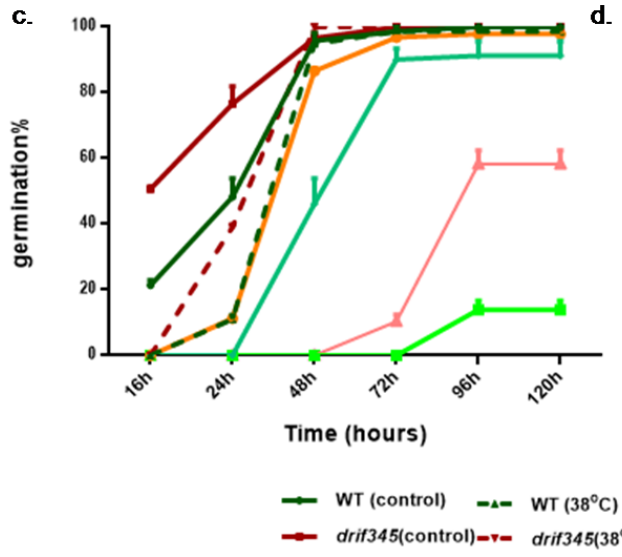
MYB transcription factors are known to regulate various developmental processes in response to abiotic stress (Jin *et al.*, 2000; Agarwal *et al.*, 2006; Lee *et al.*, 2015; Wang *et al.*, 2015; Yu *et al.*, 2016). The *RAD* and *DIV* homologs in *Arabidopsis* are functionally characterized as regulators of germination in salt stress (Fang *et al.*, 2018; Yang *et al.*, 2018). We initially found that the germination rate of *drif345* was faster than WT. In order to investigate the role of *DRIFs* in response to abiotic stresses; we conducted salt and heat stress assays to explore their role in germination during these responses. As shown in Fig.3a, *drif345* showed a faster germination rate in medium containing 50, 100 or 150 mM NaCl. Germination of WT seeds started after 24 h in 50 mM NaCl while in the case of *drif345* it started at 16h. The germination percentage of *drif345* at 24 h and 48 h was also significantly higher than WT in 50mM NaCl (Fig.3.b). Germination of both WT and *drif345* seeds started after 24 h in 100 mM NaCl. At 48 h, the germination percentage of *drif345* was higher than WT (Fig.3.b). In 150 mM NaCl, germination started after 48 h and the rate of germination of *drif345* was significantly higher than WT and also the total germination of *drif345* reached 90-95 % at 120 h (5th day) while the germination of WT reached only up to 80 % (Fig.3a and 3b).

The germination of seeds after heat shock treatment at 38, 45, and 50 °C (for an hour before exposure to light) caused a reduction in both the rate and total percentage of germination. After 38 °C treatment, the germination rate in WT and *drif345* was reduced when compared with non-stressed seeds still the germination rate of *drif345* was significantly higher than WT (Fig.3.c and d). After heat shock at 45 °C, the germination of WT and *drif345* seeds started after 24 h of light exposure. At 48 h the germination in *drif345* was higher than WT while the total germinated seeds after 120 h were the same in WT and *drif345* (Fig.3.c and d).



b.

Time (h)	NaCl (0 mM)	NaCl (50 mM)	NaCl (100 mM)	NaCl (150 mM)
16	***	*	-	-
24	-	**	-	-
48	-	**	**	-
72	-	*	**	***
96	-	*	*	*
120	-	*	**	*



d.

Time (h)	Control	38 °C	45 °C	50 °C
16	***	-	-	-
24	-	***	***	-
48	-	-	-	-
72	-	-	-	*
96	-	-	-	***
120	-	-	-	***

Fig.3. Germination of *drif345* in salt and heat stress. (a) Germination of *drif345* and WT in ½ MS media containing 0, 50, 100, and 150 mM NaCl, (b) difference between the germination of *drif345* and WT in different salt concentrations, (c) ½ MS media after heat shock at 38, 45, and 50 °C for 1 h, after 4 days of stratification, (d) difference between the germination of *drif345* and WT after heat shock at different temperatures. The germination rate of WT and *drif345* represents the mean of four biological replicates ± SE. P≤0.001 (***), P≤0.001 (**), and P≤0.05 (*).

The germination of *drif345* seeds started after 48 h when seeds were subjected to 50 °C for an hour before germination while the germination of WT started only after 72 h in the same condition. Total germinated seeds in *drif345* were around 50 percent while in the case of WT, total germination percentage was only around 20 percent (Fig.3.c and d).

In all the germination assays, the germination rate of *drif345* was always higher than WT irrespective of salt and heat stress during germination. These results indicated that *DRIF3*, *DRIF4*, and

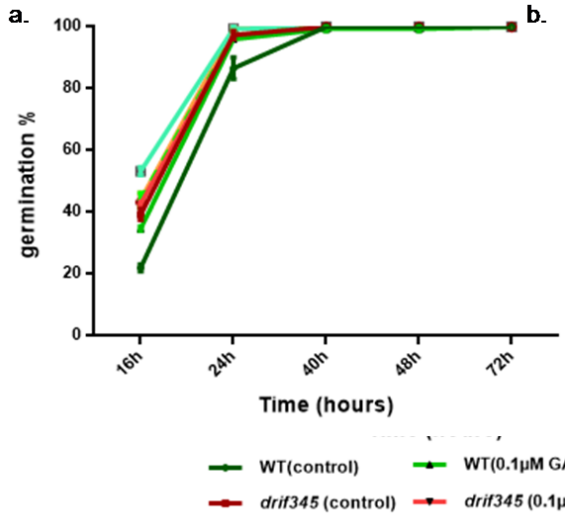
DRIF5 are actively involved in seed dormancy and germination. Salt and heat stress delayed the germination of WT and also affected the total number of seeds germinated in 150 mM NaCl and heat stress at 50 °C for an hour (Fig.3). In the case of *drif345*, the negative effect of salt and heat stress on seed germination was less significant than WT.

4.4.4 Effect of ABA, GA3, and PAC on *drif345* germination

It is well established that seed germination is regulated by GA and ABA levels. These phytohormones are also known to regulate seed germination in response to salt and osmotic stresses (reviewed in Vishal and Kumar, 2018). In order to understand the role of ABA and GA in the early germination phenotype of *drif345*, we tested different concentrations of ABA, GA3, and PAC to compensate for the early germination phenotype of mutant.

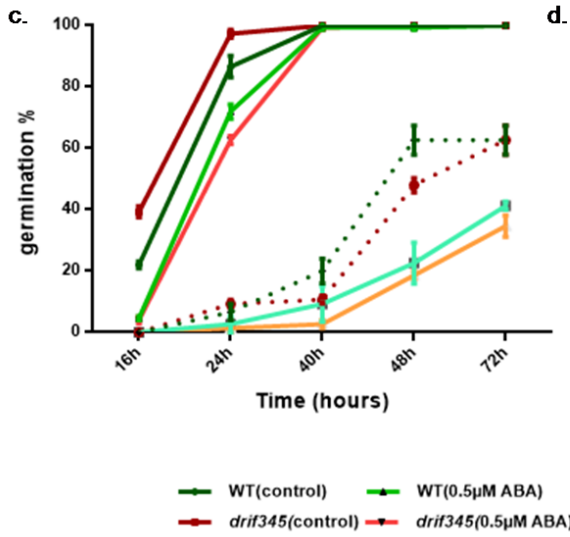
First, we tested three concentrations of GA3 (0.1, 1, and 10 μM), in order to see the effect of GA on the germination rate of WT and *drif345*. We found that the percentage of germination of WT increased by 0.1 μM GA3 at 16 h and it was similar to that of the germination percentage of *drif345* in the absence of GA3 (Fig.4.a). By increasing the concentration of GA3, the rate of germination was not much affected, further (Fig.4.b). There was no significant difference in the germination of *drif345* in all the concentrations of GA3 (Fig.4.b). We speculated that *drif345* may have a higher level of GA that may cause early germination (Fig.4.a and b).

Further, we tested the germination of *drif345* and WT in three concentrations (0.5, 1, and 5 μM) of ABA. The germination in control (ethanol 0.1% v/v) started at 16 h but in ABA, germination of both WT and *drif345* seeds began only after 24 h. We observed that the *drif345* mutant started behaving similarly to WT in 0.5 and 5 μM ABA at 24, 40 and 48 h (Fig.4.c). When we compared the germination of *drif345* in control and in ABA, we found that the rate of germination of *drif345* was significantly affected 16 and 24 h (Fig.4.c and d). These results showed that the germination of *drif345* is ABA sensitive. By introducing a small amount of ABA (0.5 μM), we observed a significant decrease in *drif345* germination and in 1 μM ABA the germination of *drif345* was more affected than WT.



b.

Time (h)	*drif345*			WT		
	GA3 (0.1 μM)	GA3 (1 μM)	GA3 (10 μM)	GA3 (0.1 μM)	GA3 (1 μM)	GA3 (10 μM)
16	-	-	-	**	**	***
24	-	-	-	-	-	-
40	-	-	-	-	-	-
48	-	-	-	-	-	-
72	-	-	-	-	-	-



d.

Time (h)	*drif345*			WT		
	ABA (0.5 μM)	ABA (1 μM)	ABA (5 μM)	ABA (0.5 μM)	ABA (1 μM)	ABA (5 μM)
16	***	***	***	**	***	***
24	***	***	***	-	***	***
40	-	***	***	-	***	***
48	-	***	***	-	*	***
72	-	*	***	-	*	***

Fig.4. Germination of *drif345* in response to GA3 and ABA. (a) Germination of *drif345* and WT in 0, 0.1, 1, and 10 μM GA3, (b) germination of *drif345* in GA3 compared to *drif345* without GA3 and WT in GA3 compared to WT without GA3, (c) germination of *drif345* and WT in 0, 0.5, 1 and 5 μM ABA, and (d) germination of *drif345* in ABA compared to *drif345* without ABA and WT in ABA compared to WT without ABA. The germination rate of WT and *drif345* represents the mean of four experimental replicates. The bar represents SE and for statistical significance, multiple t-tests were performed. The germination rate of WT and *drif345* represents the mean of four biological replicates ± SE. P≤0.001 (***), P≤0.001 (**), and P≤0.05 (*).

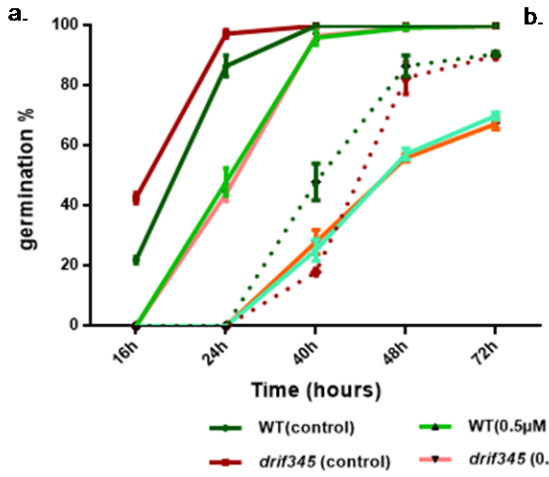
Seed germination is promoted by GA while ABA antagonizes the effect of GA (reviewed in Penfield and King, 2009). In our study, we found that the germination difference between *drif345* and WT could be diminished by GA. This result indicated that faster germination in *drif345* maybe because of the higher level of GA in *drif345* mutant. To check whether the early germination phenotype of *drif345* is GA dependent, we used GA antagonizing chemical paclobutrazol (PAC) in the medium to

assay the germination (Fig.5.a). Germination of *drif345* and WT started in 0.5 μM PAC after 16 h in light and the rate of germination of *drif345* was similar to WT. In a higher concentration of PAC (5 μM), the germination of WT and *drif345* seeds started after 24 h and the germination rate of *drif345* became similar to WT. In 1 μM of PAC, the germination of *drif345* was significantly reduced (Fig.5.b).

We investigated the GA sensitivity of *drif345* during the germination after blocking the endogenous synthesis of GA with 1 μM PAC. We found that the germination of *drif345* responded in a similar manner as WT in 0.1 and 1 μM GA3 and the rate of germination of *drif345* and WT was similar in both the concentrations of GA3 (Fig.5.c). When we compared the germination of *drif345* in 0.1 and 1 μM of GA3 +1 μM PAC to the germination of *drif345* in 1 μM PAC, we found that the germination of *drif345* was more responsive to GA3 (Fig.5.d). When 10 μM GA3 was used along with 1 μM PAC, the germination rate of *drif345* was slightly less than WT at 24 h (Fig.5.c) and the value was statistically less significant ($p \leq 0.05$) (Fig.5.d). After 24 h the rate of germination of *drif345* and WT was similar in the same condition.

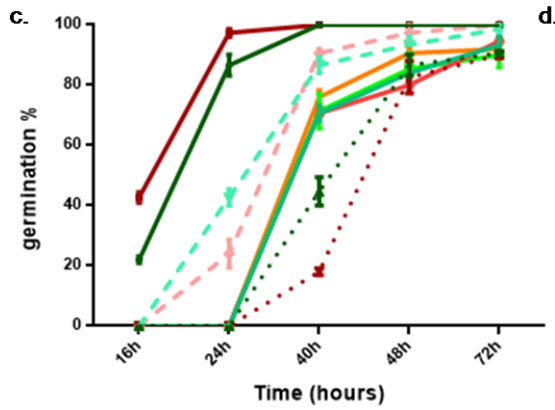
These results combinedly showed that the early germination phenotype of the *drif345* mutant is ABA and GA dependent. The difference in the germination of *drif345* and WT was diminished by the external application of GA3 and ABA. GA3 positively affected the rate of germination of WT while the germination of *drif345* was not significantly affected. Further, when endogenous GA was blocked by PAC, the germination rate of *drif345* was reduced and became similar to the one of WT. When GA3 was introduced along with PAC, the germination of *drif345* was more responsive to GA3 in 0.1 and 1 μM concentrations of GA3 (Fig.5.d).

The results suggested that *DRIFs* might regulate the overall ABA/GA ratio by integrating their signaling pathways and limit the rate of germination. In the absence of *DRIFs*, the ABA/GA balance might change that lead to faster germination in *drif345*.



b.

Time (h)	drif345			WT		
	PAC (0.5 μM)	PAC (1 μM)	PAC (5 μM)	PAC (0.5 μM)	PAC (1 μM)	PAC (5 μM)
16	***	***	***	***	***	***
24	***	***	***	*	***	***
40	-	***	***	-	**	***
48	-	*	***	-	*	***
72	-	**	***	-	*	***



d.

Time (h)	drif345			WT		
	1 μM PAC+ 0.1 μM GA3	1 μM PAC+ 1 μM GA3	1 μM PAC+ 10 μM GA3	1 μM PAC+ 0.1 μM GA3	1 μM PAC+ 1 μM GA3	1 μM PAC+ 10 μM GA3
16	-	-	-	-	-	-
24	-	-	*	-	-	***
40	***	***	***	*	-	*
48	-	-	-	-	-	-
72	-	-	**	-	-	*

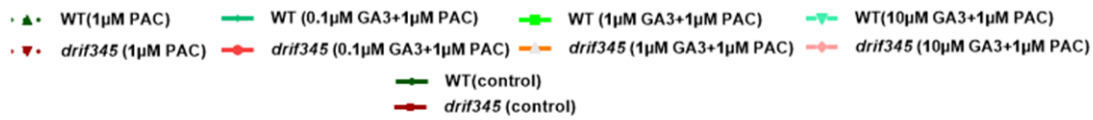


Fig.5. Germination of *drif345* in response to PAC and GA3. (a) Germination of *drif345* and WT in 0, 0.5, 1, and 5 μM PAC, (b) germination of *drif345* in PAC compared to *drif345* without PAC and WT in PAC compared to WT without PAC, (c) germination of *drif345* and WT in 0.1, 1, and 10 μM GA3 along with 1μM PAC, and (d) germination of *drif345* in response to GA3 compared to *drif345* in 1 μM PAC and WT in response to GA3 compared to WT in 1 μM PAC. The germination rate of WT and *drif345* represents the mean of four biological replicates. The bar represents SE and for statistical significance multiple t-test was performed. The germination rate of WT and *drif345* represents the mean of four biological replicates ± SE. P≤0.001 (***), P≤0.001 (**), and P≤0.05 (*).

4.4.5. Effect of *drifs* mutation on ABA and GA biosynthesis and signaling genes during germination

From the germination assays,we speculated that the *drif345* triple mutant may have altered ABA/GA ratio. To test this hypothesis, we performedsemi-quantitative PCR to analyse the expression of genes related to ABA and GA biosynthesis and signaling. after 24 h of exposure of seeds in LD, we tested the expression of three ABA biosynthesis genes *NCED5*, *ABA1* and *ABA3* in germinating seeds of

drif345 and WT (Fig. 6). We found that the expression of *ABA1* and *ABA3* was reduced in *drif345* while there was no difference in the expression of *NCED5* in WT and *drif345*.

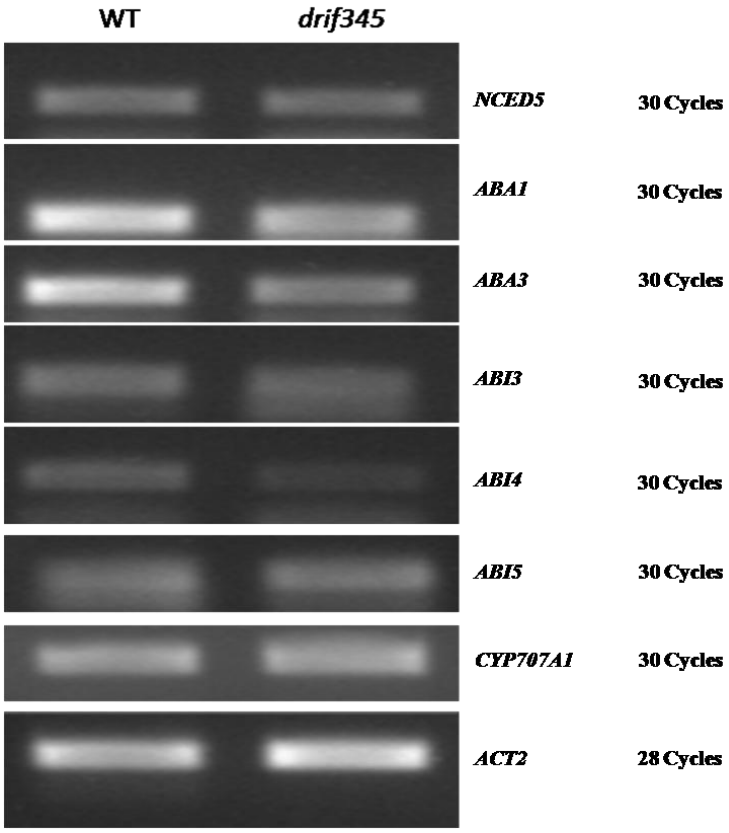


Fig.6. RT-PCR analysis of genes involved in ABA synthesis and signaling. Expression of *NCED5*, *ABA1*, *ABA3*, *ABI3*, *ABI4*, *ABI5*, and *CYP707A1* was analyzed using imbibed seed of WT and *drif345* after 24 h exposure in LD photoperiod. *ACTIN2* (*ACT2*) was used as an internal control.

We further examined the expression of three ABA signaling genes; *ABI3*, *ABI4*, and *ABI5* that are involved in positive ABA response and negative regulation of seed germination. We observed that there was a slight reduction in the expression of *ABI4* in *drif345* while the expression of *ABI3* and *ABI5* was similar to WT.The expression of *CYP707A1* was not changed in *drif345*. Expression analysis of ABA biosynthesis genes showed that there may be a decrease in ABA synthesis.

We checked the expression of GA biosynthetic genes; *GA20ox1* and *GA3ox1*, negative regulators of GA response; *GAI*, *RGA*, and *RGL2* along with *PIF1*, *PIF3*, and *PIF4*, which are known to involved in germination in response to GA (Fig.7). In this experiment, we found that the expression of *GA3ox1* was increased in the *drif345* mutant (Fig.7). We also found that the expression of *RGL2* was slightly reduced while the expression of *GAI* and *RGA* was not changed. There was no change in *PIF1*, *PIF3* and *PIF4* expression levels in the germinating seeds of *drif345* and WT.

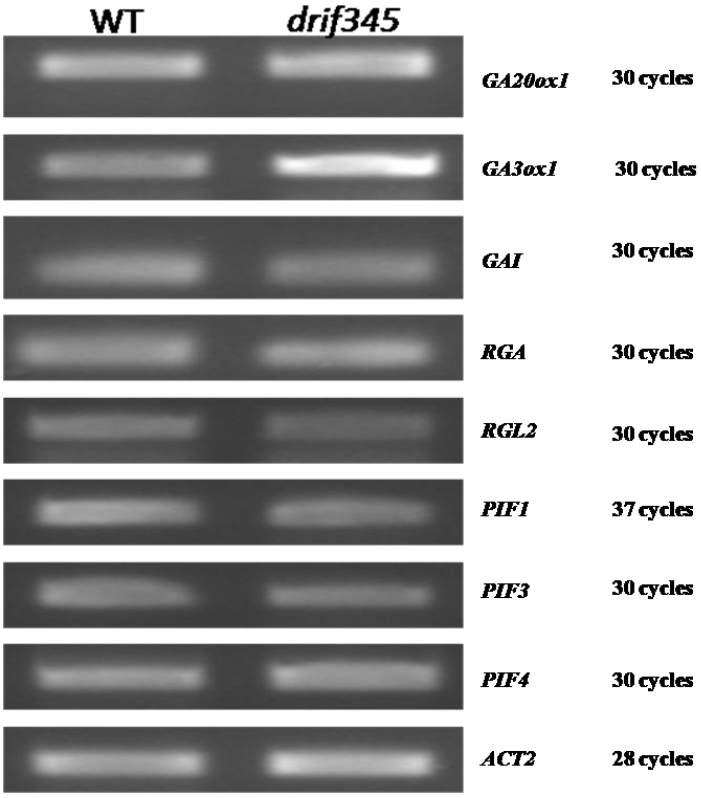


Fig.7. RT-PCR analysis of genes involved in GA synthesis and signaling. Expression of *GA20ox, GA3ox, GAI, RGA, RGL2, PIF1, PIF3, and PIF4* was analyzed using the imbibed seed of WT and *drif345* after 24 h exposure in LD photoperiod. *ACTIN2* (*ACT2*) was used as an internal control.

These results combinedly suggest that in *drif345* a reduced biosynthesis and signaling of ABA and increased biosynthesis of GA may contribute to the early germination phenotype of *drif345.*

4.5 Discussion

Seed germination is a crucial developmental process that is tightly regulated by the coordinated regulation of genes network during the transition of seed from dormant to non-dormant phase (Bassel *et al.*, 2011). In this study, we provide experimental evidence that the three *Arabidopsis MYB* genes, *DRIF3*, *DRIF4*, and *DRIF5* play an important role in seed germination. We observed the expression of *DRIF5* in green siliques and dry seeds while the expression of *DRIF4* was negligible in green siliques and dry seeds and started increasing only after imbibition and maintained during the germination. The relative expression of *DRIF3* was almost similar in green siliques, dry seeds and during the germination of seeds.

The early germination phenotype of *drif345* shows that *DRIF3*, *DRIF4*, and *DRIF5* acts as negative regulators of seed germination. None of the single mutants showed a significant difference in germination when grown in similar conditions. This result suggests the functional redundancy among these homologs in regulating the germination. The expression pattern of these three MYB homologs is similar to the expression of GA receptors *GID1a*, *GID1b*, and *GID1c* in siliques, dry seeds and during the germination of seed, which acts as positive regulators of GA signaling and reported as positive regulators of seed germination (Griffiths *et al.*, 2006). In the case of *GID1*, the non-germinating phenotype was observed only in triple mutant because of the functional redundancy of these genes (Griffiths *et al.*, 2006). But, recently it has been found that these *GID1* homologs are involved in positive and negative regulation of germination (Ge and Steber, 2018). In that study, the phenotypic difference in the germination of single and double *GID1* mutants was observed only in fresh seeds of less than 2 weeks old. In our study, we used three-month-old after-ripened seeds, which might be a reason that we could not see any significant difference in the germination rate of single and double mutants of *drifs*. In future, the germination assays with fresh seeds of *drif* single and double mutants will help to characterize the role of these homologs in detail.

During seed germination under favorable environmental conditions, the level of ABA decreases and the level of GA increases (Seo *et al.*, 2006) facilitating the germination. ABA plays a key role in the suppression of GA biosynthesis by forming a feedback loop with *ABIs* and *DELLAs*. When non-dormant seeds get exposed to a stressful environment, the endogenous level of ABA increases and postpones the germination (Kazachkova *et al.*, 2016). Germination of seeds in salt and osmotic stresses is related to the primary dormancy of seeds. Loss of dormancy opens the germination window and permits the germination of seeds in a wide environmental range (Finch-Savage and Leubner-Metzger, 2006; Yazdanpanah *et al.*, 2017). Here, when we assessed the germination in NaCl or subjected the seed to a higher temperature before germination, we found that the WT seeds showed delayed germination and the overall percentage of germination was also reduced. There were around 95 percent seeds of *drif345* germinated in 150 mM NaCl while only around 75 percent seeds of WT were germinated in this condition. After 50 °C heat shock for an hour, the total germination of *drif345* was higher than WT. These results show that these *DRIFs* may be involved in salt and heat response and delay the germination during the stress conditions. The germination of seeds during stress is also reduced due to higher synthesis of ABA in response to stresses. These results support the idea that *DRIFs* are likely to be involved in maintaining seed dormancy by influencing the level of ABA.

Loss of dormacy of seed is related with lower level of ABA and higher level of GA in seed. The genes involved in ABA and GA signaling are also important for germinatination and further growth of plants (Steber *et al.*, 1998; Lopez-Molina *et al.*, 2001; Ge and Steber, 2018). The DELLAs proteins are negative regulators of GA response and play a crucial role in mediating GA and ABA signaling (Penfield *et al.*, 2006a; Achard and Genschik, 2009; Liu and Hou, 2018), while *ABI3*, *ABI4,* and *ABI5* are positive regulators of ABA signaling (Shu *et al.*, 2013). Among five *DELLAs* in *Arabidopsis*, *RGL2* was reported as major inhibitor of seed germination (Lee *et al.*, 2002; Cao *et al.*, 2005; Ravindran *et al.*, 2017). *ABI4* has been reported as a repressor of seed germination and positive regulator of early seed dormancy (Penfield *et al.*, 2006b; Shu *et al.*, 2013). Gemination of *drif345* in ABA, GA, and PAC showed that unlike *ABI* mutants, germination of *drif345* is ABA sensitive. Also, when we blocked the endogenous synthesis of GA, the germination rate of *drif345* reduced, which suggests that the early germination phenotype of *drif345* is related to the ABA/GA amount rather than downstream signaling pathway of these hormones. The higher expression of *GA3ox1* in *drif345* and the lower expression of *ABA1* and *ABA3* suports this hypothesis. It may be a possibility that the*DRIFs* are involved in balancing ABA/GA ratio by positively regulating ABA biosynthesis and negatively regulating GA biosynthesis. Therefore, in the absence of *DRIFs,* the expression of ABA synthesizing genes, *ABA1* and *ABA3* decrease and the one of GA synthesizing gene, *GA3ox1* increases, leading to earlier germination in *drif345* plants.

MYB proteins are involved in GA-ABA biosynthesis and signaling to mediate various abiotic stresses (Ambawat *et al.*, 2013; Roy, 2016). In *Arabidopsis*, several reports on *RAD* and *DIV* homologs show the role of *Arabidopsis RAD* and *DIV* genes in seed germination by changing the expression of ABA synthesis and its sensitivity (Chen *et al.*, 2017; Fang *et al.*, 2018; Yang *et al.*, 2018). The seeds of *Atdiv2* mutants show less germination percentage and higher level of ABA (Fang *et al.*, 2018) while the *RSM1/AtRL1* is involved in activation of *ABI5* and overexpression *RSM1* induces salt tolerance in seedlings (Yang *et al.*, 2018). In this study, we found *DRIFs* as a novel players in seed germination and may involved in the maintinance of the ABA/GA ratio in seeds. In future, the quantification of ABA and GA hormones in *drif* mutants will help to evaluate our hypothesis and further investigation of the role of *DRIFs* in germination and plant development.

In this study, we report three MYB homologs, *DRIF3*, *DRIF4*, and *DRIF5* regulators germination in functionally redundant manner. They may involve in the regulation of ABA/GA ratio and maintenance of fine-tune of the gene regulatory network involved in seed dormancy and germination. It will be

interesting to gain more information about the *DRIF* homologs and the connection of these MYB transcription factors and ABA/GA hormones will help to understand their role in plant life cycle.

4.6 References

Achard, P. & Genschik, P. (2009). Releasing the brakes of plant growth: how GAs shutdown DELLA proteins. *Journal of Experimental Botany* 60(4): 1085-1092.

Agarwal, M., Hao, Y., Kapoor, A., Dong, C. H., Fujii, H., Zheng, X. & Zhu, J. K. (2006). A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *J Biol Chem* 281(49): 37636-37645.

Ali-Rachedi, S., Bouinot, D., Wagner, M. H., Bonnet, M., Sotta, B., Grappin, P. & Jullien, M. (2004). Changes in endogenous abscisic acid levels during dormancy release and maintenance of mature seeds: studies with the Cape Verde Islands ecotype, the dormant model of Arabidopsis thaliana. *Planta* 219(3): 479-488.

Ambawat, S., Sharma, P., Yadav, N. R. & Yadav, R. C. (2013). MYB transcription factor genes as regulators for plant responses: an overview. *Physiol Mol Biol Plants* 19(3): 307-321.

Bassel, G. W., Lan, H., Glaab, E., Gibbs, D. J., Gerjets, T., Krasnogor, N., Bonner, A. J., Holdsworth, M. J. & Provart, N. J. (2011). Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions. *Proceedings of the National Academy of Sciences* 108(23): 9709-9714.

Bewley, J. D. (1997). Seed Germination and Dormancy. *The Plant Cell* 9(7): 1055-1066.

Bittner, F., Oreb, M. & Mendel, R. R. (2001). ABA3 is a molybdenum cofactor sulfurase required for activation of aldehyde oxidase and xanthine dehydrogenase in Arabidopsis thaliana. *J Biol Chem* 276(44): 40381-40384.

Brocard, I. M., Lynch, T. J. & Finkelstein, R. R. (2002). Regulation and role of the Arabidopsis abscisic acid-insensitive 5 gene in abscisic acid, sugar, and stress response. *Plant Physiol* 129(4): 1533-1543.

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.

Chen, M., MacGregor, D. R., Dave, A., Florance, H., Moore, K., Paszkiewicz, K., Smirnoff, N., Graham, I. A. & Penfield, S. (2014). Maternal temperature history activates Flowering Locus T in fruits to control progeny dormancy according to time of year. *Proceedings of the National Academy of Sciences* 111(52): 18787-18792.

Chen, Y. S., Chao, Y. C., Tseng, T. W., Huang, C. K., Lo, P. C. & Lu, C. A. (2017). Two MYB-related transcription factors play opposite roles in sugar signaling in Arabidopsis. *Plant Mol Biol* 93(3): 299-311.

Colebrook, E. H., Thomas, S. G., Phillips, A. L. & Hedden, P. (2014). The role of gibberellin signalling in plant responses to abiotic stress. *The Journal of Experimental Biology* 217(1): 67-75.

de Lucas, M., Daviere, J. M., Rodriguez-Falcon, M., Pontin, M., Iglesias-Pedraz, J. M., Lorrain, S., Fankhauser, C., Blazquez, M. A., Titarenko, E. & Prat, S. (2008). A molecular framework for light and gibberellin control of cell elongation. *Nature* 451(7177): 480-484.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. & Lepiniec, L. (2010). MYB transcription factors in Arabidopsis. *Trends Plant Sci* 15(10): 573-581.

Fang, J. & Chu, C. (2008). Abscisic acid and the pre-harvest sprouting in cereals. *Plant Signal Behav* 3(12): 1046-1048.

Fang, Q., Wang, Q., Mao, H., Xu, J., Wang, Y., Hu, H., He, S., Tu, J., Cheng, C., Tian, G., Wang, X., Liu, X., Zhang, C. & Luo, K. (2018). AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Rep* 37(11): 1499-1511.

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

Finkelstein, R., Reeves, W., Ariizumi, T. & Steber, C. (2008). Molecular aspects of seed dormancy. *Annu Rev Plant Biol* 59: 387-415.

Frey, A., Effroy, D., Lefebvre, V., Seo, M., Perreau, F., Berger, A., Sechet, J., To, A., North, H. M. & Marion-Poll, A. (2012). Epoxycarotenoid cleavage by NCED5 fine-tunes ABA accumulation and affects seed dormancy and drought tolerance with other NCED family members. *The Plant Journal* 70(3): 501-512.

Ge, W. & Steber, C. M. (2018). Positive and negative regulation of seed germination by the Arabidopsis GA hormone receptors, GID1a, b, and c. *Plant Direct* 2(9): e00083.

González-Guzmán, M., Apostolova, N., Bellés, J. M., Barrero, J. M., Piqueras, P., Ponce, M. R., Micol, J. L., Serrano, R. & Rodríguez, P. L. (2002). The Short-Chain Alcohol Dehydrogenase ABA2 Catalyzes the Conversion of Xanthoxin to Abscisic Aldehyde. *The Plant Cell* 14(8): 1833-1846.

Griffiths, J., Murase, K., Rieu, I., Zentella, R., Zhang, Z.-L., Powers, S. J., Gong, F., Phillips, A. L., Hedden, P., Sun, T.-p. & Thomas, S. G. (2006). Genetic Characterization and Functional Analysis of the GID1 Gibberellin Receptors in Arabidopsis. *The Plant Cell* 18(12): 3399-3414.

Gubler, F., Millar, A. A. & Jacobsen, J. V. (2005). Dormancy release, ABA and pre-harvest sprouting. *Curr Opin Plant Biol* 8(2): 183-187.

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., Tabata, S., Kakubari, Y., Yamaguchi-Shinozaki, K. & Shinozaki, K. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *Plant J* 27(4): 325-333.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. & Martin, C. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *Embo j* 19(22): 6150-6161.

Kazachkova, Y., Khan, A., Acuña, T., López-Díaz, I., Carrera, E., Khozin-Goldberg, I., Fait, A. & Barak, S. (2016). Salt Induces Features of a Dormancy-Like State in Seeds of Eutrema (Thellungiella) salsugineum, a Halophytic Relative of Arabidopsis. *Frontiers in plant science* 7: 1071.

Kim, J., Yi, H., Choi, G., Shin, B., Song, P.-S. & Choi, G. (2003). Functional Characterization of Phytochrome Interacting Factor 3 in Phytochrome-Mediated Light Signal Transduction. *The Plant Cell* 15(10): 2399-2407.

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D. & Penin, A. A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *The Plant Journal* 88(6): 1058-1070.

Koornneef, M., Hanhart, C. J., Hilhorst, H. W. M. & Karssen, C. M. (1989). In Vivo Inhibition of Seed Development and Reserve Protein Accumulation in Recombinants of Abscisic Acid Biosynthesis and Responsiveness Mutants in Arabidopsis thaliana. *Plant Physiology* 90(2): 463-469.

Koornneef, M., Jorna, M. L., Brinkhorst-van der Swan, D. L. C. & Karssen, C. M. (1982). The isolation of abscisic acid (ABA) deficient mutants by selection of induced revertants in non-germinating gibberellin sensitive lines of Arabidopsis thaliana (L.) heynh. *Theoretical and Applied Genetics* 61(4): 385-393.

Lee, K., Lee, H. G., Yoon, S., Kim, H. U. & Seo, P. J. (2015). The Arabidopsis MYB96 Transcription Factor Is a Positive Regulator of ABSCISIC ACID-INSENSITIVE4 in the Control of Seed Germination. *Plant Physiol* 168(2): 677-689.

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

Lefebvre, V., North, H., Frey, A., Sotta, B., Seo, M., Okamoto, M., Nambara, E. & Marion-Poll, A. (2006). Functional analysis of Arabidopsis NCED6 and NCED9 genes indicates that ABA synthesized in the endosperm is involved in the induction of seed dormancy. *Plant J* 45(3): 309-319.

Leivar, P. & Monte, E. (2014). PIFs: systems integrators in plant development. *Plant Cell* 26(1): 56-78.

Liu, X. & Hou, X. (2018). Antagonistic Regulation of ABA and GA in Metabolism and Signaling Pathways. *Frontiers in plant science* 9: 251.

Lopez-Molina, L., Mongrand, S. & Chua, N.-H. (2001). A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factor in Arabidopsis. *Proceedings of the National Academy of Sciences* 98(8): 4782-4787.

Machemer, K., Shaiman, O., Salts, Y., Shabtai, S., Sobolev, I., Belausov, E., Grotewold, E. & Barg, R. (2011). Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. *Plant J* 68(2): 337-350.

Meng, L. & Feldman, L. (2010). A rapid TRIzol-based two-step method for DNA-free RNA extraction from Arabidopsis siliques and dry seeds. *Biotechnol J* 5(2): 183-186.

Nambara, E., Naito, S. & McCourt, P. (1992). A mutant of Arabidopsis which is defective in seed development and storage protein accumulation is a new abi3 allele. *The Plant Journal* 2(4): 435-441.

Nelson, S. K., Ariizumi, T. & Steber, C. M. (2017). Biology in the Dry Seed: Transcriptome Changes Associated with Dry Seed Dormancy and Dormancy Loss in the Arabidopsis GA-Insensitive sleepy1-2 Mutant. *Frontiers in plant science* 8: 2158.

Oh, E., Kim, J., Park, E., Kim, J.-I., Kang, C. & Choi, G. (2004). PIL5, a Phytochrome-Interacting Basic Helix-Loop-Helix Protein, Is a Key Negative Regulator of Seed Germination in Arabidopsis thaliana. *The Plant Cell* 16(11): 3045-3058.

Oh, E., Yamaguchi, S., Hu, J., Yusuke, J., Jung, B., Paik, I., Lee, H.-S., Sun, T.-p., Kamiya, Y. & Choi, G. (2007). PIL5, a Phytochrome-Interacting bHLH Protein, Regulates Gibberellin Responsiveness by Binding Directly to the GAI and RGA Promoters in Arabidopsis Seeds. *The Plant Cell* 19(4): 1192-1208.

Oh, E., Yamaguchi, S., Kamiya, Y., Bae, G., Chung, W.-I. & Choi, G. (2006). Light activates the degradation of PIL5 protein to promote seed germination through gibberellin in Arabidopsis. *The Plant Journal* 47(1): 124-139.

Penfield, S., Gilday, A. D., Halliday, K. J. & Graham, I. A. (2006a). DELLA-mediated cotyledon expansion breaks coat-imposed seed dormancy. *Curr Biol* 16(23): 2366-2370.

Penfield, S. & Hall, A. (2009a). A role for multiple circadian clock genes in the response to signals that break seed dormancy in Arabidopsis. *Plant Cell* 21(6): 1722-1732.

Penfield, S. & Hall, A. (2009b). A Role for Multiple Circadian Clock Genes in the Response to Signals That Break Seed Dormancy in Arabidopsis. *The Plant Cell* 21(6): 1722-1732.

Penfield, S., Josse, E. M., Kannangara, R., Gilday, A. D., Halliday, K. J. & Graham, I. A. (2005). Cold and light control seed germination through the bHLH transcription factor SPATULA. *Curr Biol* 15(22): 1998-2006.

Penfield, S. & King, J. (2009). Towards a systems biology approach to understanding seed dormancy and germination. *Proceedings of the Royal Society B: Biological Sciences* 276(1673): 3561-3569.

Penfield, S., Li, Y., Gilday, A. D., Graham, S. & Graham, I. A. (2006b). Arabidopsis ABA INSENSITIVE4 regulates lipid mobilization in the embryo and reveals repression of seed germination by the endosperm. *Plant Cell* 18(8): 1887-1899.

Peng, J., Carol, P., Richards, D. E., King, K. E., Cowling, R. J., Murphy, G. P. & Harberd, N. P. (1997). The Arabidopsis GAI gene defines a signaling pathway that negatively regulates gibberellin responses. *Genes Dev* 11(23): 3194-3205.

Pham, V. N., Kathare, P. K. & Huq, E. (2018). Phytochromes and Phytochrome Interacting Factors. *Plant Physiology* 176(2): 1025-1038.

Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., Coen, E. & Costa, M. M. R. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. *The Plant Journal* 75(4): 527-538.

Raimundo, J., Sobral, R., Laranjeira, S. & Costa, M. M. R. (2018). Successive Domain Rearrangements Underlie the Evolution of a Regulatory Module Controlled by a Small Interfering Peptide. *Mol Biol Evol* 35(12): 2873-2885.

Ravindran, P., Verma, V., Stamm, P. & Kumar, P. P. (2017). A Novel RGL2–DOF6 Complex Contributes to Primary Seed Dormancy in Arabidopsis thaliana by Regulating a GATA Transcription Factor. *Molecular Plant* 10(10): 1307-1320.

Roy, S. (2016). Function of MYB domain transcription factors in abiotic stress and epigenetic control of stress response in plant genome. *Plant Signal Behav* 11(1): e1117723.

Saito, S., Hirai, N., Matsumoto, C., Ohigashi, H., Ohta, D., Sakata, K. & Mizutani, M. (2004). Arabidopsis CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiol* 134(4): 1439-1449.

Schwartz, S. H., Tan, B. C., Gage, D. A., Zeevaart, J. A. D. & McCarty, D. R. (1997). Specific Oxidative Cleavage of Carotenoids by VP14 of Maize. *Science* 276(5320): 1872-1874.

Seo, M., Hanada, A., Kuwahara, A., Endo, A., Okamoto, M., Yamauchi, Y., North, H., Marion-Poll, A., Sun, T.-p., Koshiba, T., Kamiya, Y., Yamaguchi, S. & Nambara, E. (2006). Regulation of hormone metabolism in Arabidopsis seeds: phytochrome regulation of abscisic acid metabolism and abscisic acid regulation of gibberellin metabolism. *The Plant Journal* 48(3): 354-366.

Shu, K., Chen, Q., Wu, Y., Liu, R., Zhang, H., Wang, P., Li, Y., Wang, S., Tang, S., Liu, C., Yang, W., Cao, X., Serino, G. & Xie, Q. (2016). ABI4 mediates antagonistic effects of abscisic acid and gibberellins at transcript and protein levels. *Plant J* 85(3): 348-361.

Shu, K., Zhang, H., Wang, S., Chen, M., Wu, Y., Tang, S., Liu, C., Feng, Y., Cao, X. & Xie, Q. (2013). ABI4 Regulates Primary Seed Dormancy by Regulating the Biogenesis of Abscisic Acid and Gibberellins in Arabidopsis. *PLOS Genetics* 9(6): e1003577.

Shu, K., Zhou, W., Chen, F., Luo, X. & Yang, W. (2018). Abscisic Acid and Gibberellins Antagonistically Mediate Plant Development and Abiotic Stress Responses. *Frontiers in plant science* 9: 416.

Silverstone, A. L., Ciampaglio, C. N. & Sun, T.-p. (1998). The Arabidopsis Gene Encodes a Transcriptional Regulator Repressing the Gibberellin Signal Transduction Pathway. *The Plant Cell* 10(2): 155-169.

Steber, C. M., Cooney, S. E. & McCourt, P. (1998). Isolation of the GA-response mutant sly1 as a suppressor of ABI1-1 in Arabidopsis thaliana. *Genetics* 149(2): 509-521.

Sun, T. P. (2008). Gibberellin metabolism, perception and signaling pathways in Arabidopsis. *Arabidopsis Book* 6: e0103.

Vishal, B. & Kumar, P. P. (2018). Regulation of Seed Germination and Abiotic Stresses by Gibberellins and Abscisic Acid. *Frontiers in plant science* 9: 838.

Vleeshouwers, L. M., Bouwmeester, H. J. & Karssen, C. M. (1995). Redefining Seed Dormancy: An Attempt to Integrate Physiology and Ecology. *Journal of Ecology* 83(6): 1031-1037.

Wang, T., Tohge, T., Ivakov, A., Mueller-Roeber, B., Fernie, A. R., Mutwil, M., Schippers, J. H. & Persson, S. (2015). Salt-Related MYB1 Coordinates Abscisic Acid Biosynthesis and Signaling during Salt Stress in Arabidopsis. *Plant Physiol* 169(2): 1027-1041.

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-e718.

Yang, B., Song, Z., Li, C., Jiang, J., Zhou, Y., Wang, R., Wang, Q., Ni, C., Liang, Q., Chen, H. & Fan, L.-M. (2018). RSM1, an Arabidopsis MYB protein, interacts with HY5/HYH to modulate seed

germination and seedling development in response to abscisic acid and salinity. *PLOS Genetics* 14(12): e1007839.

Yazdanpanah, F., Hanson, J., Hilhorst, H. W. M. & Bentsink, L. (2017). Differentially expressed genes during the imbibition of dormant and after-ripened seeds - a reverse genetics approach. *BMC Plant Biol* 17(1): 151.

Yu, Y.-T., Wu, Z., Lu, K., Bi, C., Liang, S., Wang, X.-F. & Zhang, D.-P. (2016). Overexpression of the MYB37 transcription factor enhances abscisic acid sensitivity, and improves both drought tolerance and seed productivity in Arabidopsis thaliana. *Plant Mol Biol* 90(3): 267-279.

4.7 Supplementary data

Table.4.1 Primers used in this study

Primer name	Sequence (5′-3′)	Application
UMC8 (DRIF3R)	GAG GAC TTT GCA ACC GTT TTG T	amplification for genotyping SAIL_636_A08 (*DRIF3*)
UMC9 (DRIF3F)	CGCCTTTCGAAGGTTTAAG	amplification for genotyping SAIL_636_A08 (*DRIF3*)
UMC14 (DRIF4F)	CGAAGGTTAGTTATAGAAAGGTAGTGG	amplification for genotyping SALK_204381C (*DRIF4*)
UMC15 (DRIF4R)	CACAGAAAGGTTTGAGGAGATT	amplification for genotyping SALK_204381C (*DRIF4*)
UMC405 (DRIF5F)	TCAAAATTGAACAAATGGCAAG	amplification for genotyping SALK_121145 (*DRIF5*)
UMC406 (DRIF5)	CATGGCGTAAGATGATTCACC	amplification for genotyping SALK_121145 (*DRIF5*)
UMC76 (LB3)	TAGCATCTGAATTTCATAACCAATCTCGATACAC	genotype SAIL lines (*DRIF3*)
UMC237 (LBb1.3)	ATTTTGCCGATTTCGGAAC	genotype SALK lines (*DRIF4*, and *DRIF5*)
UMC264 (ACT2F)	CCCTCAGCACATTCCAGCAG	RT-PCR for *ACT2*
UMC265 (ACT2R)	TCCCATTCATAAAACCCCAGC	RT-PCR for *ACT2*
UMC672(qDRIF3F)	AGGTGTTTGCTTGCGTTCTC	qPCR for *DRIF3*
UMC673(qDRIF3R)	GCTCAGATTTGGCATCAACA	qPCR for *DRIF3*
UMC674 (qDRIF4F)	CTTTCAAAATTTAAAGATGAACCACA	qPCR for *DRIF4*
UMC675 (qDRIF4R)	TTTCTTCGCTTTTGCGTCA	qPCR for *DRIF4*
UMC676(qDRIF5F)	GTTCTAGCAAGGGCAAGCAG	qPCR for *DRIF5*

UMC677(qDRIF5R)	GCTTGTAGATGGCCCTGAGA	qPCR for *DRIF5*
UMC797 (GAI-F)	CACCTCGGCTTGGAAACTCTC	qPCR for *GAI*
UMC798 (GAI-R)	TGACAAAGGGAAAAACAGTAGGATTT	qPCR for *GAI*
UMC 818 (qACT2F)	TGGTCGTACAACCGGTATTG	qPCR for *ACT2*
UMC 819 (qACT2R)	CAAGGTCAAGACGGAGGATG	qPCR for *ACT2*
UMC820 (PIF3 F)	GAAAGGAGACGGCGTGATAG	qPCR for *PIF3*
UMC 821(PIF3R)	GCCATTGACATGATCTGCAC	qPCR for *PIF3*
UMC822 (PIF4 F)	ATCAGATGCAATCGGTAACAAG	qPCR for *PIF4*
UMC823 (PIF4)	CTGCAGTGAGGTATTAGTTCTTGC	qPCR for *PIF4*
UMC824(RGA F)	ACGTTGAGTCAATGGGGAAA	qPCR for *RGA*
UMC825(RGA R)	AGCGGAGGTGGTAATGAGTG	qPCR for *RGA*
UMC826(HLS1F)	TTCACAGGAGACAAGGGATTG	qPCR for *HLS1*
UMC827(HLS1R)	ACATTTCCCGGTGAACAAAT	qPCR for *HLS1*
UMC834 (PHYBF)	TTGAAGACGGTTCATTTGTGC	qPCR for *PHYB*
UMC835(PHYBR)	CACGGATCAGCTGAAGACCT	qPCR for *PHYB*
UMC836 (PIF1F)	CTCGCTGCAACAAGCAGAT	qPCR for *PIF1*
UMC837 (PIF1R)	CATCATTCCACATCCCATTGACA	qPCR for *PIF1*
UMC840 (NCED5F)	AGAGTTGGGAATCGGAGCTT	qPCR for *NCED5*
UMC841 (NCED5R)	GCAACCGGTTTTGGTTTAAG	qPCR for *NCED5*
UMC842 (ABI3F)	GAATTTGCGGTTTCTCTTGC	qPCR for *ABI3*
UMC843 (ABI3R)	TAGCTCCGGCAAGTGTGTCT	qPCR for *ABI3*
UMC844 (ABA1F)	TGTGACCGATAACGAAGGAA	qPCR for *ABA1*

UMC845 (ABA1R)	CCTAAACGCCGCCTTCTT	qPCR for *ABA1*
UMC48 (RGL2F)	GAGATGACTCGCCTGAAACC	qPCR for *RGL2*
UMC849 (RGL2R)	CGCACAAGGTCCAAAAAGAT	qPCR for *RGL2*
UMC850 (GA20ox1F)	TTCAGCCATTTGGGAAGGT	qPCR for *GA20ox1*
UMC851 (GA20ox1R)	TTCAGTCTCATTATTGAATCGTTTTC	qPCR for *GA20ox1*
UMC852 (GA3ox1F)	CCCCAACATCACCTCAACTAC	qPCR for *GA3ox1*
UMC853 (GA3ox1R)	CAGCTTGGGCCCAGTTTA	qPCR for *GA3ox1*
UMC858 (ABI4F)	TCCGGGTTATATGGTTGGAG	qPCR for *ABI4*
UMC859 (ABI4R)	CAATGCTCTTGACCGACCTT	qPCR for *ABI4*
UMC861 (ABI5F)	GAAGAGGAAGCAACAGTATTTTGAG	qPCR for *ABI5*
UMC862 (ABI5R)	TCATCAATGTCCGCAATCTC	qPCR for *ABI5*
UMC873 (ABA3F)	CCAAAAGAAGGTGCCTGAGA	qPCR for *ABA3*
UMC874 (ABA3R)	TCGTTCCTCGGGTTATATGAAT	qPCR for *ABA3*

CHAPTER 5
"Role of *ArabidopsisDIV-and-RAD-INTERACTING-FACTORS* (*DRIFs*) in skotomorphogenic andphotomorphogenic development of seedlings"

Role of *Arabidopsis DIV-and-RAD-interacting-factors* (*DRIFs*) in skotomorphogenic and photomorphogenic development of seedlings

Shweta Singh, Rómulo Sobral, Sara Larangeira, João Raimundo, M. Manuela R. Costa

Biosystems and Integrative Sciences Institute (BioISI), Plant Functional Biology Center, University of Minho, Campus de Gualtar,

5.1 Abstract

Light plays a crucial role in the early development of terrestrial plants and the seedling growing in light shows distinct morphology than in dark. Various families of transcription factors are known to involve in light-mediated morphological changes in plants. The MYB family transcription factors are functionally diverse and involved to perceive various environmental signals in plants. In the present study, we characterized the role of three MYB family genes, *DIV-and-RAD-interacting-factor3 (DRIF3)*, *DRIF4*, and *DRIF5* in the development of *Arabidopsis* seedling in light and dark conditions. The triple mutant seedlings *drif345* do not display apical hook in the dark. Also, these *DRIFs* are involved in hypocotyl elongation in red, far-red, blue and white light as the *drif345* mutant showed a reduced hypocotyl in these light conditions. Taken together, these results suggest that *DRIF3*, *DRIF4*, and *DRIF5* are part of an intricate gene network involved in photomorphogenic and skotomorphogenic development of seedlings. Interestingly, the expression of *HOOKLESS1* (*HLS1*) was reduced in dark-grown *drif345* seedlings and the expression of *PHYTOCHROME B* (*PHYB*) was increased in light-grown seedlings of *drif345*. These results suggest that *DRIF3*, *DRIF4* and *DRIF5* may be part of the gene regulatory network that regulates the formation of apical hook and hypocotyl elongation in response to light.

5.2 Introduction

Transcription factors (TFs) play a pivotal role to control plant development and environmental responses. The MYB family proteins represent one of the large, functionally diverse group of TF that regulates various developmental processes and environmental responses in plants (Jin *et al.*, 2000; Agarwal *et al.*, 2006; Cheng *et al.*, 2009; Mandaokar and Browse, 2009; Kwon *et al.*, 2013). DIV-and-RAD-interacting-factor (DRIF) is an MYB-family protein that interacts with other related proteins, RADIALIS (RAD) and DIVARICATA (DIV) in *Antirrhinum majus* and regulates flower zygomorphy (Raimundo *et al.*, 2013). In tomato, the homolog of *DRIF* (*FSB1*) is involved in the expansion of pericarp and regulates fruit size along with *RAD* (*FSM1*) and *DIV* homologs (*MYB1*) (Machemer *et al.*, 2011). It has been reported that the additional domain of DRIF can interact with the proteins associated with wood-formation in *Populus* that suggests the DRIF proteins may involve in diverse functions in different species (Petzold *et al.*, 2018).

Previous studies showed that some *DIV* and *RAD* homologs are associated with seed germination, seedling development and flowering in *Arabidopsis thaliana* (Hamaguchi *et al.*, 2008; Li *et al.*, 2015; Fang *et al.*, 2018; Yang *et al.*, 2018). There are five *DRIF* homologs in *A. thaliana* and these DRIF proteins are able to interact with *Arabidopsis* DIV and RAD proteins in yeast two-hybrid assay (Machemer *et al.*, 2011; Raimundo *et al.*, 2018), suggesting that the interaction between the DIV-DRIF-RAD proteins may form a regulatory module (DDR) in *Arabidopsis*, also. However, the process in which they may involve and the function of *DRIFs* is yet to be characterized. *DRIFs* is differentially expressed in the transcriptome profiling of *Arabidopsis* dry seeds, stratified seeds, and germinating seedlings in light conditions (Narsai *et al.*, 2011). Comprehensive RNAseq database also revealed that the expression of *DRIFs* is present in the hypocotyl of one-day-old seedling that suggests the involvement of *DRIF* homologs in seedling development (Klepikova *et al.*, 2016).

After germination, the seedlings follow two developmental programs, depending on whether they are growing in the light or in dark. The developmental program followed by the seedlings growing underneath the soil is called skotomorphogenesis, a dark-grown developmental program, characterized by long hypocotyl, closed cotyledons and apical hook formation in order to protect the apical meristem. When the seedling emerges out from the soil, it undergoes a different developmental program in response to light called photomorphogenesis, which is characterized by the opening of the apical hook, the greening of the cotyledons and the inhibition of hypocotyl growth (reviewed in Neff *et al.*, 2000). It has been reported that in many plants photomorphogenesis is the default developmental pathway after germination and skotomorphogenesis has evolved to partly inhibit photomorphogenesis in the dark until

the seedlings reach appropriate light conditions. This hypothesis is based on the fact that many lower plants lack etiolated development and directly undergo photomorphogenesis (Wei *et al.*, 1994; Alabadí *et al.*, 2004).

Light is known to cause massive transcriptional reprogramming in *Arabidopsis* (Mazzella et al., 2005). Microarray analysis of *Arabidopsis* has shown that around one-third of the genes exhibit altered expression in response to light (Ma *et al.*, 2001). In the same study, it was also found that the genes up-regulated or down-regulated in response to light are involved in at least 26 cellular metabolic pathways. Plants are able to detect the quality and quantity of light and use this external signal to optimize their development. There are several photoreceptors in plants to perceive various wavelengths and intensity of light, namely the phytochromes, cryptochromes, and phototropins (Quail *et al.*, 1995; Lin and Shalitin, 2003; Goh, 2009). Phytochromes are the best characterized photoreceptors having impact on the light-regulated development of plants (Schafer and Bowle, 2002; Han *et al.*, 2007). There are five genes encoding Phytochrome (PHY) proteins in *Arabidopsis*: *PHYA*, *PHYB*, *PHYC*, *PHYD*, and *PHYE* (Sharrock and Quail, 1989; Clack et al., 1994) that show individual as well as overlapping functions.

Phytochromes are chromoproteins that exist in an inactive form (Pr) in the cytosol under dark. Upon exposure to light they absorb red light and convert into an active form (Pfr) that translocate to the nucleus (Sharrock and Quail, 1989; Kircher *et al.*, 1999; Yamaguchi *et al.*, 1999). Functionally characterized transcription factors that act downstream of the phytochromes are the *PHYTOCHROME-INTERACTING-FACTORS* (*PIFs*), and belong to the bHLH family (Leivar *et al.*, 2012). There are seven homologs of *PIFs* in *Arabidopsis* out of which four- PIF1, PIF3, PIF4, and PIF5 accumulate in the dark and maintain skotomorphogenesis (Leivar *et al.*, 2008; Shin *et al.*, 2009). In light, the photoactivated form of phytochromes directly interacts with PIFs proteins in the nucleus and facilitates the degradation of PIFs (Leivar *et al.*, 2008; Ni *et al.*, 2014).

Photomorphogenesis in plants is also regulated by various hormones. It has been reported that a reduction in the endogenous levels of gibberellins (GA) in *Arabidopsis* can cause the de-etiolation of seedlings growing in dark (Alabadí *et al.*, 2004). GA has also been found to enhance the expression of *HOOKLESS1* (*HLS1*), an ethylene-induced gene essential for apical hook formation in the dark (Lehman *et al.*, 1996; Ogawa *et al.*, 2003). In the GA signaling pathway, DELLA proteins act as negative regulators of GA and repress almost all GA responses. There are five genes in *Arabidopsis* synthesizing DELLA proteins: *GA INSENSITIVE* (*GAI*), *REPESSOR OF GAI-3* (*RGA*), *RGA-LIKE1* (*RGL1*), *RGL2* and *RGL3* (Sun, 2008). De-etiolated phenotypes of dark-grown *ga1-3* mutants can be recovered in *rga* and

gai mutant background, indicating that *RGA* and *GAI* are the two main DELLA members involved in repression of GA-dependent skotomorphogenic development of seedling, in dark (Alabadí *et al.*, 2004; Achard *et al.*, 2007a). The abundance of GAI and RGA protein during the day inhibit the growth of the hypocotyl while the abundance of PIF3 during the night promotes its growth (Arana *et al.*, 2011; Soy *et al.*, 2012). DELLAs mediate the degradation of PIFs via the ubiquitin-protease system (Li *et al.*, 2016). The accumulation of the DELLA and PIFs proteins in a diurnal manner coordinates the light- and GA-mediated growth.

In the present study, we characterized the role of three *DRIF* homologs in *Arabidopsis*, *DRIF3*, *DRIF4*, and *DRIF5* in seedling development under dark and in light. We found that these three homologs are functionally redundant and involved in the formation of the apical hook in the dark. The *drif345* triple mutant showed loss of apical hook in etiolated seedlings and shorter hypocotyl in light. Further, we explored the expression of other genes, which regulate the growth and development of seedlings in dark and light and we found that expression of *HLS1* was decreased in dark-grown *drif345* seedlings while the expression of *PHYB* was increased in light-grown seedlings.

5.3 Material and Methods

5.3.1 Plant Material and Growth Conditions

The *drif3* (SAIL_636_A08), *drif4* (SALK_204381C), and *drif5* (SALK_1211145C) mutants were obtained from the Nottingham Arabidopsis Stock Centre (NASC, UK). Double mutants *drif34*, *drif35*, *drif45*, and triple mutant *drif345* were generated in our laboratory by multiple crossing and genotyped using the primers in Table.4.1. Three-month-old synchronized seeds of the *drif* mutants along with Columbia (Col-0) ecotype were used in all the experiments. Seeds were surface sterilized with 70% (v/v) ethanol, 0.5% (w/v) SDS and germinated on half-strength Murashige and Skoog (1/2 MS) salts media plates with 1.2% (w/v) agar without sucrose following stratification for 4 days at 4 °C in the dark. For the analysis of apical hook development, after stratification, plates were exposed to light at 22 °C for an hour (h) and then transferred to dark. Experiments were performed under long days (LD-16 h light/8 h dark), short days (SD-8 h light/16 h dark) with continuous white light (80 µmol $m^{-2}s^{-1}$), in continuous red light (RL) (15 µmol $m^{-2}s^{-1}$), far-red light (FR) (5 µmol $m^{-2}s^{-1}$), blue light (BL) (20 µmol $m^{-2}s^{-1}$) and in continuous dark to measure hypocotyl growth.

5.3.2 Apical Hook and Hypocotyl Measurements

Two to four-day-old dark-grown seedlings of WT and single, double, and triple *drif* mutants were used to measure the hook angle. At least 50 seedlings were imaged after 60 hours of germination and the angle between the hypocotyl axis and cotyledons was measured using ImageJ software (http://rsb.info.nih.gov/ij). The rose diagram was obtained by using online software (https://geographyfieldwork.com/). The opening of the apical hook of *drif345* and WT was measured at 48, 60, 72 and 96 h and the mean value of the hook angle of around 40 seedlings for each genotype was used to plot a graph.

Seedlings of *drif345* and WT, grown under different photoperiods, different wavelengths of light (red, far-red, and blue), and in dark were imaged every 24 h till 8 days after germination and the hypocotyl length was measured. The mean of around 40-50 seedlings for each genotype was used to compare the difference between the hypocotyl length of *drif345* and WT. For statistical significance, a t-test was performed between the mean of *drif345* and WT at each time point (Kim, 2015). Differences were considered statistically significant at $p \leq 0.05$.

5.3.3 RNA Isolation and Gene Expression Analysis

Total RNA was isolated from 48 h dark-grown seedlings, 5 and 10-day-old LD-grown seedlings using the CTAB/LiCl method with slight modification (Chang *et al.*, 1993) and quantified on the NanoDrop Microvolume Spectrophotometer (ThermoFischer). Around 500 ng of total RNA was incubated with DNase I to remove genomic DNA contamination and first-strand cDNA was synthesized using Superscript IV enzyme (Invitrogen) and oligo-dT according to the manufacturer's manual. Resulting cDNA was diluted 25x before being used for semi-quantitative and quantitative real-time polymerase chain reaction (qRT-PCR).

Expression of *DRIF3*, *DRIF4,* and *DRIF5* was analyzed on the 10th day in LD grown WT seedlings by qRT-PCR. The expression of *HLS1*, *PHYB*, *PIF1*, *PIF3*, *PIF4*, *GAI*, *RGA*, and *RGL2* was analyzed in dark and light-grown seedlings of WT and *drif345* by using semi-quantitative PCR. The list of primers used in qRT-PCT and semi-quantitative RT-PCR are mentioned in Table.4.1.

5.4 Results

5.4.1 Analysis of diurnal expression of *DRIF3*, *DRIF4*, and *DRIF5*

In *Arabidopsis*, there are five *DRIF* homologs that encode putative MYB-family transcription factors (Raimundo *et al.*, 2018). Among these homologs, *DRIF1* and *DRIF2* are expressed more in seed

and flowers, respectively, while *DRIF3*, *DRIF4,* and *DRIF5* are expressed throughout the plant development and are nested together in a phylogenetic study (Klepikova *et al.*, 2016; Raimundo *et al.*, 2018). In order to know the role of *DRIF3*, *DRIF4,* and *DRIF5* in seedling growth and development in response to light and dark we examined the expression of these three *DRIFs* in LD grown seedlings of WT during the course of a day (Fig.1).

The expression of *DRIF3* throughout the day and night was almost constant but had a peak at the end of the day at 16 h. *DRIF4* showed an increase in expression at 8 h till 16 h. The expression of *DRIF4* decreased during the night. The expression pattern of *DRIF5* was similar to the one of *DRIF4*. The change in the expression levels of *DRIF3*, *DRIF4*, and *DRIF5* in light and dark gave us a clue that they might have some function in light-mediated growth of the seedling. The overlapping pattern of expression of these homologs suggested the functional redundancy among them.

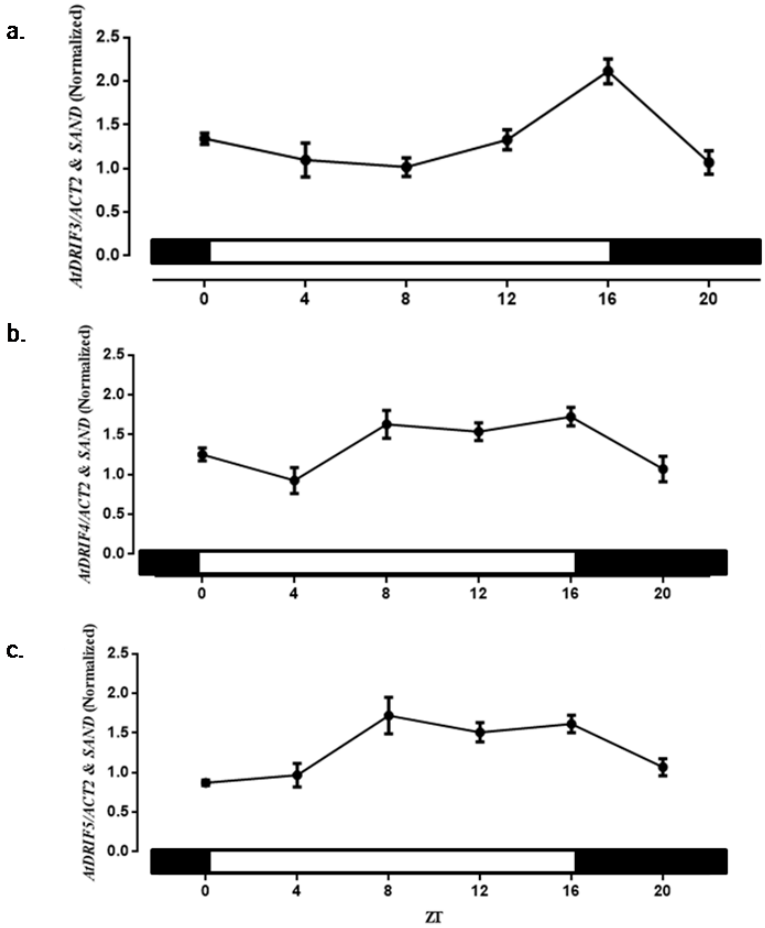


Fig.1. Relative expression pattern of *DRIFs* in *Arabidopsis thaliana* during long days. Expression of (a) *DRIF3*, (b) *DRIF4*, and (c) *DRIF5* were analyzed using 10-days old seedling grown in LD. Around 30 seedlings were collected at 0, 4, 8, 12, 16 and 20 h. *ACTIN2* was used as an internal control. Data represent the mean value of three technical replicates. Error bar indicates the SE. The black and white blocks at the x-axis represent dark and light phases, respectively.

5.4.2 *drif345* mutant showed hookless phenotype in dark

In order to know the effect of *DRIF3*, *DRIF4,* and *DRIF5* on the skotomorphogenic development of the seedling we examined the phenotype of dark-grown *drifs* mutants. After 24 h of germination in the dark, *drif* mutants and WT showed emergence of the radicle. After 48 h of germination, the emergence of the hypocotyl with closed shoot apex was observed in WT, *drif3*, *drif4*, *drif5*, *drif45*. In case of *drif34* and *drif35* double mutants, there were about 20 percent of seedlings showing opened hook with an angle varying from 90 to 30°. The apical hook of the seedlings of *drif345* triple mutant was already opened at 48 h, showing hook angle between 100 to 20° with no seedling showing a closed hook with 180° angle.

After 60 h of germination in the dark, the seedlings of WT still maintained the apical hook. Around 50 percent of seedlings of WT showed completely closed hook with the angle 180 to 175° (Fig.2.a and 3.a) while the seedlings of *drif3*, *drif4,* and *drif5* already started losing the curvature (Fig.2.b, c, and d). The apical hook angle of *drif34, drif35*, and *drif45* started reducing after 60 h of germination. The seedlings of *drif35* showed more reduction in hook angle followed by *drif34* and *drif45* (Fig.2. e, f, and g). At this time point, the majority of the seedlings of *drif345* mutant showed completely open apical hook (Fig.2.h and 3.a). These results suggested that *DRIF3*, *DRIF4* and *DRIF5* are redundant in their involved in apical hook formation.

After 72 h, the apical hook started opening in WT and after 96 h, the majority of the seedlings of WT had an opened apical hook with closed, yellow cotyledons and long elongated hypocotyls. In case of *drif345* mutant, the apical hook of all the seedlings was completely opened at 72 h (Fig.3.b). Although *drif345* showed a hookless phenotype, the hypocotyl length of *drif345* seedling was same as WT at 72 and 96 h after germination in the dark. The cotyledons of *drif345* were yellow and closed (Fig.3.a)

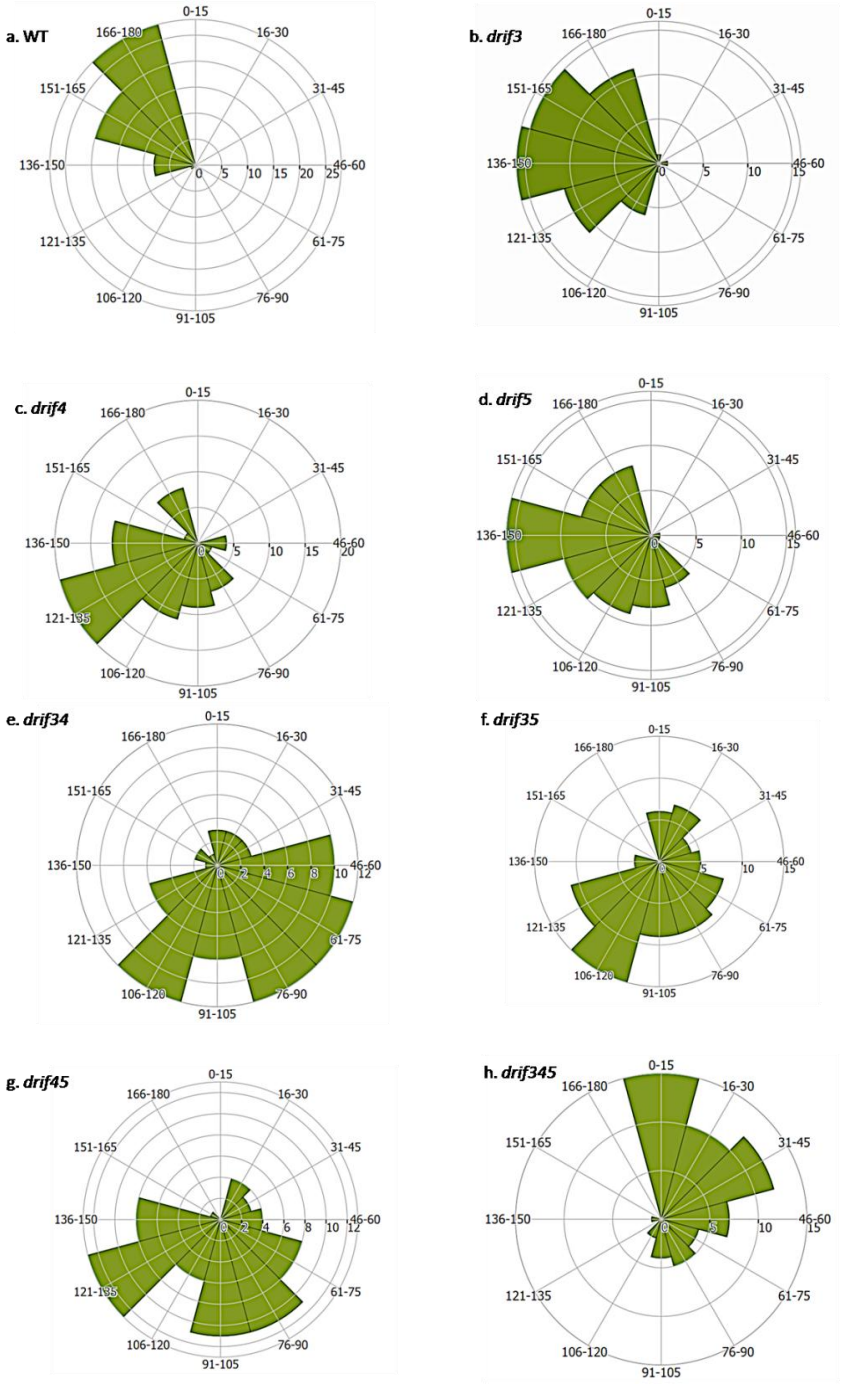


Fig.2. **Effect of *DRIF3*, *DRIF4,* and *DRIF5* in the apical hook formation.** The apical hook angle of the 60 h dark-grown seedlings of (a). WT, (b). *drif3*, (c). *drif4*, (d). *drif5*, (e). *drif34*, (f). *drif35*, (g). *drif45*, and (h). *drif345*. The numbers on the periphery represent the hook angle and the numbers on radius represent the number of seedlings.

These observations suggested that although *drif345* failed to show completely closed apical hook, it maintained the other characteristic features of etiolated seedlings, e.g. yellowish and closed cotyledons and long hypocotyl.

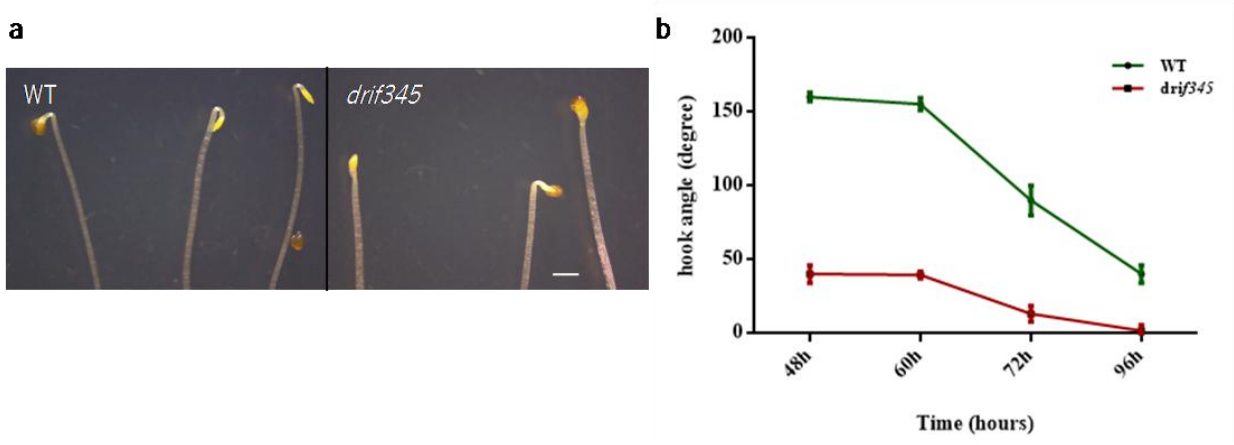


Fig.3. Phenotypic analysis of dark-grown *drif345* mutant. (a) The apical hook of WT and *drif345* grown in the dark for 60 h after germination. Scale bar 1 mm. (b) The time course of the hook opening was examined by measuring the angle between the hypocotyl axis and cotyledons. The graph represents the mean value of the hook angle of 40 seedlings of WT and *drif345* at each time point ± SEM.

5.4.3 Expression of *HLS1* was reduced in *drif345*

The apical hook structure in dark-grown seedling is the result of differential cell growth of the apical zone of the hypocotyl. *HLS1* is a key gene involved in differential growth in the apical hook in response to ethylene (Guzmán and Ecker, 1990; Lehman *et al.*, 1996). Recently, it has been reported that *PIFs* also activates the transcription of *HLS1* and acts in parallel to *ETHYLENE INSENSITIVE3* (*EIN3*), a key regulator of *HLS1* (Zhang *et al.*, 2018). The role of *PIFs* is encountered by DELLA proteins in the presence of light and low GA levels (Li *et al.*, 2016).

After the phenotypic characterization of *drif345*, we found that the *drif345* showed a hookless phenotype in the dark. In order to understand the role of *DRIFs* at the molecular level, we checked the expression of *HLS1*, along with *PHYB*, *PIF1*, *PIF3*, *PIF4*, *GAI*, *RGA*, and *RGL2* in dark-grown seedlings of *drif345* and *WT*, after 48 h of germination (Fig.4).

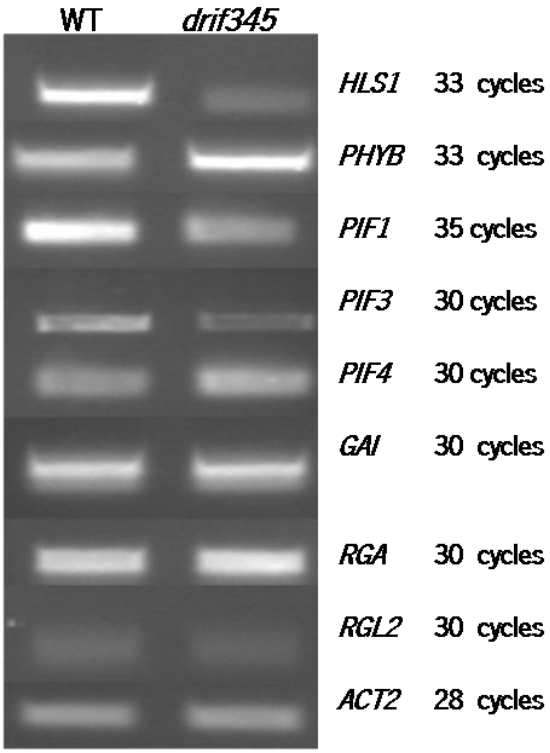


Fig.4. RT-PCR analysis of genes involved in light-mediated early seedling morphogenesis. Expression of *HLS1*, *PHYB*, *PIF1*, *PIF3*, *PIF4*, *GAI1*, *RGA*, and *RGL2* were analyzed using around 30 seedlings of WT and *drif345* after 48 h germination in dark. *ACTIN2* (*ACT2*) was used as an internal control.

We found that the expression of *HLS1* was reduced in the *drif345* mutant when compared to WT. The expression of *PHYB* was slightly higher while the expression of *PIF1* was low in *drif345*. There was no change observed in the expression level of *GAI*, *RGA*, and *RGL2*. A low level of expression of *HLS1* in *drif345* mutant can explain the hookless phenotype of *drif345*. Our results showed that *DRIFs* may be one of the molecular components that regulate the development of the apical hook in the dark.

5.4.4 Reduced hypocotyl elongation of *drif345* in light

Precise regulation of hypocotyl elongation in response to various external and internal factors is crucial for plant growth and survival. Light is one of the negative regulators of hypocotyl growth (Yang *et al.*, 2009; Liu *et al.*, 2013; Ma *et al.*, 2018). Previously, it has been reported that a *RAD* homolog-*RSM1* and a *DIV-like* gene *MYBH* regulate hypocotyl growth in response to light (Hamaguchi *et al.*, 2008; Kwon *et al.*, 2013).

In order to check the difference in diurnal growth of *drif345* and WT, we measured the hypocotyl length after the completion of each LD (16h L/8 h D), SD (8 h L/16 h D) cycle, and in dark. After 24 h there was only radicle emergence in *drif345* and WT. So, we measured the hypocotyl length of seedling only on 2nd day and continued till 8th day when hypocotyl length reached a maximum for both *drif345* and WT. In this study, we observed that the exposure of the seedling for a longer duration (LD) of light inhibits the hypocotyl growth of WT and average hypocotyl length of WT plants grown in LD was just 2-2.5 mm (Fig.5.a) in comparison to the dark-grown WT seedlings with average hypocotyl length 15-16 mm (Fig.5c). In the case of *drif345*, the length of the hypocotyl was shorter than WT in all conditions (Fig.5.a, b, and c). Further, we observed that even 8 h exposure in light during the day could significantly reduce the growth of hypocotyl in *drif345* (Fig.5b). In the case of LD the significant difference between the hypocotyl length of WT and *drif345* was observed at 5th day (Fig.5.a) while in dark, the significant difference between the hypocotyl growth was observed in 7th day (Fig.5.c). These results suggest that *DRIFs* may be involved in promoting hypocotyl elongation.

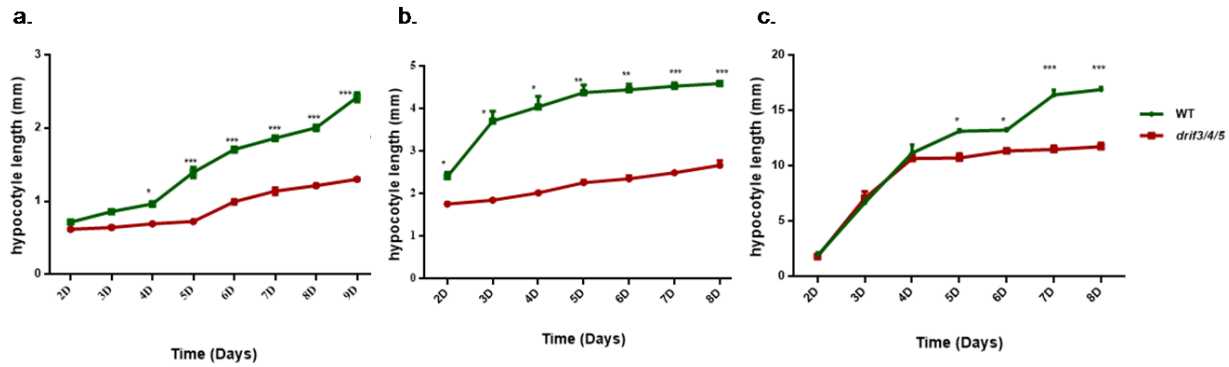


Fig.5. Hypocotyl elongation of *drif345* in different photoperiods. Seedlings of WT and *drif345* germinated in (a) Long day, (b) Short day, and (c) in dark and hypocotyl length of around 50 seedlings were measured after the establishment of seedling and continued till 8 d after every 24 h intervals. The graph represents the mean of hypocotyl length of seedlings of each genotype ± SE. t-test was performed to calculate the significant difference between the hypocotyl length of WT and *drif345*. P≤0.05 (*),P≤0.001 (**), and P≤0.0001 (***).

In order to analyze the effect of *DRIFs* on hypocotyl elongation in different wavelengths of light, we examined the growth of *drif345* and WT in continuous red (RL), blue (BL) and far-red (FR) light. In the previous experiment, we observed a significant difference between the hypocotyl length of 5-days-old seedlings of WT and *drif345* (Fig.5). So, in this experiment, we measured the hypocotyl length of *drif345* and WT on the 5th day. We observed the reduction in hypocotyl length of *drif345* in all the light conditions tested (Fig. 6). The length of the hypocotyl of *drif345* was almost half of the length of the WT hypocotyl in blue light and red light. In far-red light, the hypocotyl of *drif345* was only 1/3 of the WT

(Fig.6d and g. In dark, the difference in the hypocotyl length of WT and *drif345* was not as significant as in different light conditions. These results again supported our hypothesis about the involvement of *DRIFs* in hypocotyl elongation in response to light.

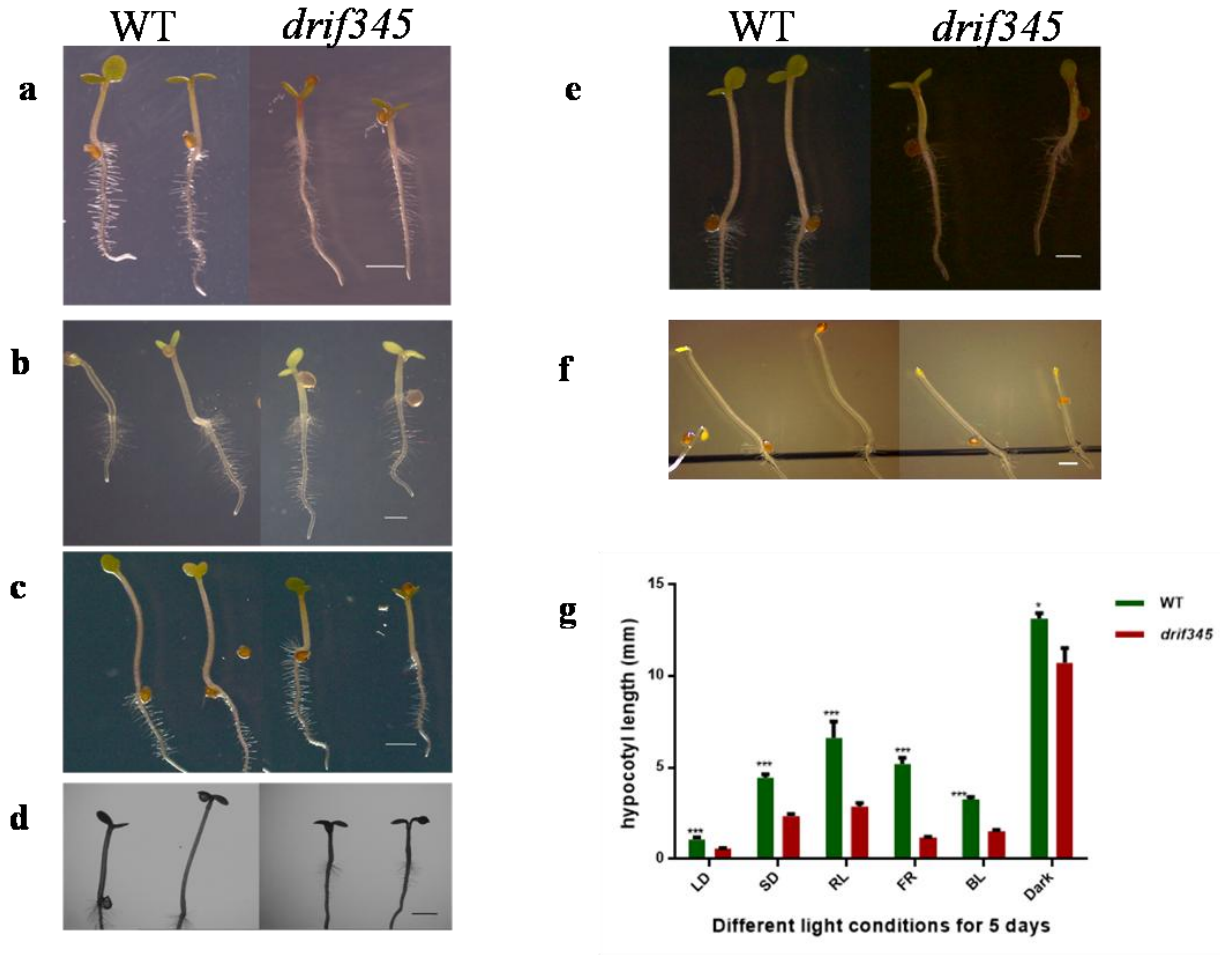


Fig.6 Hypocotyl elongation of *drif345* in different light conditions. Seedlings of WT and *drif345* germinated in (a) LD, (b) SD, (c) continuous red light (RL), (d) continuous far-red light (FR), (e) continuous blue light (BL), and (f) continuous dark for 5 days. Hypocotyl length of around 50 seedlings of each genotype in every condition was measured using ImageJ. Scale bar 1 mm. (g) The graph represents the mean of hypocotyl length ± SEM. A t-test was performed to find the significant difference between the hypocotyl lengths of WT and *drif345*. P≤0.05 (*), P≤0.001 (**), and P≤0.0001 (***)

5.4.5 Effect of application of GA3 on *drif345* seedlings

Previous studies have shown that gibberellins are involved in hypocotyl growth via cell elongation, and the shorter hypocotyl of GA-deficient mutants could be recovered by the exogenous application of GA. For instance, the application of GA can restore the hypocotyl length of*ga1* mutant (Cowling and Harberd, 1999). Therefore, to test whether exogenous GA could rescue the shorter hypocotyl phenotype of *drif345*, seeds of *drif345* and WT were sown in medium containing 0.1, 1, 10,

and 100 µM of GA3 along with control (0.1% (v/v) ethanol). Seeds were stratified for 4 days in the dark at 4 °C and then transferred to LD for 5 days.

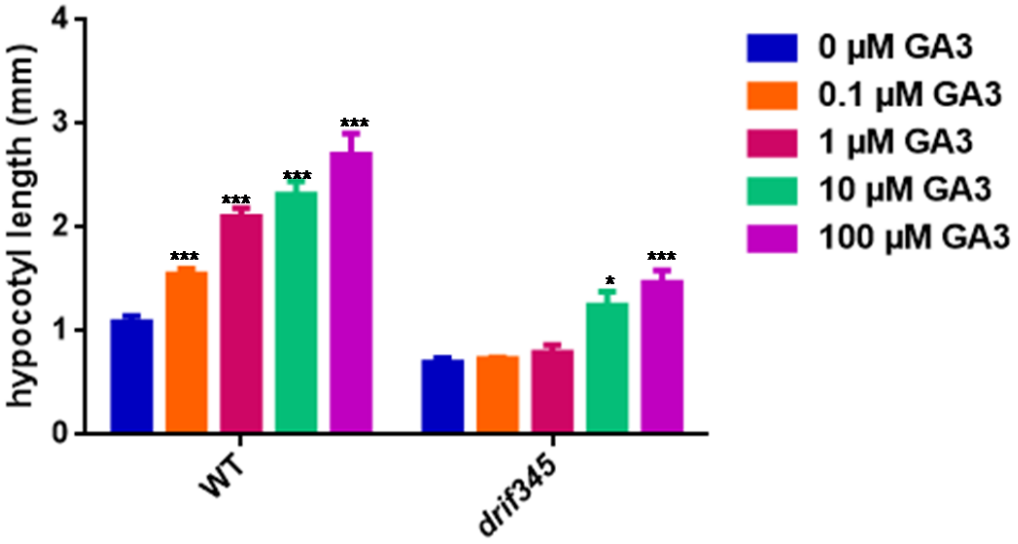


Fig.7. Effect of GA3 on hypocotyl elongation of *drif345*. Seeds of WT and *drif345* were germinated on the agar plate with ½ MS salts supplemented with (0, 0.1, 1, 10, and 100 µM) concentration of GA3 in LD photoperiod. Hypocotyl length of 5 day-old seedlings of *drif345* and WT were measured. The graph represents the mean of hypocotyl length of around 50 seedlings *drif345* and WT for each treatment ± SE. t-test was performed to find the significant difference between the hypocotyl lengths of WT and *drif345*. P≤0.05 (*), P≤0.001 (**), and P≤0.0001 (***)

We observed that very low concentrations of GA3 (0.1 and 1 µM) did not have any effect on hypocotyl growth of *drif345* while the higher concentrations of GA3 (10 and 100 µM) could increase the hypocotyl length of *drif345*. None of the GA3 concentrations, used in this experiment, could rescue the hypocotyl length of *drif345* to the same extent as WT in the same conditions, but in 10 and 100 µM GA3, the hypocotyl length of *drif345* became similar to the hypocotyl length of WT without GA3 (Fig.7).

The results showed that *drif345* mutant is not responsive to the lower concentration of GA3. A higher concentration of GA3 could recover the length of hypocotyl, which shows that the *drif345* is not completely insensitive to GA.

5.4.6 Expression of *PHYB* was increased in *drif345* seedling

Previous reports showed that overexpression of *PHYB* (Wagner *et al.*, 1991) and mutation in *pifs* exhibit reduced hypocotyl in light (Zhong *et al.*, 2012). So, we examined the expression of *PHYB*, *PIF1*, *PIF3*, *PIF4* in 5-day-old seedlings of *drif345* and WT grown in LD as the significant difference between the hypocotyl length of these two genotypes was clear at this time point. We also analyzed the expression of *GAI*, *RGA*, and *RGL2* because DELLAs are the negative regulators of *PIFs* in light. So, we wanted to know if there is any change in the DELLAs at the transcriptional level in *drif345*.

In this study, we found that the expression of *PHYB* was higher in *drif345* that WT (Fig.8). In two days old dark-grown seedling also the expression of *PHYB* was higher in *drif345*. We also observed that the expression of *PIF3* was reduced while the expression of *GAI* was increased in *drif345* (Fig.8). This expression analysis suggested a role for *DRIFs* in integration with *PHY* and *PIFs* to mediate the light-regulated growth of hypocotyl in *Arabidopsis*.

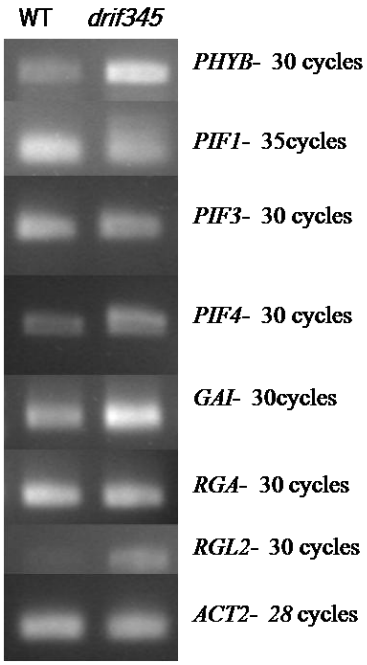


Fig.8. RT-PCR analysis of genes involved in light-mediated early seedling morphogenesis. Expression of *PHYB*, *PIF1*, *PIF3*, *PIF4*, *GAI1*, *RGA,* and *RGL2* was analyzed using around 30 seedlings of WT and *drif345* after 5 days of germinaiton in LD condition. *ACTIN2* (*ACT2*) was used as an internal control.

5.5 Discussion

In this study, we present the role of *DRIF3*, *DRIF4* and *DRIF5* during the skotomorphogenic and photomorphogenic development of seedlings. In *Arabidopsis*, there are twelve *DIV*, six *RAD*, and five

DRIF homologs but the physiological function of only few *DIVs* and *RADs* are known (Hamaguchi *et al.*, 2008; Fang *et al.*, 2018; Yang *et al.*, 2018). The current understanding of the *RAD* and *DIV* antagonism is based on their role in the bilateral symmetry of the *Antirrhinum* flower (Raimundo *et al.*, 2013). Also, overexpression of tomato *RAD* homolog *FSM1* reduced the size of fruit through the suppression of cell expansion of the pericarp, which is potentially a result of competition between FSM1 protein with tomato DIV protein MYB1, for the binding with tomato DRIF protein FSB1 (Machemer *et al.*, 2011).

The expression analysis of *DRIF3*, *DRIF4,* and *DRIF5* showed a change in expression level during the LD cycle. This result suggested that their expression may be light-regulated. Further, the hookless dark-grown seedling and shorter hypocotyl length of light-grown seedlings of *drif345* confirm the involvement of these *DRIFs* in light-regulated development of seedlings. The overlapping pattern of the expression of *DRIFs* and no significant phenotypic difference between *drif* single mutant and WT indicates functional redundancy among these genes.

In *Arabidopsis*, the hookless phenotype and shorter hypocotyl of the *drif345* mutant observed in this study are similar to the overexpression of *Arabidopsis RAD* homolog, *RSM1* (Hamaguchi *et al.*, 2008). Transgenic *Arabidopsis* overexpressing *Antirrhinum RAD,* also showed a hookless phenotype in dark-grown seedling (Costa *et al.*, 2005). Transgenic *Arabidopsis* overexpressing *RSM1* showed reduced expression of *HLS1*, but the authors could not find a direct connection between *RSM1* and *HLS1* (Hamaguchi *et al.*, 2008). In our study, we also found reduced expression of *HLS1* in the dark-grown seedlings of the *drif345* mutant. The activity of *HLS1* is required for the expression of auxin-responsive genes in the apical hook region of dark-grown seedlings of *A. thaliana* (Lehman *et al.*, 1996; Li *et al.*, 2004). In *RSM*-ox lines, the expression of *DR5*::*GUS* was not confined to the inner side of the apical curvature of dark-grown seedlings, which suggested that this small MYB-like protein is involved in the asymmetric distribution of auxin during the development of the apical hook (Hamaguchi *et al.*, 2008). In tomato, the formation of the DIV-DRIF (MYB1-FSB1) complex in competed by RAD-like protein FSM1 (Machemer *et al.*, 2011). In this way, FSM1/FSB1/MYB1 regulates the cell expansion of pericarp and fruit size in tomato (Machemer *et al.*, 2011). In *Arabidopsis, RSM1* overexpression induces hookless phenotype in dark and shorter hypocotyl in red light (Hamaguchi *et al.*, 2008). In this study, we found that *drif345* mimics the phenotype of *RSM1* overexpression. It has been reported that *Arabidopsis* DRIF homologs physically interact with RAD and DIV homologs. It may be possible that the early development of seedling is regulated by RAD/DRIF/DIV in *Arabidopsis*. In the absence of *DRIFs*, transcriptional activation of genes governed by DIV-DRIF heterodimers, perhaps *HLS1,* may be disrupted. In the case of

overexpression of *RSM1*, the DIV-DRIF heterodimers may get disrupted and therefore the overexpression of a RAD protein phenocopies the effect of the *drif345* mutations.

Overexpression of a *DIV-like* gene, *MYBH*, showed an increased hypocotyl in light conditions, but the levels of GA were not increased in these lines (Kwon *et al.*, 2013). On the contrary, the overexpression of *RSM1* showed reduction of hypocotyl length in dark and red light conditions (Hamaguchi *et al.*, 2008). Recently, it was observed that the RSM1 protein interacts with HY5 protein- a positive regulator of light, and it regulates seedling development (Yang *et al.*, 2018) while the overexpression of *MYBH* showed higher expression of *PIF3* and *PIF4*, which negatively regulates the effect of light (Huq and Quail, 2002; Kim *et al.*, 2003; Kwon *et al.*, 2013). In this study, the shorter hypocotyl length of *drif345* again mimics the *RSM1* overexpression phenotype. The lower expression of *PIF3* and higher expression of *PHYB* and *GAI* in *drif345* suggests the involvement of these *DRIF* homologs in negative regulation of light response.

The gene expression analysis showed that the expression of *PHYB* was higher in *drif345* when the seeds were germinated in dark for 48 h or in LD for 5 days. Overexpression of *PHYB* showed shorter hypocotyl in light (Wagner *et al.*, 1991). The expression of two *DELLA* genes, *GAI1* and *RGL2* in *drif345* was similar to WT when seeds were germinated in dark for 48 h. The 5 d old seedling of *drif345* showed higher level of expression than WT (Fig.4 and 8). DELLAs controls the growth of hypocotyl in response to light and its level changes in response to light and GA during day/night cycle (Achard *et al.*, 2007b). From the higher expression of *GAI* and *RGL2* in 5 days old seedling of *drif345*, we speculate that *drif345* may have higher DELLA proteins, which reduce the growth of mutant. It might be a possibility that *DRIFs* are involved somehow in DELLA degradation and the absence of these genes causes higher DELLA levels. The expression of *PIF1* was reduced in *drif345* in seeds germinated in dark and LD conditions. It has been shown that *PIFs* regulates light responses downstream of phytochromes (Leivar *et al.*, 2008; Shin *et al.*, 2009). Higher level of expression of *PHYB* in *drif345* is likely to responsible for lower expression of *PIF1*.

In future, the investigation of *DRIFs* in the mutants of *phys*, *pifs*, and *dellas* will provide us a great opportunity to understand the role of *DRIFs* during the light-mediated transition from skotomorphogenesis to photomorphogenesis. Our results along with previously characterized *RAD* and *DIV* homologs suggest that the genes of the DDR regulatory module may have been co-opted for various developmental processes in *Arabidopsis*.

5.6 References

Achard, P., Liao, L., Jiang, C., Desnos, T., Bartlett, J., Fu, X. & Harberd, N. P. (2007a). DELLAs Contribute to Plant Photomorphogenesis. *Plant Physiology* 143(3): 1163-1172.

Achard, P., Liao, L., Jiang, C., Desnos, T., Bartlett, J., Fu, X. & Harberd, N. P. (2007b). DELLAs contribute to plant photomorphogenesis. *Plant Physiol* 143(3): 1163-1172.

Agarwal, M., Hao, Y., Kapoor, A., Dong, C. H., Fujii, H., Zheng, X. & Zhu, J. K. (2006). A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *J Biol Chem* 281(49): 37636-37645.

Alabadí, D., Gil, J., Blázquez, M. A. & García-Martínez, J. L. (2004). Gibberellins Repress Photomorphogenesis in Darkness. *Plant Physiology* 134(3): 1050-1057.

Arana, M. V., Marín-de la Rosa, N., Maloof, J. N., Blázquez, M. A. & Alabadí, D. (2011). Circadian oscillation of gibberellin signaling in Arabidopsis. *Proceedings of the National Academy of Sciences* 108(22): 9292-9297.

Chang, S., Puryear, J. & Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 11(2): 113-116.

Cheng, H., Song, S., Xiao, L., Soo, H. M., Cheng, Z., Xie, D. & Peng, J. (2009). Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in Arabidopsis. *PLoS Genet* 5(3): e1000440.

Clack, T., Mathews, S. & Sharrock, R. A. (1994). The phytochrome apoprotein family inArabidopsis is encoded by five genes: the sequences and expression ofPHYD andPHYE. *Plant Mol Biol* 25(3): 413-427.

Costa, M. M. R., Fox, S., Hanna, A. I., Baxter, C. & Coen, E. (2005). Evolution of regulatory interactions controlling floral asymmetry. *Development* 132(22): 5093-5101.

Cowling, R. J. & Harberd, N. P. (1999). Gibberellins control Arabidopsis hypocotyl growth via regulation of cellular elongation. *Journal of Experimental Botany* 50(337): 1351-1357.

Fang, Q., Wang, Q., Mao, H., Xu, J., Wang, Y., Hu, H., He, S., Tu, J., Cheng, C., Tian, G., Wang, X., Liu, X., Zhang, C. & Luo, K. (2018). AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Rep* 37(11): 1499-1511.

Goh, C. H. (2009). Phototropins and chloroplast activity in plant blue light signaling. *Plant Signal Behav* 4(8): 693-695.

Guzmán, P. & Ecker, J. R. (1990). Exploiting the triple response of Arabidopsis to identify ethylene-related mutants. *The Plant Cell* 2(6): 513-523.

Hamaguchi, A., Yamashino, T., Koizumi, N., Kiba, T., Kojima, M., Sakakibara, H. & Mizuno, T. (2008). A small subfamily of Arabidopsis RADIALIS-LIKE SANT/MYB genes: a link to HOOKLESS1-mediated signal transduction during early morphogenesis. *Biosci Biotechnol Biochem* 72(10): 2687-2696.

Han, Y.-J., Song, P.-S. & Kim, J.-I. (2007). Phytochrome-mediated photomorphogenesis in plants. *Journal of Plant Biology* 50(3): 230-240.

Huq, E. & Quail, P. H. (2002). PIF4, a phytochrome-interacting bHLH factor, functions as a negative regulator of phytochrome B signaling in Arabidopsis. *Embo j* 21(10): 2441-2450.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B. & Martin, C. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *Embo j* 19(22): 6150-6161.

Kim, J., Yi, H., Choi, G., Shin, B., Song, P.-S. & Choi, G. (2003). Functional Characterization of Phytochrome Interacting Factor 3 in Phytochrome-Mediated Light Signal Transduction. *The Plant Cell* 15(10): 2399-2407.

Kim, T. K. (2015). T test as a parametric statistic. *Korean J Anesthesiol* 68(6): 540-546.

Kircher, S., Kozma-Bognar, L., Kim, L., Adam, E., Harter, K., Schäfer, E. & Nagy, F. (1999). Light Quality–Dependent Nuclear Import of the Plant Photoreceptors Phytochrome A and B. *The Plant Cell* 11(8): 1445-1456.

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D. & Penin, A. A. (2016). A high resolution map of the Arabidopsis thaliana developmental transcriptome based on RNA-seq profiling. *The Plant Journal* 88(6): 1058-1070.

Kwon, Y., Kim, J. H., Nguyen, H. N., Jikumaru, Y., Kamiya, Y., Hong, S. W. & Lee, H. (2013). A novel Arabidopsis MYB-like transcription factor, MYBH, regulates hypocotyl elongation by enhancing auxin accumulation. *J Exp Bot* 64(12): 3911-3922.

Lehman, A., Black, R. & Ecker, J. R. (1996). HOOKLESS1, an ethylene response gene, is required for differential cell elongation in the Arabidopsis hypocotyl. *Cell* 85(2): 183-194.

Leivar, P., Monte, E., Oka, Y., Liu, T., Carle, C., Castillon, A., Huq, E. & Quail, P. H. (2008). Multiple Phytochrome-Interacting bHLH Transcription Factors Repress Premature Seedling Photomorphogenesis in Darkness. *Current Biology* 18(23): 1815-1823.

Leivar, P., Tepperman, J. M., Cohn, M. M., Monte, E., Al-Sady, B., Erickson, E. & Quail, P. H. (2012). Dynamic antagonism between phytochromes and PIF family basic helix-loop-helix factors induces selective reciprocal responses to light and shade in a rapidly responsive transcriptional network in Arabidopsis. *Plant Cell* 24(4): 1398-1419.

Li, C., Zhou, Y. & Fan, L.-M. (2015). A novel repressor of floral transition, MEE3, an abiotic stress regulated protein, functions as an activator of FLC by binding to its promoter in Arabidopsis. *Environmental and Experimental Botany* 113: 1-10.

Li, H., Johnson, P., Stepanova, A., Alonso, J. M. & Ecker, J. R. (2004). Convergence of signaling pathways in the control of differential cell growth in Arabidopsis. *Dev Cell* 7(2): 193-204.

Li, K., Yu, R., Fan, L.-M., Wei, N., Chen, H. & Deng, X. W. (2016). DELLA-mediated PIF degradation contributes to coordination of light and gibberellin signalling in Arabidopsis. *Nature Communications* 7(1): 11868.

Lin, C. & Shalitin, D. (2003). Cryptochrome Structure and Signal Transduction. *Annu Rev Plant Biol* 54(1): 469-496.

Liu, X., Qin, T., Ma, Q., Sun, J., Liu, Z., Yuan, M. & Mao, T. (2013). Light-Regulated Hypocotyl Elongation Involves Proteasome-Dependent Degradation of the Microtubule Regulatory Protein WDL3 in Arabidopsis. *The Plant Cell* 25(5): 1740-1755.

Ma, L., Li, J., Qu, L., Hager, J., Chen, Z., Zhao, H. & Deng, X. W. (2001). Light Control of Arabidopsis Development Entails Coordinated Regulation of Genome Expression and Cellular Pathways. *The Plant Cell* 13(12): 2589-2607.

Ma, Q., Wang, X., Sun, J. & Mao, T. (2018). Coordinated Regulation of Hypocotyl Cell Elongation by Light and Ethylene through a Microtubule Destabilizing Protein. *Plant Physiology* 176(1): 678-690.

Machemer, K., Shaiman, O., Salts, Y., Shabtai, S., Sobolev, I., Belausov, E., Grotewold, E. & Barg, R. (2011). Interplay of MYB factors in differential cell expansion, and consequences for tomato fruit development. *Plant J* 68(2): 337-350.

Mandaokar, A. & Browse, J. (2009). MYB108 acts together with MYB24 to regulate jasmonate-mediated stamen maturation in Arabidopsis. *Plant Physiol* 149(2): 851-862.

Mazzella, M. A., Arana, M. V., Staneloni, R. J., Perelman, S., Rodriguez Batiller, M. J., Muschietti, J., Cerdán, P. D., Chen, K., Sánchez, R. A., Zhu, T., Chory, J. & Casal, J. J. (2005). Phytochrome

Control of the Transcriptome Anticipates Seedling Exposure to Light. *The Plant Cell* 17(9): 2507-2516.

Narsai, R., Law, S. R., Carrie, C., Xu, L. & Whelan, J. (2011). In-Depth Temporal Transcriptome Profiling Reveals a Crucial Developmental Switch with Roles for RNA Processing and Organelle Metabolism That Are Essential for Germination in Arabidopsis. *Plant Physiology* 157(3): 1342-1362.

Neff, M. M., Fankhauser, C. & Chory, J. (2000). Light: an indicator of time and place. *Genes Dev* 14(3): 257-271.

Ni, W., Xu, S.-L., Tepperman, J. M., Stanley, D. J., Maltby, D. A., Gross, J. D., Burlingame, A. L., Wang, Z.-Y. & Quail, P. H. (2014). A mutually assured destruction mechanism attenuates light signaling in Arabidopsis. *Science* 344(6188): 1160-1164.

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.

Petzold, H. E., Chanda, B., Zhao, C., Rigoulot, S. B., Beers, E. P. & Brunner, A. M. (2018). DIVARICATA AND RADIALIS INTERACTING FACTOR (DRIF) also interacts with WOX and KNOX proteins associated with wood formation in Populus trichocarpa. *Plant J* 93(6): 1076-1087.

Quail, P., Boylan, M., Parks, B., Short, T., Xu, Y. & Wagner, D. (1995). Phytochromes: photosensory perception and signal transduction. *Science* 268(5211): 675-680.

Raimundo, J., Sobral, R., Bailey, P., Azevedo, H., Galego, L., Almeida, J., Coen, E. & Costa, M. M. R. (2013). A subcellular tug of war involving three MYB-like proteins underlies a molecular antagonism in Antirrhinum flower asymmetry. *The Plant Journal* 75(4): 527-538.

Raimundo, J., Sobral, R., Laranjeira, S. & Costa, M. M. R. (2018). Successive Domain Rearrangements Underlie the Evolution of a Regulatory Module Controlled by a Small Interfering Peptide. *Mol Biol Evol* 35(12): 2873-2885.

Schafer, E. & Bowle, C. (2002). Phytochrome-mediated photoperception and signal transduction in higher plants. *EMBO Rep* 3(11): 1042-1048.

Sharrock, R. A. & Quail, P. H. (1989). Novel phytochrome sequences in Arabidopsis thaliana: structure, evolution, and differential expression of a plant regulatory photoreceptor family. *Genes Dev* 3(11): 1745-1757.

Shin, J., Kim, K., Kang, H., Zulfugarov, I. S., Bae, G., Lee, C.-H., Lee, D. & Choi, G. (2009). Phytochromes promote seedling light responses by inhibiting four negatively-acting phytochrome-interacting factors. *Proceedings of the National Academy of Sciences* 106(18): 7660-7665.

Soy, J., Leivar, P., González-Schain, N., Sentandreu, M., Prat, S., Quail, P. H. & Monte, E. (2012). Phytochrome-imposed oscillations in PIF3 protein abundance regulate hypocotyl growth under diurnal light/dark conditions in Arabidopsis. *The Plant Journal* 71(3): 390-401.

Sun, T. P. (2008). Gibberellin metabolism, perception and signaling pathways in Arabidopsis. *Arabidopsis Book* 6: e0103.

Wagner, D., Tepperman, J. M. & Quail, P. H. (1991). Overexpression of Phytochrome B Induces a Short Hypocotyl Phenotype in Transgenic Arabidopsis. *The Plant Cell* 3(12): 1275-1288.

Wei, N., Chamovitz, D. A. & Deng, X.-W. (1994). Arabidopsis COP9 is a component of a novel signaling complex mediating light control of development. *Cell* 78(1): 117-124.

Yamaguchi, R., Nakamura, M., Mochizuki, N., Kay, S. A. & Nagatani, A. (1999). Light-dependent Translocation of a Phytochrome B-GFP Fusion Protein to the Nucleus in Transgenic Arabidopsis. *The Journal of Cell Biology* 145(3): 437-445.

Yang, B., Song, Z., Li, C., Jiang, J., Zhou, Y., Wang, R., Wang, Q., Ni, C., Liang, Q., Chen, H. & Fan, L.-M. (2018). RSM1, an Arabidopsis MYB protein, interacts with HY5/HYH to modulate seed

germination and seedling development in response to abscisic acid and salinity. *PLOS Genetics* 14(12): e1007839.

Yang, S. W., Jang, I.-C., Henriques, R. & Chua, N.-H. (2009). FAR-RED ELONGATED HYPOCOTYL1 and FHY1-LIKE Associate with the Arabidopsis Transcription Factors LAF1 and HFR1 to Transmit Phytochrome A Signals for Inhibition of Hypocotyl Elongation. *The Plant Cell* 21(5): 1341-1359.

Zhang, X., Ji, Y., Xue, C., Ma, H., Xi, Y., Huang, P., Wang, H., An, F., Li, B., Wang, Y. & Guo, H. (2018). Integrated Regulation of Apical Hook Development by Transcriptional Coupling of EIN3/EIL1 and PIFs in Arabidopsis. *The Plant Cell* 30(9): 1971-1988.

Zhong, S., Shi, H., Xue, C., Wang, L., Xi, Y., Li, J., Quail, Peter H., Deng, Xing W. & Guo, H. (2012). A Molecular Framework of Light-Controlled Phytohormone Action in Arabidopsis. *Current Biology* 22(16): 1530-1535.

Concluding remarks

Understanding the emergence of novel morphological traits during the evolution has been a key thrust for evolutionary biologists. The molecular mechanism underlies the origin of new traits includes the network of genes/transcription factors that regulate the pattern of expression of several other downstream genes acting. Transcriptional control of these genes depends on the *cis*-regulatory elements that can change the spatial-temporal expression of genes and bring functional novelty. In the present study, the regulation of genes involved in flower zygomorphy and their functional characterization in actinomorphic species provides a way to understand how a gene regulatory network can evolve and adopt different functions in different species.

In **Chapter 2**, it is described that the *CYC-like* genes are the key players of the flower zygomorphy. The asymmetric expression of these genes in flower meristems is essential to establish dorsoventral asymmetry within the flowers of zygomorphic species. Like other legume species, *Medicago* also has two *CYC2-like* homologs that are expressed asymmetrically in the floral meristems and in dorsal petals while their expression in vegetative tissues is negligible. From this study, we also conclude that the *cis*-elements of *CYC2-like* genes in asymmetric species are different from the symmetric species. The tissue-specific expression of these genes may be determined by these *cis*-elements that may interact with different TFs present in the dorsal and ventral domain of floral meristem and act as suppressor or enhancer for *CYC2-like* genes.

In order to understand the function of *DIV*, *DRIF*, and *RAD* homologs and the possible role of the DDR module in *Arabidopsis thaliana*, an actinomorphic species, in **Chapter 3** we analyzed the expression and co-expression of these genes using various bioinformatics tools. This study showed that the *DIV*, *DRIF*, and *RAD* homologs are expressed in *Arabidopsis* during different developmental stages. We have also found that *DRIF3, DRIF4*, and *DRIF5* are expressed throughout the development of *Arabidopsis* and having overlapping expression pattern. The co-expression network of the genes expressing along with *DRIFs* in seeds and siliques are involved in hormonal signaling, light signaling and the promoter region of *DRIF* genes contains various hormones and environmental signals responsive elements. It suggested that the *DRIF* homologs may play some role in seed germination and seedling development in response to various external and internal factors.

The expression of *Arabidopsis DRIFs* in seeds and seedlings gave us a hint about their role in germination and seedling development. In **Chapter 4**, we found that *DRIF3*, *DRIF4*, and *DRIF5* are involved in the regulation of germination as the triple mutant *drif345* showed faster germination and a higher germination percentage in salt and heat stress. Further, this study showed that the in

germinating seeds of *drif345* the GA biosynthetic gene *GA3ox1* is more expressed whereas the opposite was observed for the ABA synthesis genes *ABA1* and *ABA3*. Also, the *drif345* showed less expression of *RGL2* and *ABI4*, the genes related to GA and ABA signaling. This study suggests that *DRIFs* are likely to be involved in the regulation of seed dormancy and germination by maintaining the ABA/GA ratio during the transition of seed from dormant to non-dormant germinating phase.

We further investigated the role of these *DRIF* homologs during the development of *Arabidopsis* seedlings. In **Chapter 5**, we observed that the *drif345* mutant failed to form a completely closed apical hook after germination of seeds in the dark and reduced hypocotyl length in red, blue, far-red, and white light and also in dark. The higher expression of *PHYB* in light-grown seedlings and lower level of *HLS1* in dark-grown seedlings of *drif345* suggests the possible role of these *DRIFs* homologs in photomorphogenesis.

In conclusion, this work provides information about the role of *cis*-elements in differential regulation of *CYC2-like* genes and the role of *DRIF* homologs in regulating different developmental processes in *Arabidopsis*. In *Arabidopsis*, the *DRIF* homologs are involved in seed germination in normal as well as in abiotic stress conditions and also, involved in seedling growth in various light conditions. Further investigation of the function of DDR homologs will provide a deeper understanding of underlying mechanisms controlling a network that regulates different developmental processes and help to improve production of agronomically important crop plants.

"सूर्यसलिलसंयोगात्वनस्पतिःवनायते।।
निरीक्षण परीक्षण संयोगात् अभिनिवेशःज्ञानायते"

"With the confluence of the sun and the water, the plant grows in forest. (Just as) the insight (of a researcher) grows into knowledge through observation and analysis."

- Rajeshwari Pandharipande