"To investigate the relationship between substrates of the N-end rule pathway and genes regulated by 'GCCGCC' *cis*-elements in *Arabidopsis thaliana".*

By

Geeta Prasad

A thesis submitted to The University of Nottingham for the degree of Doctor of Philosophy

July 2016

ABSTRACT

The N-end rule pathway of targeted proteolysis is regulated by Group VII ethylene response factors (ERFVII's). The aim of this research work was to analyse the relationship between substrates (ERFVII's) of the N-end rule pathway and genes, which have promoters containing a double 'GCCGCC' Ethylene-Responsive Element Binding Protein (EBP) *cis*-element.

Several genes were identified containing double EBP elements. Cloning and transformation of the promoters from five of these genes (PYL, ERD4, AT1G14810, AT3G13440 and AT5G44420) carrying two copies of the GCC-boxes present in the 5' UTR (5' untranslated region) or promoter region was conducted into *Arabidopsis* wild-type (Col-0) and *prt6-1* mutant plants. Expression driven by these promoters in the leaves and flowers of transgenic plants was analysed through GUS staining to reveal promoter activities. Enhanced promoter activity was seen in *prt6-1* lines (mutated in the E3-ligase of the N-end rule pathway) in comparison to Col-0. Further, cDNA of leaves and flowers of Col-0 and *prt6-1* were analysed by q-RT-PCR (quantitative real-time PCR) for expression of PYL7, ERD4, AT1G14810, AT3G13440; t-test analysis showed a significant difference (p-value<0.05) only in leaves of Col-0 and *prt6-1* for PYL7. Analysis of the genetic relationship between N-end rule pathway and genes containing GCC-boxes was also performed by analysing double mutant combinations of *prt6-1* and mutants of genes containing the EBP elements (*pyl7prt6-1*, *erd4prt6-1 and abi5prt6-1*) and Col-0 under different concentration of salt to determine the effect of stress due to salinity on the regulation of genes. At 125mM salt concentration significant difference was identified in highest number of mutant lines in comparison to Col-0.

An analysis of the *in-vivo* binding of the ERFVII RAP2.3 to the promoter of GCC-boxes containing genes was performed through Chromatin Immuno-precipitation assay (ChIP). The t-test analysis on qChIP-PCR data indicated significant difference between IgG and HA-IPs for both ABI5 and PYL7 performed on normoxic 35S:MA-RAP2.3-HA in Col-0 line. Further, *in-vivo* localization of ERFVII's HRE1 and RAP2.2 conditional stability was analysed using *promERFVII's:MC/MA-ERFVII's-YFP* constructs in Col-0 and *prt6-1*.

This thesis suggests that genes *PYL7, ERD4, AT1G14810, AT3G13440* that have double 'GCCGCC' EBP elements are downstream targets of the N-end rule pathway. Further

analysis *via* ChIP suggests that RAP2.3 interacts with the ‘GCCGCC’ binding site of promoters in the *ABI5* and *PYL7* genes, however further work is needed to confirm this. Additionally, sub-cellular localization of *promERFVII’s:MC/MA-ERFVII’s-YFP* studies suggest the location of HRE1 and RAP2.2 in nuclei of early stage root tips studied on 4-days old etiolated seedlings.

ACKNOWLEDGEMENT

Firstly, I would like to express my sincere gratitude to my supervisor Prof Michael Holdsworth for his guidance and support during my Ph.D study. Thank you for your guidance & help in all the time of my research and writing of this thesis.

I would like to thank our present and past lab people Dr. Daniel Gibbs, Dr. Jorge V. Conde, Dr. Guillermina Mendiondo, Dr. Nurul, Dr. Sophie Berckhan, Dr. Tinne Boeckx, Dr. Melania Ghita, Dr. Cristina Sousa Correia, Daniel Rooney, Christopher Till, Julietta Marquez, Natasha Gladstone and Charlene Dambire for all their help and sharing knowledge.

I am thankful to my parents, my two younger brothers, my relatives and my friends for encouraging and supporting me.

Finally, I would like to thank ‘Ministry of Social Justice & Empowerment, Government of India for funding my PhD. Last but not least I am thankful to God for all blessings.

Table of Contents

TABLE OF FIGURES

List of Tables:

Abbreviations

APX: ascorbate peroxidase

BBX: B-box transcription factors

bHLH: BASIC HELIX-LOOP-HELIX

BSA: Bovine serum albumin

ChIP: Chromatin Immuno Precipitate

DOA10: DEGRADATION OF ALPHA10

DREB1A: Dehydration-responsive element-binding protein 1A

EIF: eukaryotic initiation factor

GUS: β-glucuronidase

HA: Hemagglutinin

HECT: Homology to E6-AP C-Terminus

IgG: Immunoglobulin G

IP: Immuno-precipitant

JERF1: Jasmonate and Ethylene Response Factor 1

LB: Luria broth

MG132: Chemical name of geldanamycin

$NaHCO_3$: sodium bicarbonate

PBS: Phosphate buffered saline

PEG: polyethylene glycol precipitation

PGD1: phosphogluconate dehydrogenase

PIC: Protease inhibitor cocktail

PMSF: Phenylmethylsulfonyl fluoride

PI : Propidium Iodide

R1G1B: R1G1domain containing protein B

RGS: Regulator of G protein signalling

Tween 20: Polyoxyethylene (20) sorbitan monolaurate

35S-CMv: 35S-Cauliflower Mosaic Virus

cPTIO: 2-4-carboxyphenyl-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide

CHAPTER-1

Introduction

1.1 *Arabidopsis thaliana* as a plant genetic model organism

Arabidopsis thaliana belongs to the mustard family (*Brassicaceae*). It is a diploid plant having approximately 27,000 genes distributed on five chromosomes. *Arabidopsis* is considered as a model organism in plant biology and is the first plant whose genome was sequenced and was found to be one of the smallest genome in size among plants (Arabidopsis Genome Initiative, 2000). Additionally, it is easy to grow in a greenhouse or indoor growth chambers as it is small and a self-pollinator, thousands of seeds can be harvested within 6-7 weeks (Bennett *et al*., 2003; Leitch *et al*., 2005).

Arabidopsis can be easily transformed through the floral dip *Agrobacterium tumefaciens* system (Zhang *et al*., 2006; Clough *et al*, 1998). Thus, it has many advantages for genome and genetic analysis. A large number of *Arabidopsis* mutant seed stocks are available and can be obtained from the two major seed stock centres, Arabidopsis Biological Resource Centre (ABRC) and Nottingham Arabidopsis Stock Centre (NASC).

1.2 Ubiquitin-proteasome dependent N-end rule pathway

1.2.1 The Ubiquitin-26S Proteasome System

In prokaryotic and eukaryotic cells misfolded and damaged proteins are continually being degraded to small polypeptides and replaced by newly synthesized proteins. This process helps to maintain regular functioning of cells and also balance the amino acid pool through recycling (Glickman and Ciechanover, 2002). In addition, targeted proteolysis is used to control the stability of important regulatory proteins. In all tissues, most proteins are degraded by the ubiquitin-proteasome system (UPS) (Vierstra *et al.,* 2009). Ubiquitin serves as a recognition marker to the downstream 26S proteasome complex. It is composed of 76 amino acids and its primary sequence is highly conserved from yeast to mammals. Ubiquitin is covalently attached to the protein to be degraded. This allows transfer of the attached protein to the 26S proteasome for degradation. The 26S proteasome is a large complex composed of two sub complexes, a barrel-shaped proteolytic core and the 20S proteasome that is capped at one or both ends by 19S regulatory complexes, which recognize ubiquitinated proteins in both the nucleus and in the cytoplasm (Peters *et al.*, 1994).

Degradation of proteins through the UPS is performed in three steps. First ubiquitin forms a thio-ester bond between the Carboxyl-terminal (Ct.) Gly residue of it to a specific Cys of ATP dependent ubiquitin-activating enzyme-E1 (Fig: 1.1). Ubiquitin is transferred from E1

to ubiquitin conjugating enzyme-E2 through an additional high-energy thiol ester intermediate, E2-S~ubiquitin and from there to the substrate that is specifically bound to the ubiquitin protein ligase-E3. There are different types of E3 enzymes that recognize different substrate proteins. A particular E3 recognizes and binds to particular target proteins as well as E2 enzyme. This process occurs multiple times to form a chain of ubiquitin on the target protein. Such polyubiquitinated chains are recognized and degraded by the 26S proteasome. After degradation of protein the attached ubiquitins are released and are reused again for ubiquitylation (Smalle *et al*., 2004: Vierstra *et al*., 2009).

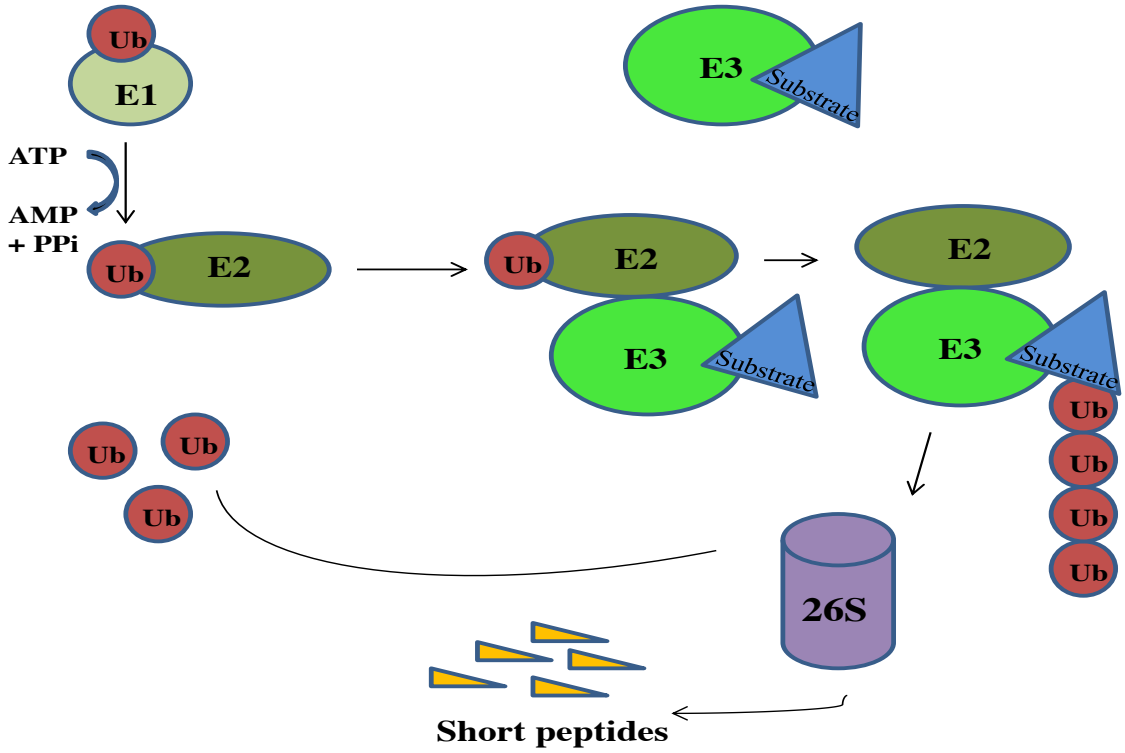


Fig: 1.1 Protein degradation by the 26S proteasome pathway

The pathway starts when E1 activates ubiquitin by forming thioester bond using energy derived from ATP cleavage. E1 transfers activated ubiquitin to E2. Here ubiquitin reforms thioester bond with E2. E2 further donates ubiquitin chain to substrate bound by E3-ligase. This ubiquitin chain mediates substrate to 26S proteasome for degradation. After degradation ubiquitin molecules move back to cytosol and are reused again for the 26S proteasome pathway. Ub-ubiquitin; E1-ubiquitin activating enzyme; E2-ubiquitin conjugating enzyme; E3-ubiquitin protein ligase; ATP-adenosine triphosphate; AMP-adenosine monophosphate. (Figure modified from Vierstra *et. al*., 2003)

1.2.2 There are many types of E3 ubiquitin ligases in plants

The plant E3 ligases are divided into three groups. The first group is HECT-type E3s (Homology to E6-Associated Carboxy-Terminus) that is the smallest E3 with seven members in *Arabidopsis*. The HECT domain consists of a 350 amino acid motif that contains both an ubiquitin-binding site and an E2-binding site. This enables the E3 to form a covalent thio-ester linked E3-ubiquitin intermediate before the transfer of ubiquitin to the substrate (Downes *et al*., 2003). The second group is RING-type (Really Interesting New Gene) and third is U-box-type E3 ligases. Both RING and U-box domains are structurally and functionally similar. RING-type and U-box-type E3s transfer ubiquitin directly from the E2 to the substrate without forming an E3-ubiquitin intermediate. In the *Arabidopsis* genome 64 genes encode for U-box-type E3s (Yee *et al*., 2009) and RING-type E3s have 469 RING domain-containing proteins found in the *Arabidopsis* genome (Stone *et al*., 2005).

Among the three types (E1-E2-E3) of ubiquitination-related enzymes, E3 ubiquitin ligase is the one that usually determines substrate specificity. Studies carried out through mutant screening have shown the importance of Skp1/Cullin/F-box (SCF), a type of multi-protein E3 ubiquitin ligase complex that catalyzes the ubiquitination of proteins destined for proteasomal degradation. It has an important role in the ubiquitination of proteins involved in the cell cycle and also marks various other cellular proteins for destruction (Morgan *et al*., 2007). The SCF (Skp1; cul1; F-box) name is derived from three of its four subunits: first, SKP1 (Skp1 which is needed for recognition and binding of the F-box. It is a horse shoe-shaped complex, along with cullin (cul1), second, Cullin (it forms the main structural scaffold of the SCF complex that links the skp1 domain with the Rbx1 domain) and third is the F-box protein (it specifically aggregates target proteins of the complex and then binding to the Skp1 component thus, allowing the protein to be brought into proximity with the functional E2 protein). Also, the F-box affinity for protein substrates can be regulated through cdk/cyclin mediated phosphorylation of target proteins. The fourth subunit of SCF is the RING finger protein RING-Box 1 (RBX1), RBX1 contains a small zinc-binding domain called the RING Finger, to which the E2-ubiquitin conjugate binds, allowing the transferring of the ubiquitin to a lysine residue on the target protein (Hershko and Ciechanover, 1998).

1.2.3 The Ubiquitin Proteasome System regulates plant growth and development

The UPS is utilized by plants to perform many functions during growth and development through maintaining protein level. Such as cell cycle control, transcriptional regulation, hormone signalling, pathogenesis, photomorphogenesis, chromatin structure regulation, responses to environmental challenges (Vierstra *et al*., 2009), self-incompatibility (Zhang *et al*., 2009), control of shoot apical meristem (Di Giacomo *et al*., 2013), and chloroplast biogenesis (Huang *et al*., 2013). Many E3 ligases have been shown to be involved in regulating plant response to many abiotic stresses such as, drought, cold, salinity, heat, radiation, and nutrient deprivation (Yee and Goring, 2009; Lyzenga and Stone, 2012).

1.2.3.1 Gibberellin signalling pathway regulation by ubiquitin proteasome pathway

Gibberellin for example, is a plant hormone that promotes seed germination, stimulation of organ elongation, and induction of flowering. The regulation of GA signalling is mediated by an SCF complex (Gao *et al*., 2011). GA is perceived by a protein called GIBBERELLIN INSENSITIVE DWARF1 (GID1). GA binds with GID1 that leads to binding to growth repressors called DELLA proteins localized in nucleus (Ueguchi-Tanaka *et al*., 2005; Nakajima *et al*., 2006; Willige *et al*., 2007). This complex of GID-GA-DELLA is targeted for recognition and ubiquitination by SCF^{SLY1}/SNZ E3 ligases, resulting in degradation of the DELLAs (Gao *et al*., 2011) (Fig: 1.2).

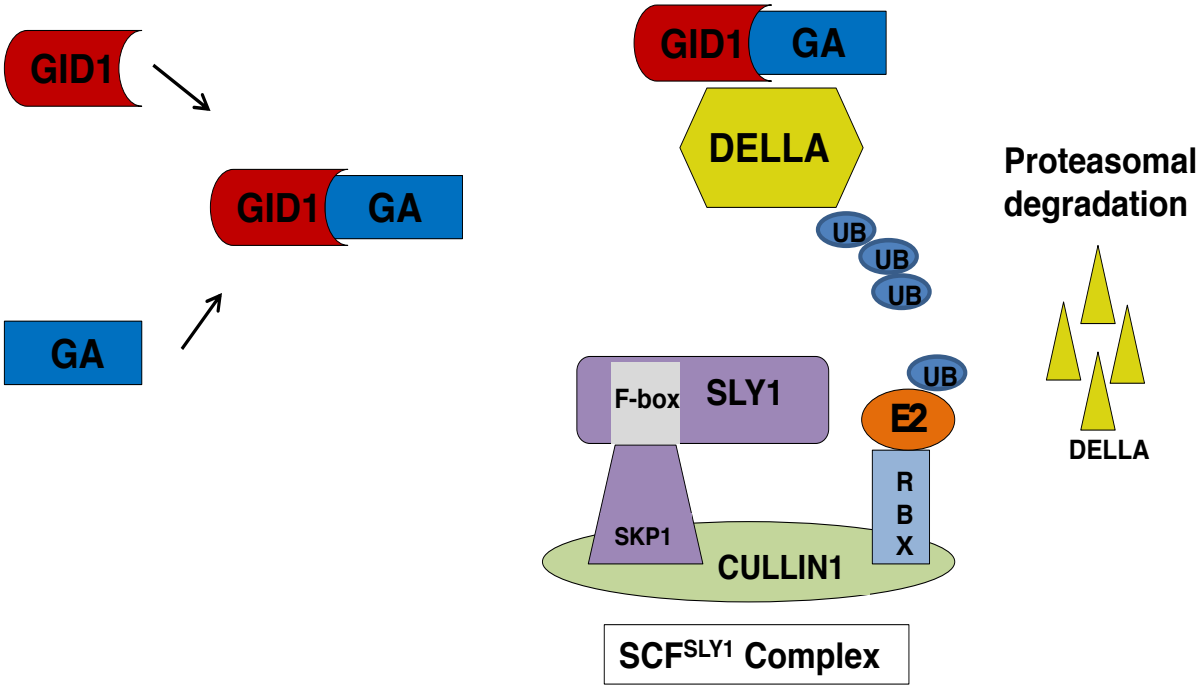


Fig: 1.2 Schematic representation of Giberellin signalling regulated by UPS

In presence of GA, DELLA is activated by SCF^{SLY1}/SNZ E3 complex. DELLA protein degradation releases growth-promoting transcription factors that promote inactive GA response genes to become activated.

(Figure modified from: Hauvermale *et al*., 2012.)

1.2.3.2 Regulation of ethylene signalling by the UPS

An important gaseous hormone, ethylene regulates hypocotyls and root elongation, apical hooks formation, and fruit ripening (Kendrick and Chang, 2008). Ethylene is regulated through five receptors that are related to bacterial two-component histidine kinases (Kendrick and Chang, 2008). In the absence of ethylene, a pair of SCF-box proteins, EIN2-TARGETING PROTEIN1 (ETP1) and EIN2-TARGETING PROTEIN2 (ETP2), promotes the degradation of the ethylene signalling protein ETHYLENE INSENSITIVE2 (EIN2), which reduces ethylene response (Qiao *et al*., 2009). In the presence of ethylene, ETP expression is repressed, thus leading to accumulation of EIN2. Downstream of EIN2 lies EIN3 and EIN3-like1 (EIL1), transcription factors that directly target ethylene-responsive genes. At low ethylene levels, EIN3 and EIL1 are targeted for ubiquitylation and degradation by another pair of F-box proteins, EIN3-BINDING F-BOX1 (EBF1) and EIN3-BINDING F-BOX2 (EBF2), which are also subject to proteasomal degradation (Guo *et al*., 2003; Potuschak *et al*., 2003; An *et al*., 2010). As ethylene levels increase, the stability of EBF1/2 decreases, leading to a build-up of EIN3 and EIL1, thus inducing transcription (An *et al*., 2010). Therefore, ethylene signalling is another example of a hormone signalling pathway that involves SCF complexes for its regulation (Fig: 1.3).

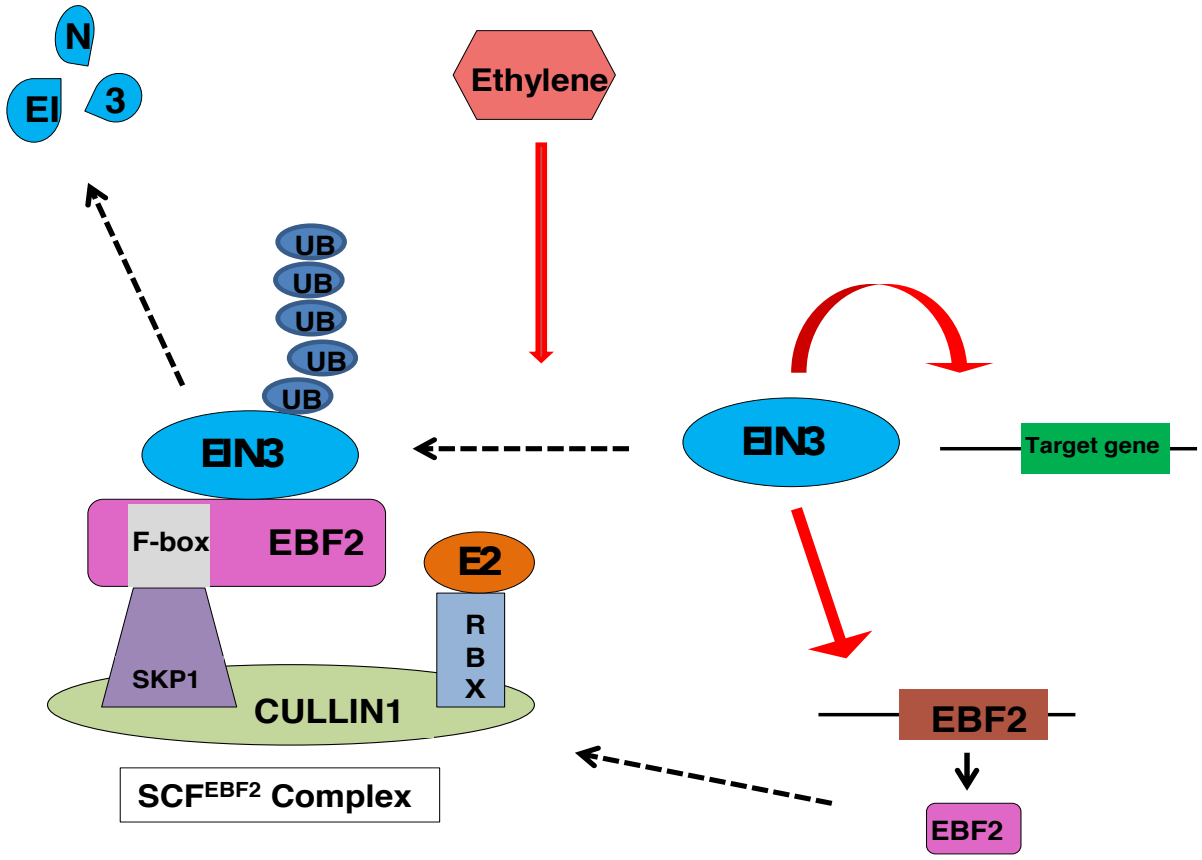


Fig: 1.3 Schematic representation of Ethylene signalling regulated by UPS

Ethylene biosynthesis enzymes are degraded by the ubiquitin proteasome pathway. In the absence of ethylene, ETP1/2 directs degradation of the membrane-bound protein EIN2. Degradation of the ETHYLENE INSENSITIVE3 (EIN3) transcription factor is controlled by EBF1 and EBF2 depending on level of ethylene.

(Figure modified from: Konishi *et al*., 2008.)

1.2.3.3 Absisic acid signalling pathway

Abscisic acid controls seed dormancy and drought response (Nambara *et al*, 2005). ABA signalling is regulated by ABA receptors *i.e*., GTG1 (GPCR-type G protein 1) and GTG2 (Pandey *et al.*, 2009) and a family of START proteins called PYR/PYL/RCAR (Ma *et al.*, 2009; Park *etal.*, 2009). Two RING E3 ligases, ABI3-Interacting Protein (AIP2) and Keep on Going (KEG) are involved in ABA signalling by regulating the abundance of ABA-responsive transcription factors, namely *ABA-INSENSITIVE 3 (ABI3)* and *ABA-INSENSITIVE 5 (ABI5)* (Zhang *et al.*, 2005; Stone *et al.*, 2006). It was observed that ABA increases AIP2 expression that leads to ubiquitination and degradation of *ABI3* (Zhang *et al.*, 2005). Whereas, *ABI5* is protected by ABA from ubiquitination and degradation by preventing recognition of *ABI5* by KEG (Stone *et al.*, 2006) (Fig: 1.4).

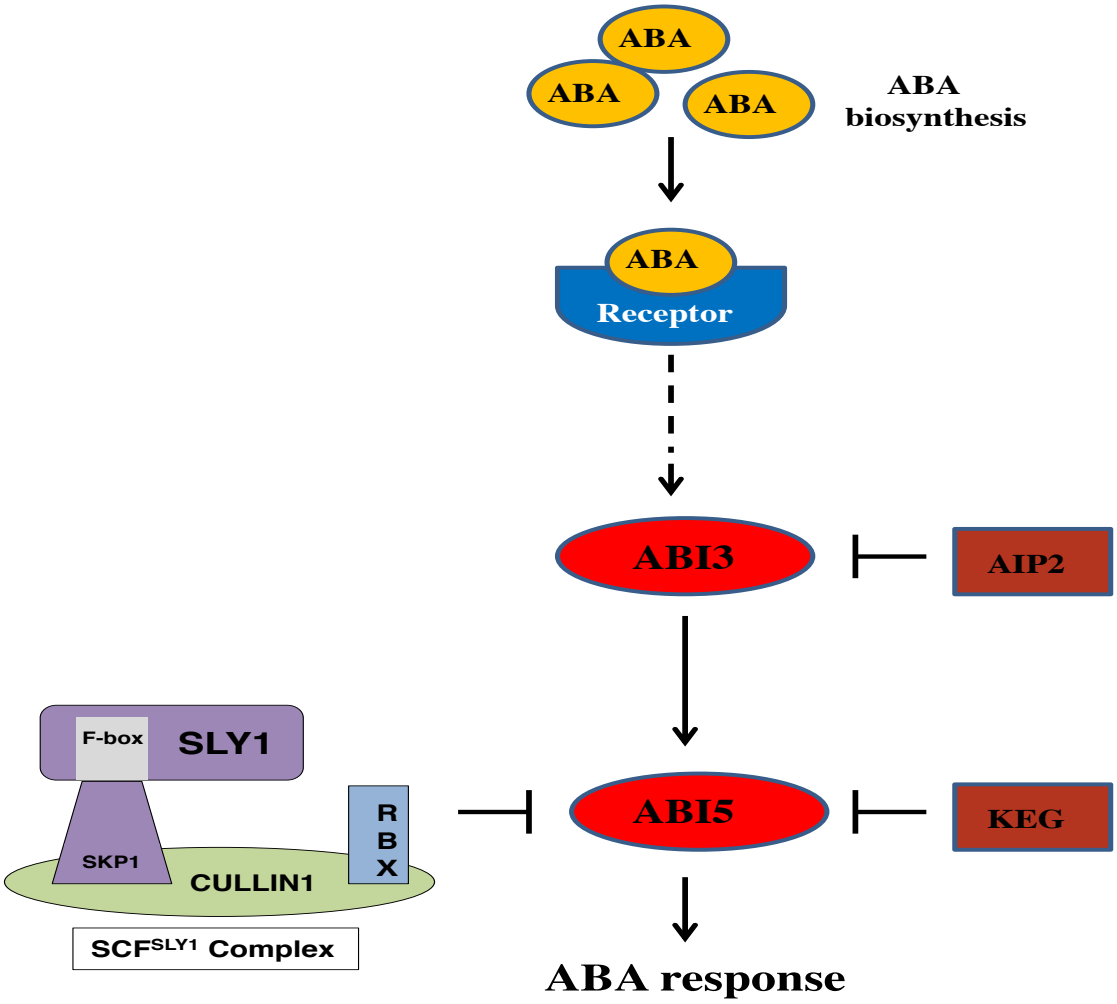


Fig: 1.4 Schematic representation of Abscisic acid signalling regulated by Ubiquitin-proteasome system

ABI3-INTERACTING PROTEIN (AIP2) and KEEP ON GOING (KEG) regulate ABA signalling. AIP2 controls the *ABA-INSENSITIVE 3 (ABI3)* and *ABA-INSENSITIVE 5 (ABI5)* by maintaining the level of ABA since it was observed that ABA increases AIP2 expression that leads to ubiquitination and degradation of ABI3. In contrast, ABA protects *ABI5* from ubiquitination and degradation by preventing recognition of *ABI5* by KEG. SCF complex: the scaffold protein cullin 1 is connected with the F-box protein through the adaptor protein SKP1 to one end and by the RING protein RBX1 to another end.

(Figure modified from: Liu *et al.,* 2011)

1.2.3.4 Regulation of light perception by the UPS

In plants the UPS is also seen playing a role as regulator of gene function through light as the physical factor required for that gene function (Sullivan *et al*., 2003). Studies have shown that light promotes the expression of some genes during seed germination to first leaf emergence (Wang *et al*., 2001). This study showed that after germination seedlings followed two developmental patterns. Skotomorphogenic (etiolated) where seedlings were kept in dark demonstrated long hypocotyls and closed unexpanded cotyledons that are protected by an apical hook. Another is photomorphogenic (de-etiolated) grown in light exhibited short hypocotyls and open and expanded cotyledons (Quail *et al*., 2002). The developmental stage of plant involves genome-wide changes in gene expression at transcription level. Many components of the signalling pathway that are downstream of photoreceptors were identified using mutant screening methods. In light the genes involved in de-etiolation are the pleiotropic COP/DET/FUS (constitutive photomorphogenesis/de-etiolated/fusca) genes, which are required for the repression of photomorphogenesis in darkness (Wei *et al*., 1996).

Previous studies demonstrated that COP1 E3-ligase functions as a light-inactivatible repressor of photomorphogenesis localized in the nucleus in the dark (v/v) light. It contains a RING domain (Deng *et al*., 1992) which is found in many ubiquitin E3 ligases. COP1 interacts with transcription factors such as HY5, HYH and LAF1, which promote photomorphogenesis (Holm *et al*., 2002: Ang *et al*., 1998). It was seen that COP1 is involved in the ubiquitylation and subsequent degradation of HY5 in darkness (Osterlund *et al*., 2002) (Fig: 1.5).

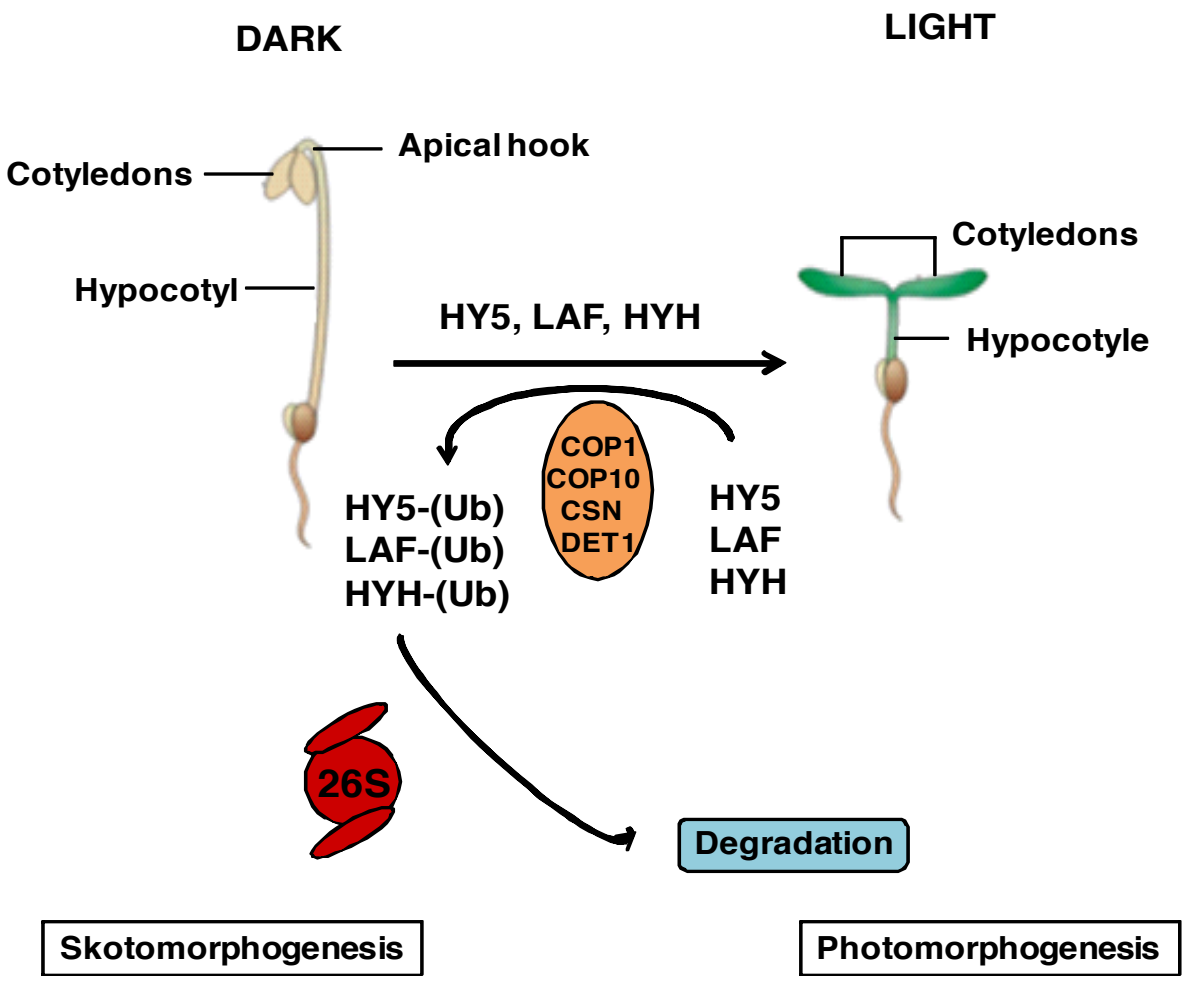


Fig: 1.5 A diagram showing effect of light on *Arabidopsis* seed germination and the degradation of proteins through the UPS.

(Figure taken from James *et al.*, 2003)

In addition to hormone and light signalling, the UPS also plays important roles in plant defence mechanics against pathogens and other microbes. For example, in plant-virus interactions, the host UPS can act in antiviral defence by targeting viral proteins for degradation (Dang *et al.*, 2001).

1.3 The N-End Rule Pathway of ubiquitin mediated targeted proteolysis

The N-end rule pathway is a ubiquitous UPS associated protein degradation pathway that is conserved from bacteria to mammals (Mogk *et al*., 2007). The pathway relates the stability of a protein to the identity of its N-terminal amino acid. A protein to be degraded bears a primary destabilizing residue at the N-terminus that functions as degradation signal. In this pathway the substrates bearing primary destabilizing residues are recognized by so called N-recognin E3-ligases (Section: 1.3.1.3). The N-recognin binds the N-terminal primary destabilizing residue and marks substrate protein with ubiquitin through covalent linkage. Subsequently, the protein is transferred to the 26S proteasome complex for degradation.

This pathway was first identified in 1986, when Alex Varshavsky and his colleagues analysed protein degradation in *Saccharomyces cerevisiae.* They found that the stability of a protein was related to the N-terminal amino-acid of that protein and therefore gave the name 'N-end rule' (Bachmair *et al*., 1986; Varshavsky*et al*., 1992, 1996, Mogk *et al*., 2007). Varshavsky and colleagues observed that some amino acids, when placed at the N-terminus of β-galactosidase, led to rapid degradation, whilst others did not. Therefore, amino-terminal residues are referred as 'destabilizing' and 'stabilizing'.

1.3.1 Components of N-end rule pathway

1.3.1.1 The N-degron

An N-degron or degradation signal, is an element within a protein at its N-terminal end that is sufficient for recognition and degradation by ubiquitin-proteasome pathway (Varshavsky *et al*., 1991). That is, the genetically engineered attachment of such sequences confers metabolic instability (a short half-life) on otherwise long-lived proteins (Hochstrasser *et al*., 1996).

An N-degron must contain one appropriate acceptor site, such as, a Lys residue, for the attachment of the polyubiquitin chain. After formation of N-degron, the complex is bound by an N-recognin (see section: 1.3.1.3) which initiates the degradation process. The N-degron of a protein to be degraded through N-end rule pathway can be produced either by the action of Methionine Amino Peptidases (MetAPs) or endo-peptidases (Tasaki and

Kwon 2007). In mammals, the removal of N-terminal Methionine by MetAPs leads to cleavage of methionine and only Cys residue at N-terminal remains. This cysteine gets oxidised and creates N-degron. Secondly, an N-degron is created through the cleavage of polypeptides by endo-peptidase resulting the C-terminal of the protein which subsequently bears an N-terminal destabilizing residue (Tasaki and Kwon 2007). In *Arabidopsis*, the secondary destabilizing residues N-terminal Asp, Gln and oxidized Cys can be arginylated by an arginyl-tRNA protein arginyl-transferase (ATE) to produce a primary destabilizing residue (Fig: 1.6) (Yoshida *et al*., 2002; Graciet *et al.,* 2010). Recent research has shown that the tertiary destabilizing residue Cys can be oxidized by oxygen and nitric oxide to form secondary destabilising residue (Gibbs *et al*., 2014). Besides this, there are also two distinct N-terminal amidohydrolase (Nt-amidase) that convert tertiary destabilizing residues Asn and Gln to secondary destabilizing residues Asp and Glu that undergo arginylation to become primary destabilizing residues, as part of N-degrons.

1.3.1.2 R-transferase

In the N-end rule pathway of targeted proteolysis, a protein with a destabilising N-terminal residue is created through a specific initial proteolytic cleavage. But can also be generated through enzymatic or chemical modifications to the N-terminus, for example, arginylation by Arg-tRNA protein transferases (ATEs). ATE is an enzyme that adds arginine to a secondary destabilizing residue to form a primary destabilizing residue. In yeast and mammals, only one ATE gene has been found, whereas in *Arabidopsis* two ATE genes are found *i.e.* AT5G05700/ATE1, AT3G11240/ATE2 (Fig:1.6) (Yoshida *et al*., 2002; Holman *et al*., 2009; Graciet *et al*., 2009).

1.3.1.3 N-recognin E3-ligases

In eukaryotes, primary destabilizing amino acids at the N-terminal end of a protein are recognized and bound by E3-ligase enzyme also known as N-recognins (Varshavsky, 1996), whereas, in prokaryotes, the N-degron is recognized by ClpS (Fig: 1.6) (Mogk *et al.*, 2007). The N-recognin E3 ligase in eukaryotes consist of a ~80 residue domain called the UBR domain (a general substrate recognition domain for destabilizing N-terminal residues) (Garzon *et al*., 2007; Tasaki *et al*., 2005).

These N-recognins are divided into two types. Type I and II. The basic amino acid residues Arg, Lys and His are recognised by Type I N-recognins and the bulky

hydrophobic amino acid residues Try, Trp, Phe, Ile and Leu are recognised by Type II N-recognins. In yeast, the N-terminal primary destabilizing residues are recognized and bound by the single N-recognin UBR1. It is a 225-kDa RING finger E3 ligase (Bachmair *et al*., 1986). UBR1 mediates substrate polyubiquitination as a complex with the Rad6/Ubc2 E2 conjugating enzyme, leading to its proteasomal degradation (Dohmen *et al*., 1991; Pickart *et al*., 1985).

In mammals, seven UBR box-containing N-recognins named UBR1 to UBR7 are involved in recognition (Tasaki and Kwon, 2007). UBR1, UBR2 and UBR3 of size about 200 KDa are conserved domain that includes UBR box and RING finger. UBR4, UBR5, UBR6 and UBR7 differ in size to each other. The UBR1, UBR2 and UBR4 can bind to both type I and type II residues. While, the UBR5 shows binding specificity for type I residues only (Tasaki *et al*., 2009). Depending on the binding specificity of UBR-boxes to substrates and their subsequent degradation, UBR1, UBR2, UBR4, and UBR5 are classified as N-recognins, and UBR3, UBR6, and UBR7 as non-N-recognins (Sriram *et al*., 2011; Tasaki *et al*., 2012).

In plants only two N-recognins are found, identified in the model plant *Arabidopsis*, PROTEOLYSIS (PRT) 1 and PROTEOLYSIS (PRT) 6 (Bachmair *et al*., 1993; Potuschak *et al*., 1998; Grazon *et al*., 2007) (Fig: 1.6) The 45 kDa protein PRT1 has two RING domains and one ZZ domain and PRT6 is a 224 kDa Nrecognin that contains the UBR box (Grazon-*et al*., 2007).

1.3.2 The Arginyl-branch of N-end rule pathway

The N-end rule pathway is a proteolytic system where destabilizing N-terminal residues function as degradation determinant (Tasaki *et al.,* 2012; Sriram *et al*., 2011; Bachmair *et al*., 1986). Post-translational conjugation of Arg to N-terminal Asp, Glu or oxidised Cys leads to generation of an N-degron with Nt-Arg (Kwon *et al*., 2002). This is called the arginylation branch of the N-end rule pathway, catalyzed by arginyl-tRNA-protein transferase (ATE1) in yeast and mammals and (ATE1ATE2) in plants (Bachmair *et al.,* 1986). The substrate protein bearing Arg is ubiquitinated and sent for proteolytic cleavage.

Research has shown that Arginyl (Arg) branch of N-end rule pathway regulates several physiological functions in plants studied through T-DNA mutant analysis. Such as,

research carried on ATE1/ATE2 and PRT6 shows that it promotes seed germination and establishment through the removal of sensitivity to the hormone ABA since *ate1ate2* double mutant showed hypersensitive germination phenotype to ABA (Holman *et al*., 2009). Further studies demonstrated that both ATE1 and ATE2 are required in the regulation of shoot and leaf development partly by repressing expression of *BP (BREVIPEDICELLUS)* since BP is important in meristem promotion (Graciet *et al*., 2010).

1.3.3 The Acetyl-branch of N end rule pathway

The Acetyl-branch of N-end rule pathway was recently identified (Hwang *et al*., 2010). The Acetyl-branch involves the transfer of acetyl groups from acetyl-coenzyme-A to the N-terminal alpha-amino group of a newly synthesized protein. Acetylation starts when the nascent chain with the initiator Met emerges from the ribosome at a length of ~25 residues or somewhat later (~50 residues) if the initiator Met has to be removed by Methionine Amino Peptidase (Tasaki *et al*., 2012). N-terminal acetylation is irreversible and involves a majority of cellular proteins and occurs at the newly exposed N-terminal residue (Ala, Val, Ser, Thr or Cys) after N-terminal cleavage by MetAPs (Hwang *et al*., 2010). In yeast, the E3 ligase Doa10 recognises the acetylated N-degrons (Hwang *et al*., 2010; Varshavsky *et al*., 2011; Tasaki *et al*., 2012).

1.3.4 The structure of the N-end rule pathway in different species

1.3.4.1 The bacterial N-end rule pathway:

In bacteria, the degradation of protein occurs in two steps. First, the Nterminal Arg, Lys and Met act as pro-N-degron (the N-terminal amino acids whose modification generates an N-degron) through the conjugation of Leu or Phe by specific aminoacyl-tRNA transferases. In a second step, the primary destabilizing residues (Leu, Phe, Trp and Tyr) including those generated through leucylation and phenylanylation are recognized and bound by ClpS, which transfers substrates to the ClpAP protease complex for degradation (Tobias *et al*, 1991; Erbse *et al*, 2006).

1.3.4.2 The Yeast N-end rule pathway:

In *Saccharomyces cerevisiae,* the N-end rule pathway is composed of three parts. Where in addition to destabilizing N-terminal 'primary' residues, 'secondary' and 'tertiary' destabilizing residues are also found at N-terminal of the substrate protein following proteolysis by endopeptidases. Asn and Gln are tertiary destabilizing residues that are deamidated into the secondary destabilizing residues Asp and Glu by the Nterminal amidase NTAN1. Asp and Glu further conjugate with Arg encoded by ATE1 gene which creates the primary destabilizing residue. The N-terminal Arg and other primary destabilizing residues are recognized and directly bound by the N-recognin Ubr1 (Ubiquitin ligase N-recognin 1), the E3 ligase. Ubr1 mediates substrate recognition and polyubiquitylation in a complex with the E2 conjugating enzyme that brings substrates to proteasome for degradation. (Fig:1.6-I).

1.3.4.3 The Mammalian N-end rule pathway:

In the mammalian N-end rule pathway, proteins with tertiary destabilizing residues Asn and Gln are deamidated into Asp and Glu by two different N-terminal amidases, NTAN1 and NTAQ1 respectively (Grigoryev *et al*., 1996; Kwon *et al*., 2000). Secondly, N-terminal Asp and Glu conjugate with Arg by R-transferases that are encoded by ATE1 and require Arg-tRNA as a cofactor (Kwon *et al*., 1999*)*. In addition to Asn and Gln, an N-terminal Cys functions as a tertiary destabilizing residue through a two-step modification involving oxidation and arginylation (Hu *et al.*, 2005; Kwon *et al*., 2002) (Fig: 1.6 II) N-terminal Arg and other type 1 and type 2 degrons can bound UBR1, UBR2 and UBR4 whereas UBR5 shows a preference for only type1 N-degrons (Tasaki *et al*., 2009). On the basis of binding and degradation assays, UBR1, UBR2, UBR4, and UBR5 are classified as N-recognins, and UBR3, UBR5, UBR6, and UBR7 as non-N-recognins (Sriram *et al*., 2011; Tasaki *et al*., 2012) (Fig: 1.6-II).

1.3.4.4 The plant Nend- rule pathway:

It was seen that plants have similar pathway as mammals except that in plants secondary destabilizing residues are arginylated by two R-transferases ATE1 and ATE2 (Yoshida *et al*., 2002). And secondly, two types of E3-ligases have been identified in plants. PRT1 (Bachmair *et al*,. 1993) and PRT6 (Garzon *et al*., 2007) which recognize primary residue for degradation (Fig: 1.6-III).

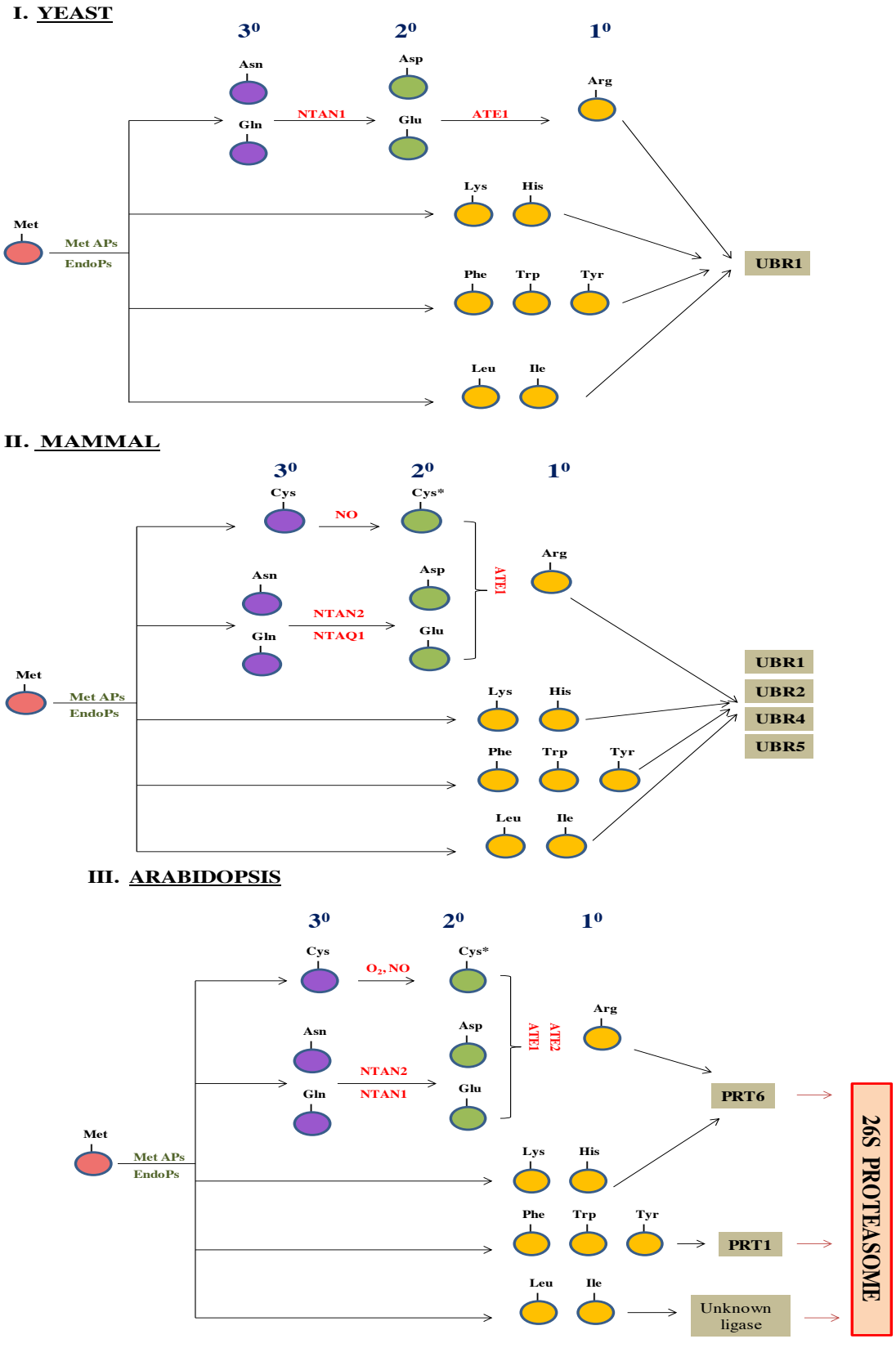


Fig: 1.6 The Arg/N-end rule pathway

The N-terminus 'NH_2-Met' of a protein is cleaved by Methionin Amino Peptidase (MAP) exposing N-terminal amino acid.

I. **Yeast:** Proteins bear three kind of destabilizing residue. After MAP action first the tertiary destabilizing residues are deamidated by a single NtNQ-amidase (NTA1) and converts to secondary. Thereafter, proteins with the secondary ATE1 that conjugates Arg to their N-termini to from primary destabilizing residue. These primary destabilizing residues are further recognized by N-recognin UBR1. Proteins carrying N-terminal acetylated residues are recognized by the Doa10 Ub ligase for degradation.

II. **Mammals:** The N-end rule pathway in mammals involves two distinct Nt-amidases (NTAN1 and NTAQ1) that deamidate N-terminal Asn, Gln and Cys tertiary residues in presence of oxygen and nitric oxide. The oxidized secondary destabilizing residues are recognized by the R-transferase ATE1. Arginylated primary residues are recognized by N-recognins (UBR1, UBR2, UBR4 and UBR5).

III. **Arabidopsis:** In plants, the N-end rule pathway is similar as in mammals having three stages for degradation of protein but the N-recognins differ *i.e*, PRT1 and PRT6 that recognizes primary destabilizing residue.

Figure modified from Graciet *et.al.,* 2010

1.3.5 The function of both oxygen and nitric oxide together in oxidation of Nt-cysteine

The Cys-Arg branch of the Arg/N-end rule pathway requires oxygen and nitric oxide together for oxidation of N-terminal Cys residue (Gibbs *et.al*., 2014). In plants, the occurrence of an N-terminal Cys residue was shown to be essential to generate a functional degron in the Group VII Ethylene Responsive Factor transcription factors (ERFVII's) (Gibbs *et. al*., 2011) (Fig: 1.7).

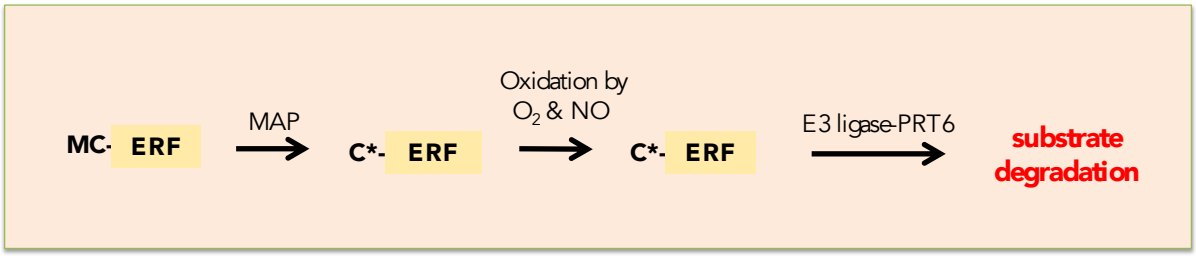


Fig: 1.7 The Cys-Arg/N-end rule pathway regulates degradation of ERFVII's

In *Arabidopsis thaliana*, the Group VII Ethylene Response Factor (ERF) transcription factors were the first identified plant substrates of the N-end rule pathway. A recent phylogenetic analysis showed that the Group VII consists of *HYPOXIA REPONSIVE 1* (*HRE1*), *HYPOXIA REPONSIVE 2* (*HRE2*), *RELATED TO AP2 2 (RAP2.2)*, *RELATED TO AP2.3/ETHYLENE-RESPONSIVE ELEMENT BINDING PROTEIN (RAP2.3), RELATED TO AP2 12 (RAP2.12)*. All these transcription factors begin with the amino acids 'NH_2-Met-Cys'.

(Licausi *et.al.,* 2011; Gibbs *et.al.,* 2011)

N-terminal Cys is oxidized to Cys-sulfinate (SO_2H) or Cys-sulfonate (SO_3H) in presence of both oxygen and nitric oxide (Hu *et.al*., 2005). The oxidized cysteine is arginylated by addition of Arg to form the primary destabilizing N-terminal residue which is recognized by E3-ligases of the N-end rule pathway followed by ubiquitination.

Previous studies showed the effect of lack of O_2 and NO on the stability of two representative physiological N-end rule substrates, the MC-initiating ERFVII's RAP2.3 and HRE2 (Gibbs *et al*., 2011) (Fig: 1.8). MC-RAP2.3-HA and MC-HRE2-HA accumulated in WT seedlings treated with hypoxia (low oxygen) and cPTIO (2-4-carboxyphenyl-4,4,5,5-

tetramethyl imidazoline-1-oxyl-3-oxide) nitric oxide scavenger. Stability of these proteins was seen enhanced under low oxygen and low nitric oxide conditions. Whereas, accumulation reduced in the presence of SNAP (NO donors S-nitroso-N-acetyl-DL-penicillamine) (Gibbs *et al*., 2014).

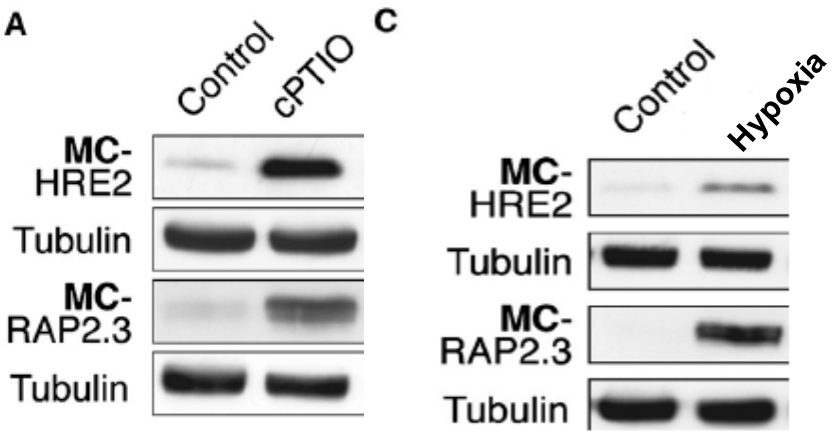


Fig: 1.8 Western results showing stabilizing of substrates under low oxygen and nitric oxide conditions

A. MC-HRE2 and MC-RAP2.3 seedling treated with cPTIO were stabilized compare to normal untreated seedlings under same condition.
B. MC-HRE2 and MC-RAP2.3 seedling treated with hypoxia were also stabilized compare to normal untreated seedlings under same conditions.

(Figure taken from Gibbs *et al*., 2014)

Therefore, these experiments demonstrate that ERFVII's that carry cysteine as a destabilizing residue at their N-terminal end require both oxygen and nitric oxide together to get oxidized and arginyl to form primary destabilizing residue.

1.3.6 The role of the N-end rule pathway in plant development:

Analysis of the N-end rule pathway in plants has shown that it plays important roles in controlling plant growth and development. Studies of the *Arabidopsis ate1* mutant showed that leaf senescence is delayed in the *ate1* mutants due to the lack of R-transferase activity (Yoshida *et al*., 2002). Therefore, this indicates that proteolysis by the N-end rule pathway has an important role in the progression of leaf senescence. The N-end rule pathway was also seen involved in regulating shoot and leaf development. Plants with *ate1ate2* double mutant and *prt6* single mutant have shown phenotypic abnormalities such as, serrated and lobed leaves, loss of apical dominance, and stem and internode elongation defects (Graciet, *et al.,* 2009).

An additional function of the plant N-end rule pathway was studied by Holman *et.al.,* (2009). They identified a mutant allele of the N-recognin PRT6 that altered seed after-ripening and germination. The high dormancy of the *prt6* mutant was found to correlate with a hypersensitive response to ABA an inhibitor of germination. This result suggested that PRT6 might be involved in the removal of ABA sensitivity prior to germination. Genetic interaction analyses with components of the ABA signalling pathway further indicated that PRT6 functions upstream of the ABA response regulator *ABA INSENSITIVE 5 (ABI5)* in the control of germination. Notably, both mutant R-transferases (i.e. *ate1ate2*) exhibit a similar ABA hypersensitive phenotype to that of *prt6* mutants, strongly suggesting that the N-end rule substrates responsible for the germination defects are first arginylated by R-transferases before they are recognized by PRT6 for degradation process.

The N-end rule pathway also functions as an oxygen sensor in plants through regulated proteolysis of hypoxia-sensitive transcription factors carrying N-terminal Cys (Gibbs *et al*., 2011; Licausi *et al.,* 2011). In Normoxia, the ERFVII's transcription factors are degraded by the N-end rule pathway through oxygen-mediated Cys-oxidation after N-terminal Met cleavage but in hypoxic condition the Cys-oxidation is repressed. The proteins therefore become stable and activate genes that promote anaerobic metabolism and survival in hypoxia.

Studies have shown that nitric oxide (NO) also plays important roles in a large number of plant signalling pathways (Grun *et al*., 2006) and in plant growth and development (Rio *et*

al., 2004; Delledonne *et al*., 2005). NO regulates seed dormancy (Bethke *et al*., 2006), flowering (He *et al*., 2004) and stomatal closure seed germination, and hypocotyl elongation (Neill *et al*., 2002; Gibbs *et al*., 2014). In mammals, it has been shown that the Cys-2 of RGS4 (Regulator of G-protein Signalling 4), RGS5 and RGS16 are oxidized by NO allowing recognition by R-transferase of the N-end rule pathway (Hu *et al*., 2005; Lee *et al.,* 2005).

1.3.7 Substrates of N-end rule pathway in eukaryotes.

The *Arabidopsis* genome encodes five ERFVII's that have been identified as substrates of the N-end rule pathway. A recent phylogenetic analysis showed that the *Arabidopsis* Group VII consists of *HYPOXIA REPONSIVE 1* (*HRE1*) and *HYPOXIA REPONSIVE 2* (*HRE2*), *RELATED TO AP2.2 (RAP2.2)*, *RELATED TO AP2.3/ETHYLENE-RESPONSIVE ELEMENT BINDING PROTEIN (RAP2.3), RELATED TO AP2.12 (RAP2.12)*. All these transcription factors begin with the amino acids 'NH_2-Met-Cys' (Licausi *et al.,* 2011; Gibbs *et al.,* 2011). The conserved N-terminus 'NH_2-Met-Cys' of the ERFVII's transcription factors are destabilized by the exposure of Cys at the N-terminus after Met cleavage by MetAP (Lee *et al*., 2005). It was shown that the following Met cleavage exposed cysteine is oxidized by either oxygen or nitric oxide, before being arginylated by ATE and then recognized and targeted for degradation by specific N-recognin (Lee *et al*., 2005).

In mammals, the substrates RGS4 (regulator of G-protein signalling 4), RGS5, (regulator of G-protein signalling 5) and RGS16 (regulator of G-protein signalling 16) were found in mice that were shown to regulate cardiovascular development (Hu *et al*., 2005; Lee *et.al*., 2005). These substrates also begin with Met-Cys and follow Met cleavage exposing cysteine to be oxidized by oxygen and nitric oxide and follow arginyl-branch of the N-end rule pathway for substrate degradation (Lee *et al*., 2005). In addition, the group ERFVII's have been shown to play an important role in adaptations to flooding in both rice and *Arabidopsis* (Bailey-Serres *et al*., 2012). Further, SNORKEL1 (SNK1) and SNORKEL2 (SNK2) were also seen transcriptionally activated during ethylene and submergence conditions consequently, responsible for a strong elongation response (Hattori *et al*. 2009). In contrast, another group VII ERF SUB1 A was found to inhibit growth and increased fermentation during submergence (Singh *et al*. 2010).

In *Drosophila melanogaster*, DIAP1 (Drosophila Inhibitor of Apoptosis 1) *i.e*., a major anti-apoptotic regulator, has been identified as a substrate of the N-end rule pathway. The DIAP1 when gets cleaved through caspase activity a C-terminal fragment that bears a destabilizing N-terminal residue is released and is targeted for degradation through the N-end rule pathway (Ditzel *et al*., 2003). In yeast, cohesion subunit SCC1 act as substrate of N end rule pathway. SCC1 when cleaved by separase produces a fragment with an N-terminal Arginine residue which is recognised and ubiquitinated by UBR1.

1.4 Identification and characterization of plant promoters

Regulation of gene expression in an organism at different growth and developmental stages occur at the transcriptional, post-transcriptional and post translational level. Mainly, gene regulation at transcriptional level plays important role in activation and suppression of a gene. This step is controlled through many *cis*-acting elements present in gene promoters. Gene promoters are the section of DNA situated in a *cis*-alignment to gene coding regions and carry *cis*-acting elements that are specific binding sites for *trans*-acting factors, "transcription factors", which may cause activation or repression of transcription. Therefore, promoter regions are essential for gene expression (Zou *et al*., 2011). Promoters are used as tools to study the regulation of gene expression at the transcriptional level in plant biotechnology (Liu *et al*., 2013).

1.4.1 Promoter structure

The promoter is the region of DNA where RNA polymerase II enzyme binds which is responsible for the generation of RNA. The promoter region contains a *cis* acting regulatory element that controls the transcription of coding region (exons) into messenger ribonucleic acid (mRNA) which is translated into proteins (Jacob and Monod, 1961; Butler *et al*., 2002). In eukaryotes, the promoter of a protein-encoding gene contains a “core promoter” that is found ~40bp upstream of the transcriptional initiation site and comprises of TATA box (Molina *et al*., 2005). The TATA box is the binding site for the transcription initiation factor TFIID TBP (TATA-box-Binding Protein) subunit. The core promoter also contains *cis*-elements that are binding sites for RNA polymerase II and its corresponding subunits (Burley *et al*., 1996).

In prokaryotes, the promoter consists of two short sequences at -10 and -35 positions upstream from the transcription start site. The sequence at **-10** (-10 element) is called the

Pribnow box, and consists of the six nucleotides **TATAAT.** The Pribnow box is absolutely essential to start transcription in prokaryotes. The other sequence at **-35** (the -35 element) usually consists of the six nucleotides **TTGACA**. Its presence allows a very high transcription rate (Ross *et al*., 1993; Estrem *et al*., 1999).

In eukaryotes the promoters are usually situated at upstream of genes. Some of the regions commonly found in nearly all promoters are:

1. **AGGA box**. A consensus sequence close to -80 bp.
2. **TATA box**. A sequence usually located around 25 bp upstream of the start point. The TATA box tends to be surrounded by GC rich sequences. The TATA box binds RNA polymerase II and a series of transcription factors.
3. **GC box**. A sequence rich in guanidine (G) and cytidine (C) nucleotides is usually found in multiple copies in the promoter region, normally surrounding the TATA box.

1.4.2 Types of promoters used to regulate gene expression

Promoters used in biotechnology are of different types according to the intended type of control of gene expression. They can be generally divided into:

1.4.2.1 Constitutive promoters

These promoters are usually derived from plant viruses or plant housekeeping genes. Constitutive promoters remain usually active and direct expression in almost all tissues continuously. They are highly active in rapidly growing meristematic tissues or vascular tissues (Odell *et al*., 1985). Viral Promoters such as the 35SCaMV promoter from the Cauliflower Mosaic Virus is frequently used in basic research and the development of transgenic plants (Benfey *et al*., 1990: Odell *et al*., 1985). High-expressing housekeeping genes that are considered strong native constitutive promoters are Ubiquitin, Actin, Tubulin, and EIF (eukaryotic initiation factor), APX (ascorbate peroxidase), PGD1 (phosphogluconate dehydrogenase), and R1G1B (R1G1domain containing protein B) genes (Park, *et al*., 2012).

1.4.2.2 Spatiotemporal promoters

The promoters direct the expression of a gene in specific tissue region or at certain stages of development. Such as, seed specific promoters were seen expressed during seed development. These kinds of promoters are therefore used in tissue-specific

targeting of industrial and pharmaceutical compounds, and development of transgenic seeds with improved nutritional quality and better functional quality of milled grain (Peremarti, *et al*., 2010; Potenza *et al*., 2004). Similarly, fruit specific promoters are used to generate fruits with improved quality and nutritional value leading to better agronomic quality and are also used for the production of antibodies, biopharmaceuticals and edible vaccines through genetic engineering (Kawakatsu *et al*., 2010; Cosgrove *et al*., 2000). Further, anther and pollen-specific promoters are used to control male sterility. For example, the RA8 promoter from rice, A9 promoter from *Arabidopsis*, and TA29 promoter from tobacco (Peremarti *et al.*, 2010; Bisht *et al*., 2007).

1.4.2.3 Inducible promoters

These promoters are responsive to environmental conditions and external stimuli that can be artificially controlled. These promoters could be used to study gene regulation under abiotic and biotic factors such as light, oxygen levels, heat, cold and wounding.

Certain genes that are differentially expressed during stress act as source of stress-responsive promoters and *cis*-elements. The promoter of two rice genes OsNCED3 and Wsi18 were seen associated in the synthesis and signalling of ABA, drought and high-salinity (Bang, *et al*., 2013; Yi, *et al*., 2011). And the Rd29A promoter of *Arabidopsis* was seen inducing drought-specific expression of DREB1A in transgenic wheat (Pellegrineschi, *et al*., 2004). However, the *cis*-regulatory elements present in stress-responsive gene promoters also provide understanding of gene regulation and plant signalling under stress conditions. For example, a stress-responsive element with dehydration-responsive element DRE (A/GCCGAC) was involved in the regulation of cold and dehydration responses in *Arabidopsis* (Yamaguchi *et al*., 1994) whereas, the low temperature responsive element C-repeat binding factor (CBF) and the ABA responsive element ABRE (ACGTGG/T) were observed regulating dehydration and salinity responses in *Arabidopsis* and *Oryza* (Jiang *et al*., 1996; Yamaguchi *et al*., 2006).

There are many defense-response promoters that are involved in development of resistant transgenic crops. Pathogen inducible promoters from defense-response associated genes become activated during diseased conditions in plants (Kovalchuk, *et al*., 2010). For example: the rice OsPR10a promoter gets induced by pathogens and defense response hormones (Hwang *et al*., 2008) and the barley Germin-Like GER4

promoter becomes highly inducible after infection by biotrophic and necrotrophic pathogens (Himmelbach *et al*., 2010)

1.4.2.4 Synthetic promoters

A synthetic promoter is the sequence of DNA that is designed to control gene expression of a target gene. Such as, *cis*-regulatory sequences derived from naturally-occurring promoter elements are used to construct many synthetic promoters that can act as a promoter enhancer; typically to initiate RNA polymerase II-mediated transcription (Rushton *et al*., 2002). One of the synthetic promoter DR5 auxin promoters is a highly active synthetic promoter that contains tandem direct repeats of the auxin responsive TGTCTC element which is used to study auxin response mechanisms in plants (Ulmasov *et al*., 1997). Another is the 35SCaMV promoter with its duplicated enhancer sequence in a synthetic context (Guerineau *et al*., 1992) is widely used to increase gene expression levels in plants.

1.5 The Group VII Ethylene Response Factor (ERF) transcription factors as substrates of N-end rule pathway in *Arabidopsis*

1.5.1 ERFs binds to ethylene-responsive binding protein (EBP) binding site "GCCGCC"

The ERF family is a large family of transcription factors and is part of the AP2/ERF super family (Riechmann *et al*., 2000). The AP2/ERF domain consists of about 60 to 70 amino acids and is involved in DNA binding. The AP2 family proteins contain two repeated AP2/ERF domains. The ERF domain was first identified as a conserved motif in four DNA-binding proteins from tobacco that are ethylene-responsive element-binding proteins 1, 2, 3, and 4 (EREBP1, EREBP 2, EREBP 3, and EREBP 4) and was shown to specifically bind to GCC-boxes, which is a DNA sequence involved in the ethylene-responsive transcription of genes (Takagi *et al*., 1995).

It has been demonstrated that *Arabidopsis* AtEBP (also named RAP2.3) was associated with OCS ELEMENT BINDING FACTOR 4 (OBF4) (Buttner and Singh, 1997). It was found that the AtEBP shares strong homology with the DNA-binding domains of tobacco EREBPs (Takagi *et al*., 1995). Electrophoretic Mobility-Shift Assay and DNase I footprint

analysis on many plant defensin genes have revealed that AtEBP can specifically bind to the GCC-box (Buttner & Singh 1997; Yang *et al*., 2009).

Further research in rice has also shown that EREBPs bind to the GCC-box DNA motif (AGCCGCC) that is located in the promoter of several pathogen related (PR) genes. The molecular and functional analysis of a rice Mitogen-Activated Protein Kinase (MAPK) BWMK1 have shown the phosphorylation of transcription factor OsEREBP1 (EREBP 1) enhanced its ability to bind to the GCC-box. Transient co-expression of the BWMK1 and OsEREBP1 in *Arabidopsis* protoplasts also showed increased expression of the β-glucuronidase reporter gene (GUS) driven by the GCC-box element (Cheong *et al*., 2003).

In Soybean also it was identified through gel retardation assays that GmERF3 binds to the *cis*-acting elements of the GCC-box and DRE/CRT element. GmERF3 contains a conserved N-terminal MCGGAII/L sequence that is required for binding to the GCC-box (Tournier *et al*., 2003). In addition to GmERF3, four new members of the ERF subfamily possessing conserved N-terminal MCGGAII/L sequence were identified. CaERFLP1 (Lee *et al*., 2004), TaERF1 (Xu *et al*., 2007), Tsi1 (Park *et al*., 2001), and CaPF1 (Yi *et al*., 2004) which specifically binds both the GCC-boxes and DRE/CRT element.

1.5.2 The ERFVII's protein AtEBP (RAP2.3) binds to EBP binding site of the *ABI5* gene

In *Arabidopsis* the Ethylene-responsive element Binding Protein (EBP) *cis*-element GCCGCC binds to the ERFVII protein, AtEBP/RAP2.3 (Yang *et al*., 2009; Jose *et al*., 2013; Cheong *et al*., 2003: Buttner *et al*., 1997). Therefore, it is concluded that genes bearing two GCC-boxes in their promoter may be bound by ERFVII's and are regulated by the substrates of N-end rule pathway. One of the best examples is *ABA-INSENSITIVE 5 (ABI5)* gene (Gibbs *et al*., 2014).

In *Arabidopsis*, ABA regulates seed development and maintains seed dormancy (Koornneef *et al*., 1982). *ABA-INSENSITIVE 5 (ABI5)* encodes a transcription factor that belongs to subfamily of bZIP transcription factors which binds to a *cis*-regulatory ABA responsive promoter element (ABRE) and regulates gene expression in seeds and in response to ABA (Finkelstein *et al*. 2000; Lopez *et al*. 2001; Kim *et al*. 2002). Previous

studies demonstrated that ABA-sensitivity of germination is regulated by the N-end rule pathway of protein degradation. Genetic analysis revealed that during seed germination the E3-ligase: PRT6, controls ABA sensitivity through *ABI5* (Holman *et al*, 2009).

Studies of the *ABI5* gene showed the presence of two *cis* regulatory elements (GCCGCC, the "EBP" element) in the 5'UTR region (Fig: 1.9). It has been demonstrated that in the absence of EBP elements the *ABI5* gene was not up-regulated in the *prt6* mutant (Fig: 1.10) (Gibbs *et al*, 2014). Hence, this suggests that genes containing EBP elements might be regulated by the substrates of the N-end rule pathway.

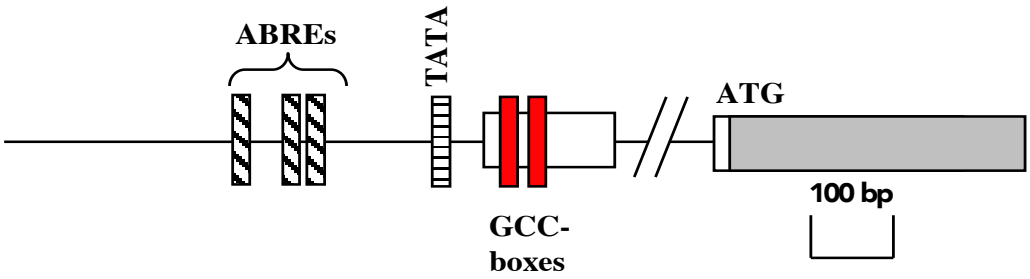


Fig: 1.9 Schematic diagram of *ABI5* gene

ABI5 promoter region showing the position of three ABA-responsive binding elements (ABREs) and two GCC-boxes in red in the 5' UTR. Scale bar is 100bps. (Gibbs *et al*., 2014; Molecular cell)

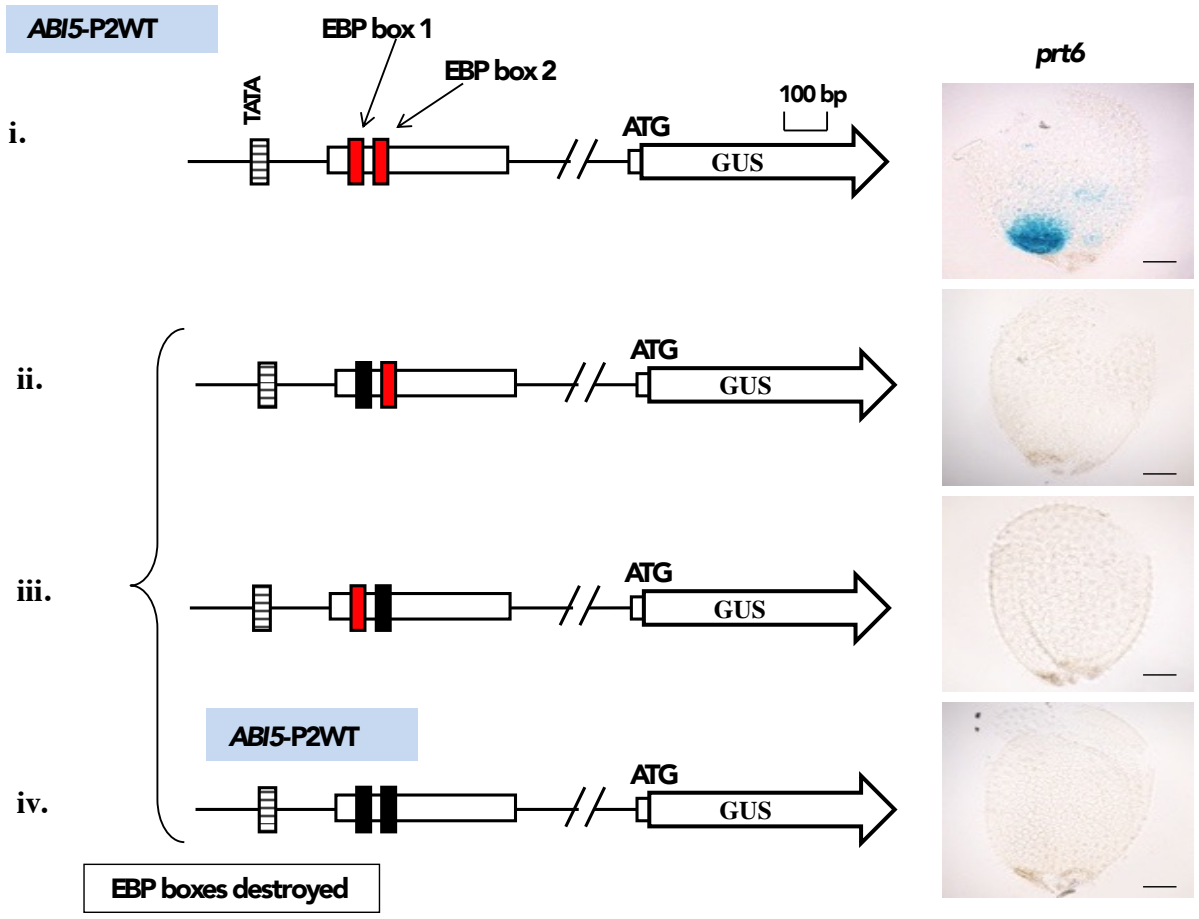


Fig: 1.10 Schematic representation of promoter activity through mutation of EBP elements present in 5' upstream region in *promABI5::GUS* in *prt6* endosperm

i. *ABI5-P2WT* containing two EBP binding elements shows GUS expression in the endosperm of *prt6* seed.

ii. *ABI5-P2WT* containing mutated EBP binding site 1 shows no GUS expression in the endosperm of *prt6* seed.

iii. *ABI5-P2WT* containing mutated EBP binding site 2 also shows no GUS expression in the endosperm of *prt6* seed.

iv. *ABI5-P2WT* containing both EBP binding sites mutated shows no GUS expression in the endosperm of *prt6* seed.

The *ABI5-P2WT* represents the minimal promoter contains two consensus-binding sites (GCCGCC-boxes) *cis*-elements for ERFVII's Red box: GCCGCC EBP site; Black box: abolished GCCGCC EBP site.

1.6 The ERFVII's control physiological processes in plants

In order to survive under adverse environmental conditions, plants have developed methods to control various stresses. Recent studies have revealed that ethylene responsive factors transcription factors are involved in responses to different biotic and abiotic stress (Buttner and Singh, 1997; Solano *et al*., 1998; Fujimoto *et al*., 2000; Tournier *et al*., 2003; Huang *et al*., 2004; Fujita *et al*., 2006).

1.6.1 The ERFVII's enhances survival and growth level in *Arabidopsis*

The five *Arabidopsis* ERFVII's have been shown to enhance plant responses in water logging, hypoxia (Licausi *et al*., 2011; Gibbs *et al*., 2011; Lee *et al*., 2011; Bailey-Serres *et al*., 2012,) and in low nitric oxide (Gibbs *et al*., 2014) conditions. Studies have shown that in *Arabidopsis* ERFVII's enhance plant survival under hypoxia by activating hypoxia-associated genes. It was seen that RAP2.3 and HRE2 remain stable under low oxygen and low nitric oxide states (Sec: 1.3.5) and help plants to survive under these condition (Gibbs *et al*., 2011; Licausi *et al*., 2011). In addition, stability of ERFVII's under low nitric oxide also showed their significant role in regulation of seed germination, hypocotyl elongation and stomatal closure (Gibbs *et al*., 2014).

A recent study has shown that interaction of RAP2.3 with DELLA has a physiological relevance in apical hook development. The development of the apical hook of etiolated seedlings is regulated by gibberellin (GAs) and ethylene (ET) (Abbas *et al*., 2013). It was previously shown that RAP2.3 expression is induced by ethylene (Buttner *et al*., 1997; Singh *et al.,* 1997). Significantly, EIN3 binds *in-vivo* to the RAP2.3 promoter (Chang *et al*., 2013) and activates its expression in etiolated seedlings suggesting that RAP2.3 could also participate in the mechanism regulating apical hook development by GAs and ethylene (Nora Marin-de *et al*., 2014).

Another study on the ERFVII's, RAP2.2, has shown its function during biotic stress. The over-expression of RAP2.2 in *Arabidopsis* induces increased resistance of plants for *Botrytis cinerea* infection (Zhao *et al*., 2012). Thus, promotes plants growth and development under stress conditions.

1.6.2 Role of ERFVII stability in other plant species under unfavourable conditions

1.6.2.1 Rice

Studies undertaken in rice (*Oryza sativa*) showed rice plants carrying *SUB1A-1* gene enhanced submergence tolerance during flooding due to the increased stability of the protein (Fukao *et al*., 2006; Bailey-Serres *et al*, 2008). Further, SUB1A is also seen involved in plant survival in drought and enhanced tolerance to oxidative stress (Takeshi *et al*., 2011). In rice, another ERF-VII protein Jasmonate and Ethylene Response Factor 1 (JERF1) showed over expression during drought. The over expression of JERF1 mediated plants to improve drought tolerance. Under drought condition over expression of JERF1 activates the expression of stress-responsive genes and increased the synthesis of the osmolyte proline by regulating the expression of OsP5CS. Since OsP5CS encodes the proline biosynthesis key enzyme deltal-pyrroline-5-carboxylate synthetase and helps transgenic rice to adapt better to drought. JERF1 also activated the expression of two ABA biosynthesis key enzyme genes, OsABA2 and Os03g0810800, and increased the synthesis of ABA in rice (Zhang *et al*., 2010). As the higher level of ABA reduces water loss from plants and also maintains growth of root to get more water from soil (Park *et.al*. 2008). Therefore, synthesis of ABA in larger quantities due to over expression of JERF1 results in greater tolerance of plants to drought stress (Schroeder *et al*. 2001; De Smet *et al*. 2006; Neill *et al*. 2008; Park *et al*. 2008).

Recent research has shown that JERF1 could enhance rice resistance to the pathogenic fungus *Rhizoctonia solani* that causes Sheath Blight (SB), a disease in rice plants that reduces yield and quality. The over-expression of JERF1 increases the expression of defence-related genes and increases the activities of lipoxygenase (LOX) and phenylalanine ammonia lyase (PAL) (Pal *et al*., 2014). Thus SB resistance is improved by modulating the expression of defence-related genes.

1.6.2.2 Barley

A study in barley (*Hordeum vulgare*) has shown that the group VII-ERF, HvRAF RNAi lines with reduced *HvPRT6* showed sustained biomass, enhanced yield, retention of chlorophyll, and enhanced induction of hypoxia-related genes in response to waterlogging and also showed reduced chlorophyll degradation in response to continued darkness. Another group VII-ERF BERF showed increased expression of hypoxia-associated genes and altered seed germination phenotypes in response to waterlogging (Mendiondo *et al*.,

2015). And was also shown to be involved in the regulation of expression of the Barley *knox3* (*Bkn3*) gene in the presence of ethylene (Osnato *et al*., 2010).

1.6.2.3 Wheat

In Wheat (*Triticum aestivum*) group-VII ERF, ethylene-responsive factor 1 (TaERF1) has demonstrated role in both abiotic and biotic stresses. Increased expression of TaERF1 activates stress-related genes, pathogen related genes (PR), COLD-REGULATED (COR) and RESPONSIVE TO DEHYDRATION (RD) genes. The activation of such genes leads to improved plant tolerance to abiotic stresses *i.e*., drought, salinity and low-temperature and also helped plants to survive under disease caused by the fungus *Blumeria graminis* (Xu *et al*., 2007).

1.6.2.4 Tobacco

In Tobacco (*Nicotiana tabacum*) it was found that JERF1 enhanced tolerance to salinity and cold stresses (Wu *et al*., 2007). Study of *Jatropha curcas,* a small oil tree showed increased transcript level of JcERF1 by ABA, osmotic stress (PEG), ethylene exposure, and salt treatment (Zhang *et al*, 2010). This suggests that JcERF1 might be involved in regulation of multiple stress signaling pathways.

1.6.2.5 Pine

In Virginia Pine (*Pinus virginiana*) a higher plant species that belongs to gymnosperm, overexpression of an ERF/AP2 pepper (*Capsicum annuum*) transcription factor (CaPF1) in Virginia Pine exhibited tolerance towards heavy metals such as, cadmium, copper, and zinc ($CdCl_2$, $CuCl_2$, or $ZnCl_2$) and in heat and cold stress. The enhanced expression of CaPF1 protects plant cells from the oxidative damage caused by stresses and improves organ growth by increasing cell numbers. Further, enhanced expression of CaPF1 induces disease tolerance against pathogens, for example, *Bacillus thuringiensis* and *Staphylococcus epidermidis (*Yi *et al*., 2004; Tang *et al*., 2005).

Therefore, it can be concluded that group VII ERFs has important functions in plant growth and development. Thus, further investigation of these proteins may help to elucidate other beneficial features of the N-end rule pathway in diverse developmental processes in plants.

1.7 AIMS & OBJECTIVES

The aim of this thesis was to find the relationship between N-end rule pathway and regulation of genes through binding of ERFVII's at two GCCGCC, *cis*-elements present in their promoter or at 5'UTR region.

The main objectives of this thesis were:

1. To analyse the relative activities of promoter of genes containing 'GCCGCC' binding sites in Col-0 and *prt6-1* tissue.

2. To determine the binding of the RAP2.3 to the GCC-boxes present in the possible downstream targets of N-end rule pathway through Chromatin Immunoprecipitation (ChIP).

3. To study *in-vivo* stabilization of promMC/MA-ERFVII's through Yellow fluorescent protein technique.

CHAPTER-2

Materials & Methods

2 MATERIALS & METHODS

2.1 Media

Half strength MS (Murashige & Skoog) medium was used for plating *A. thaliana* Col-0 and *prt6* mutant seeds. Media is composed of 0.7% (w/v) agar (Fisher Scientific, Loughborough, UK) supplemented with half-strength Murashige & Skoog salts (Duchefa Biochemie B.V., Haarlem, Netherlands, Murashige & Skoog 1962). The pH of the media was adjusted to 6.2 with potassium hydroxide before autoclaving. When the media had cooled appropriate concentration of required antibiotic was added before pouring. Media was poured in 9cm x 9cm square plates that were stored at 4^0C until use.

2.2 Seed sterilization and plating

Seeds were sterilized before plating on 0.5 MS medium to avoid bacterial or fungal growth. Seeds were put in 1.5 ml eppendorf tubes and 0.5 ml 5% (v/v) bleach (freshly diluted Parazone (JEYES, (UK)) was added and the seeds incubated for 5 minutes. Seeds were then washed three times with sterile water. Seeds were plated and plates were sealed with micropore tape and subsequently chilled at 4^0C for three days.

2.1.3. Plant Growth

Seeds on media plates were incubated at 22°C in continuous light (light intensity 60µmols $m^{-2}s^{-1}$) in a controlled environment growth chamber. Seedlings were transferred to soil when the first leaves had appeared. The soil composition was four parts compost, two parts vermiculite and one part silvaperl. Plants were grown in controlled environment rooms with a 16 hour light (23^0C) and 8 hour dark (18^0C) for the rest of their life cycle.

2.1.4 DNA extraction from leaf material for polymerase chain reaction (PCR)

Fresh plant leaf discs were collected in 1.5 ml Eppendorf tubes. 40ml of 'Extraction (E) buffer' (100mM Tris-HCl pH 9.5, 0.25M KCl, 10mM EDTA) was added and the plant material was mashed with a pipette tip for 10 seconds. Tubes were heated in a heating block for 10 minutes at 95^0C. After 10 minutes tubes were placed on ice then 40ml of 'Dilution (D) buffer' (3% (v/v) BSA in dH_2O) was added and the tubes were centrifuged (13000 rpm at room temperature for 30 seconds). The supernatant was removed and stored at 4^0C.

2.1.5 Promoter Amplification from *Arabidopsis* genomic DNA

For cloning promoter regions of fourteen different *Arabidopsis* genes PCR was performed. Every gene was tested for promoter amplification by specific primer pairs designed from their gene sequences. At all 5' end of forward primers the attB1 adapter was added (5'-GGG ACA AGT TGT ACA AAA AAG CAG GCT TC–3'), and on reverse primer attB2 adapter (5'-GGG GAC CAC TTT GTA CAA GAA AGC TGG GTC–3') was added. (Table: 2.1).

A list of the 14 genes analysed in this project (Table 2.1) with their primer pair (in capital) attached with gateway adapters (small case) and the expected length of amplified product.

PCR reactions were carried out in a 50μl volume containing 1x Phusion High-Fidelity buffer (NEB), 10mM dNTPs (Promega), 100pmol each primer (Eurofins MWG), 1.25U Phusion High-Fidelity Taq Polymerase (NEB). Thermocycler conditions were: 94^0C for 5 minutes; 40 cycles of 94^0C for 30 seconds. 54^0C-56^0C for 30 seconds, 72^0C for 1 minute, 72^0C for 7 minutes and 10^0C for continuous. PCR products were checked for amplification on 1% (w/v) agarose gel (Sigma) using 1x TBE buffer (89mM Tris base, 89mM Borate, 2mM EDTA pH 8.5). Sizes of DNA bands were compared with a 1kb ladder (Fermentas).

Gene	Symbol	Primer Sequence (5'-3')	Expected Length (bp)
AT1G13740	AFP2	F- 5' gggacaagttgtacaaaaaagcaggcttcTTGTTAATTTTTGAGGTAATAT 3' R- 5' ggggaccactttgtacaagaaagctgggtcAGATTCCGAAACAATAGAAAGA 3'	595
AT1G30820	-	F- 5' ggggacaagttgtacaaaaaagcaggcttcCAAAAGAATTAGGTTATGATAC 3' R- 5' ggggaccactttgtacaagaaagctgggtcCTTCTTCCTACTGATTTTTAAA 3'	929
AT3G13440	-	F- 5' ggggacaagttgtacaaaaaagcaggcttcATTTGCTATATCTCAACCACTA 3' R- 5' ggggaccactttgtacaagaaagctgggtcTCGTCTAAAACAACAGACAATC 3'	1541
AT3G19290	ABF4	F- 5' ggggacaagttgtacaaaaaagcaggcttcTTTAAAAAATTGACCTAAAAAC 3' R-5' ggggaccactttgtacaagaaagctgggtcAATCCTTGTATACACTCCTTCT 3'	1114
AT3G54510	ERD4	F-5' ggggacaagttgtacaaaaaagcaggcttcAGCCACTAAACAAATTCCTTTC 3' R-5' ggggaccactttgtacaagaaagctgggtcAGTTCAAAGAATCTGATAATCG 3'	697
AT3G60500	CER7	F-5' ggggacaagttgtacaaaaaagcaggcttcGTCAAAGTTAATCTTTCCAAGC 3' R-5' ggggaccactttgtacaagaaagctgggtcTATCTGCTATCACTAGTAAACT 3'	657
AT4G01026	PYL7	F-5' ggggacaagttgtacaaaaaagcaggcttcGGAGATTGAAACAATGGTGGAA 3' R-5' ggggaccactttgtacaagaaagctgggtcCTCCGACCAACCAGATATCT 3'	628
AT4G04925	-	F-5' ggggacaagtttgtacaaaaaagcaggcttcAAATATCAAAGAACGGAAATAA 3' R-5' ggggaccactttgtacaagaaagctgggtcAGCAACTTCCGCCGCAAGCC 3'	379
AT1G79630	-	F-5' ggggacaagtttgtacaaaaaagcaggcttcAAAGGTTAACACAAGTGGATAC3' R-5' ggggaccactttgtacaagaaagctgggtcGGCGTCTTGGTTAGTCCCTT 3'	827
AT2G02970	-	F-5' ggggacaagtttgtacaaaaaagcaggcttcGAATTGCATTTGGAAATCAGAA 3' R-5' ggggaccactttgtacaagaaagctgggtcCAACAATTCAAAAACCGCCA 3'	582
AT1G14810	-	F-5 ggggacaagtttgtacaaaaaagcaggcttcCGTGTAAGATTCCAAAGTTAGC 3' R-5' ggggaccactttgtacaagaaagctgggtcTGTCGCCGCCGTGAGGAAGA 3'	362
AT1G11540	-	F-5' ggggacaagtttgtacaaaaaagcaggcttcAGGAAAAATAATTGCATTGTC 3' R-5' ggggaccactttgtacaagaaagctgggtcTATCGAGAGGAAGAGTCCAC 3'	480
AT5G12330	LRP1	F-5' ggggacaagtttgtacaaaaaagcaggcttcCAAAGGATCTTGGATATTTATG 3' R-5' ggggaccactttgtacaagaaagctgggtcCTCCGTCGCCGCCGCCGAAG 3'	632
AT5G44420	PDF 1.2	F-5' ggggacaagtttgtacaaaaaagcaggcttcTACAAATTTAAATTCTGATCTC 3' R-5' ggggaccactttgtacaagaaagctgggtcGATGATTATTACTATTTTGTTT 3'	1022

Table: 2.1 Primers for Promoter region amplification

2.1.6 Promoter cloning using the Gateway System

Gateway technology (Invitrogen) is a cloning method which was used in this project to facilitate cloning of amplified promoter fragments. It was used because of its highly efficient cloning/subcloning of DNA segments. Gateway allows DNA segments to be transferred between different cloning vectors while maintaining orientation and reading frame. The Gateway reactions are of high efficiency, fast, and simple. They represent the *in-vitro* version of integration and excision reactions that occurs *in-vivo*. The two main Gateway reactions are the BP recombination reaction (PCR-product with flanking attB sites + Donor vector containing attP sites + BP clonase) and the LR recombination reaction (Entry clone containing attL sites + Destination vector containing attR sites, and promoters and tags + LR clonase) (Hartley *et al*. 2000).

2.1.6.1 The BP Recombination Reaction

An entry clone was created by using the BP recombination reaction. The reaction transfers the gene of interest to a donor vector to create an entry clone. In the BP reaction the recombination of attB DNA (*i.e.* PCR product or linearized expression clone) with an attP DNA (donor vector *i.e. pDONOR 221*) occurs to create an attL. The BP reaction is catalyzed by the BP Clonase enzyme mix. Because of the negative selection marker *ccd*B (kills background of cells with no cloned DNA) between *att*P1 and *att*P2, only transformants with the recombined vectors carrying *att*L1-ORF-*att*L2 (the 'entry clone') can grow on the selection plate containing antibiotic kanamycin (50μg/μl).

To perform the BP reaction the amplified promoter were first purified by the PEG method. DNA samples were diluted four-fold with Tris-EDTA (pH-8) then half volume of 30% (w/v) PEG solution was added. The mixture was vortexed and centrifuged at 13,000 rpm for 15 min at room temperature. The supernatant was removed and clear pellet suspended in 10μl Tris-EDTA (pH-8). The DNA concentration was checked through nanodrop (Thermo scientific). The BP reaction was setup in 10μl volume (Table: 2.2). The reaction mix was vortexed briefly and spinned down and kept for incubation at 25^0C for 1hr. Then 2μl of Proteinase K was added in the reaction mix. Each tube with reaction mix was vortexed and incubated at 37^0C for 10 min to terminate the BP reaction.

Ingredient	Final concentration	Volume in 10μl reaction
TE Buffer (pH-8)	1x	2- 8μl
PCR product	100-150ng	0.5-7μl
Donor vector (350ng)	150ng	0.5μl
BP Clonase Enzyme	2U	2μl

Table 2.2: Reaction mix for the gateway BP reaction

Amplicons were ligated to donor vector through BP Clonase Enzyme.

2.1.6.2 LR Recombination Reaction

In the LR reaction the entry clone is recombined with the destination vector to create an expression clone containing *att*B sites, flanking gene the insert of interest. The LR reaction is used to facilitate the recombination of *att*L substrate (entry clone) with an *att*R substrate (destination vector, pKGWFS7) (Fig: 12) to create an *att*B clone. The LR reaction was catalyzed by the LR Clonase enzyme. Since destination vectors also contain *ccd*B between *att*R1 and *att*R2 and have a selection marker gene that is different from the entry clone therefore, only the recombined destination vectors carrying *att*B1-ORF-*att*B2 will be selected.

The LR cloning was undertaken by using Gateway LR Clonase II enzyme mix (Invitrogen). Entry Clones were recombined with the Destination vector (pKGWFS7), a GUS-GFP fusion vector specifically designed for plant promoter analysis (Cutler *et al*., 2000) (Fig: 2.1). The reaction mix was vortexed briefly, centrifuged and incubated at 25^0C for 1 hr. Then 2μl of ProteinaseK was added to each tube and vortexed and again incubated at 37^0C for 10 minute to terminate the LR reaction. (Table: 2.3). Amplicons were ligated to the destination vector using the LR Clonase Enzyme.

Ingredient	Final concentration	Volume in 10µl reaction
TE Buffer (pH-8)	1x	2- 8µl
Entry Clone (150ng)	100-150ng	0.5-7µl
Destination Vector (pKGWFS7) (350ng)	150ng	0.5µl
LR Clonase Enzyme	2U	2µl

Table 2.3: Reaction mix for the gateway LR reaction

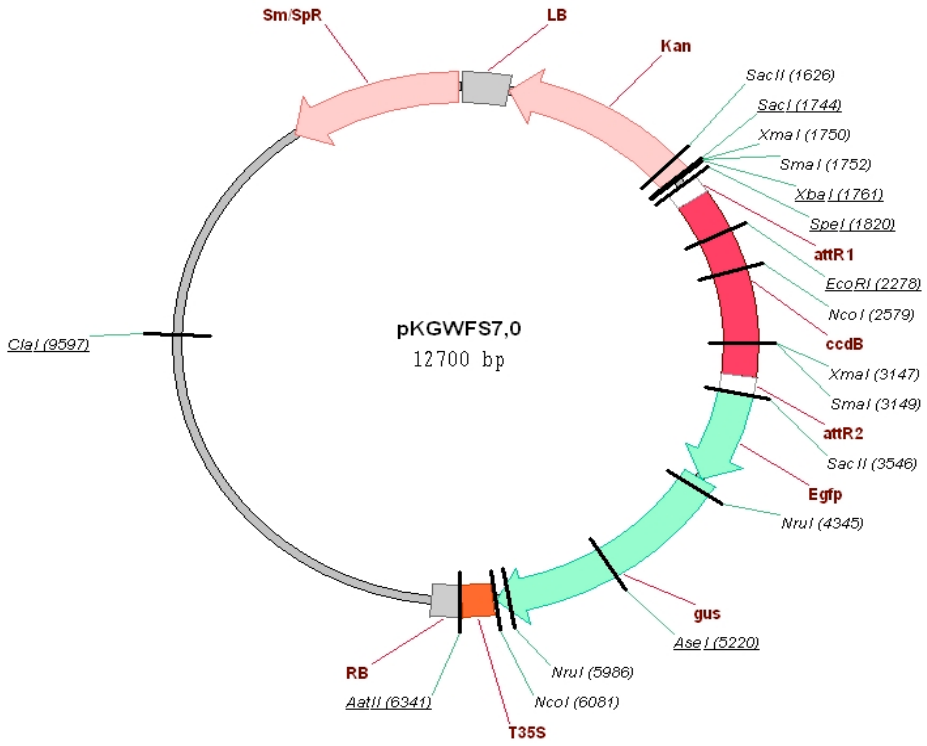


Fig: 2.1 pKGWFS7 (Destination Vector) used in LR reaction for cloning of promoter

pKGWFS7 (12700 bps) is a destination vector used for promoter analysis

LB - Left border

Sm/Spr - Spectinomycin plasmid selection marker

Kan - Kanamycin insertion selection marker

Egfp - eGFP/GUS fusion reporter protein

GUS - β-D-glucuronide gene

T35S - CaMV35S promoter

RB - Right border

(Figure is adapted from Cutler *et al*., 2000 and Karimi *et al*., 2002)

2.1.7 Transformation of DNA into *E. coli*

This step was performed for both the BP and the LR Clonase products. For transformation 2μl of cloned vector was added to 50μl of competent *E. coli* (DH5α) cells and incubated on ice for 30 min. Then cells were heat shocked by incubating at 42^0C for 45 seconds in a water bath and immediately placed on ice for 2 minutes to terminate the DNA transformation. 950μl of SOC (Super Optimal broth with Catabolite Repression) was added to reaction mixture and kept at 37^0C for 1 hour in incubator shaker. The cell suspension was then centrifuged and pellets were suspended in fresh 100μl of SOC. 100μl cell culture was plated on LB plates with kanamycin (50μg/ml) and incubated overnight at 37^0C.

2.1.8 Plasmid Isolation

Colonies from plates were picked and inoculated in 3ml Luria Broth (LB) with kanamycin (50μg/μl) and shaking overnight. Plasmids were isolated using the Qiagen Miniprep kit. 1.5ml bacterial culture was centrifuged for two minutes at 13,000 rpm at room temperature. The bacterial cell pellet was resuspended in 250μl Buffer P1 at room tempetature. Then 250μl Buffer P2 was added and mixed thoroughly by inverting the tubes 4-6 times followed by addition of 350μl Buffer N3 and mixed immediately and thoroughly by inverting tubes 4-6 times. The solution was centrifuged for 10 min at 13,000 rpm in a table-top microcentrifuge. The supernatant was transferred to the QIAprep spin column by pipetting. The columns were then centrifuged for 30–60 seconds and flow-through was discarded. The QIAprep spin column was then washed by adding 0.5ml Buffer PB and centrifuged for 30-60 seconds. The flow-through was discarded. QIAprep spin column was then washed again by adding 0.75 ml Buffer PE and centrifuged for 30–60 seconds. The flow-through was discarded and tubes were centrifuged for an additional 1 min to remove residual wash buffer. The QIAprep column was placed in a clean 1.5 ml microcentrifuge tube and 50μl of Buffer EB (10 mM Tris-Cl, pH 8.5) was added to the centre of each QIAprep to elute DNA. The columns were then centrifuged for 1 min at 13000 rpm. The solution containing purified plasmid DNA was stored at -20^0C.

2.1.9 Restriction Digestion of DNA

Plasmids were digested by the BsrG1 restriction endonuclease to detect presence of inserts. The reaction was setup in a 20μl volume containing miniprep DNA (1000ng/μl), 10x NE Buffer 2 (Biolabs), 10x Bovine serum albumin (Sigma), 5U/μl BsrGI enzyme (Bio-

labs) and 5.5μl dH_2O. Reaction mix was incubated at 37^0C for 2 hrs for enzyme reaction. BsrG1 digested DNA fragments were separated on 1% (w/v) agarose gel electrophoresis and inserts were compared with the 1kbp DNA ladder (Fermentas).

2.1.10 Transformation of Recombinant Clones from the LR reaction into *Agrobacterium tumefaciens* (GV3101 PMP90)

For transformation 2μl of the LR cloning reaction was added to thaw *Agrobacterium tumefaciens* (GV3101 PMP90) competent cells and were incubated on ice for 20 minutes. Cells were then electro pulsed using a Bio-Rad electroporator, where the programme was set as transformation *via Agrobacterium*. 900μl pre-warmed LB media was added, and the cells recovered at 28°C in a shaking incubator for 4 hours. Cells were recovered by centrifugation for 2 minutes at 3040 rpm, and then resuspended in 250μl LB media. 50μl and 200μl of the cells were plated onto LB-agar plates with 25μg/ml gentamycin (Melford, Suffolk, UK), 50μg/ml rifampicin (Melford, Suffolk, UK) to select for the *Agrobacterium* strain, and 100μg/ml spectinomycin (Sigma-Aldrich, Gilingham, UK) to select for the construct. Plates were incubated for 2 days at 28°C.

2.1.11 *Agrobacterium*-mediated *Arabidopsis* plant transformation

The *Arabidopsis* Col-0 and *prt6-1* mutant plants were grown with 4 plants per 9 cm pot. They were allowed to flower, the first bolts were removed when they reached approximately 10cm in height. 5 days later plants were used for transformation. Fresh *Agrobacterium* colonies were inoculated with a 5ml starter culture grown at 28°C overnight with appropriate antibiotics (Rifampicin (50μg/μl), Gentamycin (25μg/μl) and streptomycin (50μg/μl)). The next day the whole starter culture was inoculated into 300ml LB (containing all above three fresh antibiotics) and grown for 24 hrs (until the culture reached an OD_{600} of 0.6-0.8). Cultures were centrifuged at 5000 rpm for 10 min and the pellet was resuspended in 250ml 5% (w/v) sucrose. Cultures were transferred to a beaker and Silwet L-77 was added to a concentration of 0.05% (v/v) (50μl/100ml) and mixed. Plants were dipped in this bacterial solution for 5-20 sec, with gentle agitation. A film of liquid coating on plant was seen. Plants were covered with plastic lid to create a humid environment and left to recover overnight. The next day the cover was removed, and later seeds were harvested when mature (Clough & Bent, 1998).

2.1.12 Selection of transformed plants

To identify transformed plants, harvested seeds were sterilised and plated on an agar plate containing kanamycin (50μg/μl) and chilled for 2 days at 4^0C. Plates were transferred to the growth room and exposed to light for 4-6 hrs then covered with aluminium foil and kept in the dark for 48 hrs. Then plates were again kept in light. This procedure was undertaken for selection of the successful T_1 transformants. After 24 hrs in light, successful seedlings containing resistant gene were seen having darker green cotyledons and with longer hypocotyls than non-resistant seedlings (based on Harrison *et al*., 2006). The healthy and green transgenic (T_1) seedlings were transferred to soil in order to generate T_2 seeds to be selected to identify homozygous plants in the T_3.

2.1.13 Genotyping PCR to detect the presence of GUS gene in transformed plants

PCR was carried out for amplification of the GUS gene to verify transformation of putative transformants. Specific primers were used i.e. GUS-F: 5'-CTG GAT CGC GA ACT GCG-3 and GUS-R: 5'CCA AAG CCA GTA AAG TAG AAC-3'. The PCR reaction was carried in a 50μl volume and contained 10x Taq buffer (Bio-labs), 10mM dNTPs (Promega), 100pmol each primer (Eurofins MWG), 2U Taq Polymerase (Bio-labs). Thermocycler conditions were: 94^0C for 5 minutes; 40 cycles of 94^0C for 30 seconds. 54^0C-56^0C for 30 seconds, 72^0C for 1minute, 72^0C for 7 minutes and 10^0C for continuous. PCR products were checked for amplification on a 1% (w/v) agarose gel (Sigma) using 1x TBE buffer (89mM Tris base, 89mM Borate, 2mM EDTA pH 8.5). Bands were compared with a 1kbp DNA ladder (Fermentas).

2.1.14 Histochemical GUS staining for *β-glucuronidase* activity

Plant tissue *i.e*., leaves, flowers were taken and transferred to 1.5ml microcentrifuge tubes. They were immersed in 150-200μl of X-gluc (5-bromo-4-chloro-3-indoyl-beta-D-glucuronic acid, Sigma-Aldrich) solution (1mM X-Gluc, 20% methanol, 80% of 50mM Phosphate buffer, pH 7.0). For each of 100μl GUS solution (0.01mM), 1μl of each 100mM Potassium Ferri-cyanide and Potassium Ferro-cyanide were added. All samples (v/v) were incubated at 37°C for 16 hours. The GUS solution was removed. 70% alcohol was added to samples to remove chlorophyll. Later samples were washed with sterilized water and then mounted in Hoyer's solution (Section: 2.1.15) on slide (Jefferson *et al*., 1986). Pictures of images were taken from Leica microscope.

2.1.15 Preparation of Hoyer's solution

For 50ml final volume of Hoyer's solution:

30g gum arabic

200g chloral hydrate

20g glycerol

50mL water

i. 50ml of water was added into 500ml beaker and mixed with stirrer.

ii. 200g of chloral hydrate was added and let it dissolved slowly into the water. After all chloral hydrate was added, 30g of gum arabic was added and dissolved intothe mixing solution. This step needs to be done extremely slowly to avoid clumps of powder (this can take about 30-60 minutes).

iii. The beaker was covered with aluminium foil and left to stir overnight.

iv. 20g of glycerol was added next day and left to mix for one hour.

v. Hoyer's solution was aliquot into 50 ml of conical tubes.

2.1.16 Method for treatment of seedlings for submergence

De-oxygenated water was produced by putting freshly autoclaved water under vacuum for at least one hour. Seedlings were taken in 2ml Eppendorf tubes and immersed in de-oxygenated water. These tubes were kept in the dark (to avoid photosynthesis that may lead to generation of dissolved oxygen) for 2-4 hours at room temperature. After that de-oxygenated water was discarded through pipetting and GUS solution was added (as described in section: 2.1.14).

2.1.17 Method for treatment of seedlings for hypoxia

Seedlings were treated with nitrogen in a closed desiccator vacuum for 2-3 hrs at room temperature. The desiccator was further covered with a black bag in order to create darkness inside desiccator (to avoid photosynthesis that may lead to generation of dissolved oxygen). The nitrogen treatment in the chamber leads to hypoxic environment for the seedlings.

2.1.18 Method of cPTIO treatment of seedling

2-4-carboxyphenyl-4,4,5,5-tetramethylimidazoline-1-oxyl-3-oxide (cPTIO) was used as an NO scavenger to reduce the level of NO in the plant material (Gibbs *et al*., 2014). Seedlings were incubated in 5ml of autoclaved liquid MS media (pH-6.2) containing 500μM cPTIO in a small round petri dish. These petri dishes were kept in the growth room

in constant light (60μmols $m^{-2}s^{-1}$) at 22^0C for 4-5 hrs. Then after 4-5 hours liquid media was removed and seedling were washed with sterile water before incubating them in GUS solution at 37^0C.

2.2 Determination of protein through Western blotting

2.2.1 Plant material

The *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* Col-0 lines were used to detect RAP2.3-HA tagged proteins through Western blotting. Seeds were sterilized and plated on half-MS plates and chilled for three days at 4^0C. After 3 days plates were moved to 22°C in continuous light (light intensity 60μmols $m^{-2}s^{-1}$). Four days old seedlings were used for protein extraction.

2.2.2 Buffers used in Western blotting

2.2.2.1 Buffer composition for polyacrylamide gel preparation

Ingredients	Running gel 12%(10ml)	Stacking gel (5ml)
Water	3.3ml	3.3ml
Acryl/Bis-acryl	4.0ml	0.83ml
Tris (pH 8.8)	2.5ml	Tris (pH 6.8) 0.63
SDS (10%) (w/v)	0.1ml	0.05ml
APS (10% fresh) (w/v)	0.1ml	0.05ml
Temed	4μl	4μl

Table: 2.4 Buffer composition to prepare polyacrylamide gels

2.2.2.2 Buffer composition for membrane running in western blot

Ingredients	Amount
Water	700ml
Methanol	200ml
Tris-glycin buffer 10x	100ml

Table: 2.5 Composition of Tris-Gycine buffer

2.2.3 Western blotting technique for detection of proteins

Four days old *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* tagged Col-0 seedlings were crushed in pastle & mortar using liquid nitrogen. Powdered tissue was transferred to protein extraction buffer and mixed. The solution was then centrifuge at 13000 rpm for 10 min at 4^0C. Supernatant was transferred to new tubes. Protein concentration was detected using Bradford's method (Bradford *et al*., 1976).

2.2.3.1 Preperation of polyacrylamide gels

12% (w/v) polyacrylamide gels were prepared (Section: 2.2.2.1). Protein samples were loaded in gel and run in Tris/Glycin/SDS buffer (Geneflow) at 80 volts for 20 min. See Blue (Invitrogen) protein ladder was used for size comparison. After 20min voltage was increased to 120V. Gel was run until the loading dye was seen visible to reach near bottom of the gel (appx. time: 1 ½ hrs).

2.2.3.2 Transfer of protein from gel to membrane

Protein from gel was transferred to immune blot membrane (Bio-Rad) and run in Tris-Glycine buffer (Section: 2.2.2.2) at 80V for 2-2 ½ hrs.

2.2.3.3 Blocking

The membrane was transferred to 5% (w/v) milk (prepared with phosphate buffered saline (PBS) (Fisher-bio reagent) and 0.1% TWEEN (Fisher-bio reagent)). After 1 hr fresh milk containing first antibody α-HA (Sigma-Aldrich) 1:2500 was added onto membrane and kept overnight at 4^0C.

2.2.3.4 Washing of membrane

Next day, membrane was washed three times with 5% (w/v) milk solution. Thereafter, the second antibody Goat α–mouse (Sigma-Aldrich) was added in proportion of 1:10,000 and incubated for 1 hr. The membrane was the washed again three times with PBS solution.

2.2.3.5 Visualization

Membrane containing HA-tagged proteins was treated with the Pierce EVL western blotting substrate (Thermo Scientific). Membrane was the taken to dark room. Membrane was first placed on the film (High Performance Chemiluminescence Film). The film was then exposed for 2hrs and then developed.

2.3 Chromatin Immuno-precipitation (ChIP) technique

2.3.1 Plant material

The *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* tagged Col-0 lines were used for ChIP analysis. Seeds were sterilized and plated on half-MS plates and chilled for three days at 4^0C. After 3 days plates were moved to 22°C in continuous light (light intensity 60µmols $m^{-2}s^{-1}$). Four days old seedlings were used for ChIP experiment.

2.3.2 Buffers used in ChIP

2.3.2.1 Extraction buffer 1 (50 ml)

Ingredients	Amount from stock
0.4 M Sucrose	10ml 2M Sucrose
10 mM Tris-HCl, pH 8.0	1ml 500mM Tris-HCl, pH 8.0
10 mM $MgCl_2$	0.5ml 1M MgCl2
5 mM β-mercaptoethanol	17.4µl 14.3M β-mercaptoethanol
Protease inhibitor cocktail (PIC)	5 tablets
1mM PMSF	250µl 0.2M PMSF
1% w/v Triton X-100	30µl 100% Triton
Water	38.23ml

Table: 2.6 Composition of Buffer I for ChIP DNA extraction

2.3.2.2 Extraction buffer 2(3ml)

Ingredients	Amount from stock
0.25 M Sucrose	0.375ml 2M Sucrose
10 mM Tris-HCl, pH 8.0	60µl 500mM Tris-HCl, pH 8.0
10 mM $MgCl_2$	30µl 1M MgCl2
5 mM β-mercaptoethanol	1.05µl 14.3M β-mercaptoethanol
Protease inhibitor cocktail (PIC)	428µl 7X PIC
1mM PMSF	15µl 0.2M PMSF
1% w/v Triton X-100	30µl 100% Triton
Water	2.061ml

Table: 2.7 Composition of Buffer II for ChIP DNA extraction

2.3.2.3 Extraction buffer 3 (3ml)

Ingredients	Amount from stock
1.7 M Sucrose	2.55ml 2M Sucrose
10 mM Tris-HCl, pH 8.0	60μl 500mM Tris-HCl, pH 8.0
10 mM $MgCl_2$	6μl 1M MgCl2
5 mM β-mercaptoethanol	1.05μl 14.3M β-mercaptoethanol
Protease inhibitor cocktail	428μl 7X PIC
1mM PMSF	15μl 0.2M PMSF
0.15% Triton X-100	4.5μl 100% Triton

Table: 2.8 Composition of Buffer III for ChIP DNA extraction

2.3.2.4 Nuclei lysis buffer (1.5ml per sample)

Ingredients	Amount from stock
50 mM Tris-HCl, pH 8.0	150μl 500mM Tris
10 mM EDTA	300μl 50mM EDTA
1% SDS	150μl 10% SDS
Protease inhibitor cocktail (PIC)	214.3 7X PIC
1mM PMSF	7.5μl 0.2M PMSF
Water	678.5μl

Table: 2.9 Composition of Nuclei lysis buffer

2.3.2.5 ChIP buffer (10ml)

Ingredients	Amount from stock
10 mM Tris-HCl pH 8	200μl 500mM Tris
140 mM NaCl	280μl 5M NaCl
1mM EDTA	200μl 50mM EDTA
1% Triton X-100	1% Triton X-100
0.1% Na deoxycholate	200μl 5% Na-deoxycholate
Protease inhibitor cocktail (PIC)	1 tablet
1mM PMSF	50μl 0.2M PMSF
Water	9.170ml

Table: 2.10 Composition of ChIP buffer

2.3.2.6 Elution buffer (5ml)

Ingredients	Amount from stock
10mM Tris/HCl pH 7.5	50µl 1M Tris
50mM NaCl	50µl 5M NaCl
Water	4.9ml

Table: 2.11 Composition of elution buffer

2.3.2.7 TE buffer (5ml)

Ingredients	Amount from stock
10mM Tris/HCl pH 7.5	50µl 1M Tris
1mM EDTA pH8	100µl 50mM EDTA
Water	4.85ml

Table: 2.12 Composition of Tris-EDTA buffer

2.3.3 Process of ChIP

The ChIP was carried out following the protocol proposed by O'Neill *et al*., (2006).

2.3.3.1 Cross-linking of DNA and protein

Four days old *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* tagged Col-0 seedlings were crushed in pastle & mortar using liquid nitrogen. Powdered tissue was added to 25ml buffer I mixed with 1% formaldehyde (Sigma-Aldrich, UK) and incubated for 5min at room temperature. Formaldehyde was used to cross-link the proteins to the DNA. After 5 min 1.7ml 2M glycine was added to stop the cross-linking. The solution was filtered through miracloth filter.

2.3.3.2 Chromatin purification by a series of extraction buffers

Extract was centrifuged in a swinging bucket rotor at 4000 rpm for 15 min at 4^0C. After centrifugation a pellet was seen on the bottom of tube. Without disturbing the pellet the supernatant was discarded. The pellet was resuspended in 1ml buffer 2 (Table: 2.7) in a 1.5ml Eppendorf tube. All tubes were again centrifuged at 13000 rpm for 10 min at 4^0C. Thereafter, the pellet was resuspended to 1ml buffer (Table: 2.8). Again the solution was centrifuged at 13000 rpm for 60 min on 4^0C. After centrifugation supernatant was discarded and 500µl nuclei lysis buffer was added to the pellet.

2.3.3.3 Sonication of DNA

The lysate was processed for sonication to shear DNA to an average fragment size of 200-1000 bps. The following was performed at sonicator (Soniprep) set at amplitude 3 for 5 sec, sonication 60 sec and break 1 min. Samples were kept on ice during sonication. After sonication, the solution was centrifuged for 10 min on 4°C at 8000 rpm. The supernatant was transferred to 1.5 ml siliconized tubes.

2.3.3.4 Immuno-precipitation of ChIP products with specific antibodies

2.3.3.4.1 Preparation of beads

For immune-precipitation Hemagglutinin (HA) and Immunoglobulin G (IgG) were used. Beads were washed twice with chip buffer (Section: 2.3.2.5) and diluted in 1 volume chip buffer.

2.3.3.4.1 Preclear chromatin isolation

The supernatant obtained after sonication is incubated with 40μl IgG agarose beads on shaker for 1 hr at 4°C. The solution was then centrifuged at 4000 rpm for 30 sec to pellet beads. The supernatant is termed the precleared chromatin.

2.3.3.4.2 Immunopecipitation

140μl of precleared chromatin per IP (Immunoprecipitant) was transferred to a siliconized tube. One for negative control (IgG) and two for positive control (HA) and the rest of the precleared chromatin were saved as input. 20μl of HA beads were added to two “HA samples” and 20μl IgG beads were added to “IgG samples”. Samples with beads were incubated on rotator for overnight at 4°C.

2.3.3.5 Purification of immune complexes after immune-precipitation

2.3.3.5. 1 Washing steps for purification of immune-precipitants

The next day washing steps were followed for the purification of immunoprecipitant.

1. All tubes were centrifuged at 4000 rpm at 4°C for 30 sec to pellet immunoprecipitant. Supernatant was discarded.
2. 500μl of chip buffer was added to each tube and was again kept on rotator for 10 min. After ten minutes tubes were again centrifuged and supernatant was discarded.

1. First wash was done by adding 500μl of 150mM NaCl to beads. Samples were again rotated for 10 min at 4°C. After ten min tubes were again centrifuged and supernatant was discarded.
2. Second wash was done by adding 500μl of 500mM NaCl to beads. After five min tubes were again centrifuged and supernatant was discarded.
3. Third wash was done by adding 500μl of 250mM LiCl to beads. Samples were again rotated for 5 min at 4°C. After five min tubes were again centrifuged and supernatant was discarded.
4. Beads were treated with 500μl of TE buffer (Sec: 2.3.2.7). All samples were again rotated for 5 min at 4°C. After five min tubes were again centrifuged and supernatant was discarded. This step was repeated again.

   All washing steps were performed at 4°C.

2.3.3.5.2 Elution

All tubes were centrifuged. Supernatant was discarded. 100μl elution buffer (Section 2.3.2.6) with 50mM $NaHCO_3$ to each tube (10μl of 0.5M $NaHCO_3$) was added to pellet that contains the beads and incubated at 65°C on heating block for 30 min. 200μl elution buffer was also added to input sample and kept with other tubes at 65°C on heating block for 10 min. After 10 min tubes containing beads were centrifuged and supernatant was transferred to new tubes. 100μL elution buffer was added to tubes containing beads and again incubated at 65°C on heating block for 10 min. After 10 min tubes were centrifuged and supernatant was transferred to first elution.

2.3.3.6 Reverse cross-linking

Reverse crossing is the process in which DNA-protein complex gets separated. For this 8μl of 5M NaCl is added to each elute and also in input sample and left at 65°C on heating block for overnight. The next day 1μl of RNAase (1μl/U) was added to each tubes and left for 30 min at 65°C on heating block to remove RNA. After 10 min 10μl of 0.25M EDTA (pH-8.0) and 20μl of 0.5 Tris (pH-6.5) was added. 1μl of proteinase K was also added to remove background protein and kept for 1 hr at 50°C on heating block.

2.3.3.7 Isolation of ChIP DNA

After 1hr an equal volume of phenol:chloroform (1:1) was added and tubes vortexed slowly. All tubes were centrifuge for 5 min at max speed. Aqueous phase was transferred

to new siliconised tubes. 2.5 volume ethanol and 1/10 volume 3M sodium acetate (pH-5.2) was added to each tube and kept at -20^{0}C for overnight. Next day, all tubes were centrifuged at maximum speed for 20 min at 4°C. Supernatant was discarded and pellet was washed with 70% (v/v) ethanol. All tubes were again centrifuged at maximum speed for 5 min at 4°C. After centrifugation supernatant was discarded and pellets were kept at room temperature and air-dried. It was then resuspended in 25μl of TE Buffer.

2.3.4 Standard PCR on immuno-precipitated DNA

A standard PCR was performed on ChIP DNA to identify *ABI5* promoter amplification using *ABI5* qPCR primers. The primer pair used was:

ABI5_q forward 5' AAAACCTTTCAGTCAAAATTCTC 3'

ABI5_q reverse 5' CGGTGGCTTTGTGTTCCTC 3'

PCR reaction setup:

Components	**50µl reaction**
10X Phusion *Taq* Reaction Buffer	10μl
10 mM dNTPs	1.0μl
10 µM Forward Primer	0.3μl
10 µM Reverse Primer	0.3μl
Template DNA	2.0μl
Phusion *Taq* Polymerase	0.1μl
Nuclease-free water	36.3μl

PCR Reaction conditions:

Steps	Temp	time
Initial Duration	98°C5 min	
	98°C	30 sec
45 cycles	54°C	30 sec
	72°C	30 sec
Extension	72°C	7 min
Hold	10°C	

2.3.5 Agarose gel electrophoresis of immuno-precipitated DNA

20μl PCR product was loaded in a 1% (w/v) agarose gel containing 2.5μl ethidium bromide (Fisher Scientific). 1kbp plus DNA ladder (Invitrogen) was used for size comparison. Electrophoresis was performed at room temperature for 60-90 min at 90 volts. Gel images were captured by UV gel-doc (Bio-Rad).

2.3.6 Quantitative real-time PCR

Real-time PCR was performed in triplicate using ChIP DNA in a 12μl reaction mixture. 1X iQ™ SYBR® Green supermix (Bio-Rad Laboratories) mastermix and each of *ABI5* forward and reverse primers (Section: 2.3.4) used was.

PCR reaction setup:

Components	50μl reaction
Master mix	6μl
Nuclease-free water	0.8μl
10 μM Forward Primer	0.1μl
10 μM Reverse Primer	0.1μl
ChIP DNA (50 mg)	5μl

Each reaction mixture contained 7μl sensiFAST SYBR*Hi-Rox kit (Bio-line), 5μl DNA, 0.1μl of each *ABI5* promoter forward and reverse primers and 0.8μl molecular gradient

water to a final volume of 12μl. For each reaction 3 technical replicates were used in a 384-well plate (Thermo Fisher scientific). PCR conditions applied were 15 min at 95°C for initial activation and then 45 cycles of 15s at 94°C, 30s at 56°C and 30s at 72°C.

2.4 *PromERFVII:MC/MA-ERFVII-YFP* Construction

Transgenic lines containing *promHRE1:MC-HRE1-YFP*, *promHRE1:MA-HRE1-YFP*, *promHRE2:MC-HRE1-YFP, promHRE2:MA-HRE1-YFP*, *promRAP2.2:MC-RAP2.2-YFP, promRAP2.2:MA-RAP2.2-YFP*, *promRAP2.12:MC-RAP2.12-YFP* and *promRAP2.12:MA-RAP2.12-YFP* were generated.

2.4.1 Cloning of MC/MA ERFVII's

Wild type plants were taken to extract DNA. The extracted DNA was processed for amplification of target gene with their promoters. Amplicons were processed for cloning through gateway cloning system (Section: 2.1.6).

2.4.2 Selection of Transformed plants

T_1 seeds of transformed plants were plated on half-MS media plates containing hygromycin for selection of transgenic plants. 12 plants per selected line were transplanted in soil. Later seeds from T_2 plants of each line were harvested and again plated on hygromycin containing MS plates for identification of homozygous lines.

2.4.3 Determination of MC/MA-ERFVII::YFP by Confocal imaging

Seeds of homozygous line were plated on half-MS media plates. Plates were chilled at 4^0C for three days. After three days plates were kept in dark at 18^0C in growth room. Four days old etiolated seedlings were used to detect fluorescence of YFP tagged ERFVII's proteins in plant cells using SP5 confocal microscope (Leica). Fiji, software based on ImageJ was used to analyse YFP images.

CHAPTER-3

Investigation of relationship between N-end rule pathway and regulation of genes through binding of ERFVII's at two GCCGCC, *cis*-elements present in their promoter or at 5'UTR region

3.1 INTRODUCTION

3.1.1 GCC-Box mediated gene expression through Ethylene Response Factors (ERFs) transcription factors

The Ethylene Response Factors (ERFs) represent one of the largest families of transcription factors in the plant kingdom. These plant specific transcription factors function as *trans*-acting regulators that specifically interact with *cis*-elements such as 'GCCGCC' in promoters of different downstream genes (Park *et al*., 2001; Huang *et al*., 2004; Nakano *et al*., 2006; Lin *et al*., 2009; Zhang *et al*., 2009; Kohali *et al*., 2013)

The Group VII Ethylene Response Factor (ERFVII) transcription factors were found to be the first physiological substrates of the N-end rule pathway in *Arabidopsis* (Gibbs *et al.* 2011; Licausi *et al* 2011). In recent studies it was seen that downstream targets that contain 'GCCGCC' *cis*-elements in their promoter or 5'UTR region are regulated by N-end rule pathway through binding of ERFVII's (Gibbs *et al*., 2014). It has been shown previously that one of the ERFVII (RAP2.3) binds to the GCCGCC *cis*-elements (Yang *et al.* 2009: Jose *et al.,* 2014). Therefore, promoters that contain two GCC-boxes are expected to be bound by the ERFVII's. Previous work analysis of *ABI5* gene has shown that it contains two copies of the 'GCCGCC' *cis*-elements in the 5'UTR and its expression is enhanced in *prt6* suggesting that ERFVII's can activate expression through this promoter (Gibbs *et al.,* 2014).

2.1.2 Investigation of expression of identified five genes carrying 'GCCGCC' *cis*-elements in their promoter or in 5'UTR region through quantitative-Real Time-Polymerase Chain Reaction (q-RT-PCR)

3.1.2.1 Quantitative real time PCR

q-RT-PCR is used to measure mRNA and can be used for detecting and quantifying expression of a gene (Ferre *et al*., 1992; Piatak *et al*. 1993; McCulloch *et al*. 1995; Raeymaekers *et al*., 1995). This technique is based on measurement of the amplification of a target DNA sequence in presence of fluorescent dye. The fluorescence (amplification) produced is measured in real time. The cycle number at which a sample begins to amplify exponentially is known as the crossing point (Cp) or threshold cycle (C_T). The number of cycles required for the fluorescence to reach a specific threshold

level of detection is inversely related to the amount of nucleic acid that was originally in the sample (Rodriquez-Lazaro and Hernandez, 2013).

3.1.2.2 Application of q-RT-PCR

q-RT-PCR was used for detection of up or down regulation of genes in Col-0 and *prt6-1* backgrounds. For q-RT-PCR RNA cannot be used as template since it is single stranded and chances of its degradation is high therefore cDNA is used as template. The data obtained from q-RT-PCR was analysed through relative quantification (Livak *et al*., 2001; Pfaffl *et al*., 2001).

Relative quantification determines the difference between levels of expression of target gene to a reference gene acting as a normalizer. Relative quantification was calculated using the mathematical formula given by Pfaffl *i.e.,*

$$RE = \frac{(E\ \text{target})^{\Delta Cp\ \text{Target}}}{(E\ \text{ref})^{\Delta Cp\ \text{Ref}}}$$

Here RE is the relative expression ratio between the target and the reference gene. Etarget is the amplification efficiency of target gene; E ref is the amplification efficiency of reference gene. ΔCp iscrossing point deviation of the testing sample verses control sample. The amplification efficiency is calculated from the slope of standard curve in the log-linear range of the amplification (Klein *et al*., 1999). The reference gene used as a standard or normalizer should always be expressed at a constant level in all relevant times and conditions tested (Giulietti *et al.,* 2001)

3.2 Results

3.2.1 Analysis of expression properties of promoters having 'GCCGCC' *cis*-elements using GUS reporter gene construct in *Arabidopsis thaliana*

3.2.1.1 *In silico* search for genes containing GCC-boxes in promoter or 5'UTR region

In order to find if more *Arabidopsis* genes that exist with a similar organization of EBP elements to *ABI5*, a genome-wide search for promoters containing two consensus GCC-boxes was performed using the pat-match tool of the TAIR webpage

(http://www.arabidopsis.org/cgi-bin/patmatch/nph-patmatch.pl) and the motif “GCCGCC” as input for pattern search. Fourteen genes (Table: 3.1) were selected that are most similar to *ABI5* having two consensus GCC-boxes in close proximity to each other between range of 40-70 bp (with the exception of *ERD4*, that had a gap of 120 bp) (Fig: 3.1 a & b).

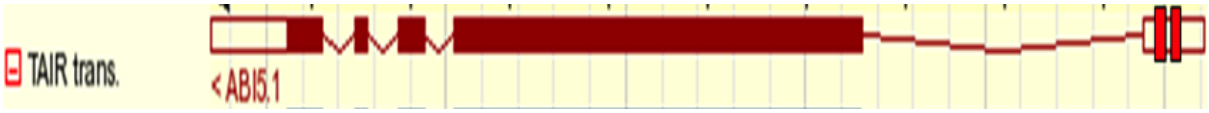
3.2.1.2 Identification of possible target promoters of the ERFVII’s

In this chapter, a search for several genes that contain two *cis*-elements in their promoter or 5’UTR region was done. The promoter that is a polynucleotide sequence enables transcription of a coding sequence in a cell. Therefore, a promoter act as a *cis*-acting transcriptional control element, including an enhancer, a promoter, a transcription terminator, an origin of replication, a chromosomal integration sequence and 5' and 3 ' untranslated regions which are involved in transcriptional regulation (Berg *et al*., 2006).

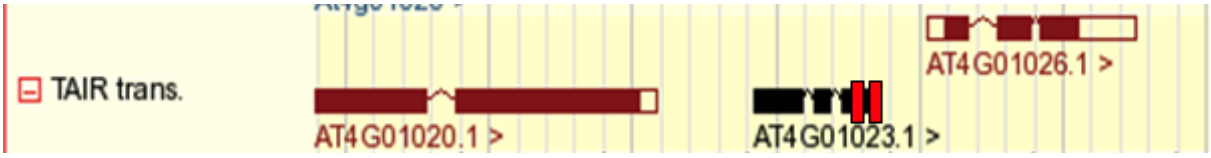
The fourteen genes that were selected have *cis*-element consists of a consensus sequence motif ‘GCCGCC’ (Yang *et al*., 2009) (Table: 3.1) in their promoter. It was found that AtEBP, an *Arabidopsis* protein binds to the GCC-boxes. Electrophoretic mobility-shift assay and DNase-I footprint analysis revealed that AtEBP could specifically bind to the GCC-boxes (Yang *et al*., 2009). Thus, it is expected that like AtEBP (RAP2.3) other ERFVII’s (HRE1, HRE2, RAP2.2 and RAP2.12) are also involved in regulation of above selected gene through N-end rule pathway by binding at two ‘GCCGCC’ *cis*-elemnts in their promoter.

To initiate the study of the regulation of the selected fourteen *Arabidopsis* genes that have two GCC-boxes in their promoter or 5’ UTR region in relation to N-end rule pathway, the promoter regions were amplified through PCR. Out of fourteen initial candidates only six of them could be amplified. Each of six amplified promoter (Fig: 3.2) was cloned through gateway cloning system. Cloned fragment sizes were verified by restriction digestion followed by sequencing. Sequencing result showed correct matches of the promoters of genes *AT4G01026, AT1G14810, AT3G54510*, *At3G13440* and *AT5G44420* while amplified promoter sequence of AT1G11540 did not tally. Finally, promoters of five genes *AT4G01026, AT1G14810, AT3G54510, At3G13440* and *AT5G44420* were cloned and transformed in Col-0 and *prt6-1* plants.

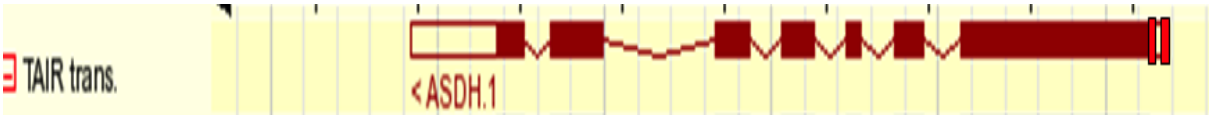
1. **AT2G36270 : ABI5 : (ABA signalling)**



2.) **AT4G01026 : PYL7 : (ABA receptor)**

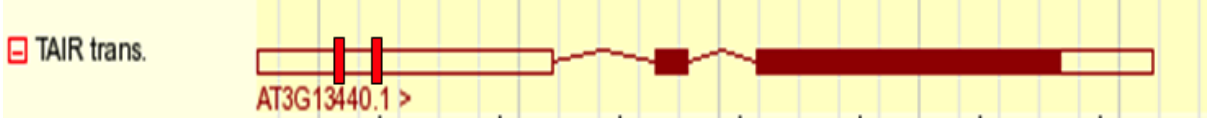


3.) **AT1G14810 : ? : (NAD-NADP binding protein)**



4.) **AT3G54510 : ERD4 : (Gibberellic acid mediated signalling pathway)**

6.) **AT3G13440 : ? : (Methyl transferase activity)**



5.) AT5G44420 : PDF 1.2 : (Pathogen induced factor)

Fig: 3.1 (a) Positioning of GCC-boxes in 5'UTR or promoter regions of target genes

1. ABI5: Two GCC-boxes are present in 5'UTR region of the gene. The difference between two boxes is 60 base pairs.

2. PYL7: Two GCC-boxes are present in 5'UTR region of the gene. The 5'UTR region of *PYL7* is situated very close to other overlapping gene. The difference between two boxes is 33 base pairs.

3. AT1G14810: Two GCC-boxes are present in 5'UTR region of the gene. The difference between two boxes is 45 base pairs.

4. ERD4: Two GCC-boxes are present in promoter region of the gene. The difference between two boxes is 120 base pairs.

5. AT3G13440: Two GCC-boxes are present in promoter region of the gene. The difference between two boxes is 66 base pairs.

6. AT5G44420: One GCC-box is present in promoter region of the gene. However, other one is the non-consensus sequence to GCCGCC and is represented with blue box.

Red-box: 'GCCGCC' *cis*-element. Blue-box: non-consensus sequence 'GCAGCC'. <and > represents the direction of gene regulation. Figures are taken from: http://atensembl.arabidopsis.info/Arabidopsis_thaliana_TAIR/index.html

Below is the sequence of promoter showing amplified segment having two GCCGCC, *cis* elements:

1. AT4G04925.1

-381 AAATATCAAAGAACGGAAATAAAAAAATAATTATTAATTAATTAAAATATTTTTGAGATA
-319 ATTACTCATTGATTGAGAAATAGTTTTTCAATCATTGTGGAGGAAATGAAAAGAAAAAAA
-288 AATTTTATTTTATTGTTTTTTCTTAAAATCTAAAAAATCTCATTTTTGTGACTTTAAAAA
-194 TTTGAACGTGAATTGAATAAAAAAGAGCAGCTATACTTTTAAAAATAAATAAAATCTTTA
-132 TGATTGATAAAATTTTGGCTTTTAATACTTTAAATTAAAACTAAAGTTTTTAAAGTGTCT
-70 TCAATCAACCCCGTATTATAGAGTGC**GGCGGC**TTGCGGCGACTTGC**GGCGGC**TTGCGGCG
-12 GAAGTTGCT**ATG**

2. AT1G79630.1

-830 AAAGGTTAACACAAGTGGATACTTCAATTATATCATTTTTTGGAAGAAAGACAAAACGGA
-770 GTTATTATATTTTTTTCCAAAACTTGGCTGGCTTCAGAAATCAGAATTTAGGTTTAATAA
-710 AGGCAGAAAAGGACTTGTGGCTGATCAATTGTGAATACTTAACTAACCATTAATTAACAA
-650 ATTCGAACCATTACACACAGACAGACGACAACTTTGACACGTGGCAAAACGTAATTGGGT
-590 TCAATATTTTCGGGTGAAATGGATGATACGATGCAGCCAAGAACCAAACCCGAGAAAATA
-530 GCAGTTTTCCCGAGAAAATAAACCAGAGGTCCCTAAAGCTAAAGCTAAAACTAAAAGCTA
-470 AAGTAGTGTGGATTTTCTAGATGCCAAGTTTCGTTTTTTGAGCCTTTTAAGCTCTTTTCT
-410 GGATTTCGTTTTTCTCGTTTCGCTCTATTCTTCGCCGAGCAAGTAATCATTGCCGACAAA
-350 CCCCCAGAAAAATCGTAGATTTCTTCAATCGGCGATTTAAAAAATTTGAACTTTTTCTTT
-290 AAAAGCTTTGAGCTCTTTTTTCTTCGGATTATTTTCAAA**ATG**GGTCTGTGTTATTCAGTA
-230 GATAGGACTACTGGGAAGGAACCTGGAAAGCAAGCTCCACCGCCACTACGGCGGAGACTG
-170 TGGAGGAGAGATCAGGCTCCGGAAGAT**GGCGGC**GGCCGAGAGATTTAAAGGGT**GGCGGC**G
-110 ATATCGAAGGGATTCCACAGGTTTTGGGTCGTTTGGTTTCGAATGGTTCAAGTAAGATTG
-50 CATGTCTTTACACACAGCAAGGCAAGAAAGGGACTAACCAAGACGCC**ATG**

3. AT2G02970.1

-583 GAATTGCATTTGGAAATCAGAAGAATCGGGTTTTGTGTATTCACAGTGTCGATTTTCGAA
-523 TGCTTAGTGAACGTTTCTGCGTGTGAGAATGATGATATTAGTGAAGCTGAAGATATCGCC
-463 ATTAAAGCTCTGATTTTGAGTTCTCATTATTCAAACTGACAAATCTTGAGGATAACGATA
-403 AGCCTGTTATGTGATCATGACCCGATTGCAAAATTAATAATCAAATAACCCGACCCGACC
-343 CGAATTGTAATTTAATAATTAGGTATAAAAATCGTAAAAGGCGCTACAACGACCAATTTA
-283 TTTTCCAATTTAATTTTGCTTTTTAATCTCTTTTATTTTTATTTTTATTTCTATCACCGA
-223 GAACGACATCGGAGACTTGAGAGC**GGCGGC**GAGCTTCACCGGCGCTGGTTAAC**GGCGGC**C
-163 CTCTCAATCTCCAGTATTCTCGGTTCAATAAAAGCTTCAAAAAAAGAAGAAAAAATCGAA
-103 AATTCACGAAAATATATATTAATTTCAATTTTCCATATTTGAATTCTTCCAGATCTTATG
-43 GTTCCTTGTTCTTCGTTCTCGTGGCGGTTTTGAATTGTTG**ATG**

3. AT1G14810.1

-365 CGTGTAAGATTCCAAAGTTAGCTATAGATGTCTTTGAGCTAAAAACCAGAATCACGAGCT
-305 CTTAAATATGATCCCTAAGCCTTGTTATATGACAAAAGACTTTTCAACAACTTAATATTT
-245 CCAAAAGAAGAAATGATGAAACACATGAGAAAGTTGTGTGAATATACTCAATCTGGTGAA
-185 TGAGCAGTTTCAAACATCATGCTGTTTCTGTATCTTAAATTAAACAATCTTGCTGTTTTG

-125 AATTTTGAAGATGTTGGTCAATTCACTAACACACTTTAAAGTCCCAAGCTTTTTTCTATC
-65 AGCCGCCTCCAAAATCTCTCTATCTCTCTGTCTCTCACTTCTTCTTCCTCACGGCGGCGA
-5 CAATG

5. AT3G60500.2

-659 GTCAAAGTTAATCTTTCCAAGCTGAAACAACAATTGCTGAGATATTAAAAAGTTCAACAT
-599 ACTACTCATTGACTTTCTTAACATTTACACCAAAAAGATAAAATAAAAAACTCATTGACG
-539 TTTTTCCCTTTGAAGTTAGTAGCCCAAAAAGCATCAAGATAATGTCTGTAAGCCCATGGG
-479 CTTATATCCATTATAATTTTTTTTGTCAAATAACAAATAATATATGAAACGACGTCGGTA
-419 ACAGCTGAAACGCGCCGGTGTTTGTTGATCCAGAACGACAGCGTTTTTAGTTATTCATCT
-359 CGAATTTTTCCGAAACCTGTGTTTTCCCATCACCAAGCTTTCTCTGAATCGCCGCCGTGT
-299 TTACATTTACTGTTGTTAAGGCCGAGAGATAGATAGAAAAAAACATCTGTGCGATTTTGC
-239 CGCCTCTCCCTAAAATCTTTGGAGCTCCTTCGTAGTTAGCCGGAGGTGAAGTACAATAAA
-179 AGAAAGAAGAAAGTCCAAGTGTCTCTTCTCATCTCTTTTCTCCAGTAGGTACGTTTTGGT
-119 TTAATTTACAGAGTTGATTCTAAGCCTAAGAGAAGGTTTTCTATCGATTAGAGTGTTTGA
-59 TTTGCTGTAAGGGTTGCTCAAAGTTTTGATTTTTAGTTTACTAGTGATAGCAGATAATG

6. AT3G13440.1

-1482 ATTTGCTATATCTCAACCACTAATTTAAGGGGCTTTCGGCCCATTACACGCATGGGAAG
-1423 ATTCAATTATGTTGATTTTGTAAGTTTTGGGGCTAAAGTGTAAATTATTGAAACTTGTA
-1364 GGCTTGATTTCATTATTCCAAGTTTCCACTTTTTCTCCTCCGCCGTAGCCGGTGCTTCT
-1305 TCTTCGTGAAAATCTTTCGTCTGAGATTTTTTCTTTACTAGTTGAATAAAATTGAAGAT
-1246 TTCGAAGACTACCCAAAACCGCGTTTTACAAAAAGCCAGATTCTTTTTTTTTTTTTTTT
-1187 TTTTTATAATTTCATTAAAACTACCCAATGGGCTTGGGCCGAAACTTACAAACAATTCG
-1128 GTCCAGTTACAAAATTAGGGTTTTCTGACACATCGCCAGGAAAAAGAAAAAAGCTCGCC
-1069 GCCTGTTGACCGTCGCTAGCAGAGCGACCAAAACCTAGATCCGGCGGACTCAAGTCCCA
-1010 CAGCTTTGTGGCCGCCTCCGAGAATCGTCCACCACCGACGAAAGATTCCGGTTTTCGTC
-951 TGTAGAGACTCAACTCCAACGCACCAACCCACCACAAGACTCCCGGAAAGCTAGCCGCT
-892 CCGACGTAGCGGTTCCACGGAACCGGATCCCGACCTAGCAACCCGAAACACTCAGCGAT
-833 AACACTTCTCTCCCATCACCGGAAAATAGCAAAGACCCAGAGCCAAACACACGTCGATC
-774 CGGACCGGTGGAGACCAAAAACGGCAGACCTAAAACTGGTGGCTCAGTCAGACTCTCTA
-715 GATCCAAAACCCGCTGCAGCCATCTAAGATCCAAACAGATCTACTGTTAACCACCGCAT
-656 CAGACAACGAACCCCCCCTTACCCCTCTCCAACTTCCGCCGTGGACGCAGCGACTAGAA
-597 TCACGGGTGAGCCTGAGAACCTCGCAGCAACCCTCAGCGTCCATAACCAAGAGACCCCC
-538 GCCGCAACCTTCTCGACCGACAAAACCAGAGCCAAATCCACATCAGAGAGAGAGCTCCA
-479 CACCACACGTTGACGGAGACAAGCATCGGCTCACTCCTGGACTTCCTTGAAGAGGCTTC
-420 AAACCGACGATGAAGAGCAGATCGGAACAAAAAGAAAAAGGAAACTAAATCTACTTACC
-361 GGAAATTGTCTTCAAGAGAGAGAGAAGGGCACGACTTTCAAGGCTACGATTGAGCCTCC
-302 ATCTCATAAAGCCAGATTCTTTATCTTCCATAATCGTATTAAAGCTACGAAAACCTCGA
-243 TTCGTGTTCAAAACTTGGATTTGCTCGCTATCGAAAGGTCTGACCTTTAATCCTTTTCA
-184 GGGTTTAAGTTTCAGATATGATTTCTCGTATCGGTATTATAGCTTAGAGGAAAAGTGTC
-105 GTAGTTTTGGCAGATTGAAATTAAAATTGGCTCTGTATTCCCAAATAGTTGAAACCCAT
-125 TTGATGAAGCTATCTTATCAGACTTTTACGCATAAGTTTTCTTGTGGTCTTGGATCGTG
-66 TCTAGTGGCAATTATCTTCTCTGAGTTCTTTTGCTTAGTTTGATTGTCTGTTGTTTTAG
-7 ACGAATG

7. AT1G30820.1

-927 GAATTAGGTTATGATACGAAATAAATCAATATTTTAAAACAAATTAGAAATTGGTATACA
-867 ATAAATTTGCTAAGTACAATTACTATCATATGATATCACAAATAAAACTAATAAATTATT
-807 TAGTAATGATTACATTAGTGATTAAAATAACTTTGTTTACATCTTCATAATGGTTTTATA
-747 GAAGGTAAATTAACTATAAAAAATTTGGTCGTATCCGCAAATATACAGAAACAATCCAAA
-687 TATAGAAAACACAATATTATAAAAATAAAAAATAGGTGGAGGGTTTACATTGGGGTTTTG
-627 AGAGAGGCGCTTCTCTGCTGCAAATTGCCTTTACTGTGTTTCTCTCTTCCCCTTTATAAA
-567 ACCGCTTCCGTCGTCAATGGCGTGAGCCTAAACAAGAAAGAAAAACCATTCTTCTCTCTT
-507 TTCGTTTTTTTCATTCTCTCTCTCTCTACTCCATCGTTTTCACACTTTTCTC**GCCGCC**CT
-447 TTATCCTTATCCGATAAACCCACAAATAAAGACAGACTCTTTTGTTCTTTAAACAGAGAT
-387 TGTAAATCCAGA**GCCGCC**GGAAAATACAACCGCCGGTTTACTCTCAGGTCTCTCTTTCAC
-327 TCTCTGTTTCTATCTCTGGTTCATACTCTTGTGTCTTCTGGGTTTGATACAAATGATAAT
-267 GCAACAGCTTTTGCTCTATTTGAATACTGATTTGAACCAATAAGTTTTTATTTTTTAATC
-207 TTTTTGGTCATCCCCTGTTTCTGATTCTGCGATAAAGAGATGTTAAATCTAGAAAAATCC
-147 AAATTCTCTCTCTGTTAAGCATTTCTCAAAAGCTTCACGCTTTCTCAGTTTCTGGTTTCT
-87 AGATTGTTGTTGACTGATTTTGGTTGGTGATATCTCTGCAGGATTAGTTGATATTATTCG
-27 ATTTTAAAAATCAGTAGGAAGAAG**ATG**

8. AT3G54510.1

-703 AGCCACTAAACAAATTCCTTTCGTTCATCAAAGTCCACAAATTCACTGAAAGTAATGATT
-643 CCCAAGCATTTTTTCCCACCAATTCAACAATTTTTAAGTACAACACTTTAGTGATTCTAA
-583 TTTCACAAAAATAACATTTTGTTGTAAAACAAATTTAGTAGTACTAGAAATTTCCTTTTC
-523 AGGGATGATGCAAAAATTCAATTATCACACTGCTATTTTGGTCGAGGTGTAATAAAAAAA
-463 CAAGGAGACAATTGCGGCGTGGAGGCTTTCTGTTTACATGCAGCACAAGAAGGGCCACAT
-403 GACATGACAGATGTGACTCTGTTTCCCATAGTCAGAGAAAAGAAAAAATATGACGCCAGA
-343 AAGCTTGTTAGTCGTCGGCGTCGATCAACATAGGACTCGCCGTCGTTGCACTCTGGCTCT
-283 TCTCCGTCCTCAAGAAGCAGCCCCGTAACGCCGTCGTATACTATGCTC**GCCGCC**TCTCCG
-223 ATCGCCACCACCACCGTCCTCTGTCTCTACACTCATCCCTTTGCCTCCCTCGTTTCCTTC
-163 CTTCCGTTGCGTGGATCCCACGCGCCTTCCGTGTCCCCGAGGACGAGATCCTCA**GCCGCC**
-103 ACGGCCTTGATGCTCTCGTCCTCATCAGGCTCTTCAAATTCGGGTAAGGAAGATTTGGTC
-43 TTATCTCGCATATAAACACGATTATCAGATTCTTTGAACT**ATG**

9. AT1G11540.1

-583 AGGAAAAATAATTGCATTTGTCATCTTTAGTTTTTTATTTCACCAAATTGTCTTAAATTA
-423 ATTAATTCCGTAATAAATACCTTGTATTAGAATTGCAGTTGACTAAACCCTTTATTGATT
-363 CATATATAAGTATAAACCAAACAATAAATGAGCATAAAATTTCTAACGTTCCAAACAAAA
-303 ATAAACCCCAAAACTTTCAAAAGGTTACAAACAACCCCCTAAGTAAAGAAAATGACGATC
-243 AATTTAGTCCCTCTTCTCCTCTCCCTAATCACCATCTTCACCATCTTCAATCCATCCGCC
-183 CTTGCAGACCAGAGTCAAATCCGCAGCACAACAAGCCAATGGCTCCATTTCTCGGAGGAC
-123 CTCAGCCAATCAACTATTGAAGTCTCCATCCCCACCATAATC**GCCGCC**GTTCTCTCGTTC
-63 TTC**GCCGCC**TCTATCTCCAGCGCCGGAGGAATC**GGCGGC**GGTGGACTCTTCCTCTCGATA
-3 **ATG**

10. AT5G12330.1

-635 CAAAGGATCTTGGATATTTATGTAGGCCACTAGACTGGTAAGTACAAAACGCGTATCTAC
-575 CGAACCATGTACCTAAAAGAATGAACCGGTGGAGCCGGTTCTGGAGAAAACAGAAGCCAA
-515 GCGCAAGTGGAGCCACGCGCACGAGCGGGGGTGGGTGTGAACTTCGTCATCTGATAAAGA
-455 TCAATAGAAAATACATAAATTATTATAAACTTCAGCTAAAATCCAAGAAAAGGAGAAGAA
-395 GGTATCATGTGCCATTTCTTTGAGCTATTATTATAAAATAAAATAAAACTATAAAATAAA
-335 GTACATTGGAGCTTTGTGCTGTGACTTTTATTTTTTTCTATTAAATTTTTTTATTTTATT
-275 TTATTTTTGGGATTAGGGTTTTGGTATAATCAGCTGTGGCTACAACGAGGAGGCATAATC
-215 TACCGACGTCGGATTCCGGTGCGTTCACCGACTG**GGCGGC**CACGACGACGACACCTTCCC
-155 GTGCCACCGAGGATCTCTCTCTTGGTTTCAATGCTGGTTCCTCCGTCATCCACGGTGGTT
-95 TAGGCTCAGCTTCCGTC**GCCGCC**GGCGTTCCATCATGGCCTCCGGGATCTTCTGTCCGTT
-35 ACGGCCTACCTTCTTC**GGCGGC**GGCGACGGAG**ATG**

11. AT3G19290.1

-1116 TTTAAAAAATTGACCTAAAAACAAAGAATAATCATTATTATGATTTAACTCCAATGAGA
-1057 AAACAATTCCAAGCACGATCTCCAACCACTCGAATCTTCACTGTCGGAAACAAACGGCC
-998 AATGTTTTACCACGTCGCATTCTTAATTGGATGTAAGTCTTTACAAGCACTTAAAACAT
-939 TAAACTCTTTTTTCTCATTCGTCCACGTTTCTTTTTAAATTAATACTGTTAACAGATAA
-880 TTATTCACAAAAAAATAATATCTGAAGAACAAAAAAAAAGGTTTCAGGGCCTTAAAATG
-821 GCTGCTCGTGTCGA**GGCGGC**AGCAACACACTTGTCGTCTCTCCCTGTAGCTTCTCTGGA
-762 GTTTTCGCAGTCACGAATCTTGTTTTGAAGAAATTAATTTGGTAAAACATATTTTCGAG
-703 CTCCTTTCTCTTTTATTATCCGTTTTCTGACTTCTGGATCTGTCTTTGTCGGATCATCC
-644 AATTTTTTTGTTGTCTTTCCTCTGTTTCGTTTATGAATCATGATGATGTACATGTTGTA
-585 ACTGAAGTAACTTTTGGTTATCAAATTAGAGTTAAAGAGAGAATTTTTACGAGATATTT
-526 GTTATAAATTTATCGATTTGATTCTTGTTGCTATTAGGGTTTTGTCTATTTGATTCGAT
-467 CGGGTTTTTTTTAAGAGAAATGGGTTGTGGAATCTGGAATAGATCCTTCTCATGCTTGT
-408 TGGGTTTTGCTGATTTTCTCGTTTTTGGTTTTATGAGTAAACTCTGTATCTGGTGTGAA
-349 TTCGTTTTCCATCAAAAAAAGTCATCAAATTACGTTTTTTTATAAACATTATCCAAAGA
-290 TCTGACCTTTATTGATTGACTTTACTGTGCTTTTACATTTGTTCGCACGCATTTATGTG
-231 GTTTGTCCAGTTTCACTTTAGTATTCTTTCTCTTATATGGATGTGAACATGCATTAGTG
-172 ACACATAGTTCACTCCTTTTTATTTTCCTCGGAAGTTCTCTTTTTCATGGATTCTTAGC
-113 ATTATGTAATATACACTCTAACTGCTTTACTTTGATTTTAGAACAAGGGTTTAGGGCTT
-54 GGATGCTTTGTTTTCATTGAAAAAGAAGTAGAAGGAGTGTATACAAGGATT**ATG**

12. AT4G01026

-630 GGAGATTGAAACAATGGTGGAAAAAGT**GGCGGC**TATGGCCACAAAGGCGGGAGAAACGGC
-570 GACAAT**GGCGGC**AGATATGGTGAAAGAAGCTGAGGAGACGATGGAAACAGCTAAAGCTAA
-510 TATGTCCAAAGCCTTTGTGGTAATGAAGTCGGTGAATTGGAACGTGTAAATCGGGTCAAA
-450 ACAGAGTTTTATTGTGATCTCATGCTGACATCAGCTACAATCTTTATCTAATAAGATAGA
-390 TTCTCACAGATTCTTTTATCTATCTACTTTTTAGAGAAGATGAGATCATACTTCGGAGAT
-330 AGATAGGTGTGTATAAAAATTGGGAAGCTTACTTGGCAGGAGAAGAGAGATATACACACG
-271 TGCTTAAAGTCAACAGCTAATCCAAAAAGGTAGACGACAAACCAGTTAAACTAAGACACG
-211 TAATCTATCTTAAAGATTTGTCGGTTCATCGTAAATCCGTTGGTGGATTAATATATGTCG
-151 GTGTCTTCGTTGATTTTCTTAGCCTAACCAACCAAACATAAATAAAAATCGAAGCTTTAC
-91 TTTGTTTGCTTCTATTATATCCACTTCCAACAAGATTCCTCAAAAGGAAAGAAGAGAAAG
-31 AAGATAAGAGATATCTGGTTGGTCGGAG**ATG**

13. AT1G13740.1

-598 TTGTTAATTTTTGAGGTAATATTTTTTTTTTTCGAGTTTGAGATTTTTTTTTTTCTCTTT
-538 TACAAAAACTTGTGCAAAAAAATTAAAAAGAAAAGGGGAAAAAGGAGAAAACGGATCTTT
-478 TTACTTGGGAGACACGTGTCGTGACACTGGATAACTCAGTCACTTTACGACACGTTTCTT
-418 GAAGCAGCACTATGGAAC**GCCGAC**GTAATCCACCGGCCACGTGTCGTTCAGGCAGCTCTC
-358 TC**GGCGGC**CACTTGCCTCCTCTCAGTCAAGGCCTGTCTCTCTTTCTTTCTTTCTTTCTTC
-298 TTCTCTCTCTTTTCCTCTAATGTCCAAAAGGGTCTTCTTCTTTGTCTCTGCTCTTCACTT
-238 TCAATTATAACCTCTTCGTCTTCTTCGCCGTCTCTGGTATTAACTCCGGCGAGTTTTTTT
-178 TTTTTTTTTACTTTGATATGAATTTGCGAGATTTACTGTTTAATTCATGATGACAATTTT
-118 GGATTTTTGATCGGATGTGAAATTCAGAGTTTTTCATTCTTCTTCATCATCTTATAACAA
-58 AGAAGAAACAAAACAAAAAACTTTTTTTTTCTTTCTTTCTATTGTTTCGGAATCT**ATG**

14. AT5G44420

-999 TACAAATTTAAATTCTGATCTCTTATAATGTTCGGTTTTTCCTTTTTTATTTATTAAGT
-940 TAGTTAAAATTTGCAGTTATTTTGTTGAATGTCGTTGTTTACGAATTTACGAATAATAC
-881 CTTTATAGCTAATCTACAAAATTTTGATGACTGACAACACCGTTAATGTTTTTTTTTAA
-822 ATTACCCTGAGCCTCTCACTTGCGGTCAGACCATGCATGTCGATAGTCCATTACGTTTA
-763 AGGCCACAATCAACTATAGTTTGTTTATCAATAGCCAACTAAGCTAACTTTTAGGTTCC
-704 TGCCCTCTCCGTTCCTCCGGTACCAATCGTTTCTTTGTCCCTTCGATAGTTTGAAAACC
-645 TACCGACGGTGAGAGCAAAATATTGATGAATCATCCAATTTTCAGTAATAGGTGTGTCC
-586 CAGGGATATATAAATGGCGAACTACGCGAGAACGGTTCCTTGTTCTGCAAACTTGGCAC
-527 TTGGCGGAACAATGCTGCTCTTGAGATCAACTGTCTAGGGGTGATCCTCTAATCGAAAA
-468 ATGTTGTATTTGTTCGACGATGACGAAGGTCAGACTATGAACTGCACAGTCTGCACTTG
-409 TCCTAACCGCGAGAATCTCTGACATCAATATACTTGTGTAACTATGGCTTGGTTAAGAT
-350 ATTATTTTCTTGAGTCTTAATCCATTCAGATTAACCA**GCCGCC**CATGTGAACGATGTAG
-291 CATTAGCTAAAAGCCGAA**GCAGCC**GCTTAGGTTACTTTAGATATCGACAGAGAAATATA
-232 TGTGGTGGAGAAACCAGCCATCAACAAACAAAAAGCAAGATCTTATCTTTTGATATTGG
-173 CTACGGGAAGATGATGTCTGTTTAATGTGTGGGGTTACCACGTTATTGTACGATGCACA
-114 AGTAGAAGATTAACCCACTACCATTTCATTATAAATAGACGTTGATCTTTGGCTTATTT
-55 CTTCACACAACACATACATCTATACATTGAAAACAAAATAGTAATAATCATC**ATG**

Fig: 3.2 (b) The sequence of fourteen genes showing pattern of two GCCGCC sites in the promoter or 5'UTR region of genes.

Two GCC-boxes are shown in highlighted in red, one non-consensus sequence (GCAGCC) in blue. The underlined ATG sequence represents the beginning of gene. Number of base pairs between two GCC-boxes represents the distance between two boxes.

The distance between two GCCGCC sites in promoter of *PYL7* is 33 bp, in *AT1G14810* is 45 bp, in *ERD4* is 120 bp, in *AT3G13440* is 66 bp and in *PDF1.2* there is only one GCCGCC site present (Fig: 3.1). In above sequences GCC-boxes are represented in red bold letters, GCAGCC (non-consensus sequence) in blue bold letters, ATG (bold UL), 5'UTR (green), and intron (blue) highlighted (Fig: 3.2)

S. no.	Gene	Function	Pat-match 1000 TAIR10 'GCCGCC'	No. of Boxes present in promoter or 5'UTR region of a gene	Distance between two GCC-boxes (in bp)	Region of GCC-boxes	Expected size of fragment to be amplified
1.	AT1G13740	AFP2 ABI5 binding protein 2	yes	2	38	promoter	595bp
2.	AT1G30820	CTP synthase family protein	yes	2	74	5'UTR	929BP
3.	AT3G13440	S-adenosyl-L-methionine-dependent methyltransferases	yes	2	66	promoter	1,541bp
4.	AT3G19290	ABF4, AREB2 ABRE binding factor 4	yes	1	-	promoter	1,114bp
5.	AT3G54510	ERD 4 Early-responsive to dehydration stress protein	yes	2	120	promoter	697bp
6.	AT3G60500	CER7 3'-5'-exoribonuclease family protein	yes	2	62	promoter	657bp
7.	AT4G01026	PYL7, RCAR2 ABA receptor protein	yes	2	33	5'UTR	629bp
8.	AT4G04925	molecular function unknown	yes	2	14	5'UTR	379bp
9.	AT1G79630	Protein phosphatase 2C family protein	yes	2	20	5'UTR	827bp
10.	AT2G02970	GDA1/CD39 nucleoside phosphatase family protein	yes	2	23	5'UTR	582bp
11.	AT1G14810	Semialdehyde dehydrogenase family protein	yes	2	45	5'UTR	362bp

12.	AT1G11540	Sulfite exporter TauE/SafE family protein	yes	3	15, 24	5'UTR	480bp
13.	AT5G12330	LRP1 Lateral root primordium (LRP) protein-related	yes	3	97, 53	5'UTR	632bp
14.	AT5G44420	PDF1.2, PDF1.2A, LCR77 plant defensin 1.2	yes	1	0	promoter	1,022bp

Table: 3.1 List of fourteen genes selected from pat-match tool of the TAIR.

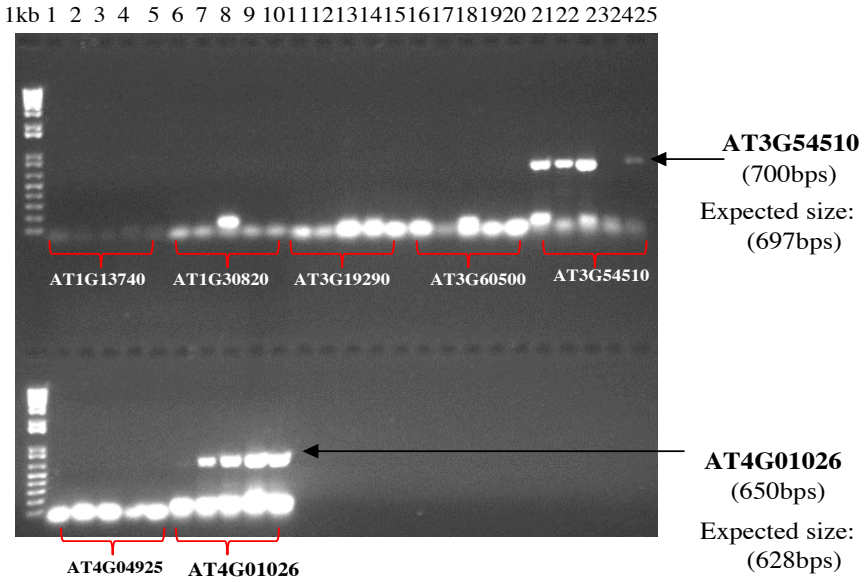
A table describing the fourteen genes function, the number of GCC-boxes present in the promoter or 5'UTR region of that gene with the number of base pairs difference between the two GCC-boxes. The region of two GCCGCC site and the expected size of fragments to be amplified for cloning.

3.2.2 Analysis of promoter::GUS activity

3.2.2.1 Agarose gel electrophoresis of amplified promoter fragments

DNA obtained after PCR promoter amplification was run on a 1% (w/v) agarose gel (Sigma) and band sizes were compared with a 1kb ladder (Fermentas). Results showed that promoters of six genes were successfully amplified (Fig: 3.2).

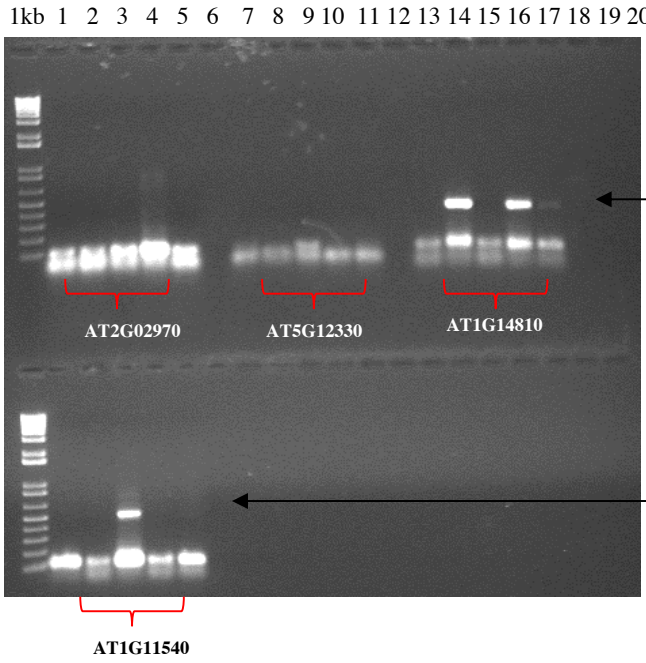
A.



AT3G54510
(700bps)
Expected size:
(697bps)

AT4G01026
(650bps)
Expected size:
(628bps)

B.



AT1G14810
(400bps)
Expected size:
(362bps)

AT1G11540
(500bps)
Expected size:
(480bps)

C.

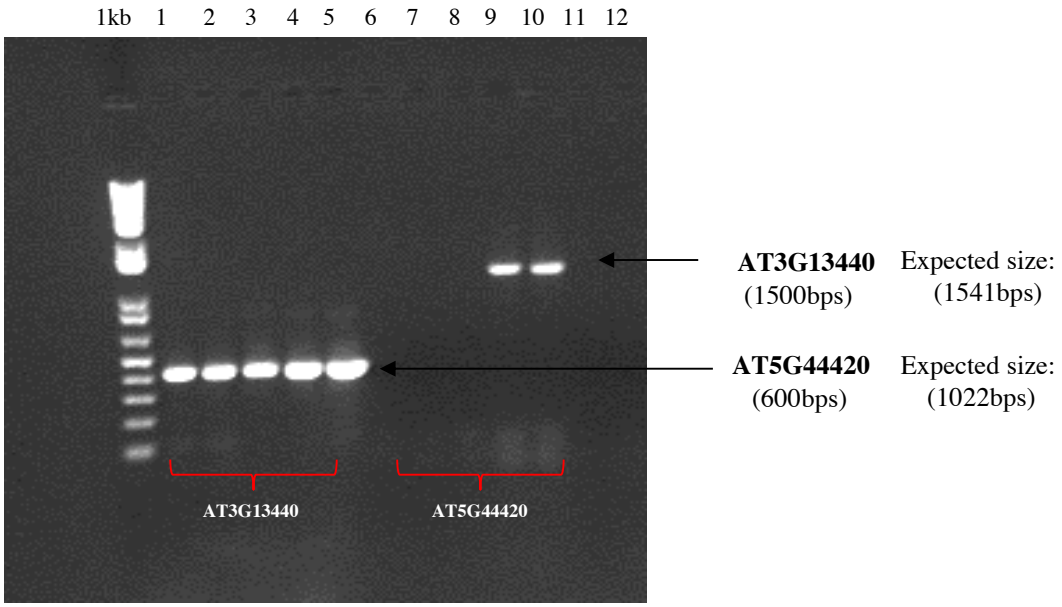


Fig: 3.2 The promoter amplification through PCR from five different *Arabidopsis* Col-0 leaf template DNA samples.

A. Gene *AT3G54510* shows amplified bands of approximately 700 bp in four Col-0 DNA samples and gene *AT4G01026* shows approximately 650bp size amplified bands in four samples.

B. Gene *AT1G14810* shows amplified bands approximately 400bp size in two DNA samples and gene *AT1G11540* shows approximately 500bp size amplified bands in one samples.

C. Gene *AT3G13440* shows amplified bands approximately 1500bp size in two DNA samples and gene *AT5G44420* shows approximately 600bp size amplified bands in all five DNA samples.1kb ladder is used for size measurement of bands and (A) 1-21, (B) 1-20 and (C) 1-14 are the numbers of wells.

3.2.2.2 Transformation of recombinant clones into competent *Agrobacterium tumefaciens*

In order to examine the difference in expression pattern driven by the *promoter::GUS* constructs, the positive clones containing *promAt3g54510::GUS*, *promPYL7::GUS*, *promAt1g14810::GUS*, *promAt3g13440::GUS* and *promPDF1.2::GUS* were separately transformed into both Col-0 and *prt6-1* plants (as described in section 4.9).

3.2.2.3 Analysis of GUS activity of transformed *promoter::GUS* constructs

The flowers and leaves from independent transgenic lines of Col-0 and *prt6-1* segregating T_2 plants were first confirmed for presence of GUS gene through genotyping. Plants that showed positive GUS amplification by genotyping PCR result were stained for GUS activity to analyse expression patterns driven by the transformed *promoter::GUS* constructs. For the *promPYL7::GUS,* 5 independent lines of Col-0 and 3 independent lines of *prt6-1* were tested, whilst for *promAt1g14810::GUS*, 5 independent lines were tested for Col-0 and 2 independent lines were tested for *prt6-1*. For *promERD4::GUS,* 8 independent lines of Col-0 and 5 independent lines of *prt6-1* and of *promAt3g13440::GUS,* 8 independent lines of Col-0 and 7 independent lines of *prt6-1* whereas, for *promPDF1.2::GUS* 8 independent lines of Col-0 and 7 independent lines of *prt6-1* were tested.

	Col-0	***prt6-1***
PYL7	5	3
ERD4	8	5
AT1G14810	5	2
AT3G13440	8	7
PDF1.2	8	7

Table 3.2 Number of lines in T_2 generation used for promoter::GUS analysis

The leaves of transgenic *A. thaliana* carrying *promPYL7::GUS* showed staining in the vasculature of the leaf in both Col-0 and *prt6-1*. However, in *prt6-1* the vasculature of the entire leaf showed stronger blue colouring, while in flowers GUS activity was observed in stigma and sepal regions in both Col-0 and *prt6-1*. Again staining was seen to be more intense in case of *prt6-1* (Fig: 3.3).

In contrast, histochemical GUS staining of *promAt1g14810::GUS* in leaves and flowers showed a different result. The Col-0 leaves only showed staining at the tip area whereas in *prt6-1* GUS activity staining was observed in the veins of leaves. In flowers no staining was observed in Col-0 and only basal part of stigma was seen stained in *prt6-1* (Fig: 3.4). Analysis of GUS staining of *promERD4::GUS* construct in the leaves of Col-0 and *prt6-1* revealed no difference in expression pattern. In contrast expression was observed in flowers. The Col-0 flowers showed GUS activity in the pistil region where as in *prt6-1* expression was observed in both pistil and filament (Fig: 3.5). The promoter activity of *promAt3g13440::GUS* is relatively high in both *prt6-1* leaves and flowers in comparison to Col-0 (Fig: 3.6). Promoter driven GUS expression occurred throughout in *prt6-1* leaves where as in Col-0 expression is more at basal part than tip region. While in flower expression was much higher in *prt6-1* than Col-0. In case of *promPDF1.2::GUS*, the activity was high in leaves and flowers showing undistinguishable differentiation between Col-0 and *prt6-1* (Fig: 3.6).

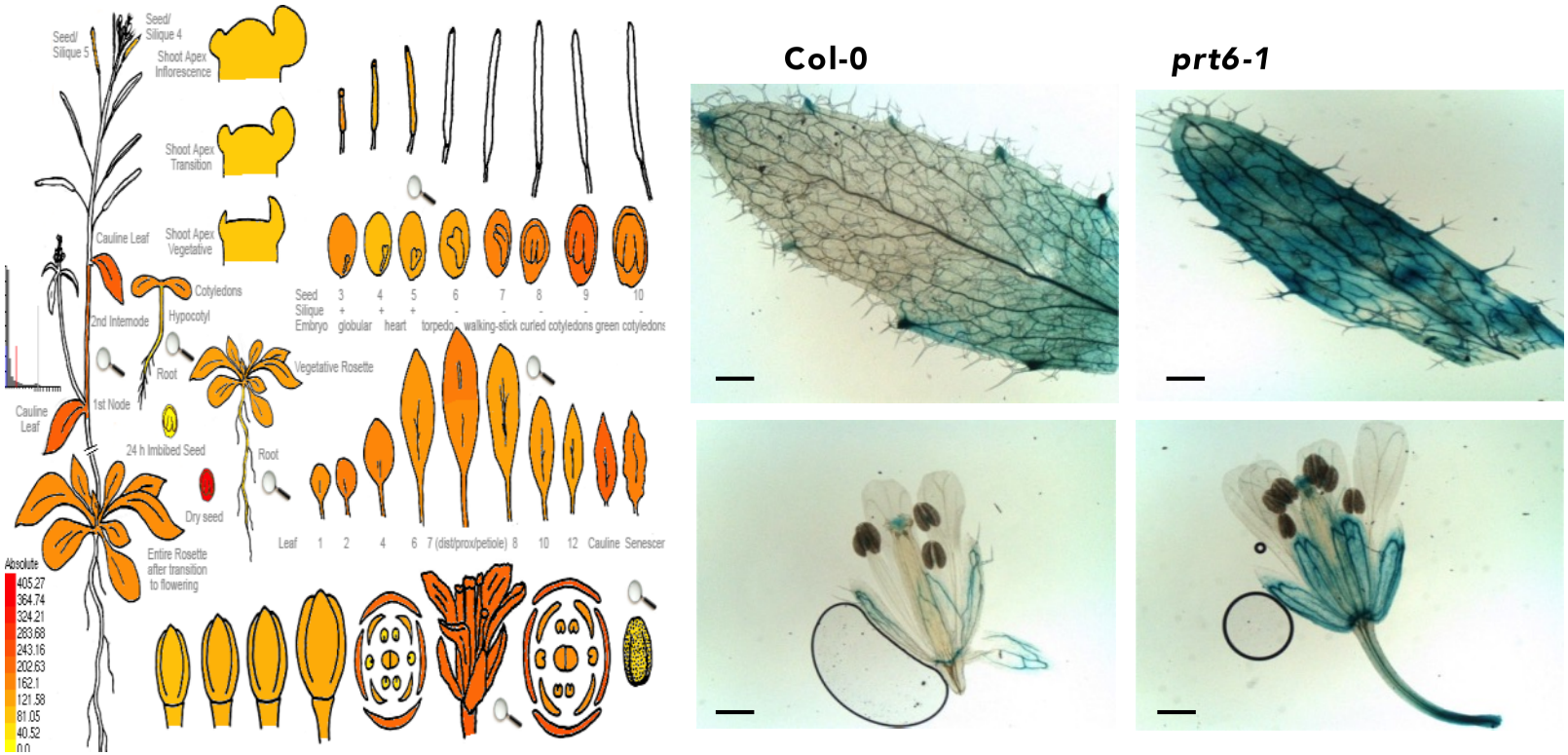


Fig: 3.3 Representative images of *PYL7* expression in different parts of plant obtained from eFP browser and the GUS activity driven by the *promPYL7::GUS* in leaf & flower of an independent line of Col-0 and *prt6-1*

The eFP browser image shows expression of gene in different parts of plant. Level of expression is analysed by colour scale (highest is represented by red and lowest by yellow colour).

GUS staining showing *promPYL7::GUS* expression in the leaves of Col-0 and *prt6-1*. Expression is high in *prt6-1* compare to Col-0.

GUS staining in flower also shows high expression of *promPYL7::GUS* in *prt6-1* compare to Col-0 in stigma & sepals part.

The scale bar represents 200µm.

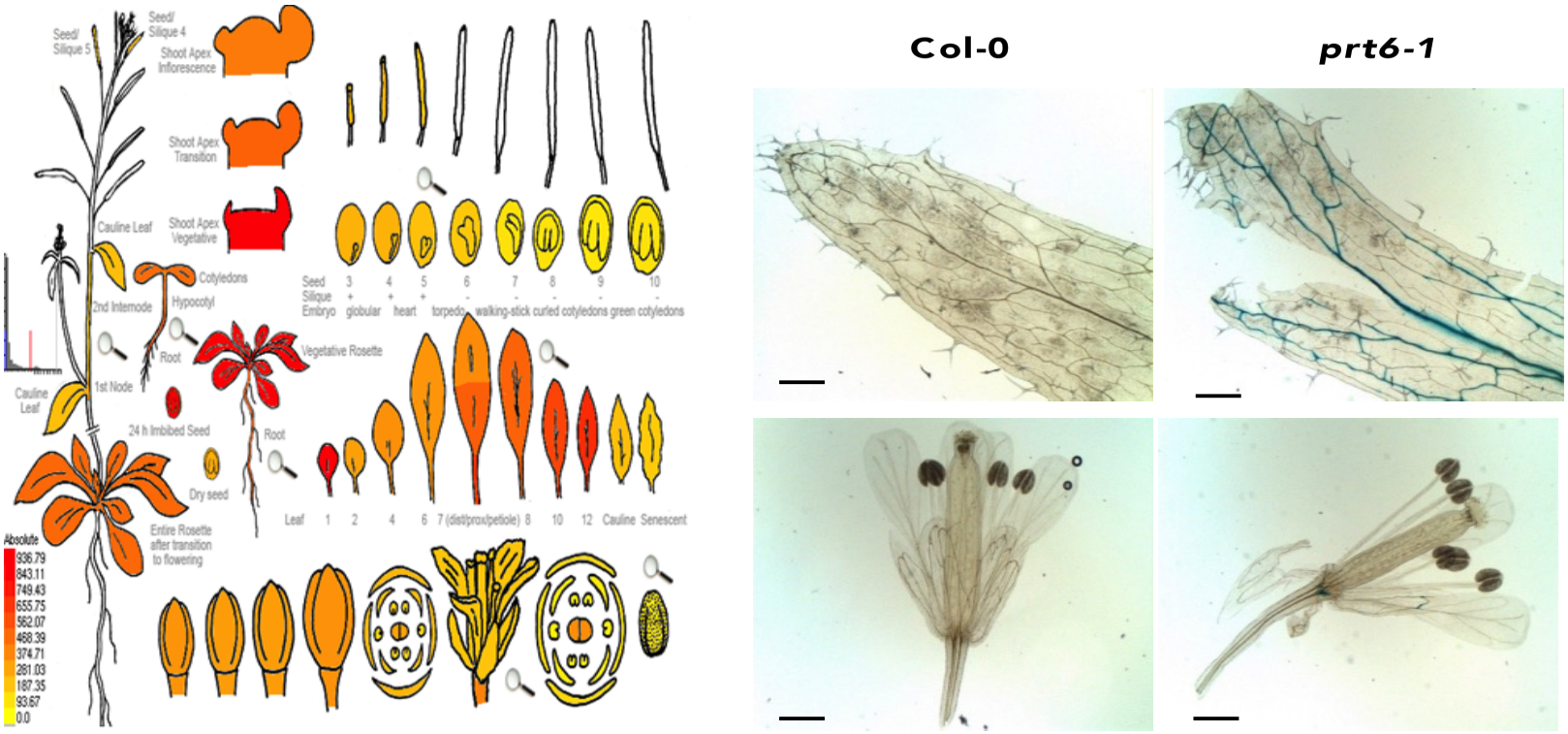


Fig: 3.4 Representative images of *At1g14810* expression in different parts of plant obtained from eFP browser and the GUS activity driven by the *promAt1g14810::GUS* in leaf & flower of an independent line of Col-0 and *prt6-1*

The eFP browser image shows expression of gene in different parts of plant. Level of expression is analysed by colour scale (highest is represented by red and lowest by yellow colour).
The leaves with prom*At1g14810::GUS* showed staining in the veins of leaves of *prt6-1* in comparison to Col-0. The flower of Col-0 is showing no GUS staining and very low in *prt6-1*. The scale bar represents 200µm

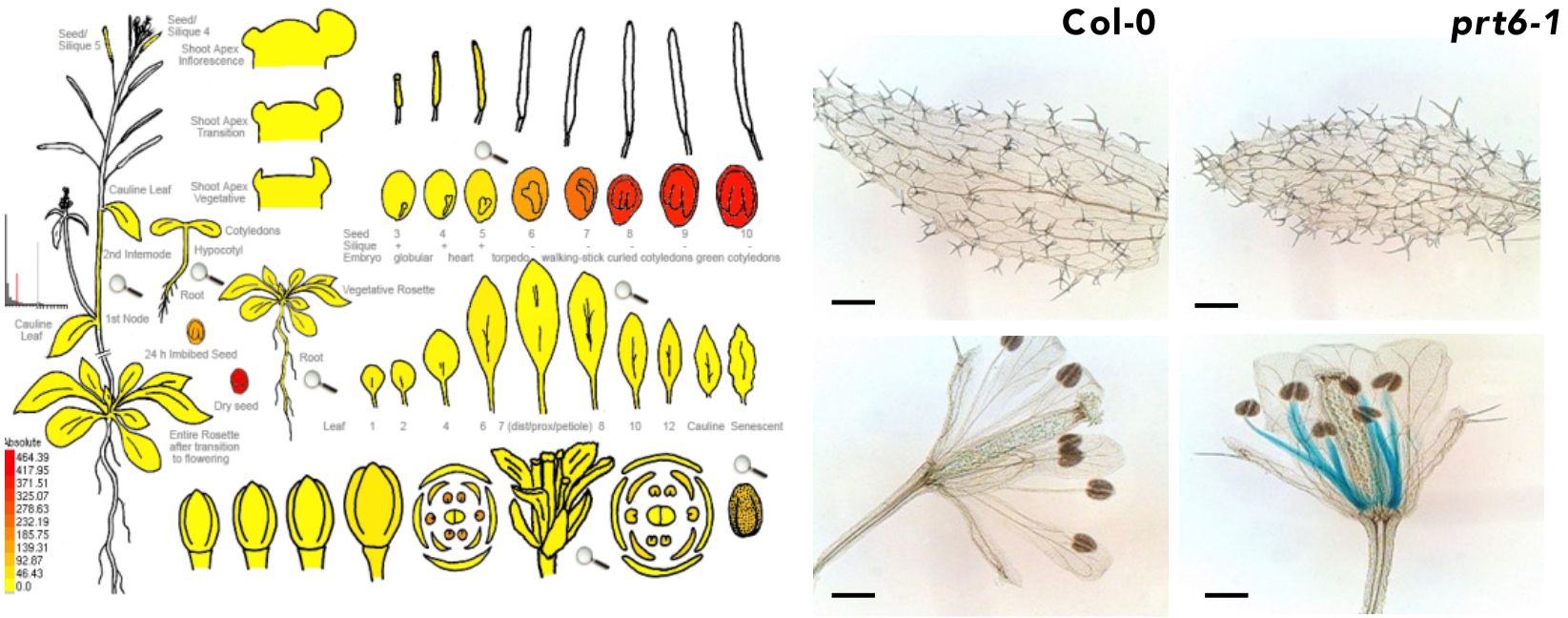


Fig: 3.5 Representative images of *ERD4* expression in different parts of plant obtained from eFP browser and the GUS activity driven by the *promERD4::GUS* in leaf & flower of an independent line of Col-0 and *prt6-1*

The eFP browser image shows expression of gene in different parts of plant. Level of expression is analysed by colour scale (highest is represented by red and lowest by yellow colour).
The *promERD4::GUS* show a higher amount of staining in the leaves of *prt*6-1 while very low in Col-0. The flower showing low GUS staining at pistil region in Col-0 but in *prt6*-1 staining is higher and can be seen at both pistil and in filaments.
The scale bar represents 200µm

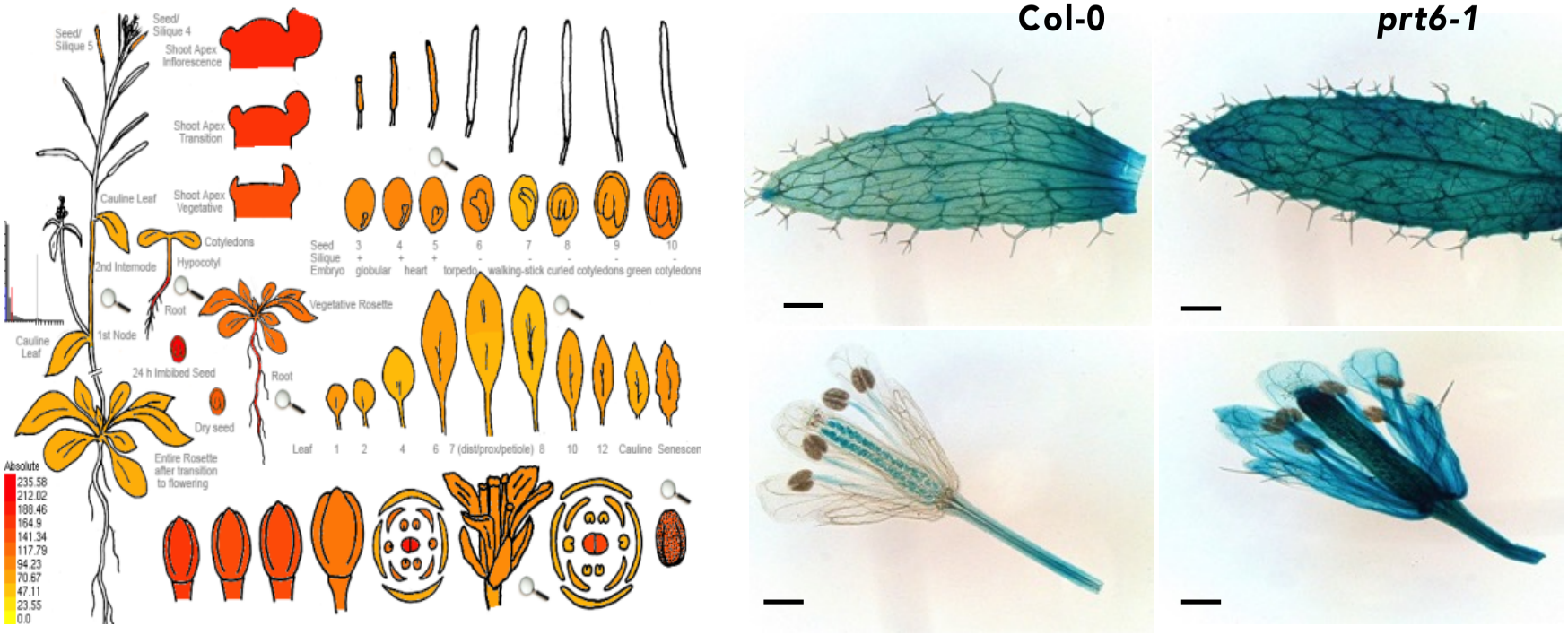


Fig: 3.6 Representative images of *At3g13440* expression in different parts of plant obtained from eFP browser *At3g13440* and the GUS activity driven by the *promAt3g13440::GUS* in leaf & flower of an independent line of Col-0 and *prt6-1*

The eFP browser image shows expression of gene in different parts of plant. Level of expression is analysed by colour scale (highest is represented by red and lowest by yellow colour).
The GUS expression is high in *prt6-1* leaves than Col-0. Similarly *prt6-1* flower is also showing high GUS expression in comparison to Col-0.
The scale bar represents 200µm.

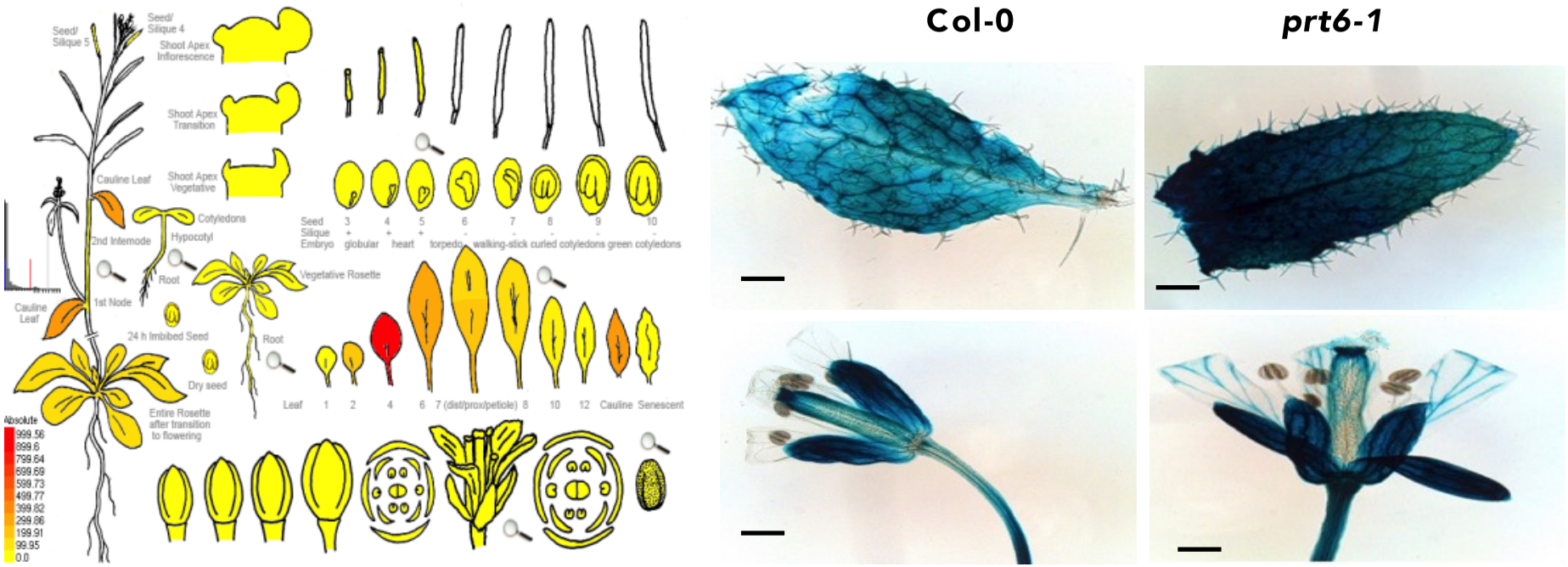


Fig: 3.7 Representative images of *PDF1.2* expression in different parts of plant obtained from eFP browser *PDF1.2* and the GUS activity drive by the *promPDF1.2::GUS* in leaf & flower of an independent line of Col-0 and *prt6-1*

The eFP browser image shows expression of gene in different parts of plant. Level of expression is analysed by colour scale (highest is represented by red and lowest by yellow colour).
The level of GUS gene expression is nearly same in the leaves of Col-0 and *prt6-1*. And also in flowers very less difference in staining between both *prt6-1* and Col-0 can be seen.
The scale bar represents 200µm

The segregating transgenic lines were screened on the basis of difference in GUS expression pattern between Col-0 and *prt6* flowers and leaf tissues tested through GUS staining. In T_3 generation three independent homozygous lines of Col-0 and two of *prt6-1* carrying *promPYL7::GUS* was obtained. Furthermore, four independent Col-0 and one *prt6-1* lines of *promAt3g13440::GUS* and two independent lines of Col-0 and two *prt6-1* lines containing *promERD4::GUS* constructs.

The GUS staining results of *promPDF1.2::GUS* showed little difference in both flowers and leaves of Col-0 and *prt6-1* (Fig: 3.7). It was concluded that the promoter is equally active in both Col-0 and *prt6-1* plants. Promoter analysis of *PDF1.2* identified one GCCGCC (EBP) *cis*-elements in the promoter region. Previous studies on *ABI5* gene regulation have shown that both *cis*-elements are required for binding of ERFVII's. Therefore, in absence of any one of GCC-box leads to no ERFVII's binding hence no gene regulation. So, it might be possible that there are other transcription factors that regulate the gene independent of N-end rule pathway. Therefore, based on the results observed this promoter was not considered ideal to carry further work on it.

3.2.2.4 Analysis of promoter activity of *prt6-1* promoter::GUS lines crossed to Col-0

For further analysis of promoter activity of transformed promoter constructs, the *promoter::GUS* fusion lines of *promERD4::GUS* and *promAt3g13440::GUS* in *prt6-1* background that showed enhanced promoter activity were back-crossed to Col-0. This strategy was followed to allow analysis of the same transgene insertion in both genotypes.

In this experiment seven days old F_3 Col-0 homozygous seedlings with *promERD4::GUS* and *promAt3g13440::GUS* were taken. T_3 homozygous Col-0 and *prt6-1* carrying *promERD4::GUS* and *promAt3g13440::GUS* were taken as controls for comparison. Seedlings were processed for GUS staining. The GUS analysis of seedlings containing *promERD4::GUS* in Col-0 reveals that GUS gene has expressed very less whereas, in *prt6-1* seedlings high GUS expression was seen. Seedlings of the same *prt6-1* line crossed to Col-0 also shows similar GUS expression pattern as was seen in Col-0 (Fig: 3.8).

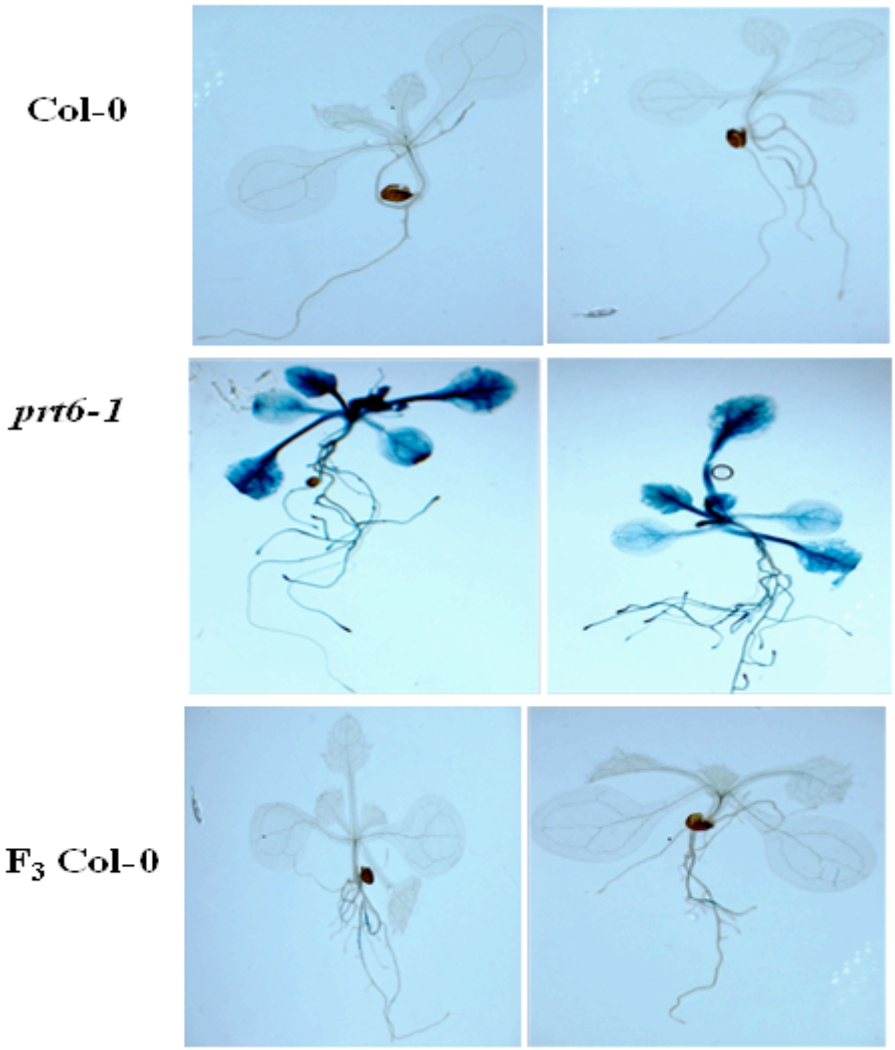


Fig: 3.8 Comparison of promoter activity of *ERD4* gene in Col-0, *prt6-1* and F3 Col-0 seedlings

Promoter activity of *promERD4::GUS* was analysed in Col-0, *prt6-1* and F_3 Col-0. In Col-0 the GUS staining is very low in whole seedling and *prt6-1* seedlings show relatively high GUS staining in all parts of seedling. The back crossed seedlings also doesn't show any staining in cotyledons but in a very small area of roots.

Later, flowers and leaves of lines having *promERD4::GUS* and *promAt3g13440::GUS* in Col-0 & *prt6-1* and F_3 Col-0 background were also processed for GUS staining to test the promoter activity under different backgrounds. The GUS staining result of flowers and leaves of Col-0 having *promERD4::GUS* demonstrates that GUS activity is very low in comparison to *prt6-1* and the flowers and leaves of back crossed *prt6-1* to Col-0 also shows similar kind of GUS activity as was seen in Col-0 (Fig: 3.9).

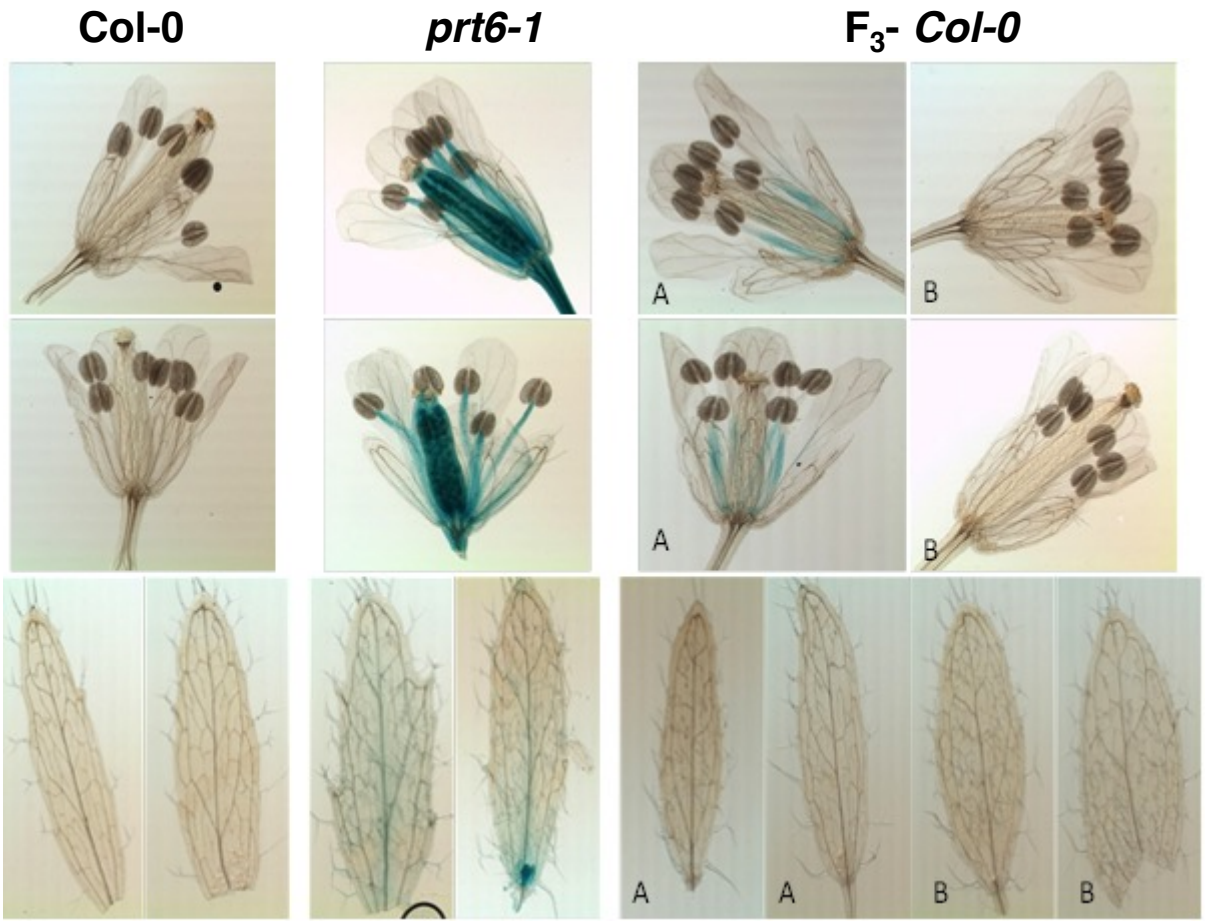


Fig:3.9 Representative images of flowers and leaves showing comparison of promoter activity of *ERD4* gene in Col-0, *prt6-1* and F_3 Col-0

The images of Col-0 flowers and leaves are showing no GUS staining whereas in *prt6-1* flower the amount of staining very high and is seen in whole part of pistil and stamen and in leaves at veins and also in some leaves at the edge near petiole region. On the other hand, in the flowers and leaves of F_3 Col-0 no staining was seen.

The GUS analysis of seedlings containing *promAt3g13440::GUS* in *prt6-1* shows high GUS expression comparison to *prt6-1*. However, in *prt6-1* line crossed to Col-0 less GUS expression appears. The Col-0 also shows much less GUS activity (Fig: 3.10).

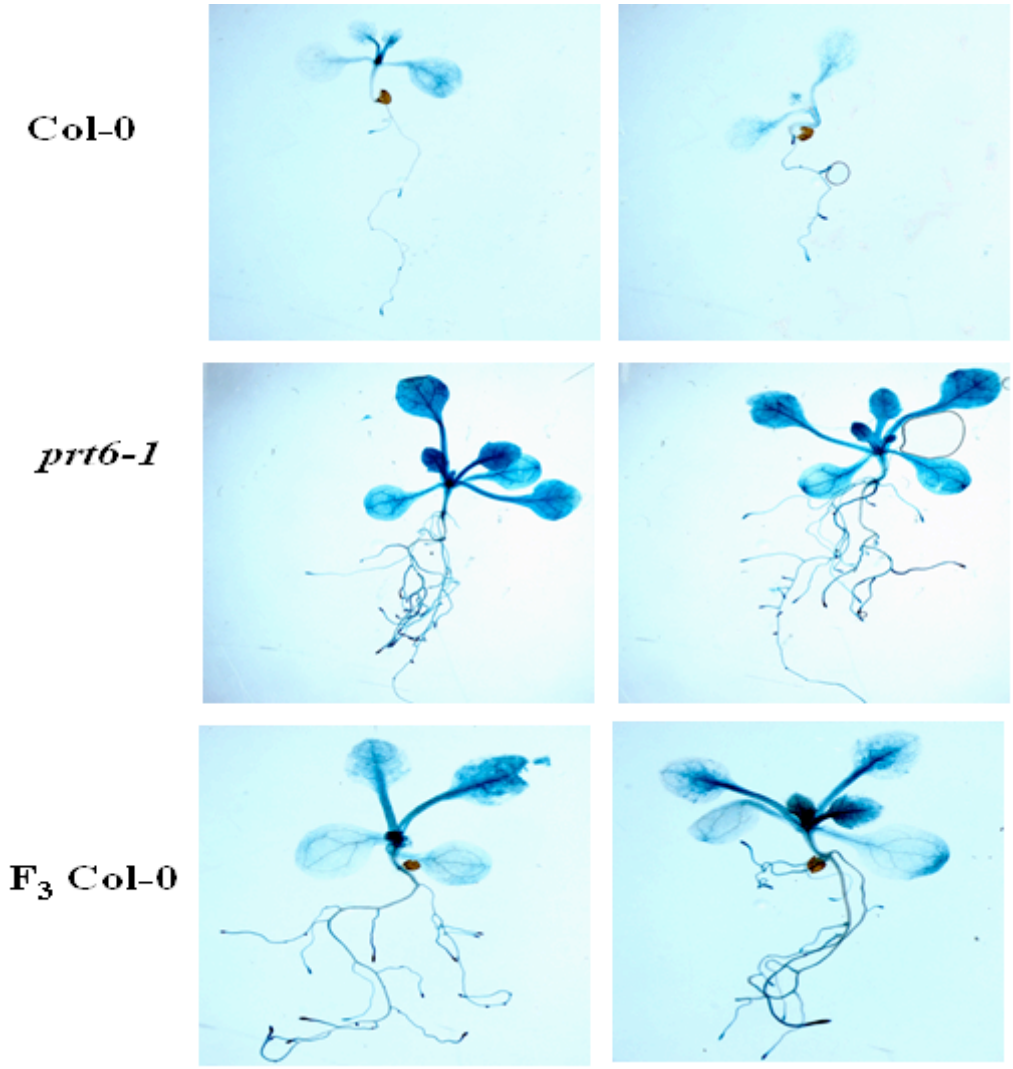


Fig:3.10 Comparison of promoter activity of AT3G13440 gene in Col-0, *prt6-1* and F_3 Col-0 seedlings

Promoter activity of *promAt3g13440::GUS* was analysed in Col-0, *prt6-1* and F_3 Col-0. A very low GUS staining is seen in Col-0 seedlings then *prt6-1*. However, in back crossed seedlings staining are again less compare to *prt6-1* and near to Col-0.

Likewise, the GUS activity in the flowers and leaves of Col-0 having *prom13440::GUS* was less comparison to *prt6-1* as well as the flowers and leaves of F_3 Col-0 also shows less GUS activity in comparison to *prt6-1* (Fig: 3.11).

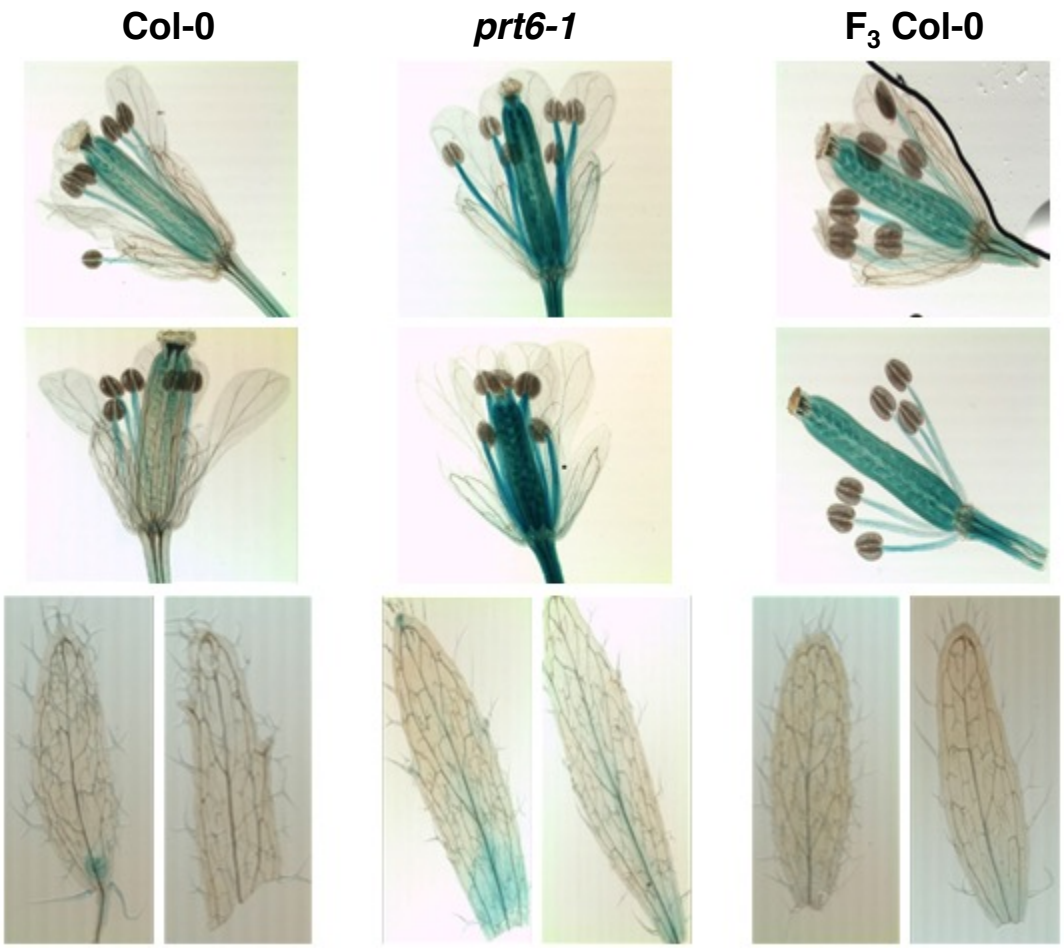


Fig: 3.11 Representative images of flowers and leaves showing comparison of promoter activity of *AT3G13440*gene in Col-0, *prt6-1* and F_3 Col-0

In the flowers images GUS staining can be seen in the stamens and pistil region in both Col-0 and *prt6-1* as well as also in F_3 Col-0 but the amount of staining is quite high in *prt6-1* compare to Col-0 and F_3 Col-0. And in leaves also the staining is more in *prt6-1* than Col-0 and F_3 Col-0.

3.2.2.5 Expression analysis of genes containing 'GCCGCC' *cis*-elements in their promoters by q-RT-PCR

To further confirm the expression of genes at transcriptional level, q-RT-PCR was conducted and the relative expression of selected genes (*PYL7, ERD4, AT1G14810* and *AT3G13440*) whose promoters (that carry *cis*-elements) were transformed in Col-0 and *prt6-1* was measured.

This experiment was done on cDNA samples derived from Col-0 and *prt6-1* untreated leaves and flowers. For this purpose first RNA was extracted from three biological replicate of Col-0 and *prt6-1* leaves and flowers using RNA extraction kit (Qiagen). Reverse transcription was carried out on RNA (500ng) using Superscript III reverse transcriptase (Invitrogen). PCR reaction was setup using sensiFAST SYBR*Hi-Rox kit (Bio-line) on three technical repeat of each biological sample in 384-well optical reaction plates. A housekeeping gene (*Arabidopsis* control 3: a housekeeping gene) *AtCTL3* (De Rybel *et al.*, 2010) was used as an internal reference. Data obtained from q-RT-PCR was prepared in an excel sheet for analysis. Relative quantification of expression level of each gene was calculated first by normalizing against *AtCTL3* internal control gene (section 3.1.2.2) and later calculated using (2^-Cp) (Delta-delta method for comparing relative expression.

The data obtained from q-RT-PCR performed on leaves and flowers of Col-0 and *prt6-1* were analysed statistically through t-test to calculate the significant value ($p<0.05$).

The analysis done on leaves of Col-0 and *prt6-1* for *PYL7* shows the p-value=0.025776 therefore, the statistical difference between Col-0 and *prt6-1* is significant ($p<0.05$) (Fig 3.12). However, no statistical difference is seen for ERD4, AT3G13440 and AT1G14810.

On the other hand, the analysis done on flower of Col-0 and *prt6-1* for *PYL7*, ERD4, AT3G13440 and AT1G14810 does not show any significant difference between Col-0 and *prt6-1* (Fig 3.12).

Therefore, the difference in level of expression can only be seen between Col-0 and *prt6-1* leaves in case of PYL7 whereas, in case of ERD4, AT3G13440 and AT1G14810 there is no significant difference between Col-0 and *prt6-1* either in leaves or in flowers.

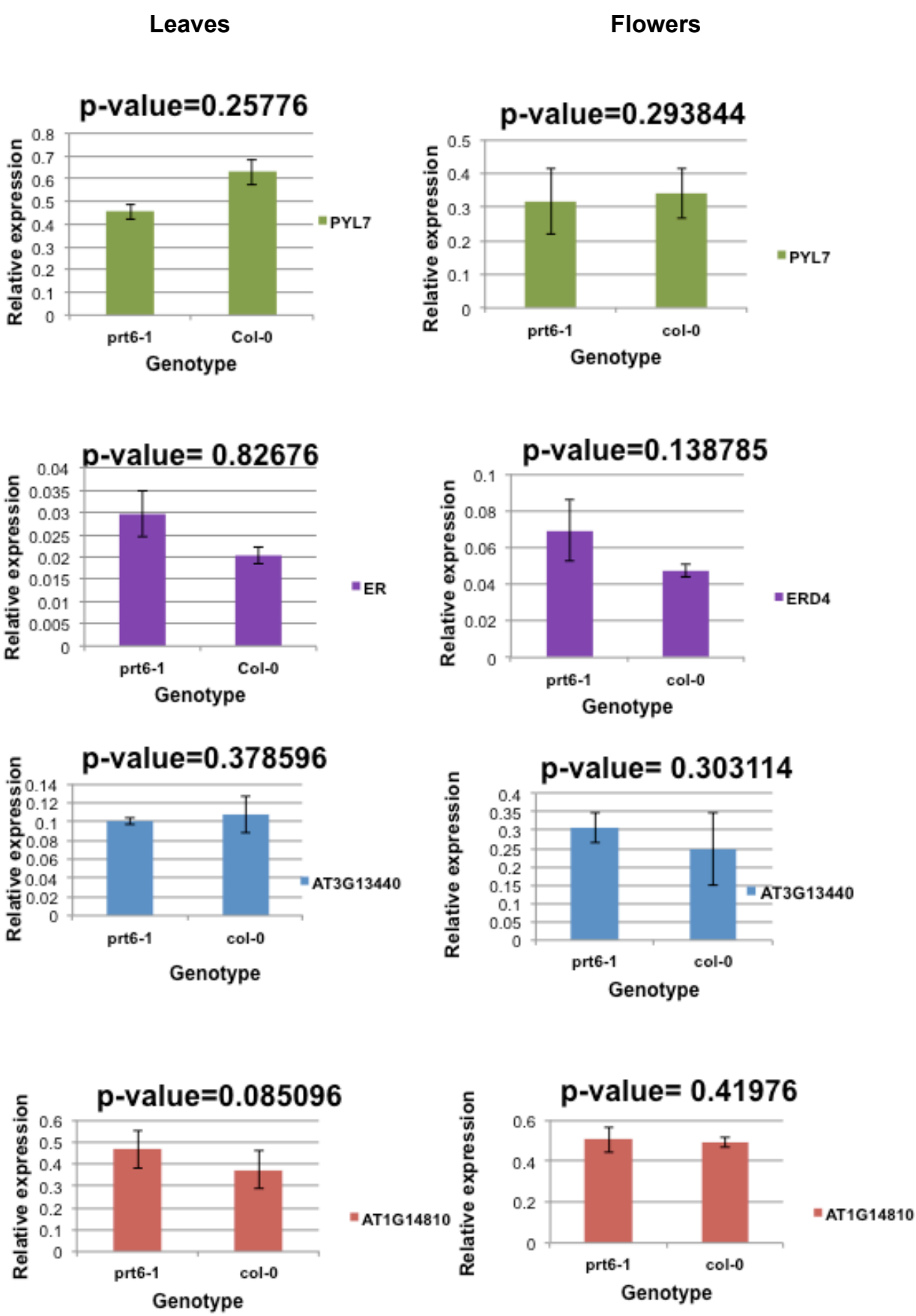


Fig 3.12 Quantitative RT-PCR data analysis of *PYL7, ERD4, AT3G13440* and *AT1G14810* gene.

Bar graphs representing the relative gene expression levels of transcripts *PYL7, ERD4, AT3G13440* and *AT1G14810* in Col-0 and *prt6-1* cDNA samples of leaves and flowers. Bars represent the average value of three technical repeats of triplicate biological samples of each genotype, error bars represent their SE. P-value is the representation of significant values ($p<0.05$). Higher value (p value>0.05) indicates no difference between the two individual samples and ($p<0.05$) indicates difference between individuals.

1.) *PYL7*: q-RT-PCR analysis for *PYL7* gene expression on Col-0 and *prt6-1*cDNA samples of leaves and flowers. Significant value is less than 0.05 i.e., p-value=0.025776 in leaves but p-value is more than 0.05 *i.e*., p-value=0.293844 in flower.

2.) *ERD4*: q-RT-PCR analysis for *ERD4* gene expression on Col-0 and *prt6-1* cDNA samples of leaves and flowers. The significant values for Col-0 and *prt6-1* leaves *i.e*., p-value=0.82676 and for flower *i.e*., p-value=0.138785.

3.) *AT3G13440*: q-RT-PCR analysis for *AT3G13440* gene expression on Col-0 and *prt6-1* cDNA samples of leaves and flowers. The significant values for Col-0 and *prt6-1* leaves *i.e*., p-value=0.378596 and for flower *i.e*., p-value=0.303114.

4.) *AT1G14810*: q-RT-PCR analysis for *AT1G14810* gene expression on Col-0 and *prt6-1* cDNA samples of leaves and flowers. The significant values for Col-0 and *prt6-1* leaves *i.e*., p-value=0.085096 and for flower *i.e*., p-value=0.41976.

3.3 N-end rule pathway mediated salt stress tolerance

Salt stress is one ofthe major abiotic factors that inhibit the seed germination and plant growth leading to low crop yield (Munns *et al*., 2002). High salinity affects plants in several ways such as: water stress, nutritional disorders, oxidative stress, alteration of metabolic processes, reduction of cell-division (Hasegawa *et al*., 2000; Munns *et al*., 2002; Zhu *et al*., 2007). All together these cause reduced plant growth and survival.

Recent studies have shown the role of N-end rule pathway in coping with different stresses through its substrate stability. In *Arabidopsis* substrates of the N-end rule

pathway are recognized by the E3-ligase PRT6 that transfers them to 26S proteasome for degradation. However in *prt6* mutant substrates remain stable. Previous studies by Holman *et a.l* in 2009 have shown the importance of PRT6 in the removal of ABA sensitivity during seed germination. Thus it is expected that plants lacking the substrates would be less sensitive to ABA.

The information provided by Holman *et al*., indicates that PRT6 and N-end rule pathway play a major role in seed ABA signalling during the seed to seedling transition. Therefore, this evidence suggests that N-end rule pathway might also have role in salt stress response. It is hypothesized that stability of ERFVII's in *prt6-1* mutants induces plants survival in high salinity condition.

3.3.1 Assessment of effect of salinity on root length of seedlings in wild-type and *prt6-1* genetic backgrounds tested on different salt concentrations

To investigate a possible role for the N-end rule pathway in regulating salt stress tolerance, a salt signalling response was studied. For this purpose a direct measurement of root length of T-DNA insertional mutant lines of *prt6-1*, *prt6-1pyl7, prt6-1abi5-8, prt6-1abi5-8pyl7, prt6-1erd4, erd4, pyl7, abi5-8* and wild-type background Col-0, were taken. Root length measurement was done on 5 different concentrations of NaCl (0mM, 75mM, 100mM, 125mM and 150mM). Initially, all nine genotype seeds were first plated on half-MS plates and kept on 4^0C for chilling. After chilling plates for three days they were kept in light for further 3 days. On fourth day, seedlings (20 seedlings of each genotype) were plated on half-MS plates containing 0mM, 75mM, 100mM, 125mM and 150mM salt with the help of sterilized forceps. All plates were kept again in light in vertical position. On seventh day, images of seedlings on each plate were taken. Measurement of root length of each line was done through ImageJ software. Measurement data was prepared on excel sheet. Average of roots of 20 seedlings of each line was taken followed by calculation of standard error and a graph was plotted. Seedlings that were on 150mM salt containing plates were transferred to 0mM salt containing half-MS plates. This procedure was followed to see the recovery ability of plants in normal condition after having high salt stress.

The root measurement data was normalized by taking percentage of whole data of 0mM salt of nine lines as 100%. And value for rest of the concentrations was calculated in comparison to 0mM 100% value (Fig: 3.13).

The t-test analysis was done for 100mM, 125mM and 150mM salt. The highest number of lines showed significant difference at 100mM salt. *prt6-1, erd4, abi5-8prt6-1* and *pyl7prt6-1* when compared with Col-0 showed significant difference (p-value<0.05) (Fig:3.13).

At 125mM salt significant difference was seen in *prt6-1, erd4prt6-1* and *erd4* in comparison with Col-0 (Fig: 3.13). The significant difference at 150 mM salt was seen in *abi5-8prt6-1, abi5* and *erd4*. Seedlings at 150mm salt go through highest level of salt stress. Therefore, it is expected that all seedlings will show less growth at 150mM of salt. Hence, less significant difference appears between lines.

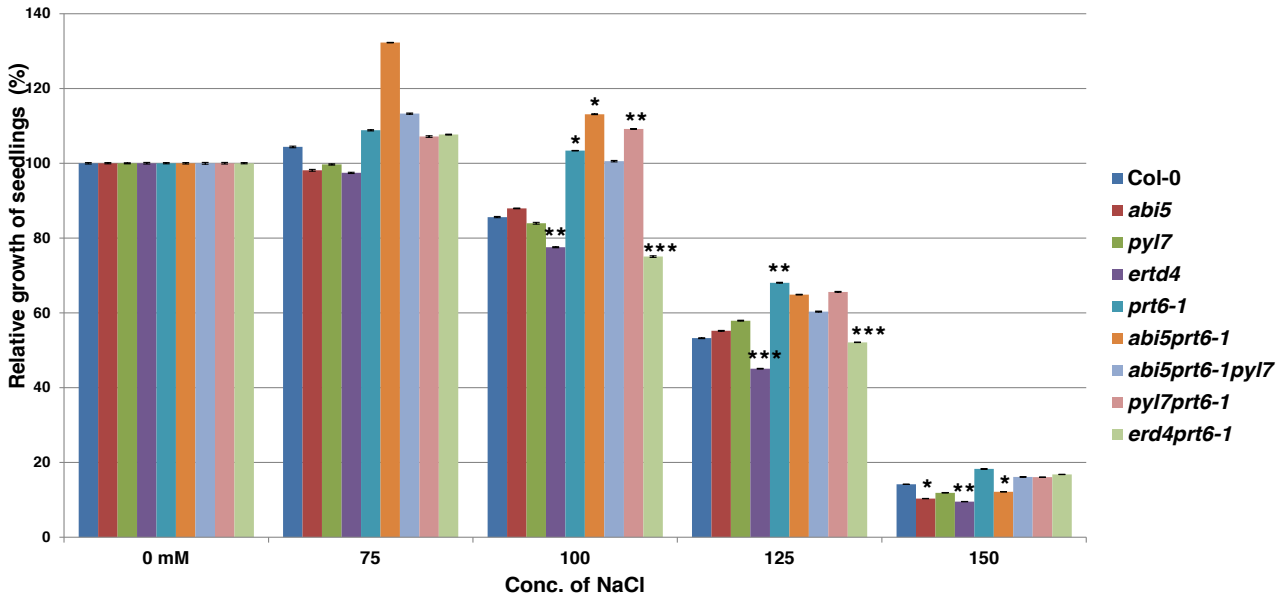


Fig: 3.13 Representation of relative growth of seedlings analysed on the basis of root length of *prt6-1* and its combination mutant lines and wild-types on different salt concentrations

Root measurement data on 0mM, 75mM, 100mM, 125mM and 150mM are represented as percentage of whole root length of a line. Error bars correspond to the SD of twenty independent seedlings. Number of stars (*) represent p-values: single star= p-value<0.05, double star= p-value<0.001 and triple star= p-value<0.005.

Seedlings that were transferred to 0mM salt plates from 150mM salt were again examined on fifth day and images were taken. From the images it could be seen that *prt6-1* mutant seedlings were better recovered than Col-0. Now measurement of weight of all 20 seedlings of each line was done and a graph was plotted on that basis. Results of this analysis demonstrate that *prt6-1* and its mutant combinations showed overall higher level of growth in comparison to wildtype genotypes (Fig: 3.14).

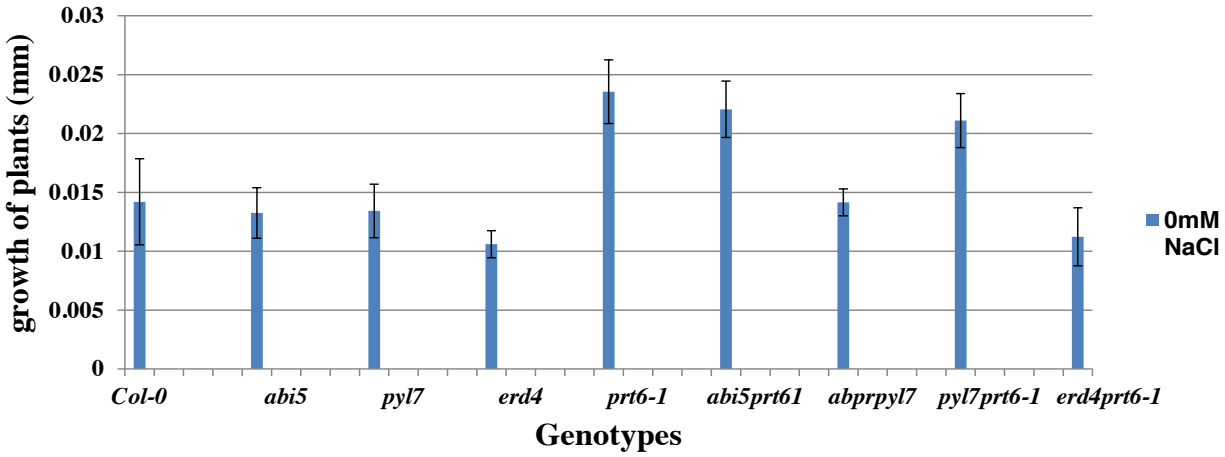


Fig: 3.14 Representation of relative growth of seedlings transferred from 150mM salt 0mM salt measured on the basis of average weight of each line.

Graph represents the average weight of 20 seedlings of each line recovered from high salt treatment to normal. Seedlings in *prt6-1* background appear showing better recovery than wild-types.

3.4 DISCUSSION

3.4.1 Identification of downstream targets of the ERFVIIs

The aim of this research work was to analyse the relationship between the ERFVII's, and genes containing multiple copies of the EBP (GCCGCC) *cis*-element in their promoter region. This was carried out in order to understand how the N-end rule pathway regulates plant growth and development through ERFVII's. The work programme involved discovery and analysis of reporter expression driven by promoters containing multiple EBP *cis*-elements in wild-type (Col-0) and *prt6-1* plants.

3.4.1.1 Determining the presence of GCC-boxes, *cis*-elements within promoters

Initially, a bioinformatics approach was used to identify the *cis*-regulatory element that is bound by the ERFVII's. This element was designated as Ethylene Responsive Element binding factor domain (EBP) and consists of a consensus sequence motif GCCGCC (Yang *et al*., 2009). Fourteen genes were selected from TAIR using pat-match tool and considered to be putative downstream targets of the ERFVII's (Table: 3.1). Bioinformatics analyses of all fourteen genes showed the presence of GCCGCC consensus sequence in the 5'UTR or in promoter region.

Previous study on *ABI5* promoter having GCC-boxes has shown that expression of *ABI5* was specifically enhanced at the micropylar endosperm region in the *prt6* mutant seed (Fig: 1.11). This suggests that the stability of the N-end rule substrates might have a role in regulating the activity of the promoter through the GCC-boxes. Therefore, the identified promoters in this project could be the putative downstream targets for the ERFVII's MC-ERFs *via* GCC-boxes, such as in *ABI5.*

3.4.1.2 Investigation of *promoter::GUS* expression in Col-0 and *prt6-1* genetic backgrounds

The activity of each promoter was analysed by GUS staining. It was seen that the promoters were more active in the tissue of transgenic *prt6-1* than Col-0 plants except for *promPDF1.2::GUS* (Section 3.1.2.3). Further, an analysis of the pattern of *promoter::GUS* expression was compared to the *eFP* browser webpage (Fig: 3.4, 3.5, 3.6, 3.7, and 3.8). The results demonstrate that each promoter was active in different region of the same tissue. Such as, in the *prt6-1* mutant flower carrying *promPYL7::GUS* showed activity in stigma and sepals but the activity of *promERD4::GUS* was in pistil and filament. Likewise, activity of *promAt1g14810::GUS* was restricted to the basal part of stigma while *promAT3G13440::GUS* showed promoter activity in whole flower. Enhanced *promoter::GUS* expression in flowers of *prt6-1* compared to wild type indicates that the N-end rule pathway play a role in flower development.

The question arises, how can the promoters be expressed in different regions if they are all regulated by ERFVII's?

As it is evident that many genes involves combined effort of other transcription factors and might also involves other factors like, enhancers, co-factors for their expression at different loci, at certain developmental stage or in response of external or internal signals (Lee *et al*., 2002; Milo *et al*., 2002; Shen-Orr *et al*., 2002). Also there is possibility that other transcription factors activity depends on binding of a particular transcription factor.

Therefore, genes that are regulated by binding of ERFVII's at two 'GCCGCC' sites also involves binding of other transcription factors or any internal signal that are responsible of expression of that gene in a particular region. This, combined action of ERFVII's binding and other gene regulation related factors lead to tissue specific differential expression of that gene.

In this study *promoter::GUS* expression was seen maintained and continuously being expressed in *prt6-1* tissue in contrast to Col-0. Previous studies on *ABI5* promoter with two EBP binding sites were shown bound by N-end rule pathway substrates (Gibbs *et al*., 2011). Therefore, the binding of the N-end rule substrates to the EBP binding element regulate *ABI5* expression. Similarly in this study, enhanced expression of *PYL7, ERD4, AT3G13440* and *AT1G14810 promoters::GUS* constructs in the *prt6-1* mutant was observed that indicates that these genes are also regulated by N-end rule pathway through binding of ERFVII's.

Alternatively, the backcross approach taken in case of *ERD4* and *AT3G13440* also suggests that absence of N-end rule pathway component *prt6-1* leads to enhanced promoter activity however same transgene in Col-0 performs *vice-versa*. Thus, this is another step in confirming the hypothesis, that the ERFVII's transcription factors regulate the expression of GCC-boxes containing genes and therefore N-end rule pathway.

3.4.1.3 Determination of gene regulation at transcriptional level through q-RT-PCR

Initially the promoter::GUS approach was adapted to test whether genes that contain two GCCGCC sites are regulated through N-end rule pathway or not. Results of this experiment demonstrated that lack of PRT6, an N-end rule pathway component resulted in stabilization of ERFVII's when studied through GUS staining. This illustrates that selected genes are regulated by binding of ERFVII's followed by N-end rule pathway. This

was one of the steps to see the promoter activity of genes. In order to further confirm that selected genes are expressed at transcriptional level, q-RT-PCR analysis was used.

Four genes *PYL7, ERD4, AT3G13440* and *AT1G14810* that were shown to be regulated in *prt6-1* compared to Col-0 through promoter::GUS analysis were investigated by q-RT-PCR for relative expression on cDNA of untreated Col-0 and *prt6-1* leaves and flowers tissue. The results of this experiment showed that the significant difference can only be seen in leaves of Col-0 and *prt6-1* for *PYL7*. And there is no significant difference between flowers of Col-0 and *prt6-1* for *PYL7* and leaves and flowers of *AT3G13440* and *AT1G14810*.

3.4.1.4 Relation between N-end rule pathway and genes carrying GCCGCC *cis*-elements in promoter or 5'UTR region

The promoter::GUS analysis suggests that identified genes are regulated by N-end rule pathway. Among these *PYL7* and *ERD4* are considered as abiotic stress-associated genes. Studies by Holman *et al.,* 2009 showed the link between N-end rule pathway and ABA signalling during seed germination. As it is known that two GCC-boxes are present in the 5'UTR region of *ABI5* gene and this gene is responsible for encoding a transcription factor that promotes enhanced sensitivity of seeds for ABA. The *ABI5* gene is been regulated by ERFVII's binding to GCCGCC site. ERFVII's that are the substrates of N-end rule pathway begins with M-C at their N-terminal (Neill *et al*., 2003). Like *ABI5*, the two GCC-boxes are also present in the 5'UTR or promoter of other genes that have been identified as regulated by ERFVII's such as *PYL7, ERD4*.

PYL7 has been identified to encode pyrabactin resistance (regulatory component of ABA receptor) family proteins and has been shown to interact with ABI1, ABI2. The PYL-ABA-PP2C interaction blocks PP2C activity and allows downstream targets, such as the Suc nonfermenting-related kinases group 2 (SnRK2s) and other PP2C substrates, to remain phosphorylated (Ma *et al.,* 2009; Park *et al.,* 2009; Umezawa *et al.,* 2009; Vlad *et al.,* 2009). This, in turn, leads to the phosphorylation of ABF/AREB/ABI5 bZIP family transcription factors and the activation of ABA-induced gene expression (Fujii *et al.,* 2009) or the phosphorylation of Slow Anion Channel1 in guard cells (Geiger *et al.,* 2009, 2010; Lee *et al*., 2009). The *ERD4* protein is localized in chloroplast membrane. The function of

highly conserved *ERD4* protein may thus be associated with its RNA-binding ability during the stress response (Rai *et al.,* 2012).

Therefore, it can be considered N-end rule pathway have key role in functioning of identified genes during ABA signalling through binding of ERFVII's to GCCGCC site.

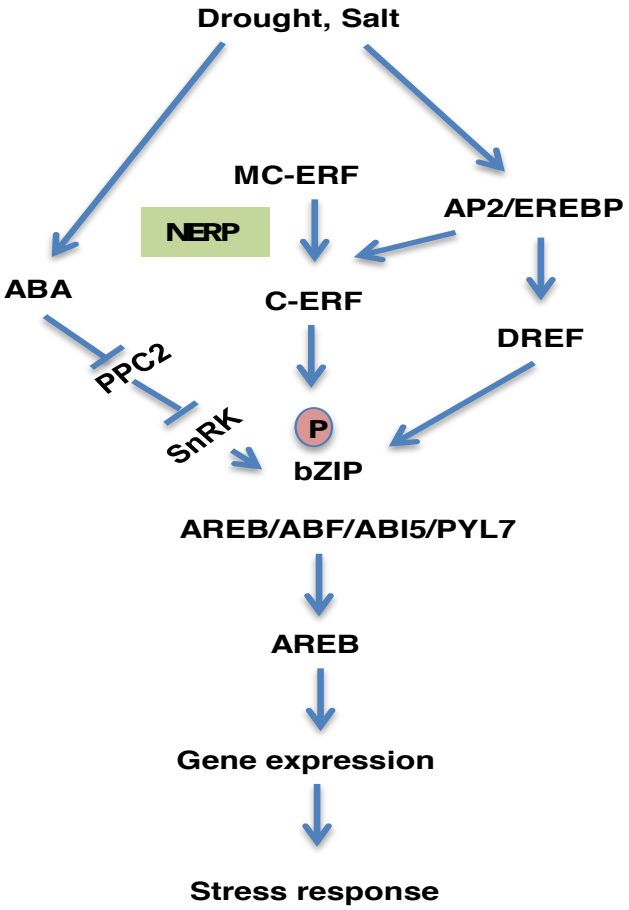


Fig:3.15 Schematic diagram of ABA signalling pathway in relation with N-end rule pathway.

ABA signalling regulated through AREB/ABF/ABI5/PYL7 transcription factor. These transcription factors contain 'GCCGCC' sites for binding of ERFVII's hence involves N-end rule pathway for gene regulation and stress responses *via* stability of ERFVII's that begins with M-C and follow Arg/N-end rule pathway for degradation. Figure is taken from Lee *et al.,* 2011.

3.4.2 The N-end rule pathway of protein degradation regulates salt stress tolerance

The result presented in section 3.2.3 showed that N-recognin PRT6 lacking seedlings have salinity hypersensitivity. However, when the same seedlings were grown in normal conditions on 0mM salt they showed better recovery and increased weight mass in comparison to wild-type. This suggests that substrates stability in *prt6-1* mutant during stress condition activates other genes that help plants to persist. Therefore, later in normal condition *prt6-1* mutants can recover better. This shows the involvement of the N-end rule pathway in plants survival in response to salt stress.

3.4.3 Future prospect of study of regulation of genes having 'GCCGCC' site using *erfVIIprt6-1* mutant combination

The present study was aimed for the identification of those genes that are regulated by binding of ERFVII's to two GCC-boxes present in their promoter. In relation with this another experiment can be performed that involves study of promoter activity of identified genes that have two GCCGCC sites through *erfVIIprt6-1* mutant combinations. This study will help to find two objectives:

First, to investigate the promoter activity of above five identified genes in different combinations of the five ERF mutants in *prt6-1* background. As it is hypothesized that in *erfVIIprt6-1* mutant, there would be no promoter activity due to lack of the N-end rule pathway substrates hence there will be no promoter activity. Also this will further confirm that these genes are regulated *via* ERFVII's (Fig: 3.16 A).

Second, there is possibility that there are other transcription factors that operate those genes regulation by binding to GCCGCC site. Such as newly identified N-end rule pathway substrates (BBX and bHLH). Or there might N-end rule pathway independent regulation of that gene by binding of unknown transcription factors (Fig: 3.16 B).

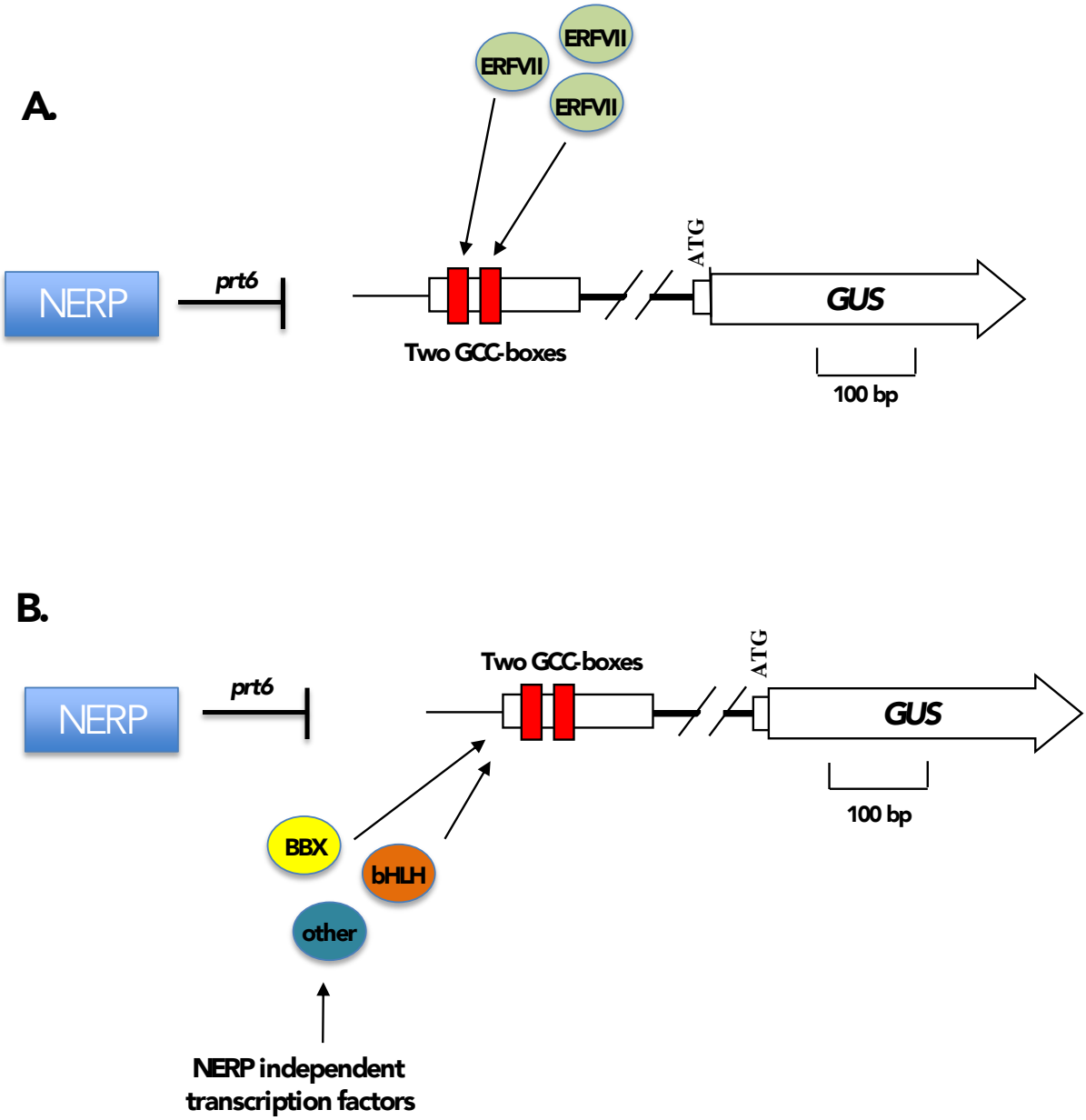


Fig: 3.16 Hypothetical model showing regulation of genes having two GCC-boxes in two different conditions

(A) Regulation of genes by binding of ERFVII's on two GCC-boxes present in promoter or 5'UTR region of genes directed by N-end rule pathway.

(B) Regulation of genes by binding of recently identified substrates of N-end rule pathway i.e., BBX or bHLH and follows N-end rule pathway. Or can be regulated in N-end rule pathway independent manner.

GGCGGC site: binding element, NERP: N-end rule pathway, Scale:100 bp

In conclusion, promoter analysis studies revealed that Ethylene Binding Protein (EBP) *cis*-element have binding specificity for ERFVII's and are operated by N end rule pathway. Further, examination of *prt6-1* in comparison to Col-0 genome suggests role of N-end rule pathway in plant survival in response to high salinity.

Chapter-4

Investigation of *in-vivo* binding of RAP2.3 protein to promoter of genes containing the 'GCCGCC' *cis*-element through Chromatin immunoprecipitation (ChIP)

4.1 Introduction

4.1.1 Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation is a technique that facilitates the identification of the *in-vivo* interaction of a protein to a specific segment of DNA at the chromatin level (Aparicio *et al*., 2004). This method was first developed by Alexander Varshavsky whilst studying histone distribution in Drosophila (Solomon and Varshavsky, 1985).

In this technique, cross-linking of proteins to DNA strands is facilitated by formaldehyde. The cross-linked chromatin is fragmented by sonication into small segments. The DNA-protein complex is subsequently immunoprecipitated with an antibody against the cross-linked protein to allow selective enrichment of the immunoprecipitated sample. The captured DNA-protein complex is reversed and after reversal of cross-links, the recovered DNA can be tested *via* amplification through quantitative PCR using primers specific to the binding site. The amount of PCR product is indicative of the relative abundance of the DNA-protein complex *in vivo*. Further, ChIP DNA can be processed for microarray analysis to understand DNA-binding proteins at the genome level (Aparicio *et al*., 2004). Also, analysis of ChIP DNA can be done *via* sequencing to verify presence of protein binding site on DNA. Therefore, the ChIP assay is considered a good tool to examine direct binding of *cis*-acting targets by specific transcription factors *in-vivo*.

4.1.2 Study of ERFVII's binding to DNA through ChIP

It has been seen that RAP2.3, an ERFVII binds *in-vitro* and *in-vivo* to the GCC-boxes (Buttner and Singh, 1997; Yang *et al*., 2009; Gibbs *et al*., 2014) and subsequently activates gene transcription for example: *OBF4, ABI5* in *Arabidopsis*. So, the ChIP approach can be used to determine *in-vivo* binding of RAP2.3 to the promoter regions of those genes that have the *cis* element 'GCCGCC', such as *ABI5, PYL7, ERD4*. Subsequently, the quantitative evaluation of the amount of DNA associated with an immunoprecipitated protein can be done by qPCR (Dahl *et al*., 2007; Das *et al.,* 2004).

4.2 Results

4.2.1 Confirmation of MA-RAP2.3 and MC-RAP2.3 protein stability *in-vivo* through Western blotting

To analyze the binding of RAP2.3 to 'GCCGCC' *cis*-element, homozygous lines of *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* were used. For ChIP experiment I used *35S::MC/MA-RAP2.3-HA* seedlings in normoxia, *35S::MC-RAP2.3-HA* hypoxia (3 hrs nitrogen treated) seedlings and submerged (5 hrs in water in dark) seedlings and etiolated *35S::MA-RAP2.3-HA* seedlings.

Before processing to ChIP experiment first confirmation of the presence of RAP2.3 protein in plants through western blotting (Section: 2.2) was conducted. For this purpose, first extraction of protein from 4 days old light grown seedlings (normoxia, hypoxia or submerged treated) and processed them for western blotting.

100mg of protein extract was separated by 12% (v/v) polyacrylamide gel electrophoresis. Later proteins were transferred to a nitrocellulose membrane. The membrane was treated with the first antibody; mouse α-HA (Sigma-Aldrich) 1:2500 and kept overnight at 4^0C. This process leads to binding of primary antibody to RAP2.3 HA tagged protein on the membrane. Thereafter, the second antibody Goat α mouse (Sigma-Aldrich) was added in proportion of 1:10,000. The secondary antibody binds to the primary antibody that has bound with RAP2.3-HA protein. The membrane was then treated with Pierce EVL western blotting substrate (Thermo Scientific). The substrate reacts with the secondary antibody to generate a coloured substance. The size of identified protein bands was determined by comparison with a molecular weight marker.

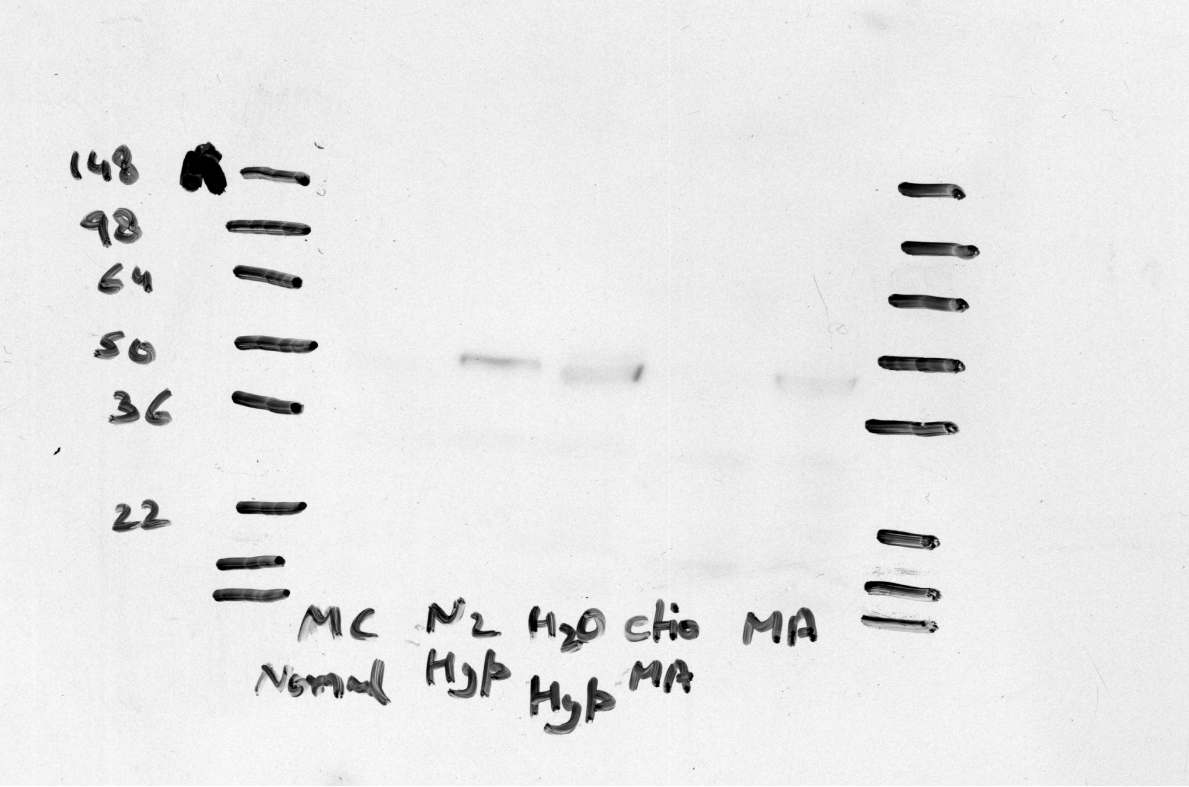


Fig: 4.1 The image of film showing the conditional stabilityof MC/MA-RAP2.3-HA protein when determined through western blotting

The image shows the presence of RAP2.3-HA tagged protein. 35S::MC-RAP2.3-HA treated under hypoxia and submerged and in 35S::MA-RAP2.3-HA normal condition. The approximate size of RAP2.3-HA is around ~33 kDa. However the expected band size always appears higher.

Lane 1 & 7 contains molecular weight marker (in KDa, marked with a marker pen)

2. 35S::MC-RAP2.3-HA (seedlings in normoxia)

3. 35S::MC-RAP2.3-HA (seedlings treated with hypoxia)

4. 35S::MC-RAP2.3-HA (seedlings treated for submergence)

5. 35S::MA-RAP2.3-HA (etiolated seedlings)

6. 35S::MA-RAP2.3-HA (normoxia seedlings)

The Western blot result (Fig: 4.1) shows that in normoxia the 35S::MC-RAP2.3-HA line doesn't show protein accumulation. The probable reason for this is RAP2.3 protein begins with the amino acids 'NH_2-Met-Cys' (Licausi *et al.,* 2011; Gibbs *et al.,* 2011). In the presence of oxygen and nitric oxide the Cys at the N-terminus after Met cleavage by MetAP (Lee *et al*., 2005) is oxidized and is arginylated by ATE followed by E3 ligation and recognition leading to degradation of protein (Lee *et al*., 2005). Therefore, it is expected that MC-RAP2.3 will destabilize in normoxia. However, in the case of hypoxia and submerged conditions stabilization of protein from MC-RAP2.3-HA line was shown (Fig: 4.1). On the other hand, RAP2.3 protein shows presence in the 35S::MA-RAP2.3-HA line. This is because alanine doesn't undergo any oxidation leading to no degradation of proteins and hence gets stabilized.

Thus, through western blot, I confirmed the N-end rule and O_2 conditional stability of RAP2.3 protein in both 35S-MC/MA-RAP2.3-HA lines.

4.2.2 Chromatin immunoprecipitation assay to determine binding of RAP2.3 protein to 'GCCGCC' *cis*-elements *in-vivo*

4.2.2.1 To verify the target promoters for RAP2.3 binding *via* by ChIP

ChIP was performed using hypoxic Col-0 lines of *35S::MC-RAP2.3-HA* and *35S::MA-RAP2.3-HA* four day old seedlings grown in light. Chromatin was extracted and used for cross-linking. Glycine was used to stop the cross-linking. After sonication, each sample was divided into four parts. One part was treated with IgG antibody. The IgG antibody acts as negative control. The DNA obtained from the IgG treated samples is considered as background signal, which indicate the level of non-specific binding by IgG. The 2nd and 3rd part were treated with HA-antibodies. The HA-antibody will bind to the HA tagged RAP2.3 protein cross-linked with DNA. And the 4^{rth} part is the input DNA (no antibodies treatment). Further steps were followed to finally extract DNA as described in Chapter 2.

The DNA obtained from the ChIP experiment was first analysed for the amplification of *ABI5* promoter. The DNA obtained after ChIP experiment was first tested *via* semi-quantative PCR. qPCR primers for *ABI5* promoter amplification were used and the PCR product was run on 2% agarose gel. The expected size of *ABI5* amplified DNA is 99 bp. The electrophoresis result shows the bands of right size in all the samples (lane 3, 4, 5 and 6) (Fig: 4.2). The lane 2, which has samples, treated with IgG antibody during ChIP

experiment showed no band. Lane 3 and 4 that has HA antibody treated samples show bands due to amplification of *ABI5* promoter site on DNA that must have been bound by RAP2.3-HA protein during cross-linking. Lane 5 and 6 has input and sonicated samples that also shows band. Since both these samples contain mix of DNA fragments that have *ABI5* binding site as well as non-specific site. The intensity of band also appears high in these lanes because primers bound more number of DNA fragments since unwashed input and sonicated samples have additional amount of DNA hence, high amplification of promoter comparison to other samples.

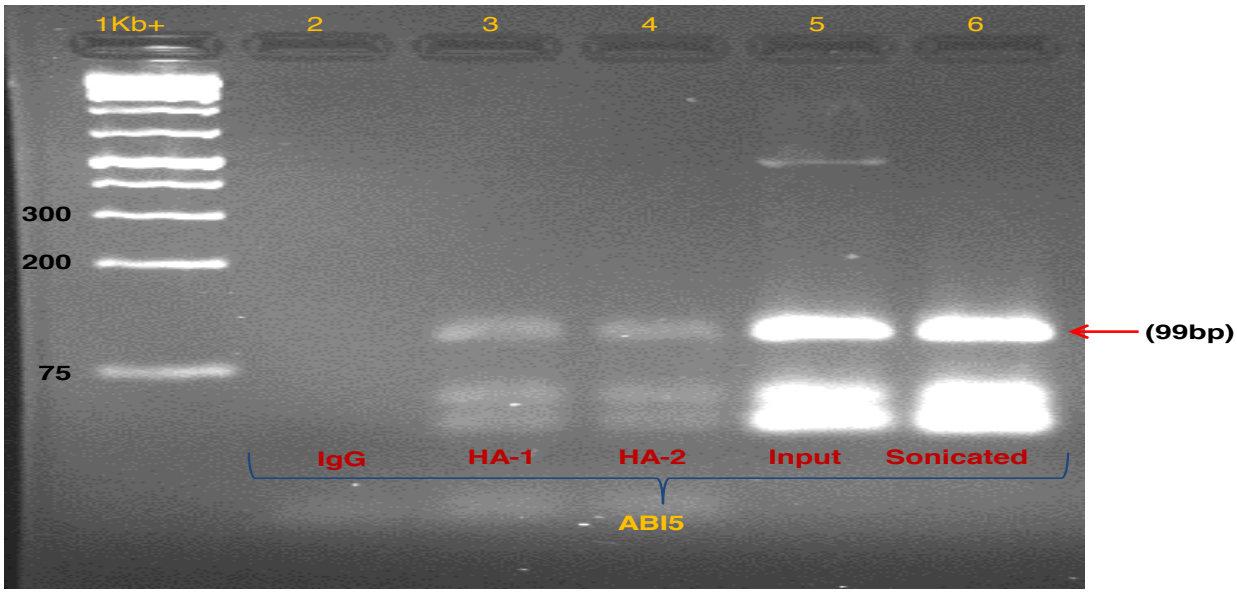


Fig: 4.2 Amplified bands of *ABI5* promoter tested on ChIP DNA derived from MC-RAP2.3-HA containing seedlings.

Agarose gel (2%) electrophoresis showing amplified fragments of *ABI5* promoter site. First lane has 1Kbp marker for size comparison. The approximate size of amplified product is 99 bp. Lane 2: Amplified *ABI5* promoter fragments from DNA obtained from IgG antibody treated sample (negative control). Lane 3&4: *ABI5* promoter amplification from HA-antibody treated samples (Positive control). Lane 5&6: Input and sonicated samples showing *ABI5* promoter fragment amplification.

IgG: Immunoglobulin G antibody treated samples.

HA : Hemagglutinin antibody treated samples.

Input : All protein specific and non-specific bound product after nuclei lysis.

Sonicated : Fragmented DNA in between size range of 50-500 bps.

After confirmation of *ABI5* promoter amplification on *35S::MC-RAP2.3-HA* lines, I further treated a set of *35S::MC-RAP2.3-HA* again with hypoxia and also submergence together with *35S::MC/MA-RAP2.3-HA* untreatedlinesand *35S::MA-RAP2.3-HA* etiolated seedlings and Col-0 as a control. The aim of this is to examine whether these different conditions also result in RAP2.3 binding to the *ABI5* promoter.

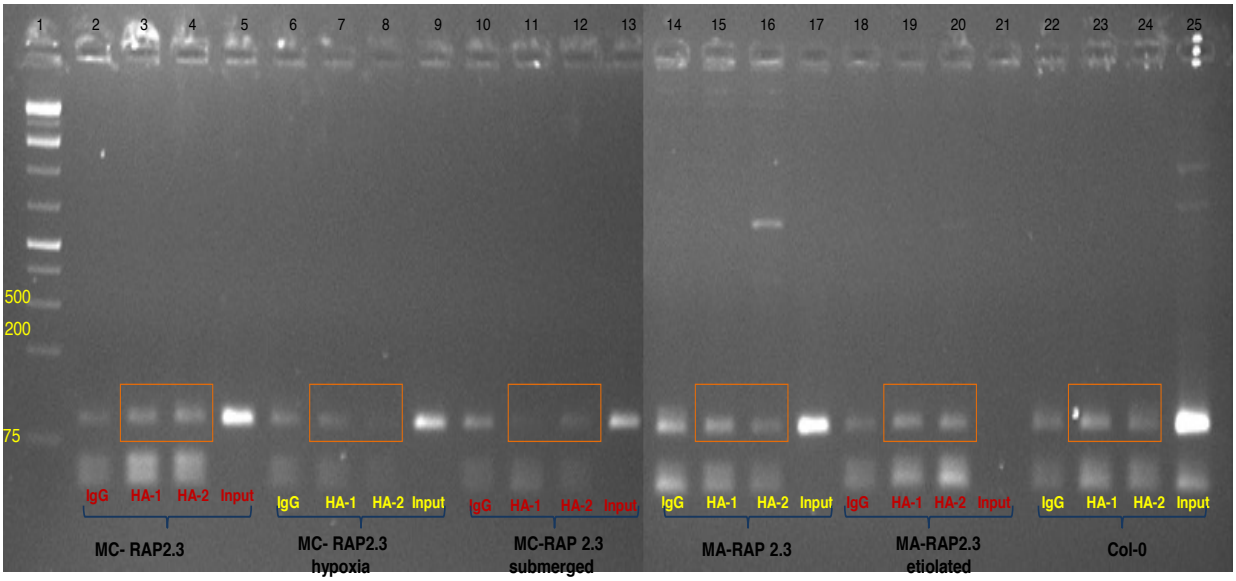


Fig: 4.3 Agarose gel (2%) with amplified PCR product of *ABI5* promoter

First well has 1Kbp marker for size comparison. The approximate size of amplified product is 99 bp. Lane1: DNA ladder (1Kbp+). Lane 2-5: MC-RAP2.3-HA (seedlings in normoxia). Lane 6-9: MC-RAP2.3-HA (seedlings treated with hypoxia). Lane 10-13: MC-RAP2.3-HA (seedlings treated for submergence). Lane 14-17: MA-RAP2.3-HA (normoxia). Lane 18-21: MA-RAP2.3-HA (etiolated seedlings). Lane 22-25: Col-0 (untreated normoxia)

IgG: Immunoglobulin G antibody treated samples.
Orange boxes represent all the samples treated with HA antibodies showing *ABI5* promoter amplification.
Input: All protein specific and non-specific bound product after nuclei lysis.

All isolated DNA samples after ChIP procedures were subjected to semi-quantative PCR. PCR products were loaded on 2% agarose gels and run at 70 volts. The results show the successful amplification of the *ABI5* promoter in all the HA and input samples isolated from MC-RAP2.3-HA (normoxia, hypoxia and submerged) and MA-RAP2.3-HA (normoxia and etiolated) (Fig: 4.3). However IgG treated samples and control Col-0 also show bands. I assume that these are unspecific band. Since IgG is a negative control here so it is expected that IgG binds to the unspecific sites on DNA and that gives the band.

4.2.2.2 Detection of amplification of promoters having GCCGCC *cis*-element from ChIP DNA through quantitative PCR (qPCR)

qPCR is a technique and based on the basic principle of normal PCR. The difference is that in qPCR, the amplified DNA is detected from the beginning of reaction *i.e.* in real time using fluorescent technology and quantification of amplified DNA is achieved, while in normal PCR the result of the reaction is detected only after the reaction is complete (Kochanowski and Reischl, 1999).

The use of qPCR also has several advantages over the normal PCR including the relatively small amount of sample required for analysis, the ability to reproduce rapid and accurate data, and the ability for analysing more than one gene at a time. Consequently, qPCR has become a standard tool in molecular biology.

4.2.2.2.1 The principle of qPCR

In qPCR a DNA-binding dye such as SYBR green is used which binds to DNA and produces fluorescence as the product forms. An increase in DNA product during PCR therefore leads to an increase in fluorescence intensity. This intensity of fluorescence is measured at each cycle. The fluorescence signal is directly proportional to DNA concentration and the linear correlation between PCR products. Thus, the fluorescence intensity is used to calculate the amount of template present at the beginning of the reaction. The point, at which fluorescence is first detected as statistically significant above the baseline or background, is called the Ct value (threshold cycle) or Cp value (crossing point-PCR-cycle).

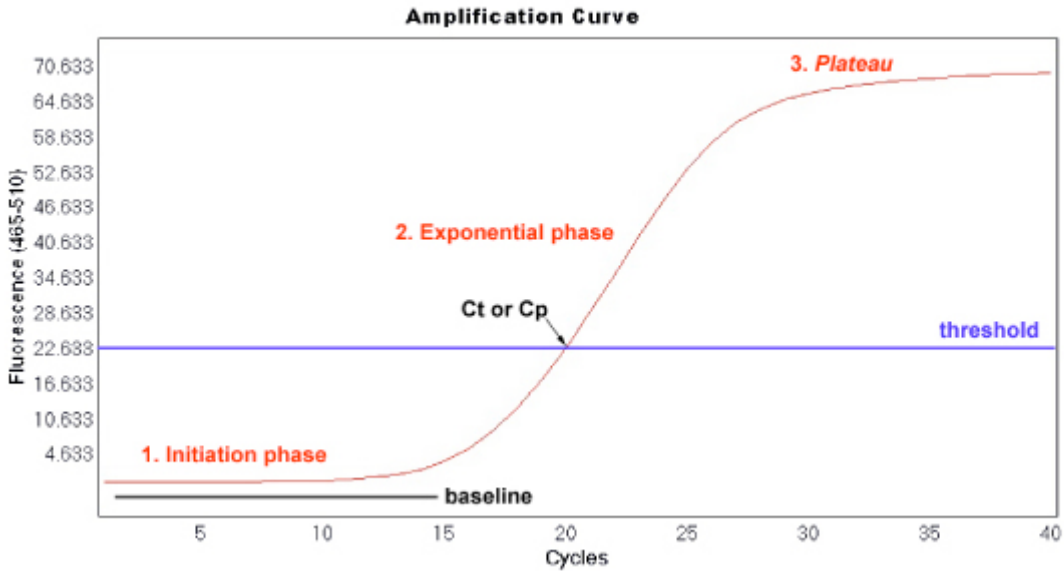


Fig: 4.4 A representative graph showing different phases of a curve during PCR amplification.

Blue line: amplification of curve of a positive sample that is Ct or Cp value.

Red line: amplification of curve.

(Figure taken from Rodriguez et al., 2013)

During qPCR an amplification curve shows different phases. The first is called the initiation phase; this is the beginning phase during the first PCR cycles where the emitted fluorescence cannot be distinguished from the baseline. Second is the exponential or log phase. During this phase, there is an exponential increase in fluorescence, before the third phase i.e., plateau phase is reached. In plateau phase, reagents start exhausting and no increase in fluorescence is observed. Quantification of amplified product occurs in exponential phase. The Ct value is the number of cycles required to attain a fixed threshold fluorescence signal level. The Ct value obtained can be measured through absolute quantification or relative quantification (Wong *et al*., 2005). Later, the quantification results can be tested for statistical significances.

In my project, I used a relative quantification method to analyze the data. The relative quantification is based on input data to determine fold-differences in expression of the target gene (Haring*et al*., 2007). In this method, the Cp value from the ChIP samples is divided by the Cp value from the input sample. It is expected that signals form ChIP samples and No antibody (input sample) signal levels are directly related to the amount of input chromatin.

4.2.2.2.2 Analysis of ChIP samples for determination of putative genes for ERFVII RAP2.3 binding in their promoter regions

Isolated ChIP DNA was processed for qPCR. For that purpose, PCR sensiFAST SYBR*Hi-Rox kit (Bio-line) was used. The primers used for *ABI5* promoter amplification for carrying out qPCR were designed on the basis of these conditions: Firstly, The qPCR primer length should range from 100-150 bps. Secondly, the melting temperatureof primers must be around 58-60°C. Thirdly, the target region should possess two 'GCCGCC' sites. The sequence details of qPCR primers are listed in Appendix: 2.2.

Data obtained from ChIP-qPCR was prepared in an Excel sheet for analysis. The qPCR results for *ABI5* promoter amplification in all samples were first calculated using (2^-Cp) (Delta-delta method for comparing relative expression). Then the average and standard deviation of 3 technical replicates of each samples was obtained. Further, both average values were used to calculate normalized values (see section: 4.2.2.2.1) for IgG and HA samples of each line. Graphs were plotted on the basis of normalized values.

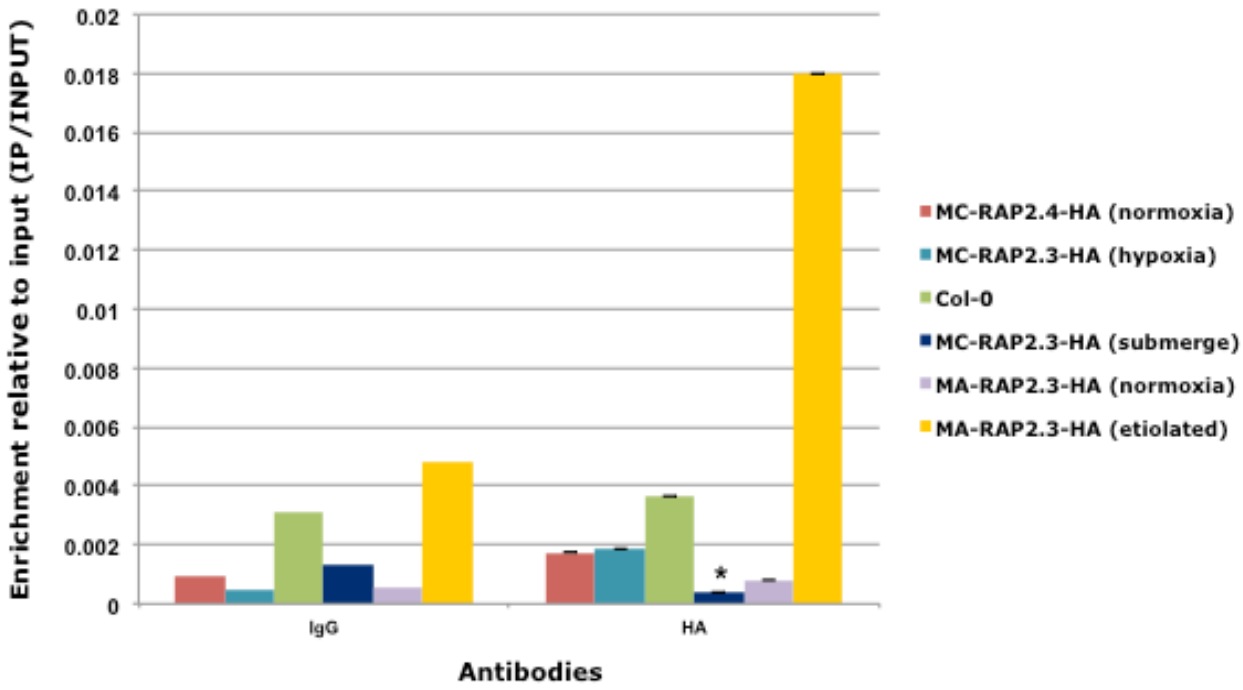


Fig:4.5 qPCR analysis of ChIP DNA obtained from MC\MA-RAP2.3-HA seedlings under normal and stress conditions. Relative fold enrichment for *ABI5* promoter obtained following HA antibody immunoprecipitated and corresponding IgG control ChIP samples.

The normalized relative enrichment of the *ABI5* promoter obtained by ChIP-qPCR is represented by the bar graphs. The error bars represent standard deviation (SD). The asterisks represent a significant difference between the Col-0 HA-normalized value to other transgenic 35S-MC/MA-RAP2.3-HA lines under normal and treated conditionusing t-test (* represents $P < 0.05$).

The relative enrichment level in each individual line was calculated as per normalized values and a graph was plotted accordingly (Fig: 4.5). However, the t-test analysis showed no statistical difference between IgG and HA-IPs (Fig: 4.5). Further on comparing Col-0 HA-normalized values to other transgenic 35S-MC/MA-RAP2.3-HA lines normal and treated, the significant difference was only seen in 35S-MC -RAP2.3-HA (submerged).

Therefore, in order to gain a definite understanding of enrichment of targeted gene, I again performed ChIP experiment only on 35S::MA-RAP2.3-HA (normoxia) seedlings. As described previously, in MA lines Ala replaces Cys-2. Therefore, protein bearing Ala in place of Cys at their N-terminal is not oxidised so it can't be converted into a secondary residue hence no degradation of protein takes place. Therefore, in MA lines stable proteins lead to upregulation of target genes in comparison to MC lines under normal growth conditions (*i.e.,* normoxia in the presence of O_2 and NO). Thus, using MA line for ChIP analysis can better confirm transcription factor binding to a target gene. As hypothetically, RAP2.3 protein will be stabilized in MA lines and that stableprotein will bind specifically to the promoter region on DNA fragments during cross-linking. After immunoprecipitation recoverd DNA will be tested for promoter amplification. So, it is expected that enriched region is the promoter section that was bound by RAP2.3.

The ChIP samples derived from MA-RAP2.3-HA Col-0 line were analysed by qPCR and the data obtained was again processed for normalisation (section: 4.2.2.2.1). A graph was made on the basis of normalized data. The graph shows enrichment for *ABI5* promoter region in HA immuno-precipitated samples. The t-test analysis showed statistical difference between IgG and HAs less than p-value<0.5. This result suggests that there is significant difference between IgG and HA-IPs (Fig: 4.6).

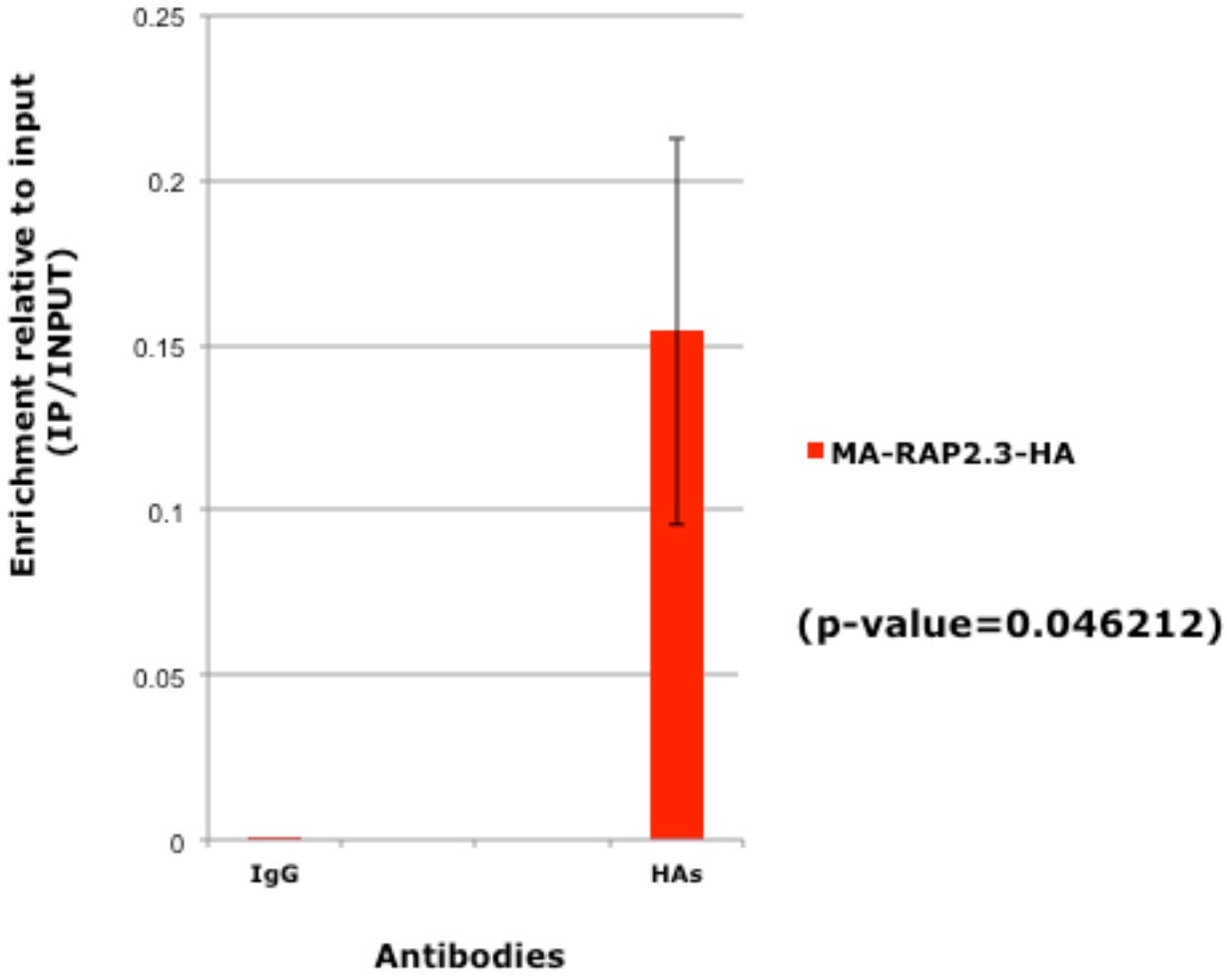


Fig: 4.6 Graph showing relative enrichment of *ABI5* promoter on ChIP-DNA from MA-RAP2.3-HA line

qPCR analysis for *ABI5* promoter region on ChIP-DNA from 4 day old normoxic etiolated seedlings containing MA-RAP2.3-HA. Normalization of IgG and HA-IP with input showing high enrichment level in HAs (average of two HA-IP values). The error bars represent the standard error between the triplicate qPCR reactions for each sample. p-value represents significant difference value.

There are several other genes that have 'GCCGCC' binding site in their promoter region and are regulated by ERFVII's such as the ABA receptor *PYL7* (Ref: Chapter-3). In order to examine binding of RAP2.3 at the PYL7 promoter, I carried out ChIP-qPCR using *PYL7* promoter specific primers (Appendix: 2.2) on 4-day old normoxic etiolated seedlings containing MA-RAP2.3-HA. qPCR data was quantified *via* normalization method and graph was ploted. This graph represented the greater enrichment for *PYL7* promoter region in HA-IPs (positive control). The t-test analysis showed statistical difference between IgG and HAs less than p-value<0.5 suggesting that there is significant difference between IgG and HA-IPs (Fig: 4.7).

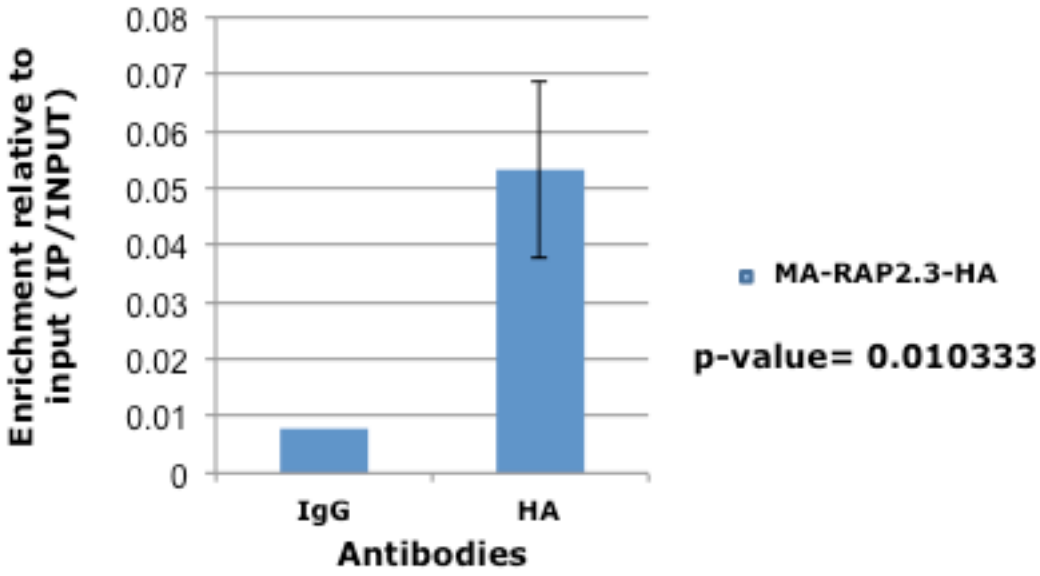


Fig: 4.7 Graph showing enrichment of *PYL7* promoter region tested on ChIP-DNA extracted from MA-RAP2.3-HA line.

qPCR analysis for *PYL7* promoter region on ChIP-DNA from 4 day old normoxic etiolated seedlings containing MA-RAP2.3-HA. Normalization of IgG and HA-IPs with input shows high enrichment level in HAs. The error bars represent the standard error between the triplicate qPCR reactions for each sample. p-value represents significant difference value.

4.3 Discussion

Initially in my project I looked for promoters that have two closely localised 'GCCGCC' *cis*-elements and therefore may be regulated by ERFVII's. To confirm the *in-vivo* specificity of binding of ERFVII's to such genes, I have used chromatin immunoprecipitation approach to identify *in-vivo* binding of the RAP2.3 to promoter fragments containing *cis*-elements. As this technique gives results based on quntitative analysis so information gained will be more precise.

4.3.1 Analysis of qPCR data

For the purpose of ChIP experiment I used 35S::MC/MA-RAP2.3-HA Col-0 lines. The experiment was performed on six individual lines and different treatment combinations such as: {MC-RAP2.3-HA (Normoxia *i.e*., in the presence of O_2 and NO), MC-RAP2.3-HA (hypoxia), MC-RAP2.3-HA (submerged), MA-RAP2.3-HA (normal), MA-RAP2.3-HA (etiolated) and Col-0}. The ChIP-qPCR data obtained were first normalized. The t-test analysis done on HA-IPs normalized data between the control Col-0 (normal) and MC/MA-RAP2.3-HA under normal and stress conditions shown a significant difference only in MC-RAP2.3-HA (submerged). Thus I used only normoxic MA-RAP2.3-HA line to further conduct ChIP for *ABI5* promoter amplification. The t-test analysis revealed significant difference between IgG and HA-IPs for *ABI5* (Fig: 4.6).

It has been shown in my previous results that *PYL7* also carries two 'GCCGCC' site at 5'UTR region like *ABI5*. Therefore, I also tested the ChIP-DNA obtained from MA-RAP2.3-HA in Col-0 seedlings (normoxia) for *PYL7* promoter region. qPCR analysis showed a high enrichment of *PYL7* promoter region in HA-IPs. The t-test analysis also showed significant difference between IgG and HA-IPs for *PYL7* (Fig:4.7).

Therefore, all these results suggest that RAP2.3 protein binds to the two *cis*-elements 'GCCGCC' binding site in *ABI5* and *PYL7*.

4.3.2 Role of genes identified from ChIP experiment for RAP2.3 binding at their promoter region in plants

The genes *ABI5* and *PYL7* that are identified to be regulated by ERFVII *RAP2.3* binding at their promoter region from above ChIP results play important roles during different stages of plants growth and development.

ABI5, a bZIP transcription factor binds ABA responsive *cis*-element (ABRE) and activates ABA-mediated signalling during seed germination process (Giraudat *et al*., 1994; Hobo *et al*., 1999). While, *PYL7* functions as an abscisic acid sensor (ABA receptor). Recent studies have shown that *PYL7* mediates ABA-dependent regulation of ABI1 and ABI2 (Zhang X, *et al*. 2013; Lee *et al*. 2013). There are 13 *PYL* genes in *Arabidopsis* that have similarity to PYR1 and are named as PYR1 LIKE1-13 (PYL1-13) (Umezawa *et al* 2010; Park *et al*., 2009; Ma *et al*., 2009). These genes are also called as RCARs (REGULATORY COMPONENTS OF ABA RECEPTOR) (Ma *et al*., 2009).

In 2009, two independent research groups demonstrated that ABA-dependent inhibition of cladeA PP2C (Protein phosphatase 2C encoded by ABI1 and ABI2) by PYR/PYL/RCARs induces repression of SnRK2s (Ma *et al*. 2009; Park *et al*, 2009). PP2C is an enzyme that removes a phosphate group from its substrate and acts as a negative regulator (Leung *et al*., 1997; Meyer *et al*., 1994; Leung *et al*., 1994). Whereas, SnRK2s (SNF-1 related protein kinase 2, enzyme adds phosphate group to proteins) are called positive regulators of ABA signaling which act downstream of the PP2Cs.

Recently, a molecular mechanism from gaseous signal sensing by ERFVII's through nitric oxide to the regulation of the key germination repressor, *ABI5*, in the seed endosperm has been identified (Gibbs *et al*., 2014). As *ABI5* is regulated by ERFVII's through binding to 'GCCGCC' at 5'UTR. However, in the absence of nitric oxide Cys2 of ERFVII's does not undergo oxidation hence remains stable resulting in enhanced *ABI5* activity. Thus, ERFVII's regulate ABA signaling through *ABI5* depending on presence of Nitric oxide. This new insight suggests downstream targets of N-end rule pathway need both oxygen and nitric oxide for their regulation.

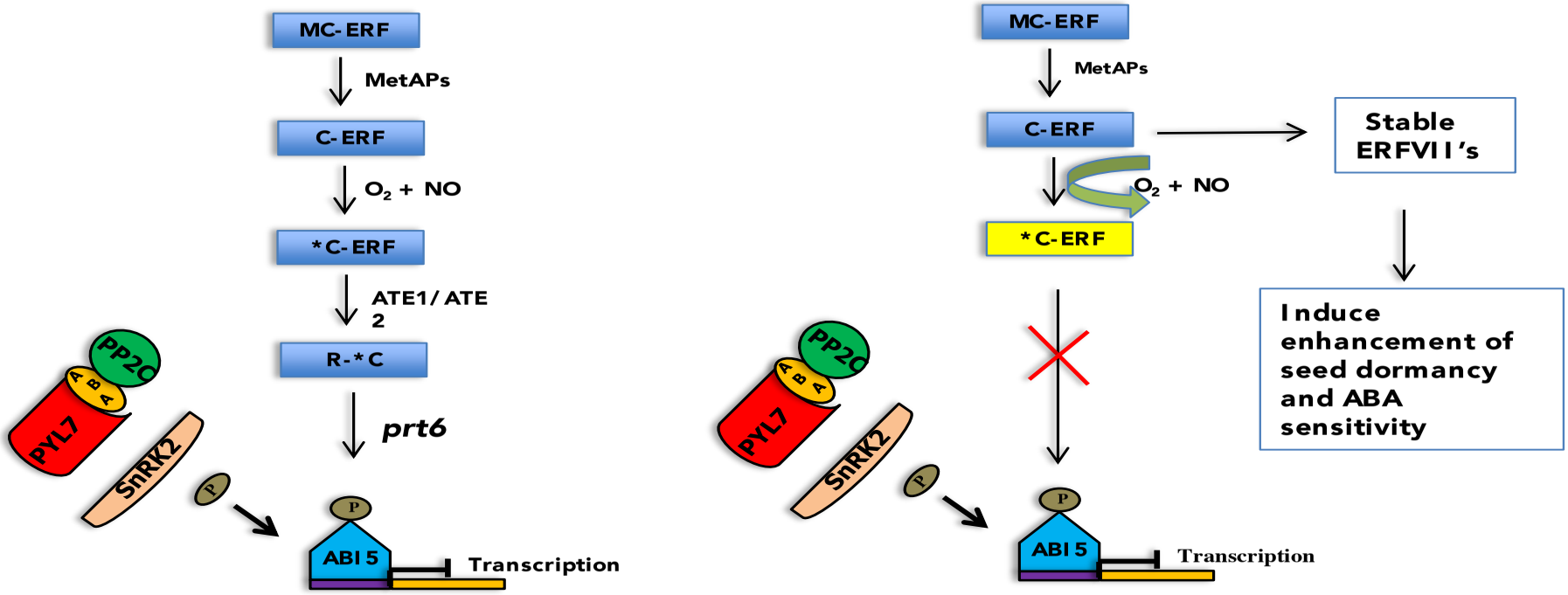


Fig:4.8 A hypothetical model showing a link between N-end rule pathway and the regulation of a downstream ABA signalling component *ABI5*

ABI5 protein activity is activated through phosphorylation by SnRK2 kinase and in addition, that ABA also stabilized the ABI5 protein. However, in absence of oxygen or nitric oxide ERFVII's doesn't undergo oxidation of cysteine leading to accumulation of substrates. Thus, *ABI5* promoter activity is enhanced by the stability of N-end rule substrates. This indicates that the ERFVII's regulate ABA signalling through controlling the expression of *ABA INSENSITIVE 5 (ABI5)*

4.3.3 Another approach for determination of ERFVII RAP2.3 binding at the promoter region of genes carrying GCCGCC site

Chromatin Immunoprecipitation technique is considered one of the best method to detect *in-vivo* protein-DNA interaction. Besides this other approach such as Electro Mobility Shift Assay (EMSA) can also be used to detect protein-DNA interaction *in-vitro*. This technique is based on electrophoretic separation of a protein-DNA complex on a polyacrylamide or agarose gel. On the basis of the speed of protein-DNA complex that run more slowly than the corresponding free unbound DNA. It could be identified that there is binding of DNA with specific protein, which leads to slowing of protein-DNA complex in gel. Once DNA-protein binding is determined *in vitro*, further visualization of the complex can be achieved by labelling nucleic acid fragments by radioactive probes followed by autoradiography.

CHAPTER-5

In-vivo detection of protein stability and localization in *Arabidopsis thaliana* cells using fluorescence imaging

5.1 Introduction

5.1.1 Fluorescent proteins

There are three amino acids in proteins *i.e*., tryptophan, tyrosine and phenylalanine that are responsible for the absorption and emission of fluorescence. Such proteins that have tryptophan, tyrosine and phenylalanine are called fluorescent proteins (Shimomura *et al*. 1962). The first green fluorescent protein (GFP) was isolated from jellyfish, *Aequorea victoria*. For the discovery of GFP and its development Roger Tsien, Osamu Shimomura and Martin Chalfie received the Nobel Prize in 2008. The protein was named "green fluorescent protein" by Morin and Hastings (1971) because it transduces blue light to green. GFP is composed of 238 amino acids i.e., nearly 27 kDa with self-folding ability into a barrel like structure (Ormo *et al*., 1996). GFP absorbs light with short wavelengths and immediately returns it with a different colour that has a longer wavelength. Such as, blue emission of light as a result of a chemical reaction of the protein *aequorin* changes into green fluorescent light. (Yang *et al*., 1996; Tsien *et al*., 1998). GFP has two excitation peaks, a major one at 395nm and a smaller one at 475nm (blue) and an emission peak at 509nm (green) (Tsien *et al*., 1998).

Currently, GFP is used as a reporter gene because fluorescent protein tagging allows proteins to be directly visualized for detection within living cells. GFP-tagged proteins have revealed gene expression and the location of a protein in a cell and these features aid in the identification of the level of gene regulation. Introducing mutations in the peptide sequence of the original fluorescent protein has led to production of many other fluorescent proteins. These fluorescent proteins are of different colours and most of them have single excitation and emission spectra (Cubitt *et al*., 1995). The most common are blue (BFP, EBFP), cyan (CFP, ECFP) (Sawano *et al*., 2000; Patterson *et al*., 2001) and yellow fluorescent proteins (YFP, EYFP, Venus, Citrine) (Griesbeck *et al*., 2001; Patterson *et al*., 2001; Nagai *et al*., 2002).

5.1.2 Yellow fluorescent protein (YFP)

Yellow fluorescent protein (YFP) is a genetic mutant of green fluorescent protein. Its excitation peak is 514 nm and its emission peak is 527 nm. Further modifications resulted into the development of the enhanced yellow fluorescent protein (EYFP). This form of YFP is the brightest and being highly used at present. GFP is considered a suitable

in-vivo search tool, but still it has some drawbacks for example: Firstly, it can induce cell death. The intense excitation of the protein during microscopic studies for extended periods led to free radicals that are quite toxic to cells resulting in cell death (Liu *et al*., 1999). Secondly, GFP has been more successfully used as a possible marker for mostly short-time scale experiments since long-term GFP-expressing cell strains have been mostly unsuccessful (Hong *et al*., 2001; Chalfie *et al*., 1994; Lybarger *et al*., 1996). Therefore, YFP a variant of GFP has been constructed through mutation. It is less toxic, more stable and allows long-term fluorescence analysis. Thus, YFP is considered more suitable marker for *in-vivo* studies than GFP (Heim *et al*., 1994).

5.1.3 Using *promERFVII's:MC/MA-ERFVII's-YFP* constructs to determine stability and locus of N-end rule pathway substrates

Transgenic lines of each ERFVII driven by their own promoter in MC (unstable) in Col-0 and *prt6-1* and MA (stable) versions in Col-0 were generated. The expression of each ERFVII gene with endogenous promoter would allow analyse the function of each ERFVII through its own promoter during any process studies. This work was carried out by Dr. George Bassel and Dr. Sophie Berkhan in our lab. Later, I continued the screening of T_2 transgenic lines to identify homozygous ERFVII's-YFP lines to use for investigation of protein stability and localization using confocal microscopy.

Diagrams representing promERFVII's:MC/MA-ERFVII's-YFP constructs and expected state of stability of proteins into Col-0 and *prt6-1* are shown in Fig: 5.1. ERFVII's under their own promoters contain a C-terminal YFP.

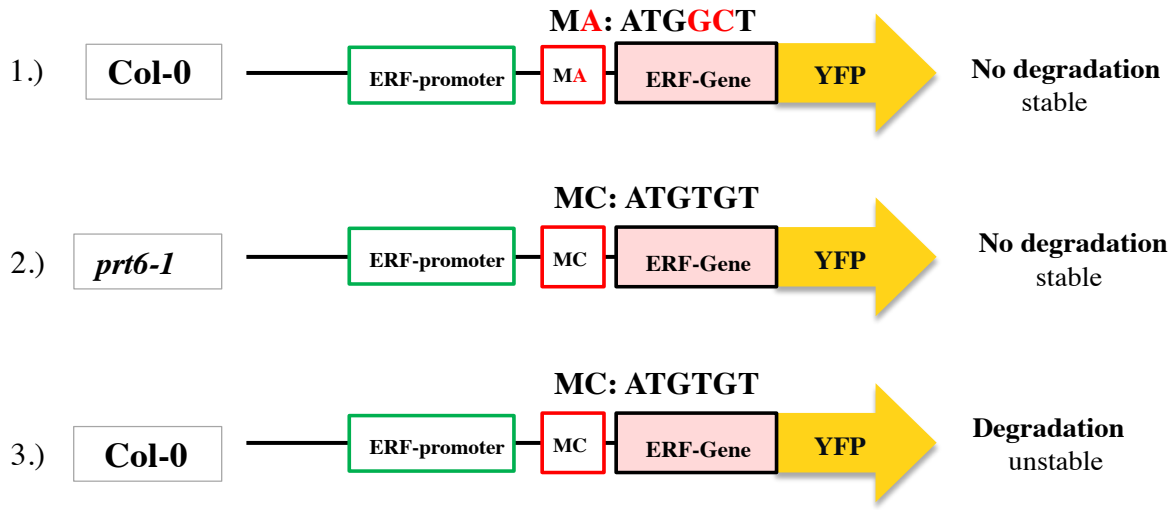


Fig:5.1 Diagram showing *promERFVII's:MC/MA-ERFVII's-YFP* in Col-0 and *prt6-1* genotypes

1. *promERFVII's:MA-ERFVII's-YFP* under their own promoter in Col-0. MA-ERFVII's-YFP could not degrade due to presence of alanine and remain stable. Alanine an analog of cysteine cannot oxidize hence show no degradation.
2. *promERFVII's:MC-ERFVII's-YFP* under their own promoter in *prt6-1*. MC-ERFVII's-YFP is expected to be stable in *prt6-1* background. Since in absence of E3-ligase: *prt6-1*, ERFVII's are not sent for degradation to 26S proteasome. Hence remains stable.
3. *promERFVII's:MC-ERFVII's-YFP* under their own promoter in *prt6-1* were crossed to Col-0. MC-ERFVII's-YFP will show degradation due to presence of E3 ligase *PRT6. PRT6* will recognize the ERFs and send it to 26S proteasome for degradation.

5.2 Results

5.2.1 Use of Yellow Fluorescent Protein (YFP) tagging to determine the expression and subcellular localization of ERFVII's proteins in *Arabidopsis thaliana*

5.2.1.1 Screening of homozygous plants containing *promERFVII's:MC/MA-ERFVII's-YFP*

The T_2 lines of *promERFVII's:MC-ERFVII's-YFP* in Col-0 and *prt6-1* and *promERFVII:MA-ERFVII-YFP* in Col-0 were preliminary generated by Dr. Gorge Bassel and Dr. Sophie Berkhan. Further, parental Col-0 plants containing *promERFVII's:MC-ERFVII's-YFP*were crossed to *prt6-1* by Dr. Sophie Berkham and F_1 lines were obtained. Later, I proceeded for selection of transgenic lines to obtain homozygous genetic material.

5.2.1.1.1 *promERFVII's:MC-ERFVII's-YFP* in *prt6-1*

T_2 seeds of five *prt6-1* lines containing *promHRE1:MC-HRE1-YFP and* three *prt6-1* lines containing *promRAP2.2:MC-RAP2.2-YFP* together with 10 transgenic lines with promHRE1:MA-HRE1-YFP and promRAP2.2:MA-RAP2.2-YFP in Col-0 were plated on half-MS plates containing hygromycin for selection of transgenic plants. Plates having 3 days old etiolated seedlings were then observed under fluorescence stereomicroscope. 12 resistant plants of each line that showed fluorescence were selected and transferred to soil. T_3 seeds of all 12 plants were collected and plated again on hygromycin (30ug/ml) (w/v) for selection of homozygous lines. Three seedlings of each five *prt6-1* lines containing *promHRE1:MC-HRE1-YFPand* three *prt6-1* lines containing *promRAP2.2:MC-RAP2.2-YFP* and promHRE1:MA-HRE1-YFP and promRAP2.2:MA-RAP2.2-YFP were observed under confocal microscope. Finally, one transgenic line with *promHRE1:MC-HRE1-YFP* in *prt6-1* and one in *promRAP2.2:MC-RAP2.2-YFP* in *prt6-1* and one of promHRE1:MA-HRE1-YFP and one of promRAP2.2:MA-RAP2.2-YFP were identified as homozygous line.

	promHRE1:MC-HRE1-YFP	***promRAP2.2:MC-RAP2.2-YFP***	**promHRE1:MA-HRE1-YFP**	**promRAP2.2:MA-RAP2.2-YFP**
T_2	5	3	10	10
T_3	1	1	1	1

Table:5.1 Summary of number of lines used for *promERFVII's:MC/MA-ERFVII's-YFP* analysis

The image analysis of root-tips of each of the *promERFVII's:MC/MA-ERFVII's-YFP* lines was done by first treating seedlings with propidium iodide (PI) (1μg/ml) and incubation for 5-7 min in dark to stain cell-walls of the tissue. PI solution was decanted and samples were washed four times with water. Later, samples were then mounted on a glass slide and were observed under the confocal microscope at 488nm. The results show fluorescence in nuclei at root tip region of seedlings. In all three seedlings of each *prt6-1* line containing *promHRE1:MC-HRE1-YFP* the fluorescence in nuclei were seen (Fig: 5.2).

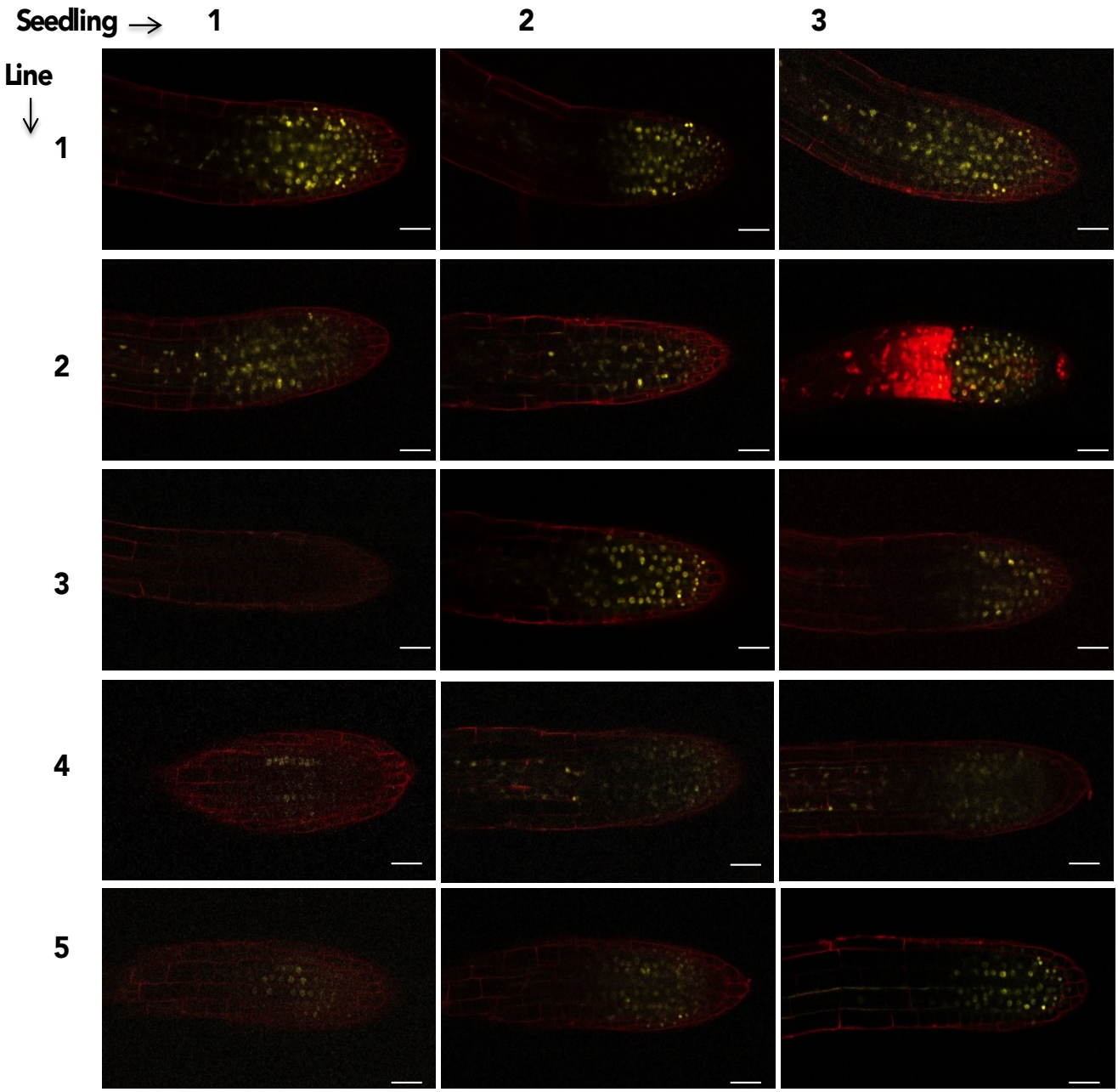


Fig: 5.2 Three representative seedlings of five *prt6-1* lines showing expression of *promHRE1:MC-HRE1-YFP* in a bright field image.

3 days old *prt6-1* seedling (1-3) of five lines carrying *promHRE1:MC-HRE1-YFP* are showing yellow fluorescence in the nuclei of their root tips.

Yellow fluorescence in images denotes HRE1 expression in nuclei, red fluorescence is due to propidium iodide (1μg/ml) staining in cell-wall.

20x object lenses and 488nm wavelength were used to obtain the images *via* confocal microscope.

Scale bars = 50 μm.

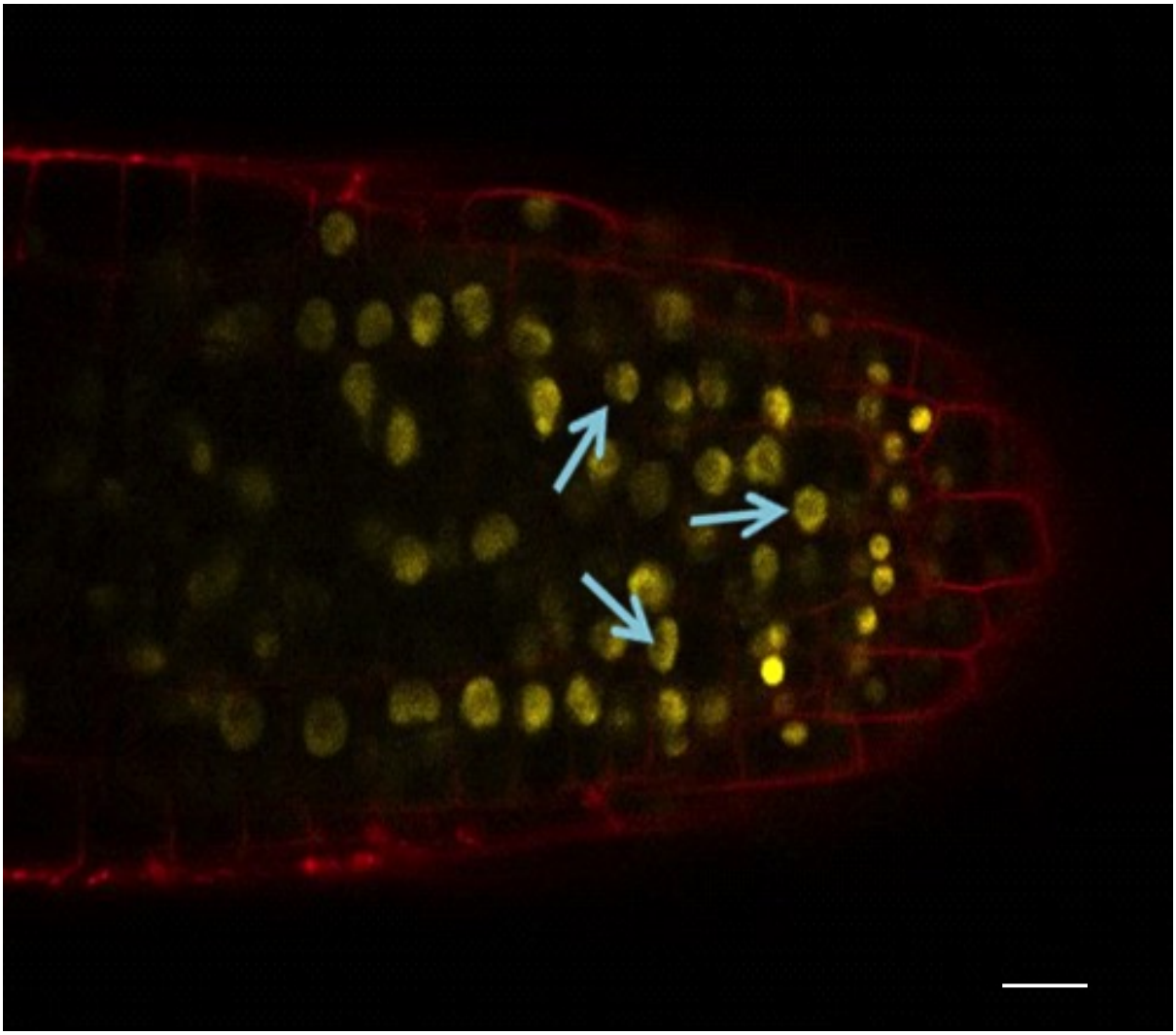


Fig: 5.3 Enlarged image of root-tip of one seedling containing *promHRE1:MC-HRE1YFP* showing fluorescence of YFP tagged protein in nuclei denoted by arrows.

The three *prt6-1* lines containing *promRAP2.2:MC-RAP2.2-YFP* were also showing fluorescence in nuclei at root tip region. However, the the levels of fluorescence is low in comparison to *promRAP2.2:MC-RAP2.2-YFP* (Fig:5.3).

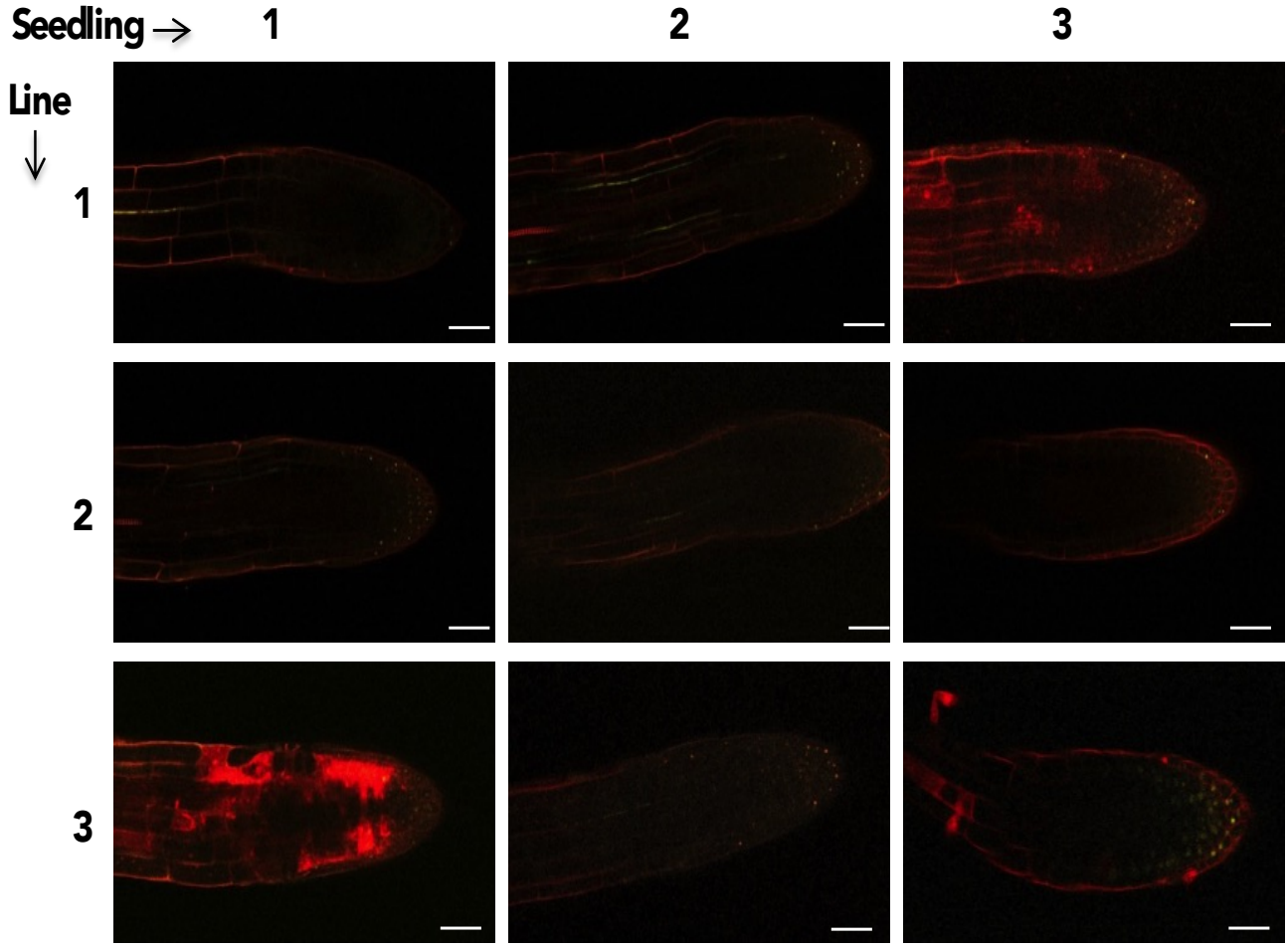


Fig:5.4 Three representative seedlings of three *prt6-1* lines showing expression of *promRAP2.2:MC-RAP2.2-YFP* through confocal microscope.

3 days old *prt6-1* seedling (1-3) in three lines carrying *promRAP2.2:MC RAP2.2-YFP* are showing yellow fluorescence in the nuclei of their root tip.

Yellow fluorescence denotes RAP2.2 expression in nuclei, red fluorescence is due to propidium iodide (1μg/ml) staining in cell-wall.

20x object lenses and 488nm wavelength were used to obtain the images *via* confocal microscope.

Scale bars = 50 μm.

Finally, on the basis of image analysis one *prt6-1* line containing *promHRE1:MC-HRE1-YFP* and one *prt6-1* lines containing *promRAP2.2:MC-RAP2.2-YFP* were selected as homozygous lines.

5.2.1.1.2 *promERFVII's:MA-ERFVII's-YFP* in Col-0

T_2 seeds of six Col-0 lines containing *promHRE1:MA-HRE1-YFPand* four Col-0 lines containing *promRAP2.2:MA-RAP2.2-YFP* were plated on half-MS plate containing hygromycin for selection of transgenic plants. 12 resistant plants of each line were selected and transferred to soil. T_3 seeds of all 12 lines were collected and plated again on hygromycin (30ug/ml) (w/v) for selection of homozygous lines. Three seedlings of each six Col-0 lines containing *promHRE1:MA-HRE1-YFP* andthree Col-0 lines containing *promRAP2.2:MA-RAP2.2-YFP* were observed under the confocal microscope.

In-vivo fluorescence was observed in the root tip of all three representative seedlings of each line with *promHREI:MA-HRE1-YFP* construct. Expression of *promHRE1:MA-HRE1-YFP* at nuclear region in tip of roots occurs as fluorescence can be seen at that region. Though fluorescence also appears in seedlings of each line with *promRAP2.2:MA-RAP2.2-YFP* (Fig 5.5) but the level is low comparison to Col-0 lines with *promHRE1:MA-HRE1-YFP* (Fig 5.4).

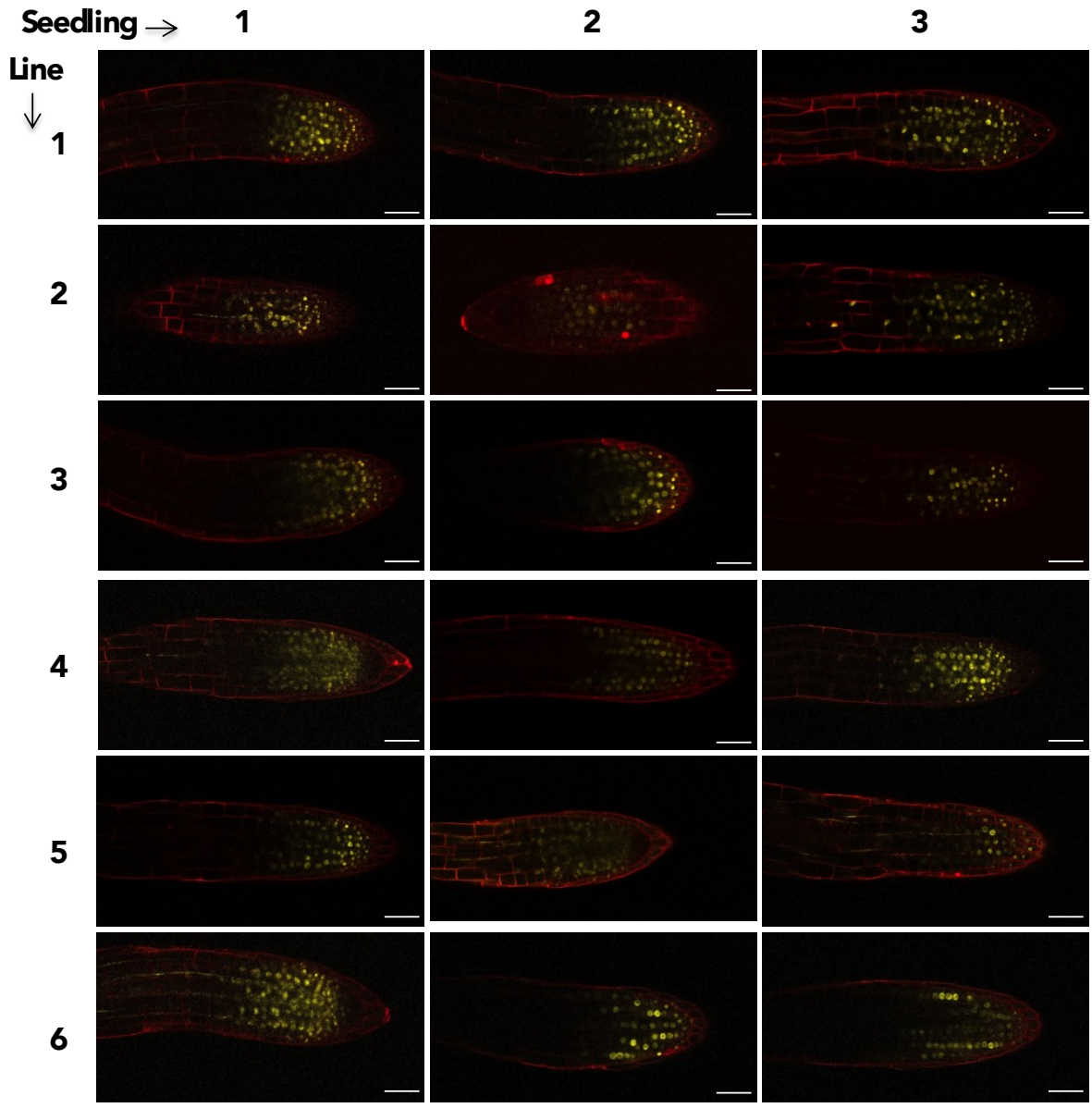


Fig:5.5 Three representative seedlings of six Col-0 lines showing expression of *promHRE1:MA-HRE1-YFP*under the confocal microscope.

3 days old etiolated Col-0 seedling (1-3) in six lines (1-6) carrying *promHRE1:MA-HRE1-YFP* in Col-0 background are showing yellow fluorescence in the nuclei of their root tip. Cell wall of each seedling appears in red due to propidium-iodide (1μg/ml).

20x object lenses and 488nm wavelength were used to obtain the images through confocal microscope. Scale bars = 50 μm.

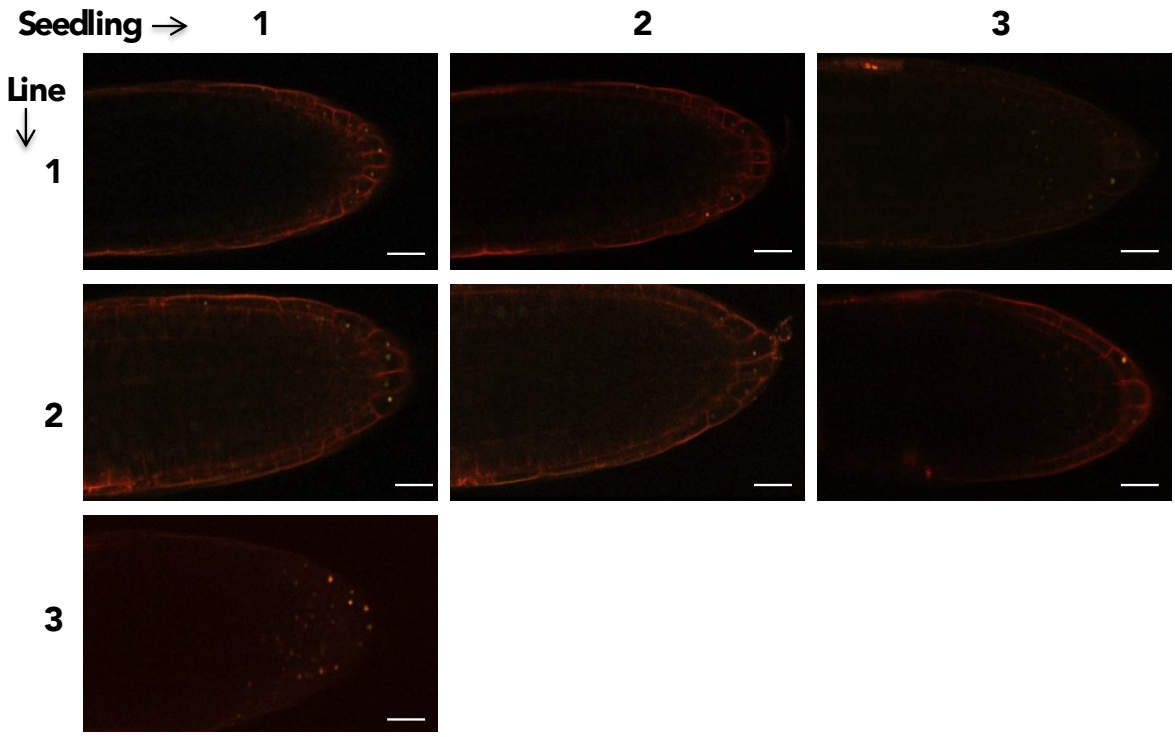


Fig:5.6 Three representative seedlings of two Col-0lines showing expression of *promRAP2.2:MA-RAP2.2-YFP* under the confocal microscope.

3 days old Col-0 seedling (1-3) in three lines (1-3) carrying *promMA-RAP2.2::YFP* are showing yellow fluorescence in the nuclei of their root tip. Cell wall of each seedling appears in red due to propidium iodide (1μg/ml).

20x object lenses and 488nm wavelength were used to obtain the images *via* confocal microscope. Scale bars = 50 μm.

One homozygous Col-0 line containing *promHRE1:MA-HRE1-YFP* and one *Col-0* line containing *promRAP2.2:MA-RAP2.2-YFP* were selected as homozygous line.

5.2.1.1.3 Analysis of *promERFVII's:MC-ERFVII's-YFP* in Col-0

F_2 seeds of Col-0 plants (result of crossing *prt6-1* containing *promHRE1:MC-HRE1-YFP* or *promRAP2.2:MC-RAP2.2-YFP* to Col-0) were plated on hygromycin (30ug/ml). 86 transgenic seedlings with *promHRE1:MC-HRE1-YFP* and 81 seedlings with *promRAP2.2:MC-RAP2.2-YFP* were selected and planted on soil. All the seedlings of

both lines were processed for genotyping to identify PRT6 wildtype and *prt6-1* mutant alleles.

	promHRE1:MC-HRE1-YFP	promRAP2.2:MC-RAP2.2-YFP
F_2	86	81
F_3	3	4

Table: 5.2 Summary of number of wild-type *PRT6* for *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* lines used for analysis

5.2.1.2 Analyzing protein stabilization and localization in different stress conditions

The homozygous F_3 seedlings containing *promHRE1:MC-HRE1-YFP* and *promRAP2.2-:MC-RAP2.2-YFP* in wild-type *PRT6* background were tested under different stress conditions such as: treatment with cPTIO (to provide low nitric oxide in cells) in order to identify any changes taking place during gene expression, stabilization and their localization. These transgenic lines were also tested for MG132 treatment. MG132 or geldanamycin (GDA) is a 26S proteasome inhibitor (Palombella *et al*., 1994). It is expected that the level of protein stability will increase by the treatment with MG132.

5.2.1.2.1 Fluorescent Microscopy analysis of *promMC-ERFVII's::YFP* in Col-0 treated with cPTIO

In order to see the effect of low nitric oxide condition on seedlings with *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* that are in Col-0 background, seedlings were treated with cPTIO, a chemical used to reduce level of nitric oxide in the tissue. It is expected that in a low nitric oxide condition the ERFVII's will be stable, as Cys2 cannot undergo oxidation.

The result of cPTIO treatment experiment did not show much difference between untreated and treated seedlings (Fig: 5.6; 5.7). I therefore analyse that the cPTIO treatment didn't work properly. The reason I concluded could be that in many papers it has been shown that cPTIO, may permit formation of N_2O_3 that leads to increase in the rate of NO oxidation (Arita *et al*., 2006). According to this equation: (NO + cPTIO → NO2 + cPTIO), which in turn reacts with NO to form N_2O_3: (NO_2 +NO→N_2O_3) (D'Alessandro *et al*., 2013).

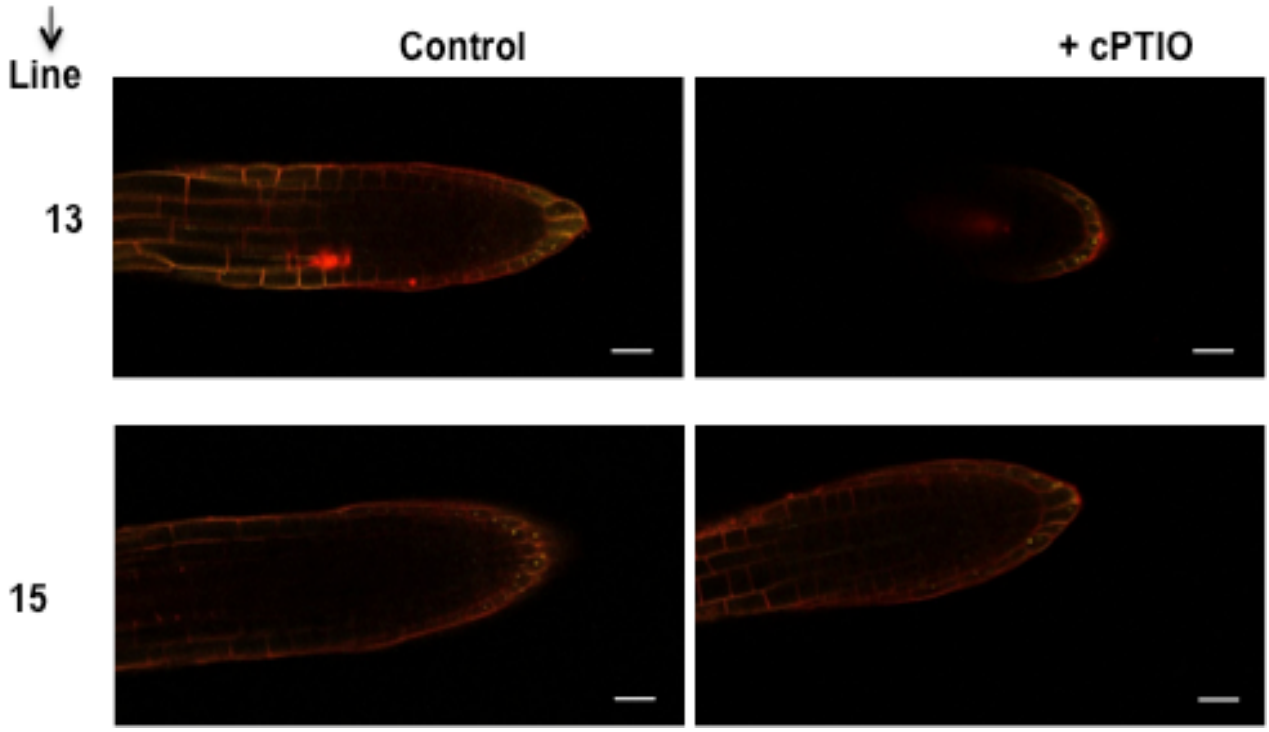


Fig: 5.7 Two representative seedlings of two Col-0 lines showing expression of p*romHRE1:MC-HRE1-YFP* in un-treated and cPTIO treated condition at root-tip region.

Four days old etiolated seedlings of Col-0 lines carrying *promHRE1:MC-HRE1 YFP* treated with cPTIO were compared with non-cPTIO treated seedlings on confocal microscope. Similar pattern of fluorescence at the root-tip area can be seen on both treated and non-treated seedlings of both lines. The cell wall of each seedling appears in red due to propidium iodide (1mg/ml). 20x object lenses and 488nm wavelength were used to obtain the images through confocal microscope.

Scale bars = 50 µm.

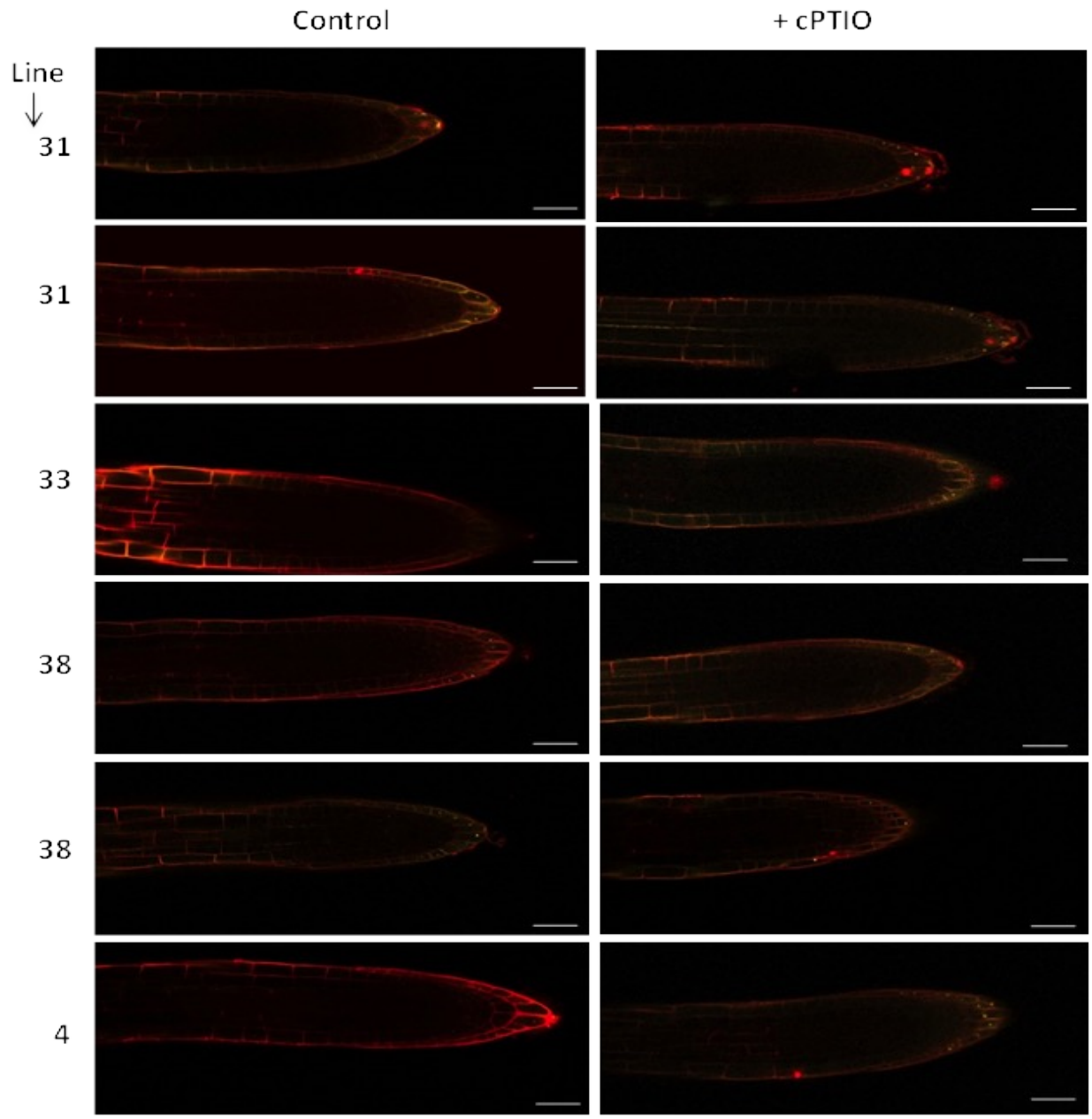


Fig: 5.8 Two representative seedlings of four Col-0 lines showing expression of *promMC-RAP2.2::YFP* in un-treated and treated cPTIO condition at root–tip region.

Comparison of 4 days old etiolated seedlings of Col-0 lines carrying *promRAP2.2:MC-RAP2.2-YFP* cPTIO treated and un-treated seedlings on confocal microscope. No difference in level of fluorescence can be seen at root-tip area on both treated and un-treated seedlings of all four lines.

The cell wall of each seedling appears in red due to propidium iodide (1μg/ml). 20x object lenses and 488nm wavelength were used to obtain the image throughconfocal microscope. Scale bars = 50 μm.

5.2.1.2.2 Analysis of effect of MG132 on *promERFVII's:MC-ERFVII's-YFP* in Col-0 through fluorescent microscopy

The proteasomal inhibitor MG132 treatment was performed on etiolated Col-0 seedlings with *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* to determine the differences in the level of fluorescent signal between treated and un-treated samples. Four days old seedlings of both lines were treated with MG132 for 5 hrs and counter-stained with propidium iodide (1μg/ml) and imaged using the confocal microscope. The levels of fluorescence were compared between MG132 and un-treated seedlings.

Results showed that treatment with MG132 caused a slight increase in fluorescence level at root-tip region in seedlings with *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* comparison to un-treated seedlings (Fig: 5.8; 5.9).

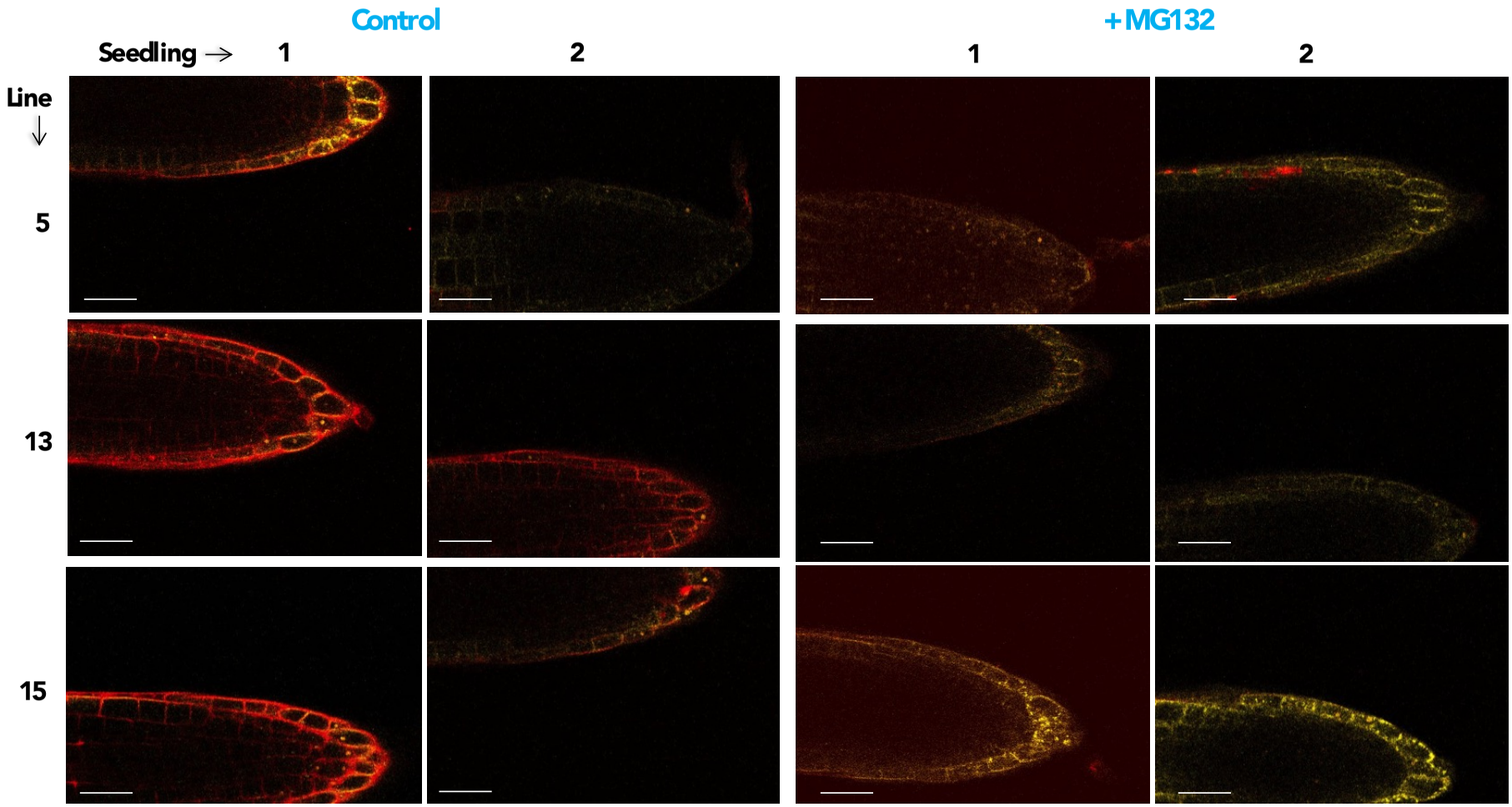


Fig: 5.9 Images of etiolated seedlings of Col-0 lines carrying *promMC-HRE1::YFP* examined for MG132 treatment.

MG132 treated and un-treated roots of Col-0 seedlings with *promMC-HRE1::YFP* showing fluorescence at nuclear region in root cells. MG132 treated seedlings appear showing a small enhanced yellow fluorescence compare to un-treated MG132. Scale bar= 50 µm

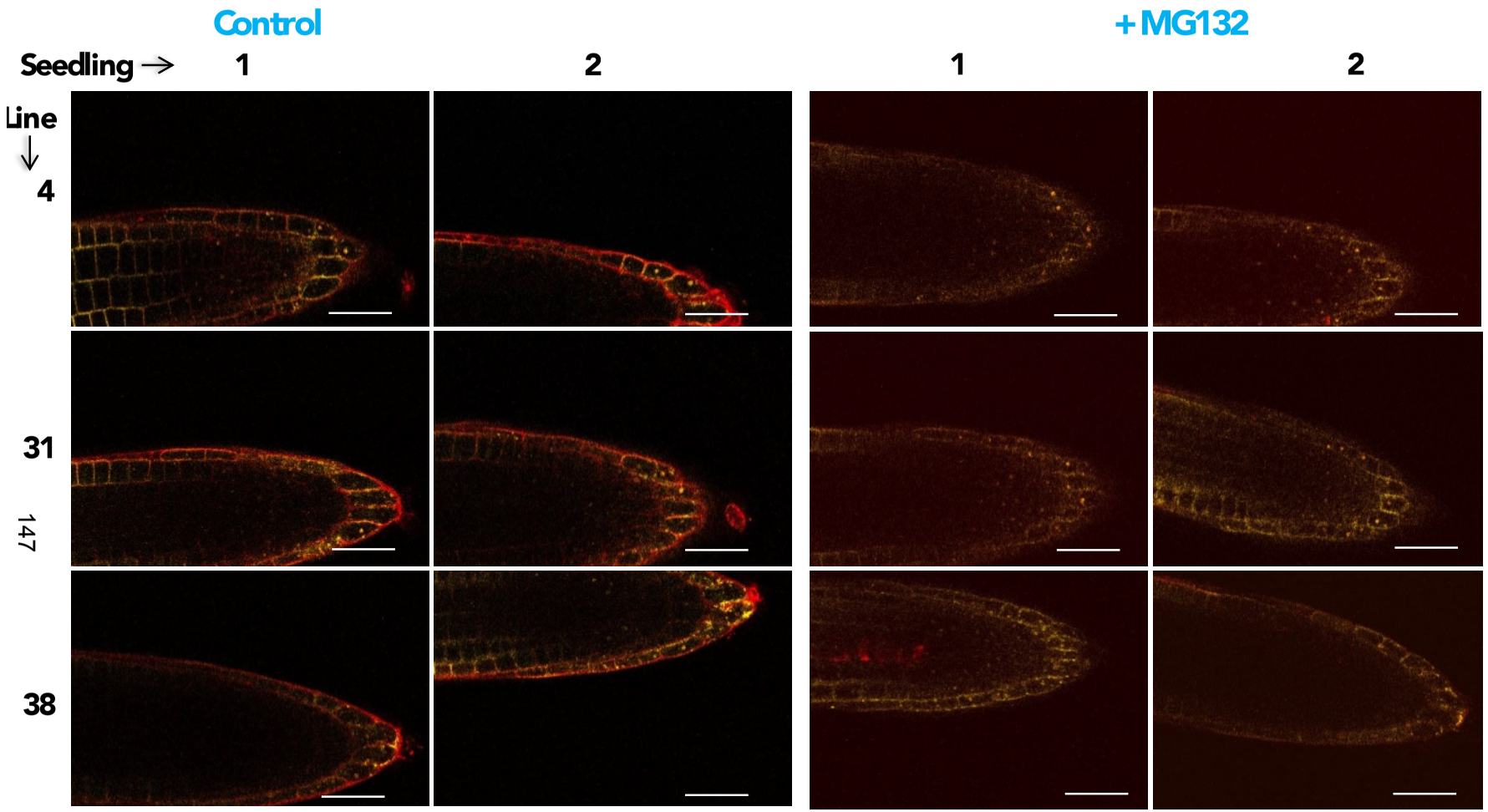


Fig: 5.10 Images of etiolated seedlings of Col-0 lines carrying *promMC-RAP2.2::YFP* examined for MG132 treatment.

MG132 treated and un-treated roots of Col-0 seedlings with *promMC-RAP2.2::YFP* showing fluorescence in root cells in nuclei. MG132 treated seedlings are showing more yellow fluorescence at small scale compare to un-treated MG132. Scale bar= 50μM

5.3 Discussion

Fluorescent-protein tagging has now become the widely used strategy for identification and imaging of protein stability in living cell. In this study, development of ERFVII's lines with fluorescent-protein tagging in *Arabidopsis* was carried out. The aim for this was to examine the localisation and expression of ERFVII's *in-vivo*. Analysis of proteins localization and stabilization was performed by visualizing the fluorescing proteins in the cells followed by imaging *via* confocal microscopy.

5.3.1 Subcellular localization of YFP tagged *promERFVII's:MC/MA-ERFVII's*

Determination of subcellular localization of *promERFVII's:MC-ERFVII's-YFP* in Col-0 and *prt6-1* and *promERFVII's:MA-ERFVII's-YFP* in Col-0 will help to understand the role of ERFVII's in regulation of gene. It is hypothesized that the stabilized ERFVII's regulate gene regulatory networks associated with anaerobic metabolism in *Arabidopsis* (Licausi *et al* in 2011; Gibbs *et al*., 2014). In the study by Licausi, it was found that RAP2.12-GFP tagged protein that was localized at plasma membrane in normoxic condition translocate to nucleus during hypoxia. Since, GFP accumulation in hypoxia appeared to be residing into the nucleus (Fig: 5.1). And in a recent study also it was found that RAP2.3 (35S:YFP-RAP2.3) in etiolated seedlings under hypoxia moves to nucleus form cytoplasm (Abbas *et al*., 2015). However, under normoxia and in light the stabilized RAP2.3 gets degraded. Further, *HRE1* and *HRE2* GFP-tagged constructs have also been seen localized in the nucleus (Licausi et al., 2010)

Likewise, it is expected that *promERFVII's:MC/MA-ERFVII's-YFP* proteinmight also show similar behavior. Therefore, investigation of ERFVII's expression and stability site under different stress conditions was conducted.

YFP is expressed under the control of *promERFVII's:MC/MA-ERFVII's-YFP* during the early stages of plant growth.The expression and stabilization of *promERFVII's:MC/MA-ERFVII's-YFP* was analysed in roots of 3-days old etiolated seedlings. During analysis fluorescence was seen greatest in the nuclear region suggesting that they are transcription factors (Fig: 5.3; 5.4; 5.3; 5.5; 5.6).

5.3.1.1 Investigation of expression patterns of *promERFVII's:MC/MA-ERFVII's YFP* individually in Col-0 and *prt6-1* root cells and under various abiotic stresses

In-vivo analysis of *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* in Col-0 (backcrossing of *promMC-ERFVII's::YFP* in *prt6-1* to Col-0 was performed to confirm the same transgenic event in both genotypes) and *prt6-1* (proteins are expected to be stable due to lack of E3 ligase-PRT6) etiolated seedlings was performed under the confocal microscope. HRE1-YFP and RAP2.2-YFP were seen stable in the nucleus of root cells of both Col-0 and *prt6-1* (Fig: 5.1; 5.2; 5.5; 5.6). However, proteins appear more stable in *prt6-1* background comparison to Col-0. Further, strong fluorescence of stable *promHRE1:MA-HRE1-YFP* and *promRAP2.2:MA-RAP2.2-YFP* in Col-0 was also seen in nucleus (Fig: 5.3; 5.4). There was no fluorescence seen at plasma membrane site. In a recent study, etiolated seedlings (seedlings kept in dark for continuous three days) under hypoxic conditions have seen showing enhanced stability of an ERFVII's. Further, etiolated seedlings under hypoxia have shown sub-cellular localization of RAP2.3-YFP in nucleus (Abbas *et al*., 2015). This suggests that *promERFVII's:MC/MA-ERFVII's-YFP* in etiolated conditions have moved ERFVII's to nucleus from cytoplasm. Since, etiolated state is acting as an abiotic stress factor that leads to movement of ERFVII's to nuclear region from cytoplasm.

Another study was done to analyse the effect of exogenous low nitric oxide and proteasome inhibitor-MG132 for stabilization of proteins. For that purpose I used *promERFVII's:MC-ERFsVII's-YFP* constructs in Col-0. Effect of these conditions was analysed on the basis of fluorescence intensity between treated and untreated seedlings.

Although protein was present in nucleus but there was no difference in the level of fluorescence between treated and untreated seedlings (Fig: 5.5; 5.6). On the other hand, the MG132 (Carbobenzoxy-Leu-Leu-leucinal, a specific 26S proteasome inhibitor) led to stabilization of proteins in nuclei. Difference in fluorescence between treated and untreated was visible (Fig: 5.7; 5.8).

These results obtained provide information of stabilization of HRE1 and RAP2.2 proteins during abiotic stress conditions in root tips. This study showed localization of these ERFVII's in nucleus as seen by Kosmacz *et al*., 2014; Licausi *et al* 2015. Hence, this

stabilization of proteins indicates its important role in root development during early stage of plants growth regulated by N-end rule pathway.

5.3.1.2 Dynamics of stability and destability of RAP2.12-GFP

A recent study by Licausi *et al*., 2015 and Kosmacz *et al*., 2014 have shown oxygen-dependent stability of the RAP2.12 and its degradation through the N-end rule pathway (Gibbs *et al*. 2011; Kosmacz *et al*., 2014; Licausi *et al*. 2015). The transcription factor RAP2.12 is one of the member of ERFVII's (Nakano *et al*. 2006). It's role was seen in activation of anaerobic response in hypoxia (Licausi *et al*., 2011). All ERFVII's posses a highly conserved N-terminal Cys that is oxidised in presence of oxygen and nitric oxide and further pursue arginyle branch of N-end rule pathway for degradation (Gibbs *et al*., 2014). So, in normoxia the cysteine residue present at the N-terminus gets oxidised with the help of plant cystein oxidases (PCOs). This led to arginylation and then recognition by E3-ligase PRT6 and degradation *via* the 26S proteasome pathway (Varshavsky *et al*., 2011; Graciet *et al*. 2010; Weits *et al*. 2014).

A study was performed for investigation role of ERFVII's in activation of hypoxia responsive genes. This study demontrated that RAP2.12 protein is protected from degradation in normal conditions through plasma membrane, where it binds to ACYL-CoA BINDING PROTEIN 1 and 2 (ACBP1 and ACBP2) (Licausi *et al*. 2011). However, studies with RAP2.12-GFP shown RAP2.12 shifting from plasma membrane to nucleus in hypoxia where it promotes activation of hypoxic responsive genes (Licausi *et al*., 2011) (Fig:5.10).

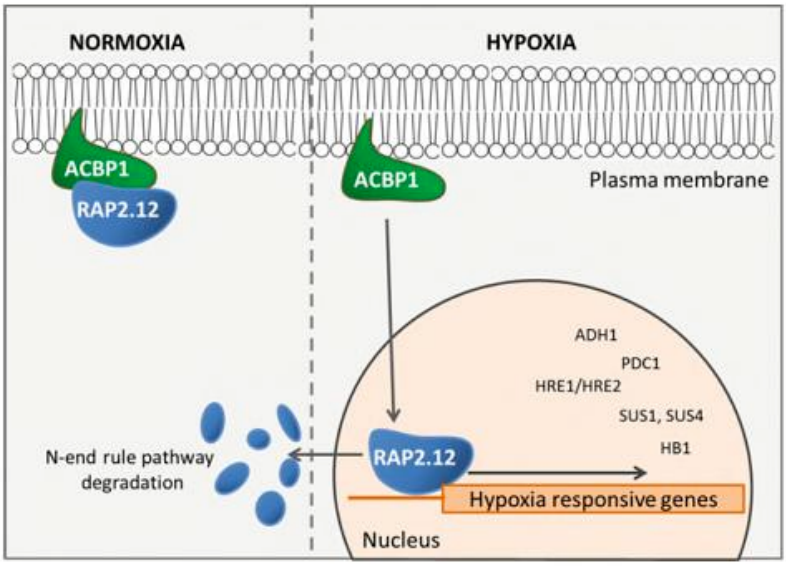


Fig: 5.11 Figure showing the movement of RAP2.12 in hypoxia from the plasma membrane to nucleus

In normoxia, RAP2.12-GFP remains attached to plasma membrane bound with ACBP1. In hypoxia, RAP2.12-GFP detaches with ACBP1 and moves to nucleus. In nucleus it activates hypoxia responsive genes. When re-oxygenated, RAP2.12 gets degraded through N-end rule pathway. (Figure is taken from Licausi *et al*., 2011)

Recently, a new study has also shown the oxygen concentration requirement and time needed for *RAP2.12* localization. In this study RAP2.12-GFP in *Arabidopsis* plant was exposed to 1% oxygen (v/v) for different time periods (0, 30, 60, 120, 180, and 240 min) and the localization of RAP2.12-GFP through confocal laser scanning microscopy. As well as, analysis of expression level of hypoxia marker genes {*ALCOHOL DEHYDROGENASE1 (ADH1)*, *PYRUVATE DECARBOXYLASE1 (PDC1)*, *HEMOGLOBIN1 (HB1)*, *SUCROSE SYNTHASE 1(SUS1)*, *LOB DOMAIN PROTEIN 41 (LBD41)*, *PLANTCYSTEINE OXYDASE 1 (PCO1)* and *2 (PCO2)*, and the *HYPOXIA-RESPONSIVE UNKNOWN PROTEIN (HUP9)}* was also carried simultaneously. It was seen that RAP2.12-GFP that was present at the plasma membrane migrates to nucleus after 3 hrs. However, induction of the hypoxia marker genes was observed much earlier since their expression was seen within 30 min after treatment. It was concluded that stabilization of RAP2.12 requires greater time than induction of hypoxia genes. Subsequently, study on the oxygen concentration for RAP2.12:GFP stability in the nucleus in response to hypoxia was also carried on different concentrations of oxygen (0,

1, 3, 5, 10, 12.5, 15 and 21% O_2 (v/v) in air). It was found that at 10% oxygen and below it, RAP2.12 translocate from membrane to nucleus. While above 10% oxygen RAP2.12 protein was still attached to membrane. Whereas, induction and expression of hypoxic genes was observed significantly at 1% oxygen in air.

Further investigation on RAP2.12-GFP also revealed that *de-novo* RAP2.12 protein occurs that translocate to nucleus during hypoxia. And after re-oxygenation, RAP2.12:GFP protein was observed to disappear from the nucleus and the expression of low oxygen genes decreased to normoxic levels (Licausi *et al*. 2011b). In order to investigate the time required to degrade nuclear RAP2.12, plants expressing RAP2.12:GFP were first treated for 4 h with 1% oxygen and subsequently exposed to a normoxic atmosphere. At regular time intervals during 8 h of re-oxygenation, the nuclear localization of the fluorescent signal and the expression levels of the hypoxia responsive genes were measured. Between 3 and 4 h of re-oxygenation, RAP2.12-GFP disappeared from the nucleus (Fig. 4a,b). Gene expression analysis showed that already after 15 min of re-oxygenation, the expression of the hypoxia responsive genes declined and progressively reached normoxic levels after 4 h of re-oxygenation.

Thus, the *promERFVII's:MC/MA-ERFVII's-YFP* constructs can be used to study the role of ACBP1/ACBP2. Since, studies of Licausi *et al.* have indicated that ACBPs might be involved in attachment of ERFVII's to the plasma membrane. Still it is unclear hence, studies in area using ERFVII's-YFP constructs can be helpful in future. Further, investigation of gene network regulated by ERFVII's binding to *cis*-regulatory elements 'GCCGCC' in low oxygen and low nitric oxide could be performed.

5.3.1.3 Stability of ERFVII's regulate plant growth and development

In past years several studies have been done to see the effect of stability of proteins on plants behaviour. It has already shown that ERFVII's act as hypoxia and low nitric oxide sensors (Gibbs et al., 2011; Gibbs et al, 2015). ERFVII's get stabilized under these conditions and help plants survival through activating anaerobic genes in nuclei. Plants recovered after survival show better development. A recent study has also shown that etiolated seedlings under hypoxic conditions have enhanced stability of ERFVII's, which leads to survival of seedlings and establishment (Abbas *et al*., 2015).

CHAPTER-6

General Discussion

6. GENERAL DISCUSSION

The first aim of my work was to identify and analyse genes that are similar to *ABI5* in a manner showing 'GCCGCC' *cis*-elements present in their promoter or 5' UTR region. Previous studies on *ABI5* have shown that it is a direct downstream target of the N-end rule pathway regulated by binding of RAP2.3 to 'GCCGCC' site in 5'UTR region and has important functions in ABA signaling (Gibbs *et al.,* 2014). Therefore, identification of other putative downstream targets of N-end rule pathway can help to find novel regulators of ABA signaling and their role in response to different stresses.

Initially, I used bioinformatics approaches to find genes that have two 'GCCGCC' sites in their promoter or 5' UTR region. I selected 14 such genes and processed them for promoter amplification. Finally promoter PCR amplification of five genes showed amplification. I cloned each promoter using the Gateway system in wild-type and *prt6-1* plants and later analyzed the promoter activity of all promoters through GUS staining. Histochemical analysis showed enhanced promoter activity in *prt6-1* compared with Col-0. This result gave a hint that identified genes might be regulated *via* the N-end rule pathway. I further performed q-RT-PCR in order to test the expression of these genes at the level of RNA and to identify the significant difference between two genotypes for each gene.

The second part of my work involved the use of the ChIP technique, followed by qPCR, to search for binding of ERFVII, *RAP2.3* to 'GCCGCC' sites of identified genes *in-vivo*. I was able to detect *ABI5* and *PYL7* showing high enrichment in HA-IPs. This indicates that binding of *RAP2.3* at 'GCCGCC' sites regulates *ABI5* and *PYL7*. These results support my initial hypothesis that genes whose promoters were cloned follow N-end rule pathway regulation. The third part of my work was based on cell-biology. Here I looked for stability and expression of two of the ERFVII, *HRE1* and *RAP2.2* proteins through yellow fluorescent technique *in-vivo*.

6.1 Characterization of genes having two 'GCCGCC' *cis*-elements in promoter or 5'UTR

6.1.1 The common features among four identified genes

In Chapter 3, I identified four genes (*PYL, ERD4, AT3G13440* and *AT1G14810*). Firstly,

these genes have a similar feature that all four genes bear two consensus GCCGCC *cis*-elements in the promoter or in 5'UTR and are the sites for ERFVII's binding (Yang *et al*., 2009; Gibbs *et al*., 2014). Secondly, promoter analysis of the four genes in flowers and leaves of wild-type and *prt6-1* have shown enhanced promoter activity in *prt6-1* tissue compare to Col-0. GUS activity for all four genes was higher in *prt6-1* tissue in comparison to wild type. However, the regions of GUS expression were different within the same tissues for each promoter. So, *promPYL7:GUS* was seen showing expression in stigma and sepal regions of the flower while *promERD4:GUS* showed expression at pistil and filament regions of flowers. *promAT1G14810::GUS* has expressed in basal part of the stigma and *promAT3G13440:GUS* in the stigma, anthers and petals. The reason behind expression at different regions in similar tissue is discussed in Section: 3.3.1.2. Thirdly, the GUS analysis of transformed promoters suggests that all four genes have a role in flower and leaves development.

6.1.2 The role of ERFVII's in mediating a specific function through regulation of multiple genes

ERFVII's are involved in a variety of protein and gene interaction complexes to mediate diverse developmental processes, stress responses and hormonal signaling such as, gibberellin and abscisic acid. Many such responses are facilitated by ERFVII's by binding to GCCGCC sites (Buttner and Singh, 1997; Zhang *et al*., 2004; Gibbs *et al*., 2014; Gibbs *et al*., 2015) suggesting that they have multiple gene targets. Furthermore, the posttranslational phosphorylation by kinases might also affect ERFVII's activity (Cheong *et al*., 2003; Xu *et al*., 2007). These report, therefore indicate that ERFVII's regulate diverse signal and context specific responses.

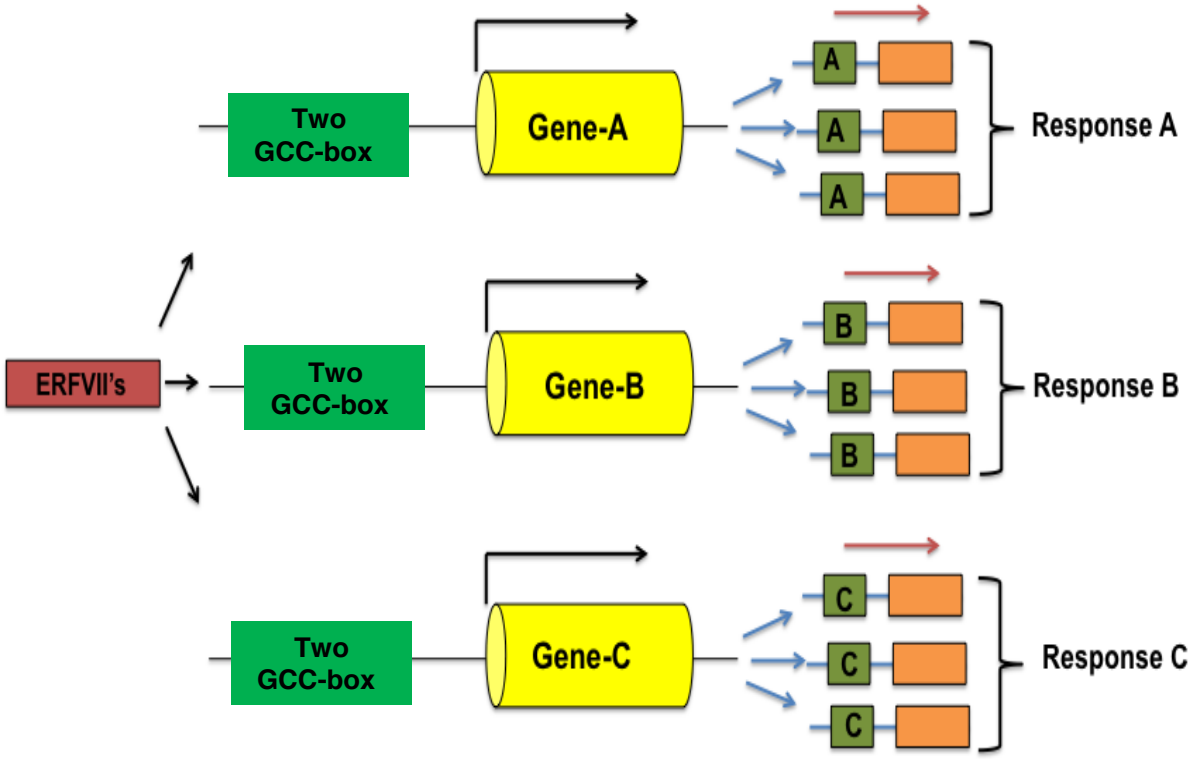


Fig: 6.1 Hypothetical model showing regulation of different responses as a result of multi gene activity *via* binding of ERFVII's to GCCGCC site.
This model represents the binding of ERFVII's at GCCGCC site of genes to activate secondary targets and trigger conditional responses.
(Model is based on Jung *et al*., 2010)

Different studies have shown the role of ERFVII's in regulating many functions in collaboration with other genes, signals or gaseous-sensing linked mechanism. For example: 1. Recently, it was demonstrated that N-end rule pathway regulates seed germination through *ABI5.* ERFVII's bind to the 'GCCGCC' *cis*-elements in the 5'UTR region of the *ABI5* promoter. Genetic study on *prt6-1abi5-8* suggests that accumulated ERFVII's might be responsible for the expression of other factors that inhibit germination, since there is no *ABI5* protein in *prt6-1abi5-8* double mutants. Hence, ERFVII's are stable and may activate other genes that inhibit germination. There are also other genes in *Arabidopsis* that contain the 'GCCGCC' *cis*-element in their 5'UTR that may also be N-end rule pathway targets. My work has identified such genes that are up-regulated in *prt6-1.* Thus, these candidate genes might also be involved together with *ABI5* for promotion of seed germination.

2. ERFVII's control seed germination by regulating the expression of *ABI5* through nitric oxide sensing (Gibbs *et al*., 2014). *ABI5* is a basic lucine zipper transcription factor that controls seed germination. In the presence of ABA, *ABI5* accumulates and vice–versa (Lopez-Molina *et al*., 2001). It was found that *ABI5* is expressed in the endosperm of seeds. The expression of *ABI5* in the endosperm was abolished after the addition of the nitric oxide donor SNP indicating that stable substrates induce *ABI5* expression in untreated controls. As the *ABI5* promoter contains binding sites for *RAP2.3*, therefore ERFVII's play a role during seed germination depending on action of the gaseous molecule, nitric oxide.

3. A major role for ERFVII's was also seen in controlling flooding and hypoxia tolerance in plants. It was seen that *hre1hre2* mutant seedlings showed highly reduced survival in anoxia (0% oxygen). However, ectopic overexpression of *HRE1* enhanced survival by increasing the expression of core hypoxia-responsive genes such as, *ALCOHOL DEHYDROGENASE1 (ADH1)* (Licausi *et al*., 2010; Hess *et al*., 2011). Similarly ectopic expression of RAP2.2 increased ADH1 transcripts in hypoxia (5% oxygen), while *rap2.2* mutants showed reduced survival under this stress (Hinz *et al*., 2010). Other studies in *Arabidopsis* have also pointed to key roles for ERFVII's in controlling hypoxia and ethylene-regulated submergence responses (Papdi *et al*., 2008; Yang *et al*., 2011). A recent study has shown that stabilized ERFVII's in hypoxia can actively maintain a closed apical hook, and repress chlorophyll biosynthesis and cotyledon greening (Abbas *et al*., 2015). This demonstrates an active role for ERFVII's through oxygen sensing in protecting them from photo-oxidative damage following extended darkness and enhancing survival rates and regulating general plant growth and development.

6.1.3 Activation of hypoxia responsive genes is induced by ERFVII's

Plants rely on molecular oxygen for respiratory energy production. During hypoxic conditions such as, flooding, during developmental processes or seed germination plants undergo a hypoxic response. This involves a modulation of gene expression leading to various biochemicals, physiological, and morphological changes that help plants to survive in that condition. ERFVII's were identified as direct sensors for hypoxia. In *Arabidopsis RAP2.12*, *RAP2.2*, and *RAP2.3* are all constitutively expressed, whereas *HRE1* and *HRE2* are hypoxia inducible. This suggests that there is a cascade of transcription and stabilization in response to low oxygen and those individual ERFVII's

has different contributions to the response (Licausi *et al*., 2010; Bui *et al*., 2015). This suggests that ERFVII's induce secondary targets *i.e*., hypoxia responsive genes in response to hypoxia. It is hypothesized that ERFVII's bind to downstream targets (genes that have two GCCGCC sites like: *ABI5, PYL7, ERD4*). These targets might in turn be involved directly or indirectly in activation of stress tolerance genes (Fig: 6.1).

6.2 Use of ChIP technique to determine *in-vivo* binding of ERFVII's to 'GCCGCC' *cis*-elements targets

6.2.1 ChIP technique for identification of other ERFVII's interaction with genes containing 'GCCGCC' *cis*-elements

Chromatin immuno-precipitation (ChIP) is one of the methods of studying genes that are targets of transcription factors. Chapter 4 of this thesis described this approach for identification of direct downstream targets. In 2014, Gibbs *et al.* showed using ChIP that two GCC-boxes (GCCGCC) are present in the 5'UTR region of *ABI5* that is controlled by substrates of the N-end rule pathway. I further investigated the binding of RAP2.3, one of the ERFVII to *ABI5* and *PYL7*. Both these genes have two copies of the GCC-box in the 5´UTR (Chapter 3).

The experiment was performed on 4-day-old untreated etiolated seedlings of 35S-MA-RAP2.3 lines. The results presented in Chapter 4 showed that the stabilized MA-RAP2.3 (Fig: 4.6, 4.7) was able to bind specifically to the promoter region of *ABI5* and *PYL7* containing the two GCC-boxes.

With respect to preliminary ChIP experiments, analysis of other ERFVII's (RAP2.2, RAP2.12, HRE1 and HRE2) for their binding to *ABI5* and *PYL7* will be interesting. This future experiment will confirm whether all MC-ERFVII's can specifically bind to GCCGCC, the *cis*-element present in the 5'UTR or promoter of genes. Together with this, study of interaction of all five MC-ERFVII's on four genes (*PYL, ERD4, AT3G13440* and *AT1G14810*) can be done. These studies will be helpful in establishing a link between N-end rule pathway and the regulation of downstream genes through ERFVII's.

The ChIP carried on 4-day-old hypoxia treated etiolated seedlings with 35S-MA-RAP2.3 showed enhanced promoter activity due to stabilized MC-RAP2.3 in comparison to

untreated 35S-MC-RAP2.3 (Fig: 4.5). This result further suggests that ChIP analysis can be accomplished on selected genes to examine the function of stability of N-end rule pathway substrates under different stress conditions.

6.2.2 Other ChIP related approaches to determine DNA-protein interaction

In addition to chromatin immunoprecipitation (ChIP) assays there are two other most widely used approaches for genome-wide identification and characterization of *in-vivo* protein-DNA interactions. The ChIP sequencing (ChIP-Seq) that provides higher spatial resolution, dynamic range, and genomic coverage, allowing it to have higher sensitivity and specificity (Goren *et al*., 2010) and microarray hybridization (ChIP-chip) (Barski *et al*, 2007; Valouev *et al*, 2008; Park *et al*., 2009). These two techniques can be further used to analyze other ERFVII interaction with DNA during the biological processes and in response to the environmental stresses.

6.3 Analysis of ERFVII's expression and localization using Fluorescent Protein technique

Regulation of gene expression is an important component of living organisms. Transcription factors are proteins that can regulate gene expression through protein-DNA interactions (Stracke *et al*. 2001). In a cell, mostly transcription factors are localized to either the cytosol or the nucleus (Hoppe *et al*. 2001).

6.3.1 *In-vivo* analysis of *promERFVII:MA/MC-ERFVII-YFP* stabilityin root of etiolated seedlings

In order to investigate the location of expression and stability of ERFVII's in *Arabidopsis* tissue, I used the yellow fluorescent protein. I used Col-0 and *prt6-1* transgenic lines with ERFVII-YFP constructs driven by their own promoters. The region of ERFVII expression and stability was analysed in roots of 4-day old etiolated seedlings containing *promERFVII:MA/MC-ERFVII-YFP*. The analysis of *HRE1* and *RAP2.2* expression and their stability was performed through *promHRE1:MA-HRE1-YFP*, and *promRAP2.2:MA-RAP2.2-YFP* in wild-type and *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* in *prt6-1* where proteins will be stable. *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP* in Col-0 were also subjected for conditional stability of

proteins. Such as under low nitric oxide condition and MG132 treatment.

Results showed that HRE1 and RAP2.2 were localized and expressed in the nucleus of root tip of etiolated seedlings. However, stability appeared higher in *prt6-1* in comparison to wild-type on the basis of intensity of fluorescence of protein. Further, low nitric oxide treatment did not result in any difference in the level of fluorescence between treated and untreated wild type seedlings of *promHRE1:MC-HRE1-YFP* and *promRAP2.2:MC-RAP2.2-YFP.* MG132 treatment however was shown to have an effect on MC-HRE1-YFP and MC-RAP2.2-YFP in root tips. This result indicates that degradation of HRE1 and RAP2.2 is controlled by Arg/N-end rule pathway through 26S proteasome system. In addition, protein expression in the nuclear region of root signifies their root development associated function in early stages of plant development.

6.3.2 The stability and localization of ERFVII's are dynamically regulated by oxygen

ERFVII's are constitutively degraded in the presence of both oxygen and nitric oxide but get stabilized under either low oxygen or low nitric oxide. Therefore, ERFVII's are now considered as oxygen and nitric oxide sensors in response to hypoxia, submergence and low nitric oxide states in plants (Gibbs *et al* 2011; Licausi *et al*., 2011, Gibbs *et al*., 2014). A recent study by Licausi *et al* in 2015 found that *RAP2.12* is membrane localized and re-localizes in the nucleus upon hypoxia. *promRAP2.12:GFP* was restricted in the plasma membrane through binding with ACYL-CoA BINDING PROTEIN 1 and 2 (ACBP1 and ACBP2) (Licausi *et al*. 2011). However, in hypoxia GFP-tagged *RAP2.12* was shown to move its localization from the plasma membrane to the nucleus. There it activates the hypoxic responsive genes (Licausi *et al*. 2011) that make plants survive in low oxygen condition. Whereas, on re-oxygenation RAP2.12 is rapidly degraded *via* the N-end rule pathway and proteasome-mediated proteolysis to inhibit the hypoxic response.

Recently it was recognized that 4 hrs of 1% (v/v) oxygen in air promoted RAP2.12:GFP relocalization into the nucleus. After re-oxygenation, degradation of RAP2.12 in the nucleus was observed within 3 hrs, resulting in reduction in hypoxia responsive gene transcripts to normoxic levels (Kosmacz *et al*., 2014). It was also identified that nuclear localization of RAP2.12 may not only depend on re-localization of existing protein, but involves de novo synthesis of the transcription factor as well (Kosmacz *et al*., 2014).
These results highlight a few questions:

1. Do all ERFVII get stabilized under hypoxia and induce hypoxia responsive genes that help plants to survive in such conditions?
2. Whether localization of ERFVII is co-ordinated by other factors?
3. Is there difference in time for stability of all five ERFVIIs during oxygen depletion and de-stability in re-oxygenated conditions?

Firstly, it is known that RAP2.12 that is localized to the plasma membrane through interaction with acyl-CoA binding proteins (ACBP1 & 2) shows a localization shift from the plasma membrane to the nucleus in hypoxia and triggers induction of hypoxia responsive gene expression (Kosmacz *et al*., 2014; Licausi *et al*., 2015). RAP2.3 was also identified as a direct interaction partner of ACBP2 (Li and Chye, 2004). Therefore, the plasma membrane localization of RAP2.12 is considered as a key component of the hypoxia-sensing mechanism. This could be used as a basis to study behavior of the other members of the ERFVII family for interactions with soluble ACBP proteins and their stability and destability in response to oxygen level. What role does nitric oxide play and is this mechanism conserved across species?

Secondly, accumulation of ERFVII's in the nucleus, switches an essential anaerobic response genes (Mustroph *et al*., 2009). However, expression of many of these genes is regulated by counterbalance of ERFVII's in co-ordination with other transcription factors. For example, recent data provide information that light is a factor responsible for degradation of protein. It was seen that an N-terminal RAP2.3-YFP fusion showed degradation in response to light (Abbas *et al*., 2015). In another study, the trihelix transcription factor *HRA1* plays such a role together with RAP2.12 (Giuntoli *et al*., 2014; Fig. 3B). *HRA1* is both induced by RAP2.12 and counteracts anaerobic gene induction by interacting with RAP2.12. Importantly, *HRA1* did not interact with any of the other four *Arabidopsis* ERFVIIs, thus indicating that each family member may have specific interaction partners (Giuntoli et al., 2014). These studies further indicate the need to search for other factors associated with each member of the ERFVII family for their role in hypoxia-sensing and nitric oxide sensitivity. Additionally, RAP2.3 was shown to associate with OCTOPINE SYNTHASE GENE ELEMENT BINDING FACTOR4 (Buttner and Singh, 1997), and *RAP2.2* with SEVEN IN ABSENTIA OF ARABIDOPSIS 2 (SINAT2; Welsch *et al*., 2007). SINAT2 is a REALLY INTERESTING NEW GENE E3 ligase, and a recent study showed that GFP-RAP2.12, in which the N-degron is removed, was stabilized in SINAT1/2-silenced *Arabidopsis* lines (Papdi *et al*., 2015).

Thirdly, for the question of stability and destability of ERFs and how much time does each member of ERFVII take for this process still needs to be researched. However, it is known that RAP2.12 was degraded within 3 hrs, resulting in down-regulation of hypoxia-adaptive gene expression (Kosmacz *et al*., 2014, Licausi *et al*., 2015). As well as, Licausi *et al*, 2015 found that RAP2.2, RAP2.3 and RAP2.12 tested on a set of hypoxia-responsive promoters were the most powerful activators. So, in future, similar studies like RAP2.12 for these 3 RAPs could be performed since these 3 RAPs appear as potential ERFVII's that could trigger hypoxia-sensing mechanism.

6.4 Prospective future work

1.) The four genes that I identified in my PhD work seem to be regulated by ERFVII's through binding at 'GCCGCC' site of their promoters. It is still unknown whether combined action of all five ERFVII's is required for regulation of the identified genes or each gene regulation is restricted to particular ERFVII's binding. Therefore, an investigation of each promoter activity of above identified genes in different combinations of the five ERF mutants and *prt6-1* will help to get information regarding gene regulation. As it is hypothesized that in *erfVIIprt6-1* mutant there would be no promoter activity due to lack of the ERFVII's, hence promoter activity of these genes will be inhibited. While combinations of one or more *erfVII* with *prt6-1* can define which ERFVII is responsible for that particular genes regulation.

2.) Study for new substrates of N-end rule pathway that can regulate *PYL7, ERD4, AT1G14810* and *AT3G13440* genes. There is a possibility that there are other transcription factors such as, newly identified N-end rule pathway substrates (BBX and bHLH) that can regulate these genes. Or there can be other transcription factors that could be potential ERFVII for regulation of these genes by specifically binding to the 'GCCGCC' site.

3.) Further study can be done to look at the role of identified genes that are regulated by ERFVII's in response to drought, waterlogging and biotic stresses. These studies could be performed using single or multiple combinations of mutants of these genes in *prt6-1* background.

4.) Electrophoretic mobility shift assay (EMSA) technique can be used to investigate the *in-vitro* binding of ERFVII's to the 'GCCGCC' site of all four identified genes (*PYL7, ERD4, AT1G14810* and *AT3G13440*). This will help to analyse the binding specificity of each of the MC-ERF to the 5' UTR or promoter of these genes.

5.) *promERFVII's:MC/MA-ERFVII's-YFP* transgenic lines could be used to see the protein location and expression in hypoxia and the time duration of stability and destability of ERFVII's following a similar approach to that used by Licausi *et al*., 2015. As well as, association of ACBPs in binding of ERFVII's to plasma membrane could also be done.

6.5 Conclusion

Work carried out in this thesis demonstrates new findings that are related with the N-end rule pathway. The promoter analysis studies showed new downstream targets of N-end rule pathway that are regulated through ERFVII's. These targets could use as potential targets for further study like *ABI5*.

Further, the ChIP analysis verified the binding of RAP2.3to the promoters of *ABI5* and *PYL7*. And the MC/MA-ERFVII's:YFP studies helped in understanding the dynamics of protein expression and its location *in-vivo*. Overall, the outcomes from this thesis have provided new information in relation to the N-end rule pathway. These experiments can help to define the role of evolutionary conservation of plant N-end rule pathway and its potential use in stress tolerance and crop improvement.

REFERENCES

7. REFERENCES

Abbas, M., D. Mohamad Abbas, Sophie Berckhan, Daniel J. Rooney, Daniel J. Gibbs, Jorge Vicente Conde, Cristina Sousa Correia, George W. Bassel, Nora Marin-dela Rosa, Jose Leo and Michael J. Holdsworth. Oxygen Sensing Coordinates Photomorphogenesis to Facilitate Seedling Survival, (2015). Current Biology 1483-1488.

Aloni R, Wolf A, Feigenbaum P, Avni A, Klee HJ. The never ripe mutant provides evidence that tumor-induced ethylene controls the morphogenesis of *agrobacterium tumefaciens*-induced crown galls on tomato stems, (1998). Plant Physiology, 117:841-849.

An *et al*, Ethylene-Induced Stabilization of ETHYLENE INSENSITIVE3 and EIN3-LIKE1 is mediated by proteasomal degradation of EIN3 binding F-Box 1 and 2 that requires EIN2 in *Arabidopsis* (2010). The Plant Cell, Vol. 22: 2384–2401.

Ang L-H. *et al.* Molecular interaction between COP1 and HY5 defines a regulatory switch for light control of *Arabidopsis* development (1998). Molecular Cell1:213–222.

Aparicio *et al*., Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in-vivo (2004). Current Protocol Cell Biology. Chapter 17, unit 17.7.

A. Pellegrineschi, *et al*., Stress-induced expression in wheat of the *Arabidopsis thaliana* DREB1A gene delays water stress symptoms under green house conditions (2004). Genome 47;493–500.

A. Peremarti, *et al.,* Promoter diversity in multigene transformation (2010). Plant Molecular Biology 73;363–378.

Axel Mogk, Ronny, Schmidt, Bernd Bukau. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies (2007). Volume 17, Issue 4, Pages 165–172.

Bachmair, A., Finley, D. & Varshavsky, A. *In-vivo* half-life of a protein is a function of its amino-terminal residue (1986). Science, 234:179–186.

Balzi, E., Choder, M., Chen, W. N., Varshavsky, A., and Goffeau, A. Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of s*accharomyces cerevisiae* (1990). J. Biol. Chem.265, 7464–7471.

Brown, R.L., Kazan, K., McGrath, K.C., Maclean, D.J., and Manners, J.M. A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*, (2003). Plant Physiol. 132, 1020–1032.

Ciechanover, A. Proteolysis: from the lysosome to ubiquitin and the proteasome (2005). Nature Reviews Molecular Cell Biology 6(1), 79-87.

C Jiang, B. Iu, J. Singh, Requirement of a CCGAC *cis*-acting element for cold induction of the BN115 gene from winter *Brassica napus* (1996). Plant Molecular Biology. 30;679–684.

Dahl *et al*., Q2ChIP, a quick and quantitative chromatin immunoprecipitation assay, unravels epigenetic dynamics of developmentally regulated genes in human carcinoma cells (2007). Stem cells. 25(4): 1037-46.

Dang J.L., Jones J.D.G. Plant pathogens and integrated defence responses to infection (2001). Nature 411: 826–833

Das PM, Dahl JA. Chop it, ChIP it, check it: the current status of chromatin immunoprecipitation (2004). Biotechniques. 37:961-969.

Deng *et al*, An *Arabidopsis* Regulatory Gene, Encodes a Protein with Both a Zinc-Binding Motif and a GP Homologous Domain (1992). Cell, Vol 71, 791-601.

Dohmen R. J., Madura K., Bartel B., Varshavsky A. The N-end rule is mediated by the UBC2 (RAD6) ubiquitin-conjugating enzyme (1991). Proc. Natl. Acad. Sci. USA88: 7351–7355.

D.J. Cosgrove, Loosening of plant cell walls by expansins (2000). Nature 407;321–326.

Dougan, K. N. Truscott, K. Zeth. The bacterial N-end rule pathway: expect the unexpected, (2010). Molecular Microbiology, 545–558.

Downes BP, Stupar RM, Gingerich DJ, Vierstra RD. HECT ubiquitin-protein ligase (UPL) family in *Arabidopsis*: UPL3 has a specific role in trichome development (2003). Plant J; 35:729-42.

Emmanuelle Graciet, Frank Wellmer. The plant N-end rule pathway: structure and functions (2010). Trends Cell Biology. 447–453.

Erbse A, Schmidt R, Bornemann T, Schneider-Mergener J, Mogk A, Zahn R, Dougan DA, Bukau B. ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. Nature (2006). 439:753–756.

Finkelstein, R.R., and Lynch, T.J. The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor, (2000). Plant Cell 12:599–609.

Forster F, Lasker K, Nickell S, Sali A, Baumeister W. Toward an integrated structural model of the 26S proteasome (2010). Mol Cell Proteomics, 9:1666–1677.

Gao *et al.,* mTOR drives its own activation *via* SCFβ-TRCP-dependent degradation of the mTOR inhibitor DEPTOR (2011). Molecular Cell, 44(2): 290–303.

Garzon M, *et al*. PRT6/At5g02310 encodes an *Arabidopsis* ubiquitin ligase of the N-end rule pathway with arginine specificity and is not the CER3 locus (2007). Febs Letters, 581:3189–3196.

Gibbs *et al*., The eukaryotic N-endrule pathway: conserved mechanisms and diverse functions (2014). Trends in Cell biology, 1-9.

Gibbs *et al*., Group VII Ethylene Response Factors Coordinate Oxygen and Nitric Oxide Signal Transduction and Stress Responses in Plants (2015). Plant physiology, 169(1):23-31.

Gibbs *et al*., Nitric Oxide Sensing in Plants is Mediated by Proteolytic Control of Group VII ERF Transcription Factors, (2014). Molecular Cell 53, 1–11.

Gibbs DJ, Cho Lee S, Md Isa N, Gramuglia S, Fukao T, Bassel GW, Sousa Correia C, Corbineau F, Theodoulou FL, Bailey-Serres J, Holdsworth MJ, (2011). Homeostatic response to hypoxia is regulated by the N-end rule pathway in plants. Nature. 479:415-418.

Giuntoli B., S. C. Lee, F Licausi, M. Kosmacz, T. Oosumi, J. T. Van Dongen, J. Bailey-Serres and P. Perata (2014). A Trihelix DNA Binding Protein Counter balances Hypoxia-Responsive Transcriptional Activation in *Arabidopsis*. PLoS Biology. 12 (9), e1001950.

Giacomo Di, Serino G and Frugis G, Emerging Role of the Ubiquitin Proteasome System in the Control of Shoot Apical Meristem Function (2013). Journal of Integrative Plant Biology, 55(1): 7–20.

Graciet E, Mesiti F, Wellmer F. Structure and evolutionary conservation of the plant N-end rule pathway (2010). Plant J. 2010, 61:741–751.

Graciet, E. *et al*. The N-end rule pathway controls multiple functions during *Arabidopsis* shoot and leaf development, (2009). PNAS, 106:13618–13623.

Grigoryev S, Stewart AE, Kwon YT, Arfin SM, Bradshaw RA, Jenkins NA, Copeland NG, Varshavsky A. A mouse amidase specific for N-terminal asparagine. The gene, the enzyme, and their function in the N-end rule pathway (1996). J Biol Chem. 271:28521–28532.

Grun S, Lindermayr C, Sell S, Durner J (2006). Nitric oxide and gene regulation in plants. J Exp Bot. 57**:** 507–516.

Guo H and Joseph R. Ecker, Plant Responses to Ethylene Gas Are Mediated by SCFEBF1/EBF2-Dependent Proteolysis of EIN3 Transcription Factor (2003). Cell, Vol. 115;667–677.

Guerineau F., Lucy A., Mullineaux P., Effect of two consensus sequences pre-ceding the translation initiator codon on gene expression in plant protoplasts (1992). Plant Molecular Biology 18;815–818.

Han DC, Chen CY, Chen YF, Winans SC Altered-function mutations of the transcriptional regulatory gene virG of *Agrobacterium tumefaciens* (1992). J Bacteriol, 174:7040-43.

Hellemans J, Mortier G, de Paepe A, Speleman F, Vandesompele J. qBase relative quantification framework and software for management and automatedanalysis of real-time quantitative PCR data (2007). Genome Bio. 8(2):R19.

Himmelbach, *et al.,* Promoters of the barley germin-like GER4 gene cluster enable strong transgene expression in response to pathogen attack (2010). Plant Cell 22;937–952

Hinz M, et al. *Arabidopsis* RAP2.2: An Ethylene Response Transcription Factor that is important for hypoxia survival (2010). Plant Physiol, 153:757–772.

Holdsworth, M.J., L. Bentsink and W.J.J.Soppe. Molecular networks regulating *Arabidopsis* seed maturation, after-ripening, dormancy and germination. New Phytologist (2008). 17;33-54.

Holm, M., Ma, L-G., Qu, L-J. and Deng, X. W. Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis* (2002). Gene, Dev. 16;1247–1259.

Hochstrasser M. UBIQUITIN-DEPENDENT PROTEIN DEGRADATION (1996). Annual Review of Genetics. Vol. 30:405-439.

Holm, M., Ma, L. G., Qu, L.J. & Deng, X. W. Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in *Arabidopsis* (2002). Gene16, 1247–1259.

Hu, R-G., Sheng, X. Qi, Z.Xu, T.T. Takahashi and Varshavasky. The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators (2005). Nature, 437 (7061), 981-6.

Hua J, Meyerowitz EM. Ethylene responses are negatively regulated by a receptor gene family in *Arabidopsis thaliana*, (1998). Cell 94:261-271.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*;973-977.

J.W. Tobias *et al.*, The N-end rule in bacteria. Science (1991). 254:1374–1377.

Julia Bailey-Serres, Takeshi Fukao, Daniel J. Gibbs, Michael J. Holdsworth, Seung Cho Lee, Francesco Licausi, Pierdomenico Perata, Laurentius A.C.J. Voesenek, Joost T. van Dongen. Making sense of low oxygen sensing (2012).Trends Cell Biol.Volume 17, Issue 3, Pages 129–138.

Jose M, Franco-Zorrilla, Irene Lopez-Vidriero, Jose L. Carrasco, Marta Godoy, Pablo Vera and Roberto solano. DNA-binding specificities of plant transcription factors and their potential to define target genes (2013). PNAS, 2367-2372.

J.T. Odell, F. Nagy, N.H. Chua, Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter(1985). Nature 313;810–812.

Kendrick and Chang, Ethylene signaling: new levels of complexity and regulation (2008). Current Opinion in Plant Biology, 11:479–485.

Kaji, H., Novelli, G. D., and Kaji, A. A soluble amino acid-incorporating system from rat liver (1963). Biochim. Biophys. Acta 76;474–477.

Kim, H.-K, R.-R, Kim, J.-H. Oh, H. Cho, A. Varshavsky and C.-S. Hwang. The N-terminal Methionine of cellular proteins as a Degradation Signal (2013). Cell, 1-12.

Kwon YT, et al. An essential role of N-terminal arginylation in cardiovascular development (2002). Science, 297:96–99.

Kwon, Y. T., Kashina, A. S., and Varshavsky, A. Alternative splicing results in differential expression, activity, and localization of the two forms of arginyl-tRNA-protein transferase, a component of the N-end rule pathway (1999). Mol. Cell. Biol.19, 182–193.

K. Yamaguchi-Shinozaki, K. Shinozaki, A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress (1994). Plant Cell 6;251–264.

K. Yamaguchi-Shinozaki, K. Shinozaki, Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses (2006). Annual Review of Plant Biology 57;781–803.

Licausi Francesco, Daan A. Weits, Bikram Datt Pant, Wolf-Rudiger Scheible, Peter Geigenberger and Joost T. van Dongen. Hypoxia responsive gene expression is mediated by various subsets of transcription factors and miRNAs that are determined by the actual oxygen availability (2011). New Phytologist, 190: 442–456.

Laurentius A. C. J. Voesenek and Julia Bailey-Serres. Air conditional (2015). Nature Plants, Vol-1.

Licausi Francesco, Masaru Ohme-Takagi and Pierdomenico Perata. APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs (2013). New Phytologist, 199: 639–649.

Licausi Francesco, Monika Kosmacz, Daan A. Weits, Beatrice Giuntoli, Federico M. Giorgi, Laurentius A. C. J. Voesenek, Pierdomenico Perata & Joost T. van Dongen. Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization (2011). NATURE, 419.

Licausi F, *et al.* HRE1 and HRE2, two hypoxia-inducible ethylene response factors, affect anaerobic responses in *Arabidopsis thaliana* (2010). Plant J, 62:302–315.

Liem T. Bui, Beatrice Giuntoli, Monika Kosmacz, Sandro Parlanti, Francesco Licausi. Constitutively expressed ERF-VII transcription factors redundantly activate the core anaerobic response in *Arabidopsis thaliana* (2015). Plant Science 37–43.

Lin, C. *et al*., Blue light photoreceptors and signal transduction (2002). Plant Cell 14, S207–S225.

Lopez-Molina, L., S. Mongrand and N.H. Chua. A postgermination developmental arrest checkpoint is mediated by abscisic acid and requires the ABI5 transcription factorin *Arabidopsis* (2001). PNAS, 4782-4787.

Lorenzo O, Piqueras R, Sánchez-Serrano JJ, Solano R. ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense, (2003). Plant Cell 15:165-178.

Ma *et al*., Regulators of PP2C Phosphatase Activity Function as Abscisic Acid Sensors (2009). SCIENCE, Vol-324. 1064-1068.

Ma, L. *et al.* Light control of *Arabidopsis* development entails coordinated regulation of genome expression and cellular pathways (2001). Plant Cell 13;2589–2398.

Mogk, A., Schmidt, R. & Bukau, B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies (2007). Trends Cell Biol.17:165–172.

Molina, E. Grotewold, Genome wide analysis of *Arabidopsis* core promoters (2005). BMC Genomics 6;25.

Moller, S. G., Ingles, P. J. & Whitelam, G. C. The cell biology of phytochrome (2002). signaling. New Phytol. 154;553–590.

Monika Kosmacz, Sandro Parlanti, Markus Schwarzlander, Friedrich Kragler, Francesco Licausi and Joost T. Van Dongen. The stability and nuclear localization of the transcription factor RAP2.12 are dynamically regulated by oxygen concentration (2015). Plant, Cell and Environment, 38;1094-1103.

N.C. Bisht, A. Jagannath, P.K. Burma, A.K. Pradhan, D. Pental, Retransformation of a male sterile barnase line with the barstar gene as an efficient alternative method to identify male sterile–restorer combinations for heterosis breeding (2007). Plant Cell Reports 26;727–733.

Nambara, E., Marion-Poll, A. (2005). Abscisic acid biosynthesis and catabolism. Ann. Rev. Plant Biol. 56:165-185.

N. Kovalchuk, et al., Defensin promoters as potential tools for engineering disease resistance in cereal grains (2010). Plant Biotechnology Journal 8;47–64.

Nakano T, Suzuki K, Fujimura T, Shinshi H. Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice (2006). Plant Physiology, 140:411–432.

Neill, S., J. Bright, R. Desikan, J. Hancock, J. Harrison and I. Wilson. Nitric oxide evolution and perception (2008). Journal of Experimental Botany 59(1), 25-35.

N. Yi, *et al*., Analysis of the Wsi18, a stress-inducible promoter that is active in the whole grain of transgenic rice (2011). Transgenic Research 20;153–163.

Osterlund T. *et al* (2000). Targeted destabilization of Hy5 during light regulated development of *Arabidopsis*. Nature 405;462–466.

Pan *et al*., Expression of signalling and defence-related genes mediated by over-expression of JERF1, and increased resistance to sheath blight in rice (2014). Plant Pathology. 63;109–116

Pandey *et al*, Cold induced Botrytis cinerea enolase (BcEnol-1) functions as a transcriptional regulator and is controlled by cAMP (2009). Mol. Genet Genomics, 281:135–146.

Park *et al*, Abscisic acid inhibits PP2Cs *via* the PYR/PYL family of ABA-binding START proteins (2009). Science, 2009 May 22; 324(5930): 1068–1071.

P.N. Benfey, N.H. Chua, The cauliflower mosaic virus 35S promoter: combinatorial regulation of transcription in plants (1990). Science 250;959–966.

Pepper, A., Delaney, T., Washburn, T., Poole, D. & Chory, J. DET1, a negative regulator of light-mediated development and gene expression in *Arabidopsis*, encodes a novel nuclear localized protein (1994). Cell. 78;109–116.

Peremarti A, *et al*., Promoter diversity in multigene transformation (2010). Plant Molecular Biology, 73;363–378.

Pickart C. M., Rose I. A. Functional heterogeneity of ubiquitin carrier proteins (1985). J. Biol. Chem.260:1573–1581

Potenza, L. Aleman, C. Sengupta-Gopalan. Targeting transgene expression in research, agricultural, and environmental applications: promoters used in plant transformation (2004). In Vitro Cellular & Developmental Biology Plant. 40;1–22.

Potuschak, T., S. Stary, P. Schlogelhofer, F. Becker, V. NejinskaiAand A. Bachmair. PRT1 of *Arabidopsis thaliana* encodes a component of the N-end rule pathway (1998). Proceedings of the National Academy of Sciences, 95(14), 7904-7908.

Ohme-Takagi, M., and Shinshi, H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element, (1995). Plant Cell 7;173–182.

Osterlund, M. T., Hardtke, C., Wei, N. & Deng, X. W. Targeted destabilization of Hy5 during light-regulated development of *Arabidopsis* (2000). Nature. 405;462–466.

Qiao H, Katherine N. Chang, Junshi Yazaki, and Joseph R. Ecker. Interplay between ethylene, ETP1/ETP2 F-box proteins, and degradation of EIN2 triggers ethylene responses in *Arabidopsis* (2009). ENES & DEVELOPMENT 23:512–521.

Quail, P. H. Phytochrome photosensory signaling networks (2002). Nature Rev. Mol. Cell Biol. 3;85–93.

Rodriguenz, A., p. Chen, H. Oliver and J. M. Abrams. Unrestrained caspase-dependent cell death caused by loss of Diap1 function requires the Drosophila Apaf-1 homolog, Dark (2002). The EMBO journal, 21 (9), 2189-2197.

Rushton P.J., Reinstadler A., Lipka V., Lippok B., Somssich I.E., Synthetic plant promoters containing defined regulatory elements provide novel insights into pathogen and wound-induced signalling (2002). Plant Cell 14;749–762.

Schrader, E.K., Harstad, K.G., and Matouschek, A.Targeting proteins for degradation, (2009). Nat Chem Biol5: 815–822.

S. H. Hwang, I.A. Lee, S.W. Yie, D.J. Hwang, Identification of an OsPR10a promoter region responsive to salicylic acid (2008). Planta. 227;1141–1150.

Shashikanth M. Sriram, Bo Yeon Kim & Yong Tae Kwon. The N-end rule pathway: emerging functions and molecular principles of substrate recognition (2011). Nature Reviews Molecular Cell Biology 12:735-747.

Shriram, S. M., B.Y. Kim and Y.T. Know. The N-end rule pathway: emerging functions and molecular principles of substrates recognition (2011). Nature Reviews Molecular Cell Biology, 12;735-747.

Shriram, S. M. & Kwon, Y. T. The molecular principles of N-end rule recognition, (2010). Nature Structure. Molecular Biology,17:1164–1165.

Shen, Q., and Ho, T.H.D. Functional dissection of an abscisic acid (ABA)-inducible gene reveals two independent ABA-responsive complexes each containing a G-box and a novel *cis*-acting element (1995). Plant Cell 7;295-307.

S. H. Park, et al., Analysis of the APX, PGD1 and R1G1B constitutive gene pro-moters in various organs over three homozygous generations of transgenic rice plants (2012). Planta, 235;1397–1408.

S.K. Burley, R.G. Roeder, Biochemistry and structural biology of transcription factor IID (TFIID) (1996), Annual Review of Biochemistry 65;769–799.

Singh N, Dang TT, Vergara GV, Pandey DM, Sanchez D, Neeraja CN, Septiningsih EM, Mendioro M, Tecson-Mendoza EM, Ismail AM, *et al*. (2010). Molecular marker survey and expression analyses of the rice submergence-tolerance gene SUB1A. Theor Appl Genet 121:1441–1453.

Stary S, Xiao-jun Y, Potuschak T, Schlogelhofer P, Nizhynska V, Bachmair A. PRT1 of *Arabidopsis* is a Ubiquitin Protein Ligase of the Plant N-End Rule Pathway with specificity for Aromatic Amino-Terminal Residues (2003). Plant Physiology. 133(3):1360-1366.

Stone SL, Hauksdottir H, Herschleb J, Kraft E, Callis J. Functional analysis of the RING-type ubiquitin ligase family of *Arabidopsis* (2005). Plant Physiol, 137:13-30.

Sullivan, J. A. & Deng, X. W. From seed to seedling; the role of photoreceptors in *Arabidopsis* development (2003). Dev. Biol. 260;289–297.

Sun, L. and Chen, Z.J. The novel functions of ubiquitination in signalling (2004). Curr. Opin. Cell. Biol., 16:119-126.

Suzuki, G., Yanagawa, Y., Kwok, S. F. & Deng, X. W. *Arabidopsis* COP10 is a ubiquitin-conjugating enzyme variant that acts together with COP1 and the COP9 signalosome in repressing photomorphogenesis (2002). Genes Dev. 16;554–559.

S.W. Bang, *et al*., Characterization of the stress-inducible OsNCED3 promoter in different transgenic rice organs and over three homozygous generations (2013). Planta 237;211–224.

Takafumi Tasaki, Shashikanth M. Sriram, Kyong Soo Park, 2*,* 3 and Yong Tae Kwon. The N-End Rule Pathway(2012). Annual. Review Biochemistry. 81:26.1-26.29.

Tasaki T, Kwon YT. The mammalian N-end rule pathway: new insights into its components and physiological roles (2007). Trends in Biochemical Science, 32:520–528.

Tasaki, T., Sriram, S. M., Park, K. S., and Kwon, Y. T. The N-end rule pathway (2012). Annu. Rev. Biochem., DOI: 10.1146.

Tara J. Holman, Peter D. Jones, Laurel Russell, Anne Medhurst, Susana Ubeda Tomas, Prabhavathi Talloji, Julietta Marquez, Steven Footitt, Andreas Bachmair, Frederica L. Theodoulou and Michael J. Holdsworth. The N-end rule pathway promotes seed germination and establishment through removal of ABA sensitivity in *Arabidopsis* (2009). PNAS, vol. 106:4549–4554.

T. Kawakatsu, F. Takaiwa, Cereal seed storage protein synthesis: fundamental processes for recombinant protein production in cereal grains (2010). Plant Biotechnology Journal 8;939–953.

Teppermann, J. M., Zhu, T., Chang, H. S., Wang, X. & Quail, P. H. Multiple transcription factor genes are early targets of phytochrome A signalling (2001). Proc. National Acadamic Science, USA. 98;9437–9442.

Tobias JW, Shrader TE, Rocap G, Varshavsky A. The N-end rule in bacteria. Science (1991). 254:1374–1377.

Varshavsky A. The N-end rule pathway and regulation by proteolysis (2011). Protein Science, 20:1298–1345.

Varshavsky, A. The ubiquitin system. Trends in Biochemical Sciences (1997). 22(10), 383-387.

Varshavsky A. The N-end rule: functions, mysteries, uses (1996). PNAS, 93(22), 12142-12149.

Wang, H., K.I. Piatkov, C.S. Brower and a. Varshvasky. Glutamine-specific N-terminal amidase, a component of the N-end rule pathway (2009). Molecular cell, 34(6), 686-695.

Wang, H. & Deng, X. W. The Arabidopsis Book (2002). American Society of Plant Biologists.

Wang, H., Ma, L., Li, J., Zhao, H. & Deng, X. W. Direct interaction of *Arabidopsis* cryptochromes with COP1 in light control development (2001). *Science* 294;154–158.

Wei, N. & Deng, X. W. The role of the COP/DET/FUS genes in light control of *Arabidopsis* seedling development (1996). Plant Physiology. 112;871–878.

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

Wong ML, Medrano JF. Real-time PCR for mRNA quantitation (2005). Biotechniques, 39(1):75–85.

Xie Y. Structure, assembly and homeostatic regulation of the 26S proteasome ((2011). J Mol Cell Biol, 2, 308–317, 46.

Yamaguchi-Shinozaki, K., and Shinozaki, K. A novel *cis*-acting element in an *Arabidopsis* gene is involved in responsiveness to drought, low-temperature, or high-salt stress (1994). Plant Cell 6;251-264.

Yang S, Wang S, Liu X, Yu Y, Yue L, Wang X, Hao D. Four divergent Arabidopsisethylene-responsive element-binding factor domains bind to a target DNA motif with a universal CG step core recognition and different flanking bases preference(2009). FEBS J. 276(23):7177-86.

Yang, H. Q., Tang, R.-H. & Cashmore, A. R. The signalling mechanism if *Arabidopsis* CRY1 involves direct interaction with COP1 (2001). Plant Cell 13, 2573–2587.

Yee D and Goring R. The diversity of plant U-box E3 ubiquitin ligases: from upstream activators to downstream target substrates (2009). Journal of Experimental Botany, Vol. 60, No. 4, pp. 1109–1121.

Yoshida, S., M. Ito, J. Callis, I. Nishida and A. Watanabe. A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway (2002). The Plant journal, 32(1), 129-137.

Zhang Z, Li F, Li D, Zhang H, Huang R. Expression of ethylene response factor JERF1 in rice improves tolerance to drought (2010). Planta. 232: 765–774.

Zhao, Y., T. Wei, K.-Q. Yin, Z. Chen, H. Gu, L.-J. Qu and G. Qin. *Arabidopsis* RAP2.2 plays an important role in plants resistance to Botrytis cinerea and ethylene respone (2012). The New Phytilogist, 195 (2), 450-460.

Zou, *et al*., *Cis*-regulatory code of stress-responsive transcription in *Arabidopsis thaliana* (2011). Proceedings of the National Academy of Sciences of the United States of America, 108;14992–14997

8. APPENDIX

1. Blasted sequence of amplified promoters of genes with Arabidopsis genomic DNA

AT4G01026: Symbols: PYL7, RCAR2/ chr4:447052-448511

GGAGATTGAAACAATGGTGGAAAAAGTGGCGGCTATGGCCACAAAGGCGGGAGAAA
CGGCGACAATGGCGGCAGATATGGTGAAAGAAGCTGAGGAGACGATGGAAACAGCT
AAAGCTAATATGTCCAAAGCCTTTGTGGTAATGAAGTCGGTGAATTGGAACGTGTAAA
TCGGGTCAAAACAGAGTTTTATTGTGATCTCATGCTGACATCAGCTACAATCTTTATCT
AATAAGATAGATTCTCACAAGATTCTTTTATCTATCTACTTTTTAGAGAAGATGAGATC
ATACTTCGGAGATAGATAGGTGTCGTAAAAATTGGGAAGCTTACTTGGCAGGAGAAG
AGAGATATACACACGTGCTTAAAGTCAACAGCTAATCCAAAAAGGTAGACGACAAACC
AGTTAAACTAAGACACGTAATCTATCTTAAAGATTTGTCGGTTCATCGTAAATCCGTTG
GTGGATTAATATATGTCGGTGTCTTCGTTGATTTTCTTAGCCTAACCAACCAAACATAA
ATAAAAATCGAAGCTTTACTTTGTTTGCTTCTATTATATCCACTTCCAACAAGATTCCT
CAAAAGGAAAGAAGAGAAAGAAGATAAGAGATATCTGGTTGGTCGGAG**ATG**

BLAST RESULT:

```
Query  69      GGAGATTGAAACAATGGTGGAAAAAGTGGCGGCTATGGCCACAAAGGCGGGAGAAACGGC  128
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446552  GGAGATTGAAACAATGGTGGAAAAAGTGGCGGCTATGGCCACAAAGGCGGGAGAAACGGC  446611

Query  129     GACAATGGCGGCAGATATGGTGAAAGAAGCTGAGGAGACGATGGAAACAGCTAAAGCTAA  188
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446612  GACAATGGCGGCAGATATGGTGAAAGAAGCTGAGGAGACGATGGAAACAGCTAAAGCTAA  446671

Query  189     TATGTCCAAAGCCTTTGTGGTAATGAAGTCGGTGAATTGGAACGTGTAAATCGGGTCAAA  248
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446672  TATGTCCAAAGCCTTTGTGGTAATGAAGTCGGTGAATTGGAACGTGTAAATCGGGTCAAA  446731

Query  249     ACAGAGTTTTATTGTGATCTCATGCTGACATCAGCTACAATCTTTATCTAATAAGATAGA  308
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446732  ACAGAGTTTTATTGTGATCTCATGCTGACATCAGCTACAATCTTTATCTAATAAGATAGA  446791

Query  309     TTCTCACAAGATTCTTTTATCTATCTACTTTTTAGAGAAGATGAGATCATACTTCGGAGA  368
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446792  TTCTCACAAGATTCTTTTATCTATCTACTTTTTAGAGAAGATGAGATCATACTTCGGAGA  446851

Query  369     TAGATAGGTGTCGTAAAAATTGGGAAGCTTACTTGGCAGGAGAAGAGAGATATACACACG  428
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446852  TAGATAGGTGTCGTAAAAATTGGGAAGCTTACTTGGCAGGAGAAGAGAGATATACACACG  446911

Query  429     TGCTTAAAGTCAACAGCTAATCCAAAAAGGTAGACGACAAACCAGTTAAACTAAGACACG  488
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446912  TGCTTAAAGTCAACAGCTAATCCAAAAAGGTAGACGACAAACCAGTTAAACTAAGACACG  446971

Query  489     TAATCTATCTTAAAGATTTGTCGGTTCATCGTAAATCCGTTGGTGGATTAATATATGTCG  548
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  446972  TAATCTATCTTAAAGATTTGTCGGTTCATCGTAAATCCGTTGGTGGATTAATATATGTCG  447031

Query  549     GTGTCTTCGTTGATTTTCTTAGCCTAACCAACCAAACATAAATAAAAATCGAAGCTTTAC  608
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  447032  GTGTCTTCGTTGATTTTCTTAGCCTAACCAACCAAACATAAATAAAAATCGAAGCTTTAC  447091

Query  609     TTTGTTTGCTTCTATTATATCCACTTCCAACAAGATTCCTCAAAAGGAAAGAAGAGAAAG  668
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  447092  TTTGTTTGCTTCTATTATATCCACTTCCAACAAGATTCCTCAAAAGGAAAGAAGAGAAAG  447151

Query  669     AAGATAAGAGATATCTGGTTGGTCGGAG  696
              ||||||||||||||||||||||||||||
Sbjct  447152  AAGATAAGAGATATCTGGTTGGTCGGAG  447179
```

AT3G54510
Symbols: Early-responsive to dehydration stress protein (ERD4)

AGCCACTAAACAAATTCCTTTCGTTCATCAAAGTCCACAAATTCACTGAAAGTAATGAT
TCCAAGCATTTTTTCCCACCAATTCAACAATTTTTAAGTACAACACTTTAGTGATTCTAA
TTTACAAAAATAACATTTTGTTGTAAAACAAATTTAGTAGTACTAGAAATTTCCTTTTCA
GGGAGATGCAAAAATTCAATTATCACACTGCTATTTTGGTCGAGGTGTAATAAAAAAA
CAAGGAGACAATTGCGGCGTGGAGGCTTTCTGTTTACATGCAGCACAAGAAGGGCC
ACATGACATGACATCATGTGACTCTGTTTCCCATAGTCAGAGAAAAGAAAAAATATGA
CGCCAGAAAGCTTGTTAGCGTCGGCGTCGATCAACATAGGACTCGCCGTCGTTGCA
CTCTGGCTCTTCTCCGTCCTCAAGAAGCAGCCCCGTAACGCCGTCGTATACTATGCT
C**GCCGCC**TCTCCGATCGCCACCACCACCGTCCTCTGTCTCTACACTCATCCCTTTGC
CTCCCTCGTTTCCTTCCTTCCGTTGCGTGGATCCCACGCGCCTTCCGTGTCCCCGAG
GACGAGATCCTCA**GCCGCC**ACGGCCTTGATGCTCTCGTCCTCATCAGGCTCTTCAAA
TTCGGGTAAGGAAGATTTGGTCTTATCTCGCATATAAACACGATTATCAGATTCTTTGA
ACT**ATG**

BLAST RESULT:

```
Query  58        AGCCACTAAACAAATTCCTTTCGTTCATCAAAGTCCACAAATTCACTGAAAGTAATGATT  117
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20179992  AGCCACTAAACAAATTCCTTTCGTTCATCAAAGTCCACAAATTCACTGAAAGTAATGATT  20180051

Query  118       CCAAGCATTTTTTCCCACCAATTCAACAATTTTTAAGTACAACACTTTAGTGATTCTAAT  177
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180052  CCAAGCATTTTTTCCCACCAATTCAACAATTTTTAAGTACAACACTTTAGTGATTCTAAT  20180111

Query  178       TTACAAAAATAACATTTTGTTGTAAAACAAATTTAGTAGTACTAGAAATTTCCTTTTCAG  237
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180112  TTACAAAAATAACATTTTGTTGTAAAACAAATTTAGTAGTACTAGAAATTTCCTTTTCAG  20180171

Query  238       GGAGATGCAAAAATTCAATTATCACACTGCTATTTTGGTCGAGGTGTAATAAAAAAACAA  297
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180172  GGAGATGCAAAAATTCAATTATCACACTGCTATTTTGGTCGAGGTGTAATAAAAAAACAA  20180231

Query  298       GGAGACAATTGCGGCGTGGAGGCTTTCTGTTTACATGCAGCACAAGAAGGGCCACATGAC  357
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180232  GGAGACAATTGCGGCGTGGAGGCTTTCTGTTTACATGCAGCACAAGAAGGGCCACATGAC  20180291

Query  358       ATGACATCATGTGACTCTGTTTCCCATAGTCAGAGAAAAGAAAAAATATGACGCCAGAAA  417
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180292  ATGACATCATGTGACTCTGTTTCCCATAGTCAGAGAAAAGAAAAAATATGACGCCAGAAA  20180351

Query  418       GCTTGTTAGCGTCGGCGTCGATCAACATAGGACTCGCCGTCGTTGCACTCTGGCTCTTCT  477
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180352  GCTTGTTAGCGTCGGCGTCGATCAACATAGGACTCGCCGTCGTTGCACTCTGGCTCTTCT  20180411

Query  478       CCGTCCTCAAGAAGCAGCCCCGTAACGCCGTCGTATACTATGCTCGCCGCCTCTCCGATC  537
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180412  CCGTCCTCAAGAAGCAGCCCCGTAACGCCGTCGTATACTATGCTCGCCGCCTCTCCGATC  20180471

Query  538       GCCACCACCACCGTCCTCTGTCTCTACACTCATCCCTTTGCCTCCCTCGTTTCCTTCCTT  597
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180472  GCCACCACCACCGTCCTCTGTCTCTACACTCATCCCTTTGCCTCCCTCGTTTCCTTCCTT  20180531

Query  598       CCGTTGCGTGGATCCCACGCGCCTTCCGTGTCCCCGAGGACGAGATCCTCAGCCGCCACG  657
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180532  CCGTTGCGTGGATCCCACGCGCCTTCCGTGTCCCCGAGGACGAGATCCTCAGCCGCCACG  20180591

Query  658       GCCTTGATGCTCTCGTCCTCATCAGGCTCTTCAAATTCGGGTAAGGAAGATTTGGTCTTA  717
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  20180592  GCCTTGATGCTCTCGTCCTCATCAGGCTCTTCAAATTCGGGTAAGGAAGATTTGGTCTTA  20180651


Query  718       TCTCGCATATAAACACGATTATCAGATTCTTTGAACT  754
                  |||||||||||||||||||||||||||||||||||||
Sbjct  20180652  TCTCGCATATAAACACGATTATCAGATTCTTTGAACT  20180688
```

AT1G14810

Symbols semialdehyde dehydrogenase family protein | **chr1**:5102424-5104695

CGTGTAAGATTCCAAAGTTAGCTATAGATGTCTTTGAGCTAAAAACCAGAATCACGAG
CTCTTAAATATGATCCCTAAGCCTTGTTATATGACAAAAGACTTTTCAACAACTTAATA
TTTCCAAAAGAAGAAATGATGAAACACATGAGAAAGTTGTGTGAATATACTCAATCTG
GTGAATGAGCAGTTTCAAACATCATGCTGTTTCTGTATCTTAAATTAAACAATCTTGCT
GTTTTGAATTTTGAAGATGTTGGTCAATTCACTAACACACTTTAAAGTCCCAAGCTTTT
TTCTATCAGGCCGCCTCCAAAATCTCTCTATCTCTCTGTCTCTCACTTCTTCTTCCTCAC
GGCGGCGACAATG

BLAST RESULT:

```
Query  420   GACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGC  479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  749   GACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGC  808

Query  480   AACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC  539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  809   AACGAACAGGTCACTATCAGTCAAAATAAAATCATTATTTGCCATCCAGCTGATATCCCC  868

Query  540   TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAA  599
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  869   TATAGTGAGTCGTATTACATGGTCATAGCTGTTTCCTGGCAGCTCTGGCCCGTGTCTCAA  928

Query  600   AATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCT  659
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  929   AATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAACAATAAAACTGTCT  988

Query  660   GCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGG  719
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  989   GCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCGAGG  1048

Query  720   CCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT  779
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1049  CCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAAT  1108

Query  780   GTCGGGCAATCANGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGANTTG  839
             ||||||||||||  |||||||||||||||||||||||||||||||||||||||||  |||
Sbjct  1109  GTCGGGCAATCAGGTGCGACAATCTATCGCTTGTATGGGAAGCCCGATGCGCCAGAGTTG  1168

Query  840   TTTCTGAAACATGGCAAANGTAGCGTTGNCAATGATGTTACAGATGANATGGTCANACTA  899
             ||||||||||||||||||  |||||||||  ||||||||||||||||||  |||||||  ||||
Sbjct  1169  TTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTA  1228

Query  900   AACTGGCTGACGGAATTTATGCCTCTTCCNANCATCAAGCATTTTATCCGNACTCCTGAT  959
             ||||||||||||||||||||||||||||| | ||||||||||||||||||  |||||||||
Sbjct  1229  AACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGAT  1288

Query  960   GANGCATGGNNACTCACCACTGCNATCCCCGGaaaaa  996
             || ||||||  ||||||||||| ||||||||||||
Sbjct  1289  GATGCATGGTTACTCACCACTGCGATCCCCGGAAAAA  1325
```

AT5G44420

Symbols: **PDF1.2**,| plant defensin 1.2 | **chr5**:17907069-17907598

TACAAATTTAAATTCTGATCTCTTATAATGTTCGGTTTTTCCTTTTTTATTTATTAAGTTA
GTTAAAATTTGCAGTTATTTTGTTGAATGTCGTTGTTTACGAATTTACGAATAATACCTT
TATAGCTAATCTACAAAATTTTGATGACTGACAACACCGTTAATGTTTTTTTTTAAATTA
CCCTGAGCCTCTCACTTGCGGTCAGACCATGCATGTCGATAGTCCATTACGTTTAAG
GCCACAATCAACTATAGTTTGTTTATCAATAGCCAACTAAGCTAACTTTTAGGTTCCTG
CCCTCTCCGTTCCTCCGGTACCAATCGTTTCTTTGTCCCTTCGATAGTTTGAAAACCT
ACCGACGGTGAGAGCAAAATATTGATGAATCATCCAATTTTCAGTAATAGGTGTGTCC
CAGGGATATATAAATGGCGAAACTACGCGAGAACGGTTCCTTGTTCTGCAAACTTGG
CGGAACAATGCTGCTCTTGAGATCAACCAAACCATATGTTTAGTCCACAACGATCTAT
ATGTCTAGGGGTGATCCTCTAATCGAAAAATGTTGTATTTGTTCGACGATGACGAAGG
TCAGACTATGAACTGCACAGTCTGCACTTGTCCTAACCGCGAGAATCTCTGACATCAA
TATACTTGTGTAACTATGGCTTGGTTAAGATATTATTTTCTTGAGTCTTAATCCATTCA
GATTAACCA**GCCGCC**CATGTGAACGATGTAGCATTAGCTAAAAGCCGAAGCAGCCGC
TTAGGTTACTTTAGATATCGACAGAGAAATATATGTGGTGGAGAAACCAGCCATCAAC
AAACAAAAAGCAAGATCTTATCTTTTGATATTGGCTACGGGAAGATGATGTCTGTTTAA

TGTGTGGGGTTACCACGTTATTGTACGATGCACAAGTAGAAGATTAACCCACTACCAT
TTCATTATAAATAGACGTTGATCTTTGGCTTATTTCTTCACACAACACATACATCTATAC
ATTGAAAACAAAATAGTAATAATCATC**ATG**

BLAST RESULT:

```
Query  54        CTTCTACAAATTTAAATTCTGATCTCTTATAATGTTCGGttttttcttttttATTTATTA  113
                 ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908576  CTTTTACAAATTTAAATTCTGATCTCTTATAATGTTCGGTTTTTCCTTTTTTATTTATTA  17908517

Query  114       AGTTAGTTAAAATTTGCAGTTATTTTGTTGAATGTCGTTGTTTACGAATTTACGAATAAT  173
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908516  AGTTAGTTAAAATTTGCAGTTATTTTGTTGAATGTCGTTGTTTACGAATTTACGAATAAT  17908457

Query  174       ACCTTTATAGCTAATCTACAAAATTTTGATGACTGACAACACCGTTAATGttttttttA  233
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908456  ACCTTTATAGCTAATCTACAAAATTTTGATGACTGACAACACCGTTAATGTTTTTTTTTA  17908397

Query  234       AATTACCCTGAGCCTCTCACTTGCGGTCAGACCATGCATGTCGATAGTCCATTACGTTTA  293
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908396  AATTACCCTGAGCCTCTCACTTGCGGTCAGACCATGCATGTCGATAGTCCATTACGTTTA  17908337

Query  294       AGGCCACAATCAACTATAGTTTGTTTATCAATAGCCAACTAAGCTAACTTTTAGGTTCCT  353
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908336  AGGCCACAATCAACTATAGTTTGTTTATCAATAGCCAACTAAGCTAACTTTTAGGTTCCT  17908277

Query  354       GCCCTCTCCGTTCCTCCGGTACCAATCGTTTCTTTGTCCCTTCGATAGTTTGAAAACCTA  413
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908276  GCCCTCTCCGTTCCTCCGGTACCAATCGTTTCTTTGTCCCTTCGATAGTTTGAAAACCTA  17908217

Query  414       CCGACGGTGAGAGCAAAATATTGATGAATCATCCAATTTTCAGTAATAGGTGTGTCCCAG  473
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908216  CCGACGGTGAGAGCAAAATATTGATGAATCATCCAATTTTCAGTAATAGGTGTGTCCCAG  17908157

Query  474       GGATATATAAATGGCGAAACTACGCGAGAACGGTTCCTTGTTCTGCAAACTTGGCGGAAC  533
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908156  GGATATATAAATGGCGAAACTACGCGAGAACGGTTCCTTGTTCTGCAAACTTGGCGGAAC  17908097

Query  534       AATGCTGCTCTTGAGATCAACCAAACCATATGTTTAGTCCACAACGATCTATATGTCTAG  593
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908096  AATGCTGCTCTTGAGATCAACCAAACCATATGTTTAGTCCACAACGATCTATATGTCTAG  17908037

Query  594       GGGTGATCCTCTAATCGAAAAATGTTGTATTTGTTCGACGATGACGAAGGTCAGACTATG  653
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17908036  GGGTGATCCTCTAATCGAAAAATGTTGTATTTGTTCGACGATGACGAAGGTCAGACTATG  17907977

Query  654       AACTGCACAGTCTGCACTTGTCCTAACCGCGAGAATCTCTGACATCAATATACTTGTGTA  713
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17907976  AACTGCACAGTCTGCACTTGTCCTAACCGCGAGAATCTCTGACATCAATATACTTGTGTA  17907917

Query  714       ACTATGGCTTGGTTAAGATATTATTTTCTTGAGTCTTAATCCATTCAGATTAACCAGCCG  773
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17907916  ACTATGGCTTGGTTAAGATATTATTTTCTTGAGTCTTAATCCATTCAGATTAACCAGCCG  17907857

Query  774       CCCATGTGAACGATGTAGCATTAGCTAAAAGCCGAAGCAGCCGCTTANGTTACTTTAGAT  833
                 |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  17907856  CCCATGTGAACGATGTAGCATTAGCTAAAAGCCGAAGCAGCCGCTTAGGTTACTTTAGAT  17907797

Query  834       ATCGACAGAGAAATATATGTGGTGGAGAAACCAGCCATCAACAAACAAAAAGCAAGATCT  893
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  17907796  ATCGACAGAGAAATATATGTGGTGGAGAAACCAGCCATCAACAAACAAAAAGCAAGATCT  17907737

Query  894       TATCTTTTGATATTGGCTACGGGAAGATGATGTCTGTTTAATGNGTGGGGTTACCACGTT  953
                 |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Sbjct  17907736  TATCTTTTGATATTGGCTACGGGAAGATGATGTCTGTTTAATGTGTGGGGTTACCACGTT  17907677

Query  954       ATTGTACGATGCACAAGTAGAN-ATTAACCCACTACCATTTCATTATAAATAGAC  1007
                 ||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Sbjct  17907676  ATTGTACGATGCACAAGTAGAAGATTAACCCACTACCATTTCATTATAAATAGAC  1790762
```

AT3G13440

Symbols: S-adenosyl-L-methionine-dependent methyltransferases superfamily protein

ATTTGCTATATCTCAACCACTAATTTAAGGGGCTTTCGGCCCATTACACGCATGGGAA
GATTCAATTATGTTGATTTTGTAAGTTTTGGGGCTAAAGTGTAAATTATTGAAACTTGT
AGGCTTGATTTCATTATTCCAAGTTTCCACTTTTTCTCCTCCGCCGTAGCCGGTGCTT
CTTCTTCGTGAAAATCTTTCGTCTGAGATTTTTTCTTTACTAGTTGAATAAAATTGAAGA
TTTCGAAGACTACCCAAAACCGCGTTTTACAAAAAGCCAGATTCTTTTTTTTTTTTTTTT
TTTTTATAATTTCATTAAAACTACCCAATGGGCTTGGGCCGAAACTTACAAACAATTCG
GTCCAGTTACAAAATTAGGGTTTTCTGACACATCGCCAGGAAAAAGAAAAAAGCTC**GC**
CGCCTGTTGACCGTCGCTAGCAGAGCGACCAAAACCTAGATCCGGCGGACTCAAGT
CCCACAGCTTTGTG**GCCGCC**TCCGAGAATCGTCCACCACCGACGAAAGATTCCGGTT
TTCGTCTGTAGAGACTCAACTCCAACGCACCAACCCACCACAAGACTCCCGGAAAGC
TAGCCGCTCCGACGTAGCGGTTCCACGGAACCGGATCCCGACCTAGCAACCCGAAA
CACTCAGCGATAACACTTCTCTCCCATCACCGGAAAATAGCAAAGACCCAGAGCCAA
ACACACGTCGATCCGGACCGGTGGAGACCAAAAACGGCAGACCTAAAGACTGGTGG
CTCAGTCAGACTCTCTAGATCCAAAACCCGCTGCAGCCATCTAAGATCCAAACAGAT
CTACTGTTAACCACCGCATCAGACAACGAACCCCCCCTTACCCCTCTCCAACTTCCG
CCGTGGACGCAGCGACTAGAATCACGGGTGAGCCTGAGAACCTCGCAGCAACCCTC
AGCGTCCATAACCAAGAGACCCCCGCCGCAACCTTCTCGACCGACAAAACCAGAGC
CAAATCCACATCAGAGAGAGAGCTCCACACCACACGTTGACGGAGACAAGCATCGG
CTCACTCCTGGACTTCCTTGAAGAGGCTTCAAACCGACGATGAAGAGCAGATCGGAA
CAAAAAGAAAAAGGAAACTAAATCTACTTACCGGAAATTGTCTTCAAGAGAGAGAGAA
GGGCACGACTTTCAAGGCTACGATTGAGCCTCCATCTCATAAAGCCAGATTCTTTATC
TTCCATAATCGTATTAAAGCTACGAAAACCTCGATTTCGTGTTCAAAACTTGGATTTGC
TCGCTATCGAAAGGTCTGACCTTTAATCCTTTTCATGGGTTTAATCGTTCAGATATGAT
TTCTCGTATCGGTATTATAGCTTAGAGGAAAAGTGTCGTAGTTTTGGCAGATGAAATT
AAAATTGGCTCTGTATTCCCAAATAGTTGAAACCCATTTGATGAAGCTATCTTATCAGA
CTTTTACGCATAAGTTTTCTTGTGGTCTTGGATCGTGTCTAGTGGCAATTATCTTCTCT
GAGTTCTTTTGCTTAGTTTGATTGTCTGTTGTTTTAGACGA**ATG**

BLAST RESULT:

```
Query  62       TAAAACTACCCAATGGGCTTGGGCCGAAACTTACAAACAATTCGGTCCAGTTACAAAATT  121
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377107  TAAAACTACCCAATGGGCTTGGGCCGAAACTTACAAACAATTCGGTCCAGTTACAAAATT  4377166

Query  122      AGGGTTTTCTGACACATCGCCAGGaaaaagaaaaaaGCTCGCCGCCTGTTGACCGTCGCT  181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377167  AGGGTTTTCTGACACATCGCCAGGAAAAAGAAAAAAGCTCGCCGCCTGTTGACCGTCGCT  4377226

Query  182      AGCAGAGCGACCAAAACCTAGATCCGGCGGACTCAAGTCCCACAGCTTTGTGGCCGCCTC  241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377227  AGCAGAGCGACCAAAACCTAGATCCGGCGGACTCAAGTCCCACAGCTTTGTGGCCGCCTC  4377286

Query  242      CGAGAATCGTCCACCACCGACGAAAGATTCCGGTTTTCGTCTGTAGAGACTCAACTCCAA  301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377287  CGAGAATCGTCCACCACCGACGAAAGATTCCGGTTTTCGTCTGTAGAGACTCAACTCCAA  4377346

Query  302      CGCACCAACCCACCACAAGACTCCCGGAAAGCTAGCCGCTCCGACGTAGCGGTTCCACGG  361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377347  CGCACCAACCCACCACAAGACTCCCGGAAAGCTAGCCGCTCCGACGTAGCGGTTCCACGG  4377406

Query  362      AACCGGATCCCGACCTAGCAACCCGAAACACTCAGCGATAACACTTCTCTCCCATCACCG  421
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377407  AACCGGATCCCGACCTAGCAACCCGAAACACTCAGCGATAACACTTCTCTCCCATCACCG  4377466

Query  422      GAAAATAGCAAAGACCCAGAGCCAAACACACGTCGATCCGGACCGGTGGAGACCAAAAAC  481
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377467  GAAAATAGCAAAGACCCAGAGCCAAACACACGTCGATCCGGACCGGTGGAGACCAAAAAC  4377526

Query  482      GGCAGACCTAAAGACTGGTGGCTCAGTCAGACTCTCTAGATCCAAAACCCGCTGCAGCCA  541
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377527  GGCAGACCTAAAGACTGGTGGCTCAGTCAGACTCTCTAGATCCAAAACCCGCTGCAGCCA  4377586

Query  542      TCTAAGATCCAAACAGATCTACTGTTAACCACCGCATCAGACAACGAAccccccTTACC  601
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377587  TCTAAGATCCAAACAGATCTACTGTTAACCACCGCATCAGACAACGAACCCCCCCTTACC  4377646

Query  602      CCTCTCCAACTTCCGCCGTGGACGCAGCGACTAGAATCACGGGTGAGCCTGAGAACCTCG  661
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377647  CCTCTCCAACTTCCGCCGTGGACGCAGCGACTAGAATCACGGGTGAGCCTGAGAACCTCG  4377706

Query  662      CAGCAACCCTCAGCGTCCATAACCAAGAGACCCCCGCCGCAACCTTCTCGACCGACAAAA  721
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377707  CAGCAACCCTCAGCGTCCATAACCAAGAGACCCCCGCCGCAACCTTCTCGACCGACAAAA  4377766

Query  722      CCAGAGCCAAATCCACATCAGAGAGAGAGCTCCACACCACACGTTGACGGAGACAAGCAT  781
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377767  CCAGAGCCAAATCCACATCAGAGAGAGAGCTCCACACCACACGTTGACGGAGACAAGCAT  4377826

Query  782      CGGCTCACTCCTGGACTTCCTTGAAGAGGCTTCAAACCGACGATGAAGAGCAGATCGGAA  841
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377827  CGGCTCACTCCTGGACTTCCTTGAAGAGGCTTCAAACCGACGATGAAGAGCAGATCGGAA  4377886

Query  842      CAAAAAGAAAAAGGAAACTAAATCTACTTACCGGAAATTGTCTTCAAGAGAGAGAGNAAG  901
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Sbjct  4377887  CAAAAAGAAAAAGGAAACTAAATCTACTTACCGGAAATTGTCTTCAAGAGAGAGAG-AAG  4377945

Query  902      GGCACGACTTTCAAGGCTACGATTGAGCCTCCATCTCATAAAGCCAGATTCTTTATCTTC  961
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4377946  GGCACGACTTTCAAGGCTACGATTGAGCCTCCATCTCATAAAGCCAGATTCTTTATCTTC  4378005

Query  962      CATAN-CGTANTAAAGCTACGAAAACCTCGATTTCGTGNTCAAA-CTTNGNTTTGCTCGC  1019
                ||||  |||| |||||||||||||||||||||||||||||| ||||| ||| | |||||||||
Sbjct  4378006  CATAATCGTATTAAAGCTACGAAAACCTCGATTTCGTGTTCAAAACTTGGATTTGCTCGC  4378065

Query  1020     TATCGAANGGTCTGACCTTNNN-CCNTTTCATGGGTTNAN-CNTN-AGATATGATTTCTC  1076
                ||||||| |||||||||||    || |||||||||| |  | |  |||||||||||||||
Sbjct  4378066  TATCGAAAGGTCTGACCTTTAATCCTTTTCATGGGTTTAATCGTTCAGATATGATTTCTC  4378125

Query  1077     G  1077
                |
Sbjct  4378126  G  4378126
```

2. List of primers used:

2.1 For genotyping:-

2.1.1 ABI5:

The allele is SALK_013163

The LP and RP primers for *abi5-8*:

LP : CAATGGAAGTTCGGAATCATG
RP : CACTCGTTTTCTTCTTAAAGCG

2.1.2 PYL7:

The allele is SALK_012096

pyl7 insertion is located in the promoter

RP: TCTCATGCTGACATCAGCTAC

LP: ATCTGTATCGTCTCCTCCGATC

2.1.3 ERD4:

SALK_055548.45.45.x
PRODUCT_SIZE 1084

PAIR_ANY_COMPL 0.00 PAIR_3'_COMPL 0.00 DIFF_TM 0.26

LP: TTGAGTGCTGCTGTGTTTGAG
Len 21 TM 60.23 GC 47.62 SELF_ANY_COMPL 0.26 3'_COMPL 0.00

RP: CTTGATGCTCTCGTCCTCATC
Len 21 TM 59.97 GC 52.38 SELF_ANY_COMPL 0.26 3'_COMPL 0.00

Insertion chr3 20180959 BP+RP_PRODUCT_SIZE 475-775

2.1.4 PRT6:

prt6-1 dn3: GTTTCTTGTTCTGGGGAGGATGGTTT

prt6-1 up2: TCAAGCTCAAGTCCCTGTTCTGAC

Garlic LB1: GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC

2.2 qPCR primers:

2.2.1 ABI5:

ABI5_qfor: AAAACCTTTCAGTCAAAATTCTC

ABI5_qrev: CGGTGGCTTTGTGTTCCTC

2.2.2 ERD4:

ERD4_qfor: CTTCTCCGTCCTCAAGAAGC

ERD4_qrev: CCTCTGTCTCTACACTCATC

2.2.3 PYL7:

PYL7_qfor:GGAGATTGAAACAATGGTGGA

PYL7_qrev: CATCGTCTCCTCAGCTTCT

2.3 q-RT-PCR primers:

2.3.1 ERD4

ERD4: Rev: AGGTTGATGCCGGTTATACG

For: TGAAAGAATGGGGAGATGG

2.3.2 PYL7

PYL7: Rev: GATCAAATCTCCGCACCAGT

For: CACCACTGCAGAGAGAACCA

3.3.3 AT3G13440

AT3G13440: Rev: ATCCGCCTTATGTGCCAACA

For: ATTTTCCCCTCCTGCCCAAG

3.3.4 AT1G14810

AT1G14810: Rev: GGCCATGAATACACGGTGGA

For: CCACCAGCACTGAACAAAGC

3.3.5 Control primers q-RT-PCR:

1. ATCTRL1 FW: AGTGGAGAGGCTGCAGAAGA
   REV: CTCGGGTAGCACGAGCTTTA

2. ATCTRL3 FW: GAAGTGTCTCGACAAAGGTCGT
   REV: CCTTTTGGCACTTCTGGTG