

Identification and characterization of ERF1b transcription factor binding motifs and their target genes

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Jote Tafese Bulcha
from Addis Ababa (Ethiopia)

Berlin, 2013

1st Reviewer: Prof. Dr. Margarete Baier

2nd Reviewer: Prof. Dr. Rupert Mutzel

date of defence: 17.07.2013

Dedicated to my mother

Table of contents

SUMMARY	IX
ZUSAMMENFASSUNG	X
LIST OF ABBREVIATIONS	XI
1. INTRODUCTION	1
1.1 TRANSCRIPTIONAL REGULATION OF GENE EXPRESSION	1
1.2 TRANSCRIPTION FACTORS	3
1.3 TRANSCRIPTION FACTOR BINDING SITE	5
1.4 DEFINING TRANSCRIPTION FACTOR BINDING SITES	7
1.4.1 <i>Computational motif prediction</i>	7
1.4.1.1 <i>Computational motif prediction with co-regulated gene</i>	8
1.4.1.2 <i>Computational motif prediction with phylogenetic footprinting</i>	9
1.4.2 <i>Experimental motif determination</i>	9
1.4.2.1 <i>Chromatin immunoprecipitation followed by sequencing (ChIP-seq)</i>	9
1.4.2.2 <i>Protein Binding Microarray (PBM)</i>	11
1.4.2.3 <i>Systematic Evolution of Ligands by EXponential Enrichment (SELEX)</i>	13
1.4.2.4 <i>Other methods to define DNA binding sites</i>	14
1.5 YEAST SYSTEM TO CHARACTERIZE OF DNA-PROTEIN INTERACTIONS	15
1.5.1 <i>Yeast-one-hybrid</i>	15
1.5.1.1 <i>GAL4 system</i>	16
1.5.1.2 <i>LexA system</i>	17
1.6 PLANT TRANSCRIPTION FACTOR FAMILY	18
1.6.1 <i>ERF1b transcription factor</i>	19
1.6.2 <i>Functions for ERF1b transcription factor</i>	21
1.7 CHLOROPLAST RETROGRADE SIGNALING AND ANTIOXIDANT ENZYMES	22
1.7.1 <i>Chloroplast antioxidant enzymes</i>	24
1.7.2 <i>2-cysteine peroxiredoxin-A (2CPA)</i>	25
1.7.3 <i>Chloroplast ascorbate peroxidase</i>	25
1.8 MOTIVATION OF THE PROJECT	26
2. MATERIALS AND METHODS	28
2.1 PLANT MATERIAL	28
2.1.1 <i>Seed sterilization</i>	28

Table of contents

2.1.2	<i>Plant growth condition</i>	28
2.1.3	<i>Mature Plant growth condition</i>	29
2.1.4	<i>Cold treatment of plants</i>	29
2.2	GENE EXPRESSION ANALYSIS	30
2.2.1	<i>RNA extraction</i>	30
2.2.2	<i>First strand cDNA synthesis</i>	30
2.2.3	<i>Quantitative real-time PCR</i>	31
2.2.3.1	<i>Primer Design</i>	31
2.2.3.2	<i>Fluorometry</i>	31
2.2.3.3	<i>Standardization</i>	32
2.3	GENERAL METHODS	32
2.3.1	<i>Spectrophotometric analysis of nucleic acids</i>	32
2.3.2	<i>Ethanol precipitation of DNA</i>	33
2.3.3	<i>Purification of DNA from enzyme reaction mixtures and PCR samples</i>	33
2.3.4	<i>Extraction of DNA from agarose gels</i>	33
2.3.5	<i>Polymerase chain reaction (PCR)</i>	34
2.3.6	<i>Agarose gel electrophoresis</i>	34
2.3.7	<i>TA cloning</i>	35
2.3.8	<i>Competent cell preparation</i>	36
2.3.9	<i>Bacterial transformation</i>	36
2.3.10	<i>Quick DNA isolation from plant</i>	37
2.3.11	<i>DNA sequencing</i>	37
2.4	GENOMIC LIBRARY CONSTRUCTION	38
2.4.1	<i>Genomic DNA isolation</i>	38
2.4.2	<i>Shearing of genomic DNA</i>	39
2.4.3	<i>Testing the quality of sheared DNA</i>	39
2.4.4	<i>End repair of the sheared DNA fragment</i>	40
2.4.5	<i>Prey plasmid construction</i>	41
2.4.6	<i>Harvesting the genomic library</i>	41
2.4.7	<i>Plasmid library preparation</i>	42
2.5	BAIT PLASMID CONSTRUCTION	42
2.6	YEAST MANIPULATION	43
2.6.1	<i>Yeast strain</i>	43
2.6.2	<i>Growth condition</i>	43
2.6.3	<i>Yeast transformation</i>	44
2.6.4	<i>Library scale yeast transformation</i>	45
2.6.5	<i>Yeast colony PCR</i>	45
2.6.6	<i>Yeast plasmid isolation</i>	46
2.6.7	<i>Stringent inverse-yeast-one-hybrid experiment</i>	47

Table of contents

2.7	BIOINFORMATIC ANALYSIS	47
2.7.1	<i>Global and multiple alignment</i>	47
2.7.2	<i>Motif analysis</i>	48
2.7.3	<i>Public microarrays data</i>	48
2.8	EXPRESSION, PURIFICATION AND DETECTION OF RECOMBINANT PROTEINS	49
2.8.1	<i>Bacterial growth conditions</i>	49
2.8.2	<i>Cloning procedures</i>	49
2.8.3	<i>Heterologous expression of proteins in E.coli</i>	50
2.8.4	<i>Purification of recombinant proteins</i>	50
2.8.5	<i>SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)</i>	51
2.8.6	<i>Coomassie staining of protein gels</i>	52
2.8.7	<i>Western blotting</i>	53
2.8.7.1	<i>His-Tag detection with a Ni-NTA-HPR conjugate</i>	53
2.8.7.2	<i>Detection of proteins with specific antibody</i>	54
2.8.8	<i>Determination of protein concentration</i>	55
2.8.9	<i>Antibody development</i>	55
2.9	PLANT PROTEIN EXTRACTION.....	56
2.10	TRANSIENT EXPRESSION OF GENES IN ARABIDOPSIS SEEDLINGS	56
2.10.1	<i>Site directed mutagenesis</i>	56
2.10.2	<i>Reporter plasmid construction</i>	57
2.10.3	<i>LR Reaction</i>	58
2.10.4	<i>Seedling Transfection</i>	58
2.10.4.1	<i>Agrobacterium suspension solution preparation</i>	59
2.10.4.2	<i>Vacuum Infiltration</i>	59
2.10.4.3	<i>Luminescence measurement</i>	60
3.	RESULTS	61
3.1	GENOMIC LIBRARY DEVELOPMENT	61
3.1.1	<i>High molecular weight DNA isolation</i>	61
3.1.2	<i>Shearing genomic DNA</i>	63
3.1.3	<i>Optimization of ligation ratio</i>	66
3.1.4	<i>Sequence analysis of genomic library</i>	68
3.2	STRINGENT INVERSE YEAST-ONE-HYBRID	68
3.2.1	<i>Optimization of library scale yeast transformation</i>	68
3.2.2	<i>Screening for promoter protein binding</i>	70
3.2.3	<i>Conforming the interacting clones</i>	73
3.2.4	<i>Sequence analysis of Rap2.4a and Rap2.4d interacting clones</i>	73
3.3	IDENTIFICATION OF RAP2.4 TARGETS	74
3.4	RAP2.4A.....	75

Table of contents

3.4.1	<i>Rap2.4a</i> targets.....	75
3.4.2	<i>Rap2.4a</i> motif analysis.....	77
3.4.2.1	MEME motif analysis of <i>Rap2.4a</i>	77
3.4.2.2	SCOPE motif analysis of <i>Rap2.4a</i>	78
3.5	RAP2.4D.....	78
3.5.1	<i>Rap2.4d</i> Targets	78
3.5.2	Motif analysis of <i>Rap2.4d</i>	79
3.5.2.1	MEME motif analysis of <i>Rap2.4d</i>	79
3.5.2.2	SCOPE motif analysis of <i>Rap2.4d</i>	80
3.6	EXPRESSION ANALYSIS OF ERF1B GENES	81
3.7	PROMOTER ANALYSIS OF SELECTED GENES	84
3.7.1	Interaction of <i>ERF1b</i> with <i>tAPx</i> and <i>sAPx</i> promoter	88
3.7.2	Expression of <i>tAPx</i> and <i>sAPx</i> in the <i>Rap2.4c</i> knockdown and overexpression lines	90
3.7.3	Interaction of <i>ERF1b</i> with <i>2CPA</i> promoter	91
3.8	RAP2.4C AND RAP2.4D IN COLD REGULATION	93
3.8.1	<i>Rap2.4d</i> negatively regulates the expression of <i>COR6.6</i>	94
3.8.2	<i>Rap2.4c</i> overexpression	96
3.8.3	<i>Rap2.4c</i> negatively regulates the expression of <i>COR6.6</i> and <i>BAP1</i>	98
3.8.4	<i>Rap2.4c</i> as a general negative regulator of cold responsive genes.....	99
3.8.5	Transient expression analysis of <i>COR6.6</i> promoter	101
4.	DISCUSSION	103
4.1	STRINGENT INVERSE YEAST-ONE-HYBRID	103
4.2	SEQUENCE ANALYSES OF RAP2.4A AND RAP2.4D INTERACTING CLONES.....	104
4.3	RAP2.4A.....	105
4.3.1	Targets of <i>Rap2.4a</i>	105
4.3.2	Motif analysis of <i>Rap2.4a</i>	107
4.4	RAP2.4D.....	108
4.4.1	Targets of <i>Rap2.4d</i>	108
4.4.2	Motif analyses of <i>Rap2.4d</i>	110
4.5	EXPRESSIONAL ANALYSES OF ERF1B	110
4.6	REGULATION OF CHLOROPLAST ANTIOXIDANT ENZYMES	111
4.6.1	<i>2-Cys peroxiredoxin-A (2CPA)</i> promoter analysis	111
4.6.2	Interaction of <i>ERF1b</i> with <i>2CPA</i> promoter	112
4.6.3	Chloroplast ascorbate peroxidases promoter analysis	113
4.6.4	Interaction of <i>ERF1b</i> with chloroplast ascorbate peroxidases promoter	115
4.6.5	Regulation of <i>tAPx</i> by <i>Rap2.4c</i>	116
4.7	RAP2.4C AND RAP2.4D IN COLD REGULATION	116
4.7.1	Promoter analysis of <i>COR6.6</i> and <i>Kin 1</i>	116

Table of contents

4.7.2	<i>Negative regulation of COR6.6 by Rap2.4d</i>	117
4.7.3	<i>Negative regulation of COR6.6 and BAP1 by Rap2.4c</i>	119
4.7.4	<i>Rap2.4c as a general negative regulator of cold responsive genes</i>	120
4.7.5	<i>Direct interaction of Rap2.4c and COR6.6 promoter</i>	124
5.	REFERENCES	125
6.	APPENDIX	154
6.1	ERF1B PRIMERS	154
6.2	RAP2.4A TARGETS	155
6.3	RAP2.4D TARGETS	156

Summary

The Ethylene Response Factor (ERF) gene family is a subset of AP2 (Apetala2)/ERF superfamily. It is characterized by exhibiting an AP2/ERF domain of 60 to 70 amino acids involved in DNA binding. In *Arabidopsis thaliana* ERF1b transcription factor family contains eight members (Nakano et al. 2006). All proteins in the group are characterized by a conserved AP2 domain, but non-conserved C- and N-termini. To investigate the functional redundancy or specificity of the transcription factors, the binding sites of Rap2.4a and Rap2.4d were defined. A new method, called stringent inverse yeast-one-hybrid, was developed to screen full genomes for transcription factor binding motifs and to identify transcription factor target genes. It allowed efficient, fast and easy screening of the *A. thaliana* genome, including the organellar genomes, for binding sites of Rap2.4a and Rap2.4d. Of the Rap2.4a and Rap2.4d binding sites detected on the nuclear genome, 88 % and 94.7 %, respectively, were found in the promoter regions of genes, demonstrating a high specificity of the screen for regulatory elements.

The transcript dynamics of Rap2.4d and Rap2.4c were shown to follow similar patterns in response to cold other abiotic stresses, suggesting that the two genes are redundant in function. Expression analysis of cold regulated genes (COR6.6, COR47, COR15A, PLD1) in a Rap2.4c overexpression line showed that Rap2.4c is a general negative regulator of cold responsive genes. Transient reporter gene activity analysis demonstrated that Rap2.4c exerts its regulatory function on the promoter region of cold responsive genes. Yeast-one-hybrid assay and analysis of T-DNA insertion lines showed that Rap2.4a and Rap2.4h regulate the expression of 2CPA antagonistically, possibly by competing for the same binding site. However, Rap2.4h is a stronger competitor than Rap2.4a as demonstrated by vigorous binding of 2CPA promoter. Promoter analysis using the identified motifs, yeast-one-hybrid assay and expression analysis of T-DNA insertion lines indicated that Rap2.4a and Rap2.4d positively regulate expression sAPx and tAPx.

Zusammenfassung

Die Ethylene Response Faktoren (ERF) Gene gehören zur AP2 (Apetala2)/ERF Superfamilie. Diese ist durch eine AP2/ERF DNA Bindedomäne charakterisiert, die aus 60-70 Aminosäuren besteht. Die Familie der ERF1b Transkriptionsfaktoren beinhaltet 8 Mitglieder (Nakano et al. 2006) mit konservierter AP2 Domäne und nicht-konservierte C- und N-Termini. Um eine mögliche funktionelle Redundanz von der Spezifität der Transkriptionsfaktoren abzugrenzen, wurden die Bindemotive der Faktoren Rap2.4a und Rap2.4d identifiziert. Hierfür wurde eine neue Methode, das sogenannte „Stringente inverse Hefe-1-Hybrid-System“, entwickelt. Es ermöglicht eine genomweite Überprüfung von Transkriptionsfaktor-Bindestellen und zeichnet sich durch eine einfache und effiziente Handhabung aus. In der vorliegenden Arbeit wurde es zur Identifizierung von Rap2.4a- und Rap2.4d-Bindestellen in *Arabidopsis thaliana*, einschließlich der Organellgenome, eingesetzt.

88 % bzw. 94.7 % der in genomischer DNA identifizierten Rap2.4a- und Rap2.4d-Bindemotive liegen in Promotorabschnitten. Dies bestätigte die hohe Spezifität der Methode. Die Transkriptspiegel von Rap2.4d und Rap2.4c werden durch Kälte und anderen abiotischen Stressformen ähnlich reguliert, was für eine funktionelle Redundanz der Gene spricht. Expressionsanalysen zu den kälteregulierten Genen COR6.6, COR47, COR15A und PLD1 in Rap2.4c-Überexpressionlinien zeigten, dass der Transkriptionsfaktor einen generellen, reprimierenden Einfluss auf kälteregulierte Gene hat. Hefe-1-Hybrid Resultate sowie T-DNA Insertionsanalysen zeigten, dass Rap2.4a und Rap2.4h die Expression von 2CPA antagonistisch regulieren. Möglich wäre eine Konkurrenz um die gleiche Bindestelle auf dem 2CPA-Promotor, wobei Rap2.4h eine stärkere Bindeaffinität hat als Rap2.4a. Promotoranalysen, Hefe-1-Hybrid Experimente und Expressionsanalysen lassen darauf schließen, dass Rap2.4a und Rap2.4d die Expression von sAPx und tAPx positiv regulieren.

List of abbreviations

2CP	2-cys peroxiredoxin
2CPA	2-cys peroxiredoxin A (At3g11630)
3-AT	3-amino-1,2,4 triazole
3-PGA	3-Phosphoglyceric acid
ABA	abscisic acid
ABF	ABRE binding factor
ABI	Aba insensitive
ABRC	ABA response complexes
ABRE	ABA responsive element
AD	activation domain
AP2	Apetala 2
APx	ascorbate peroxidase
ATP	Adenosine triphosphate
BD	DNA-binding domain
bp	Base pairs
bHLH	basic helix-loop-helix
bZIP	Basic region leucine zipper
cDNA	copy-DNA
CBF	C-repeat-binding factor
CE	coupling element
CBF	cold binding factor

List of abbreviations

Col-0	<i>Arabidopsis thaliana</i> ecotype Columbia-0
COR	cold regulated
ChIP-seq	chromatin immunoprecipitation followed by sequencing
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DRE	drought response element
DREB	drought-responsive element binding factor
DTT	dithiothreitol
EDTA	ethylene diamine tetra acetic acid
EMSA	electrophoretic mobility shift assay
ERE	Ethylene response element
ERF	Ethylene response factor
ERH1	Enhancing PRW8-mediated HR
H ₂ O ₂	hydrogen peroxide
HR	hypersensitive response
HRP	horse-radish peroxidase
GATL2	Galacturonosyltransferase-like2
ICE	Inducer of CBF expression
IP	immunoprecipitation
IPTG	Isopropyl-β-D-thiogalactoside
kb	Kilo base pairs
KD	Knock-down
KDa	kilo Dalton
KO	Knock-out
LUC	Luciferase
MAPK	mitogen activated protein kinase
MCS	multiple cloning sequence
MEME	Multiple EM for Motif Elicitation
mM	millimolar
MS	Murashige and Skoog media

List of abbreviations

MSA	multiple sequence alignment
NADH	nicotinamide adenine dinucleotide
NADPH	nicotinamide adenine dinucleotide phosphate
NCBI	national center for biotechnology information
NCED	cis-epoxycarotenoid dioxygenase
NLS	nuclear localisation signal
NPQ	non-photochemical quenching
LB	laura Bertani medium
O ₂ ⁻	superoxide radical
PAGE	polyacrylamide gel electrophoresis
PBM	protein binding microarray
PCR	polymerase chain reaction
PDA	Potato dextrose agar
PEG	poly ethylene glycol
PLD1	phospholipase D1
PQ	plastoquinone
PRx	peroxiredoxin
PTI	pattern-triggered immunity
PWM	position weight matrix
qRT-PCT	quantitative real time polymerase chain reaction
ROS	reactive oxygen species
RNA	ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
sAPx	stromal ascorbate peroxidase (At4g08390)
SD	synthetic dropout
SDM	site directed mutagenesis
SDS	sodium dodecyl sulfate
SOD	superoxide dismutase
SELEX	Systematic Evolution of Ligands by Exponential Enrichment
SPR	surface plasmon resonance

List of abbreviations

SRX	Sulfiredoxin
TA	Annealing temperature
TAE	Tris acetic acid EDTA
tAPx	thylakoid-bound ascorbate peroxidase (At1g77490)
TBE	Tris-borate EDTA
TBP	TATA-binding protein
TBS	Tris buffered saline
TE	Tris-EDTA
Tris	Tris(hydroxymethyl)-aminomethane
TRX	thioredoxin
UAS	upstream activating sequence
WIND1	wound induced dedifferentiation
Y1H	yeast-one-hybrid
YEPD	Yeast Extract Peptone Dextrose
YPD	Yeast Extract-Peptone-Dextrose

1. Introduction

1.1 Transcriptional regulation of gene expression

Strict control of gene expression is the first pivotal step in realization of genomic codes into life. Gene transcription is an elaborated multistep process that involves transcription initiation, elongation and termination. As reviewed by Maniatis and Reed (2002), contrary to what used to be known as a simple and linear assembly line (Proudfoot 2000), gene transcription requires synchronization of a complex and extensively coupled network as depicted in the **figure 1**.

As gene expression can be controlled at multiple levels, integration of diverse processes is required (Maniatis and Reed 2002, Orphanides and Reinberg 2002, Roeder 1991). Despite very basic similarities between prokaryotes and eukaryotes in terms of transcriptional gene regulation, the regulation machinery in eukaryotes is far more advanced and complex. In general, realization of the wealth of information which is encoded in the genome requires three fundamental regulation steps, i.e. transcriptional, post-transcriptional, and post-translational regulation (**Figure 1**).

Introduction

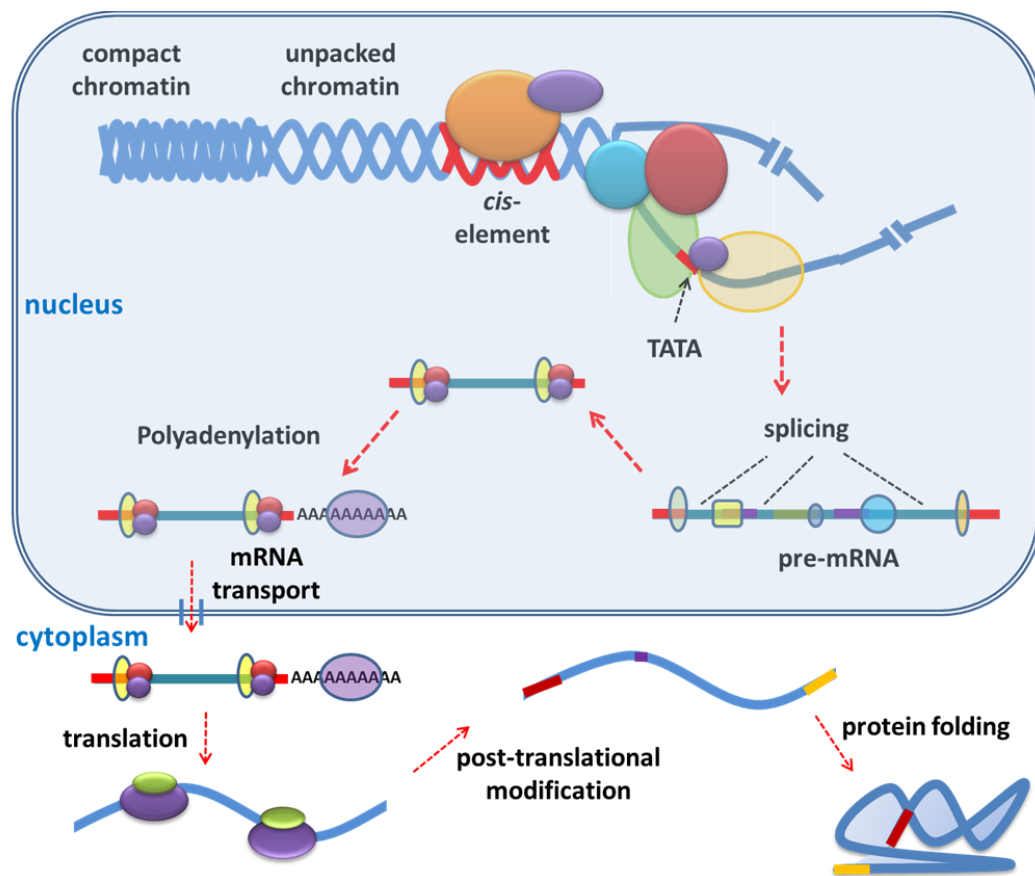


Figure 1: Depiction of the current understanding of gene expression. The red arrows indicate that each step is a part of the continuous process. Individual steps in the continuous process are subjected to regulation (transcriptional, post-transcriptional, translational and post-translational) which affect the final outcome. The steady-state transcript abundance is the outcome of gene transcription rate and mRNA decay. The mRNA available for translation is the result of mRNA processing, export and localization processes. The amount of translated protein depends on the transcript abundance and the rate of translation. The activity of translated protein is subjected to regulations (degradation, phosphorylation, glycosylation, oligomerization). The long blue lines indicate DNA, RNA or proteins. Various shape and color of circles and boxes represent a large array of proteins involved in DNA binding, protein-protein interaction, catalytic and different other activity.

Owing to much larger genomes as compared to prokaryotes, multicellular eukaryotes possess genes organized in different chromosomes whose organization itself is complex and can determine which genes can be transcribed (Schneider and Grosschedl 2007, Van Bortle and Corces 2012). Gene transcription in eukaryotic organism is mainly modulated at the transcription initiation step (Jiang *et al.* 1993, Jiang *et al.* 1996b, Roeder 1991), although

succeeding steps can be a point of regulation in some instances. Given that genes in the eukaryotic organism are constrained inside the nucleus by associating with histones to form chromatin, transcription initiation would require a mechanism to circumvent the structural hurdle (Kornberg and Lorch 2002). In contrast to naked DNA, the presence of chromatin by being a structural barrier can repress transcription, which might be essential for controlling the expression of a gene tightly *in vivo* (Ioshikhes et al. 1999). Chromatins are organized in form of nucleosomes with beads-on-a-string format which is presumed to be the transcriptionally active part of a chromosome (Venters and Pugh 2009). In a response to varieties of environmental and cellular cues, sequence specific proteins interact to their target site on the DNA. In combination with additional coactivators they form the transcription initiation complex that in turn changes the local chromatin environment and makes it accessible for transcription. The cofactors that have been recruited by the sequence specific binding factors interact directly and indirectly at the core promoter to modulate the activity of RNA polymerase II which does the actual transcription process (Kadonaga 2004). Yet another layer of regulation can be achieved by selective methylation of histones which might serve in the repression of transcription. By recruiting repressive DNA methylases and histone deacetylases that eliminate histone acetylation marks which would otherwise serve as activation mark, genes which are to be transcribed can be determined (Meyer 2000, Venters and Pugh 2009).

1.2 Transcription factors

Genome sequencing of various organisms unraveled that the genomic size and the gene number cannot alone explain the complexity of organisms as explained by C-value enigma (Gregory 2001, Gregory 2005). Eukaryotic organisms have a highly evolved mechanism to monitor the expression of their genes compared to prokaryotes. The major control of gene expression relies on a group of DNA binding proteins called transcription factors, which offer multicellular eukaryotic organism a fluidity of determining their cell types and growth patterns in a different way. 5-10 % the total protein encoded in the genome of most organisms are dedicated to proteins that regulate transcription (Levine and Tjian 2003).

Transcription factors that activate the expression of certain genes need to interact with the chromatin structure of the cognate gene to loosen the repressive chromatin structure (Narlikar *et al.* 2002). Expression of genes is therefore regulated by transcription factors that bind normally to the upstream DNA sequence of the gene to be transcribed which constitutes the regulatory segment of that particular gene (Lemon and Tjian 2000). Following the interaction of transcription factors with the promoter or the enhancer region of a particular gene, they recruit other factors and RNA polymerase II (McKenna and O'Malley 2002). Binding of transcription factors, however, can lead not only to activation but also to repression of gene expression, contingent upon the presence of other functional domains on the transcription factor and the resultant impact of the whole transcription factor and other protein complex. A typical transcription factor has multiple functional domains whose utility can be DNA binding, transcription activation, binding of signaling molecules and interaction with other proteins including RNA polymerase II. The modularity of a transcription factor allows the combination of various functionally distinct domains to form transcription factors with diverse functions (Schwechheimer *et al.* 1998). The structural motifs of transcription factors that interact with DNA are diverse. Hence, transcription factors can be classified based on the DNA binding domain they carry. As a consequence, a given transcription factor family often has highly conserved residues that are necessary for binding to the target DNA (Huang *et al.* 1996). Despite the fact that transcription factors usually have only one type of DNA binding domain, transcription factors with two different type of binding domain is rarely possible (Suzuki *et al.* 1997).

Being proteins that are encoded in the genome, transcription factors themselves are subjected to many levels of control. For instance, transcription factor autoregulation which includes positive feedback loop regulation, RNA splicing, degradation of mRNA, translation, post translation modification by phosphorylation, ligand dependent activation, protein-protein interaction (homo and heterodimerization) can be a point of regulation (Calkhoven and Ab 1996). It has been demonstrated that, for example, the activity of Rap2.4a transcription factor is determined by the redox status of a cell which affects the quaternary structure (Shaikhali *et al.* 2008).

Transcription factors have to be transferred from the cytoplasm to the nucleus which is a vital step for posttranslational control (Laskey and Dingwall 1993). Even though small proteins (<40-60 kDa) can be imported by simple diffusion, a nuclear localization signal (NLS) is needed for most transcription factors (Dingwall and Laskey 1991, Raikhel 1992). NLSs are characterized by having arginine and lysine rich residues and can be grouped into SV40 large T type antigen NLS (PKKKRKV), yeast mating type factor (Mata α 2) like NLS (KIPIK) and NLS with bipartite structure (Schwechheimer *et al.* 1998).

1.3 Transcription factor binding site

Coding regions account only for less than 2% of the total DNA in most eukaryotic organisms. The majority (80.4 %) of the non-coding part of the human genome, for instance, is involved in at least one biochemical process (Dunham *et al.* 2012). Transcription factor binding sites/motifs/*cis*-elements are functional regulatory parts of a genome that modulate the spatial and temporal transcription activity (Priest *et al.* 2009). To date, it is hardly possible to determine the *cis*-regulatory element content of a genome but it was estimated that as much as a third of the human genome actively plays a regulatory role as reviewed in Levine and Tjian (2003). The interaction of a given transcription factor to a specific region on the DNA of the genomes is a vital part of transcription orchestration. DNA sequence motifs (usually between 6 - 10 bp) have the affinity to transcription factors and often present in the flanking regions of a gene to provide specificity to the core transcriptional machinery (Wray *et al.* 2003). Upstream regulatory sequences of most multicellular eukaryotic organisms have binding sites for multiple transcription factors thereby permitting the cognate gene expression to be monitored by several signaling cascades (McKenna and O'Malley 2002). The upstream regulatory regions of a gene, also called promoter, can be arbitrarily divided into a proximal and a distal part (**Figure 2**).

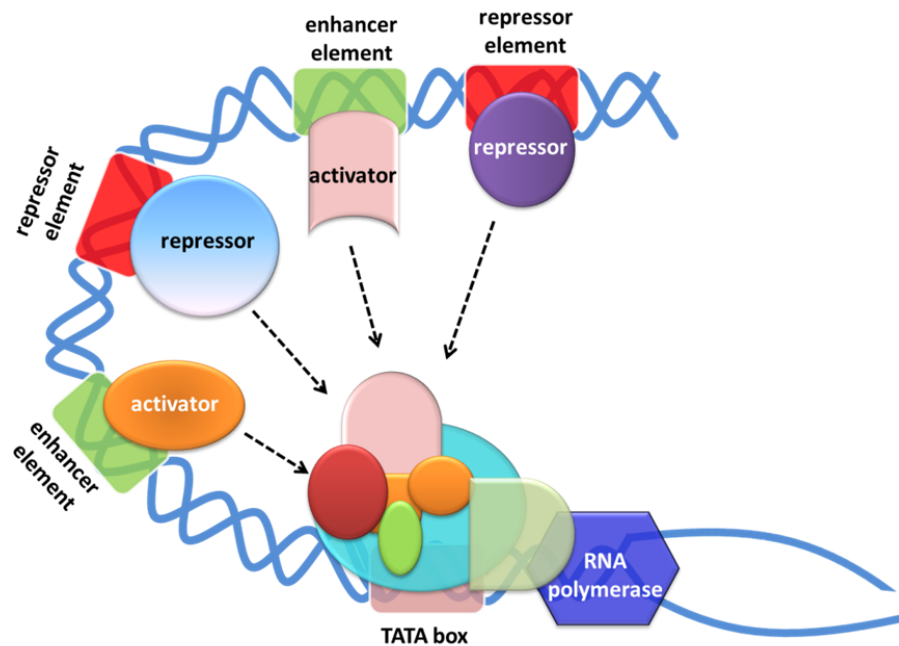


Figure 2: A complex multicellular model of transcriptional control. Many DNA-protein and protein-protein interaction is required to execute the transcription of a single gene. Transcription factors specifically bind to DNA sequences and interact with other proteins to initiate transcription. The functionally important *cis*-regulatory elements on the DNA are marked in colored boxes. Co-activators or repressors may be involved in the stabilization and destabilization, respectively, of the transcription initiation complex.

Transcription initiation complexes, which include RNA polymerase II and can drive a basal level of transcription, are presumed to be assembled on proximal part of the promoter (Nikolov and Burley 1997). The control of gene expression on the proximal promoter often relies on the DNA sequences that are present in the upstream vicinity of transcription start site which usually contain TATA elements, upon which TATA binding proteins can interact (Featherstone 2002, Struhl *et al.* 1998). The spatial and temporal gene expression pattern is regulated by the DNA elements that are present in the distal part of the promoter (Fessele *et al.* 2002, Tjian and Maniatis 1994). Additionally, many genes have one or more sequences that modulate the activity of transcription; they are known as enhancers / repressor elements. Enhancers / repressive elements can be present upstream, downstream or even within introns and can be as far as several thousand base pairs away from the gene they modulate (Bagga *et al.* 2000, Tjian and Maniatis 1994, Wasserman *et al.* 2000). For fine tuning the transcription activity, the effect of binding sites with low transcription factor

affinity can be compensated by an adjacent strong binding site as this permits the positioning of additional transcription factor via protein-protein interaction (Struhl 2001).

1.4 Defining transcription factor binding sites

Upon the culmination of Human Genome project and recently many more genome sequencing projects, the need to annotate and functionally characterize all the sequencing data has become imperative. Big leap of progress has been achieved with regard to characterization of the coding regions of a genome, which accounts merely 2 % or less (Dunham *et al.* 2012b). The challenge comes with the rest of the genome which does not code for proteins. Its characterization, however, is crucial to understand the whole architecture of the regulatory network. As part of transcriptional regulatory network, transcription factors play a very important role in temporal and spatial coordination of gene expression. Knowledge of transcription factor binding sites is of immense importance for understanding and predicting their tasks. Nevertheless, only a small number of transcription factor binding sites have been identified mainly due to technical limitations. Given regulatory sequences are evolutionarily conserved among closely related taxa, bioinformatic prediction of the *cis*-regulatory elements supported the identification of transcription factor binding site (Duret and Bucher 1997, Loots *et al.* 2000, Pennacchio and Rubin 2001). In addition, within a given organism comparison of the upstream sequences of genes with a specific spatially and/or temporally regulated expression pattern could also hint at the underlying regulatory elements (Lockhart and Winzeler 2000, Suzuki *et al.* 2005, Wasserman and Fickett 1998). Nonetheless, experimental validation of the predicted *cis*-regulatory elements and linking them to their transcription factors is cumbersome.

1.4.1 Computational motif prediction

Given the central importance of transcription factor binding sites in delineating gene regulatory pathway, there are ongoing attempts to determine the binding sites computationally (Bulyk 2004, Frith *et al.* 2004, Hannenhalli 2008, Hawkins *et al.* 2009). The

general binding site preferences of transcription factors are used to determine the binding sites computationally. These properties, such as the width of transcription factor binding site and nucleotide preference at each position, can be expressed in position weight matrix (PWM) (Stormo 2000). A PWM for a given transcription factor can be deduced by aligning experimentally determined or computationally predicted regulatory sequences that are assumed to be bound by the same transcription factor. Computational methods in predicting the underlying regulatory elements require the estimation of likelihood of occurrences of short DNA signature (usually 6- 10 bp) in a given set of DNA sequences that contain the regulatory element as compared the likelihood of such sequence in a sequence of background organism (Hudson and Quail 2003, Thijs *et al.* 2001, Tompa *et al.* 2005). Since the DNA binding motifs are usually short, it is expected that they are present at every thousands if not hundreds of base pairs throughout the genome. The biggest challenge is thus to identify the true binding site in a set of predicted sequences including many false ones (Hudson and Quail 2003). Once a DNA binding motif is predicted, it is fairly simple to scan the genome to identify the binding locus and compare with other regulatory elements which can be done with plant specific databases including PlantCARE (Lescot *et al.* 2002) and PLACE (Higo *et al.* 1998, Higo *et al.* 1999).

1.4.1.1 Computational motif prediction with co-regulated gene

Genome wide differential gene expression analysis such as microarray and RNA deep sequencing (RNA-Seq) provide data for computationally predicting the underlying *cis*-regulatory elements. Using such experiments, sets of genes that are differentially co-expressed can be determined. Assuming that co-expressed genes are co-regulated, the data can be used in analyzing the putative promoter regions of co-expressed genes to find the consensus / over represented short motif (Priest *et al.* 2009). Many computational algorithms have been developed and used to delineate putative *cis*-regulatory elements in a set of co-expressed genes (Rombauts *et al.* 2003, Tompa *et al.* 2005). However, there are drawbacks with this approach to determine the transcription factor binding site in that co-expressed genes are not necessarily co-regulated by the same transcription factor. Since

most of transcription profiling experiments measure the steady state transcript level, it is hardly possible to identify those changes caused by *cis*-regulatory element mediated transcriptional change from those caused by post-transcriptional regulation as reviewed by Priest *et al.* (2009).

1.4.1.2 Computational motif prediction with phylogenetic footprinting

The approach that aims to determine *cis*-elements by comparing the regulatory sequence of orthologous genes is called phylogenetic footprinting (Koop 1995, Wasserman *et al.* 2000). It is well established that genes belong to a larger family, which have been evolved through the process of speciation and gene duplication events. Orthologs are homologous genes that are present in separate species but often with the conserved functionality whereas paralogs are homologous genes present on the same species but usually with a different function (Mindell and Meyer 2001). As the species continue to evolve, if the regulatory sequence of orthologous gene remains to be conserved, the underlying sequences are presumed to carry the functionally important element. Hence, comparing the sequence among orthologs can predict the regulatory element (Loots *et al.* 2000). However, it is not a simple task to choose orthologous regulatory sequences for comparison, since high sequence similarity between the two sequences is less likely to help in the identification of the functional regulatory element. In addition, comparing regulatory sequences from distantly related taxa would result in missing the underlying regulatory element.

1.4.2 Experimental motif determination

1.4.2.1 Chromatin immunoprecipitation followed by sequencing (ChIP-seq)

Chromatin Immunoprecipitation followed by sequencing (ChIP-seq) is one of the newly educated and recently widely used approaches (Park 2009, Raha *et al.* 2010, Robertson *et al.* 2007), which distinguish motifs at nucleotide level resolution (Mokry *et al.* 2010). For ChIP-seq, DNA interacting proteins are crosslinked to their binding sites *in vivo* using agents such

Introduction

as formaldehyde (**Figure 3**). Once the transcription factors are fixed, a series of cell lysis and washing steps will follow to remove the contaminating cellular debris. The isolated chromatin is then mechanically sheared to a smaller fragment size (200-600 bp) to facilitate the binding site identification. The resulting protein-DNA complexes are precipitated with the help of specific antibodies against the DNA binding proteins of interest (Bortz and Wamhoff 2011). Before sequencing, the target fragments have to be amplified by constructing a library. During sequencing library construction, the immunoprecipitated DNA is subjected to size selection. Usually a range of 150-300 bp is chosen.

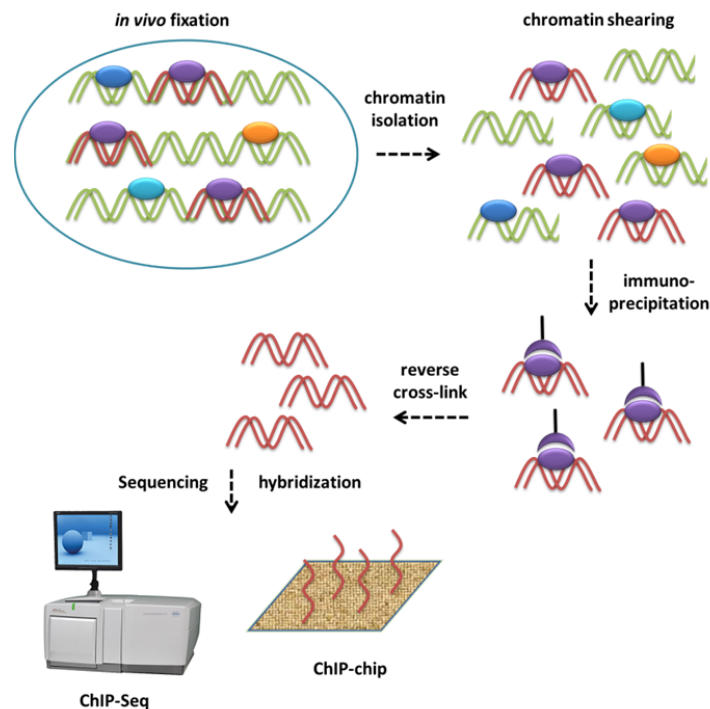


Figure 3: Workflow depiction of ChIP-seq and ChIP-chip analysis. The transcription factor of interest is fixed to its *in vivo* binding site; the chromatin is isolated and sheared to small fragment. Following immunoprecipitation of the target fragment, the DNA is reverse cross-linked and either hybridized to a chip or sequenced to identify tagged sequences.

Before the introduction of the latest steps of ChIP-seq, the fragmented and immunoprecipitated DNA is used to be hybridized to a high density microarray (ChIP-chip) that would also allow genome scale DNA-protein interactions detection (Pillai and

Chellappan 2009). In ChIP-Seq assays, the immunoprecipitated fragments are rather sequenced instead of binding them on microarray. Following immunoprecipitation, although expensive, sequencing the fragments offers much more resolution, greater coverage and fewer artifacts than detecting it on microarray as reviewed by Ho *et al.* (2011). Besides identification of the DNA binding sites, ChIP-seq can point out epigenetic landscapes, such as histone modification (Barski *et al.* 2007) and DNA methylation (Li *et al.* 2011).

Nonetheless, there are inherent technical challenges associated to the routine application of this method. For instance, ChIP-seq grade antibodies have to be developed. Their sensitivity and specificity has to be tested and ascertained beforehand as they significantly determine the enrichment level of the bound fragments and the background noise. For each transcription factor whose motif is to be elicited, one may need prior knowledge of the peak expression time and/or actively expressing tissue to predestinate the binding state of the transcription factor to its DNA target. Prior to target identification, the bound DNA sequences have to be amplified and target libraries need to be developed as well as validated. Furthermore, given the bulk amount of data, rigorous data analysis is required to differentiate directly and indirectly captured sequences. Owing to the rapid evolution of sequencing technology, the cost of running ChIP-seq experiment has decreased, although it is still expensive.

1.4.2.2 Protein Binding Microarray (PBM)

Protein Binding Microarray (PBM) is a widely used *in vitro* approach to define transcription factor binding sites (Bulyk 2007, Mukherjee *et al.* 2004b). In PBM experiments, the transcription factor of interest is fused to a protein that can be easily detected with an antibody-fluorophore conjugate (**Figure 4**). The chimeric transcription factor is hybridized to oligonucleotide microarrays or arrays of promoter (upstream) sequences. Given that the transcription factor binding sites are by average not longer than 10 bp, a universal protein binding microarray with all possible 10 bp nucleotide combination supports the identification of the binding preferences of a transcription factor (Berger *et al.* 2006,

Philippakis *et al.* 2008). In another approach to the genome wide transcription factor binding specificities, the whole-genome yeast intragenic microarrays, for instance, have been developed (Mukherjee *et al.* 2004a).

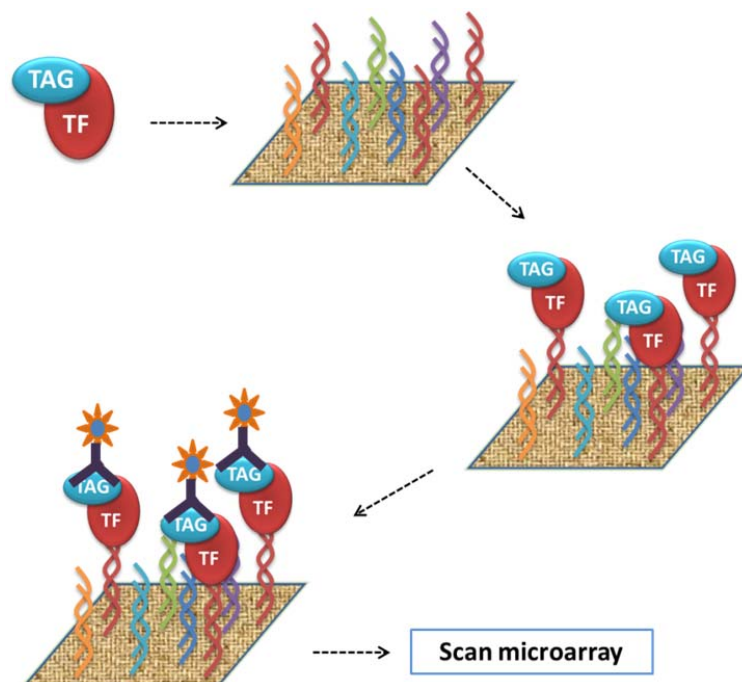


Figure 4: Depiction of the workflow of protein binding microarray. Transcription factor is expressed in the fusion format with antibody detectable tag protein. The fusion protein is hybridized to a library of double stranded oligos or promoter fragments. After a washing step, the target fusion protein is immunodetected by fluorophore bound antibodies and the chip is scanned.

It has been demonstrated that for some transcription factors the binding preferences at one position of the binding site depends on which nucleotide is present on the next position (Bulyk *et al.* 2002, Man and Stormo 2001). The binding site of a transcription factor with such property may not be adequately represented in the traditional mononucleotide position weight matrices (PWMs). Since all the nucleotide binding site combinations are equally represented in universal protein binding microarray, it is possible to detect the transcription factor binding preference dependent and independent on the nucleotide that occupies another position on the binding site (Bulyk *et al.* 2002). The importance of this

method is indispensable in determining not only the binding site but also the binding affinity toward mutations in the nucleotide sequence of a given motif, but has its own limitation. In addition to costs associated with instrumentation, sub-optimal buffer condition might affect protein folding and, consequently, the binding specificity of transcription factors. The results are not always reproducible *in vivo*, due to the typical pitfalls of an *in vitro* experiment.

1.4.2.3 Systematic Evolution of Ligands by EXponential Enrichment (SELEX)

Systematic Evolution of Ligands by EXponential Enrichment (SELEX) is another approach that determines the affinity of a transcription factor binding to a DNA in its inception (Chai *et al.* 2011, Ellington and Szostak 1990, Lorenz *et al.* 2006, Tuerk and Gold 1990). It has been an *in vitro* technique widely used to define transcription factor binding sites (Ellington and Szostak 1990, Liu and Stormo 2005). Generally SELEX is considered as an affinity chromatography technique to select specific short DNA sequences starting from a large pool of oligonucleotide library. The transcription factor of interest is bound to an inert matrix to which DNA fragments are applied in multiple rounds. In a series of washes, non-binding DNA-fragments are washed off. After each washing step, the selected oligonucleotides are amplified by PCR. One round of target binding, selection through washing and amplification via PCR is called SELEX round. The remaining fragments are finally eluted and enriched, the identity of which can then be determined by sequencing the enriched fragment. The size of starting oligonucleotide library is typically large (10^{15} - 10^{16} bp) consisting of random sequences of 20-30 bp (Djordjevic 2007). Alternatively, instead of artificially synthesized short DNA fragments as a starting material, genomic libraries with short fragment size can be used to scan the binding preference of the transcription factor at a genome scale, which is called genomic SELEX (Lorenz *et al.* 2006). Genomic SELEX was successfully employed to characterize DNA and RNA sequence targets by a transcription factor (Kim *et al.* 2003, Shimada *et al.* 2005). Despite the importance of the method, the routine application has been limited due to reproducibility problem and other difficulties inherent to the *in vitro* binding condition.

1.4.2.4 Other methods to define DNA binding sites

Other *in vitro* techniques, such as DNase I Footprinting can be used to study protein-DNA interactions (Fox *et al.* 1999, Hampshire *et al.* 2007). In the DNase I Footprinting assay, a fragment of DNA having a potential binding site is amplified by PCR with radioactive or fluorescence labeled nucleotides. The protein whose binding site is to be determined is bound to labeled DNA fragment. Cleavage agent such as DNase I is added to the mixture, so that the unbound DNA fragment is digested leaving the binding site intact. Although, theoretically simple, DNase footprinting assays that display the sequence of DNA protected from nuclease attack relative to the exposed regions of DNA are technically cumbersome and laborious.

Electrophoretic Mobility Shift Assay (EMSA) is a rapid technique to determine specific protein-DNA interaction (Garner and Revzin 1981, Garner and Revzin 1986, Hellman and Fried 2007). The principle of EMSA is based on the fact that the electrophoretic mobility of protein-DNA complexes is less than that of unbound DNA. In a typical EMSA experiment, a transcription factor is mixed with a radiolabeled DNA fragment having a putative binding site in it. The mixture is electrophoretically separated and the resulting distribution of the fragments is autoradiographically detected. EMSA is often used to qualitatively determine specific protein-DNA interactions, although binding stoichiometries can sometime be measured quantitatively (Fried 1989). The method requires optimization for each sample to be tested as chemical equilibrium is not guaranteed during electrophoresis. It has been shown that many protein-DNA complexes are significantly more stable in the gel than the actual condition *in vivo* (Fried and Bromberg 1997). Additional drawback of the method is that it does not give direct information on the exact sequences of binding nucleotides.

Surface Plasmon Resonance (SPR) can be used to test the binding affinity of a transcription factor (Jost *et al.* 1991). This is a biophysical method that relies on the fast optical systems to quantitatively determine protein-DNA association. SPR determines the interaction strength between two biomolecules by immobilizing one of them on a chip surface and detecting the

change in the refractive index of the chip as the other molecule binds. This allows the visualization of the interaction in real time (Majka and Speck 2007).

In general, these methods cannot be used in genome wide screening of binding sites, since they focus on specific DNA regions.

1.5 Yeast system to characterize of DNA-protein interactions

1.5.1 Yeast-one-hybrid

Yeast-One-Hybrid is a technique that permits the characterization of the binding specificity of heterologous transcription factors against a putative regulatory sequence of interest (Vidal and Legrain 1999). In its inception, yeast-one-hybrid is conceptually derived from the yeast-two-hybrid system. The principle of yeast-one-hybrid is based on the fact that transcriptional activation requires a complex of two physically separable proteins/domains that do the function of DNA binding and transcription activation (Hope and Struhl 1986, Keegan *et al.* 1986). In a typical yeast-one-hybrid assay, a cDNA expression library of transcription factors or other DNA binding proteins are screened for the interaction with a *cis*-acting regulatory element that is linked to the reporter gene (**Figure 5**). In its traditional way, yeast-one -hybrid is a gene centered approach in which single or multiple copies of only one regulatory sequence at a time is analyzed for its affinity towards a transcription factor. In addition the method can be utilized to map the DNA binding domain of previously uncharacterized transcription factors. The cDNA of a candidate protein to be tested or a library of proteins to be screened is usually cloned in an *E. coli*-yeast shuttle vector thereby creating a fusion protein with a strong *trans*-activation domain (AD) of usually a yeast transcription factor. As compared to the *in vitro* techniques of characterizing protein-DNA interaction, heterologous expression of the transcription factor avoids improper protein folding and allows conditions for post-translational modification, as the DNA-protein interaction is housed in the yeast system (Deplancke *et al.* 2004, Li and Herskowitz 1993, Ouwerkerk and Meijer 2001, Reece-Hoyes *et al.* 2011).

A number of yeast strains are available that are suitable for yeast-one-hybrid screening. The choice of the strain depends on the type of the reporter gene used in the vector system, and the sensitivity of the assay. Since selection of successful transformants and interacting partners is based on nutritional auxotrophy, the yeast strains used in the system were developed by deleting the auxotrophic genes, without the external supplement of which the mutant cannot survive. The common nutritional selection markers genes used in the yeast based screening system are genes involved in adenine (Ade2), histidine (His-3), leucine (Leu2), tryptophan (Trp1) and uracil (Ura3) biosynthesis.

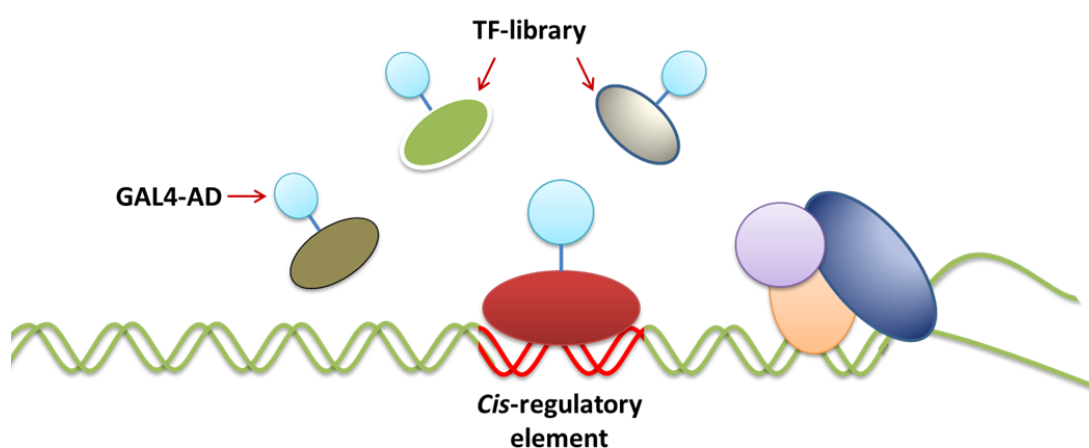


Figure 5: Schematic representation of yeast-one-hybrid. A cDNA library of transcription factors is cloned and expressed in fusion format with GAL4 activation domain. Defined *cis*-regulatory elements are cloned upstream of the reporter gene. Upon binding, the GAL4 activation domain drives the expression of the auxotrophic reporter gene.

1.5.1.1 GAL4 system

The modular nature of the yeast GAL4 transcription factor promoted the development of yeast based systems in detecting protein-protein and protein-DNA interaction. In its native function, GAL4 is a transcriptional activator of galactose/melibiose in *Saccharomyces cerevisiae* (Johnston *et al.* 1986). Therefore, the yeast strain used in the screen uses the GAL4 system and should, thus, have its GAL4 and GAL80 genes mutated. As a consequence,

such strains are slow in growth under normal growth conditions (Van Criekinge and Beyaert 1999). In GAL4 yeast-one-hybrid system, library cDNAs are cloned in frame such that the transcription factor-GAL4 activation domain fusion proteins are produced under control of the strong ADH1 yeast promoter. ADH1 promoter regulates the expression of alcohol dehydrogenase in its native function. HIS3 and LacZ are typically used auxotrophic and colorimetric interaction markers, respectively. HIS3 encodes imidazoleglycerol-phosphate dehydratase, which catalyses a key step in histidine biosynthesis. Since His3 has some level of leaky expression in most yeast strains, the background expression has to be controlled. To suppress the background expression due to non-specific protein-DNA binding and/or the leakiness of GAL4 minimal promoter by activation of yeast endogenous transcription factor, 3-amino-1,2,4-triazole (3-AT) is added to the growing media. 3-AT is a competitive inhibitor of imidazoleglycerol-phosphate dehydratase (HIS3 gene). Using His3 as an interaction marker has a merit in that by adjusting the concentration of 3-AT, it is possible to screen for weakly or strongly interacting partner proteins and *cis*-regulatory elements.

Another reporter gene usually used in combination with another auxotrophic reporter is LacZ. It is an *E. coli* gene that encodes β -galactosidase which catalyzes the cleavage of the colorless X-gal substrate to produce galactose and a blue colored product. This property is utilized to quantitatively determine the interaction strength of bait and prey construct.

1.5.1.2 LexA system

The LexA system is most frequently used in yeast-two-hybrid experiments. It uses the DNA binding domain of the LexA *E. coli* repressor protein in concert with the activation domain of the GAL4 protein (Brent and Ptashne 1985). Given that LexA is not endogenous to yeast it gives less false positive as compared to the GAL4 system. When using this system, the nuclear localization signal has to be cloned in frame as LexA is a bacterial protein and does not have such tag as in the case of GAL4 protein.

1.6 Plant transcription factor family

Transcription factors play a pivotal role in multicellular eukaryotic organisms, with respect to coordinated temporal and spatial gene expression regulation and in specifying the identity of a cell and its response to the multitudes of biotic and abiotic cues. Being central to the regulatory process of gene expression, transcription factors are important targets of adaptation. Evolutionary forces led to the formation of distinct transcription families that vary in size and identity among different organisms (Riechmann *et al.* 2000, Wray *et al.* 2003). The wide range of differences in the type and proportion of transcription factors displayed in different organisms is attributed to differences in organismal complexity (Levine and Tjian 2003). *Arabidopsis thaliana* has more than 1500 (5.9 % of the total gene estimate) transcription factors of which 45 % are shared by transcription factor families in *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans* (Riechmann *et al.* 2000, Xiong *et al.* 2005). Some of the major plant specific transcription factor families are AP2/ERF, B3, WRKY, NAC and SBP (Yamasaki *et al.* 2013). Recently, the number of predicted *A. thaliana* transcription factors rose to nearly 2000 (Guo *et al.* 2005, Iida *et al.* 2005, Riano-Pachon *et al.* 2007). According to Riano-Pachon *et al.* (2007), *Arabidopsis* transcription factors are categorized into 64 groups. In comparison of 19 transcription factor families that are shared by plant and animal, more than 14 families are found to be larger in plants than their animal counterparts (Guo *et al.* 2005). Some of the transcription factor families present in plants, that are also shared by the other organisms, are highly elaborated by forming a subgroup compared to the other organisms. This indicates that they are differentially evolved to perform plant specific function (Riechmann *et al.* 2000). The proportions of transcription factor in *Arabidopsis* as compared to *Drosophila* and *C. elegans* are 1.3 and 1.7 fold, respectively, hinting that transcriptional regulation plays more of a role in plants than animals (Mitsuda and Ohme-Takagi 2009, Riechmann *et al.* 2000). In *Arabidopsis* lineage on the other hand, genes that perform basic functions like gene transcription and signal transduction are preserved in different taxa after genome duplication (Seoighe and Gehring 2004).

1.6.1 ERF1b transcription factor

The Ethylene Response Factor (ERF) gene family is a subset of the AP2 (Apetala2)/ERF superfamily (Riechmann *et al.* 2000) which is characterized by AP2/ERF domain (Jofuku *et al.* 1994). The AP2 domain is assumed to be evolved from bacterial and viral HNH-AP2 endonucleases by horizontal gene transfer (Magnani *et al.* 2004). As in the case of other transcription factor families, the evolution of the AP2 transcription factors has been facilitated by multiple gene duplication and diversification (Kim *et al.* 2006). The AP2/ERF superfamily is further divided into three subgroups, named as AP2 sub-family with two tandemly repeated AP2/ERF domain, ERF sub-family with a single AP2/ERF domain and RAV sub-family with B3 domain in addition to a single AP2/ERF domain (Nakano *et al.* 2006). ERF sub-family can be again subdivided into the ERF-type and the CBF/DREB subgroup (Sakuma *et al.* 2002). The ERF domain is demonstrated to preferentially bind to the GCC box, that is a regulatory element with a vital function in transcription of ethylene responsive genes (Ohmetakagi and Shinshi 1995). After Arabidopsis genome sequencing came to an end, it was predicted that AP2/ERF super family contains 145 genes of which 121 belong to ERF sub-family (Sakuma *et al.* 2002). A more recent update according to Nakano *et al.* (2006), the superfamily has 147 genes of which 122 form the ERF sub-family.

In *Arabidopsis thaliana* ERF1b transcription factor family contains eight members, alternatively named as Rap2.4a-h, while ERF1a family has got only two (Nakano *et al.*, 2006). Seven members are proposed to be nuclear targeted, one to be post-transcriptionally translocated into chloroplasts (Schwacke *et al.* 2007). The sequence alignment of the ERF1b transcription factors was done by Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) multiple sequence alignment (MSA) tool which basically employs profile hidden Markov models (HMMs) and seeded guide trees to generate biologically meaningful motifs of divergent protein sequences (Sievers *et al.*, 2011 and Goujon *et al.*, 2010). The alignment unveiled that all proteins in the group are characterized by an identical and structurally strictly conserved AP2/ERF DNA-binding domain (aa 250-319), but non-conserved C- and N-termini (**Figure 6**).

Introduction

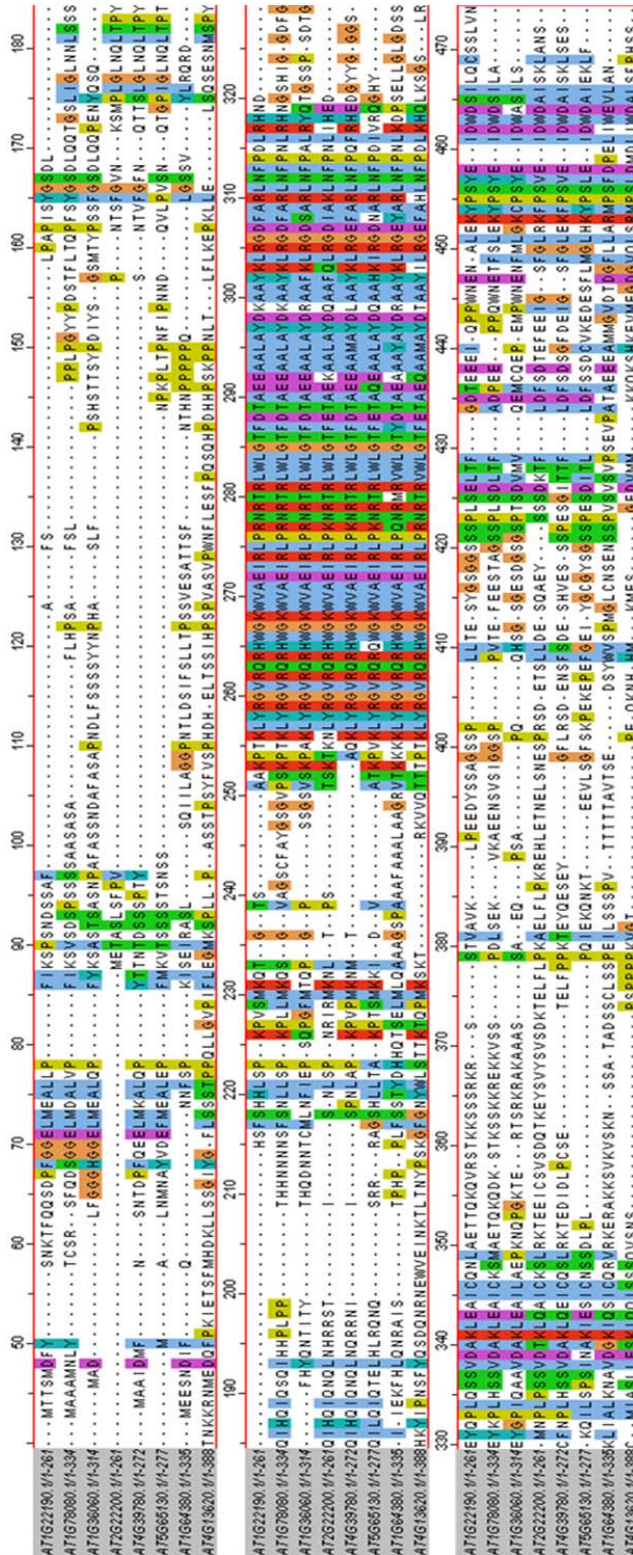


Figure 6: Amino acid sequence comparison of the eight ERF1b proteins using Clustal Omega multiple sequence alignment tool.

It is generally assumed that residues outside of the DNA binding domain play a critical role in protein-protein interaction to recruit other factors and for nuclear localization (Liu et al. 1999). To this end, the alignment also identified some loosely conserved regions, such as aa 330-349 and 451-466, outside the DNA binding domain. The highly consensus DNA-binding motif of the ERF1b contains YRG- and RAYD elements (Okamuro et al., 1997).

Generally, the super family where in ERF1b belongs to is characterized by exhibiting an AP2/ERF domain, 60 to 70 amino acids, involved in DNA binding (Nakano et al., 2006). Structurally AP2/ERF domain has three anti-parallel β -sheet and one α -helix which resemble the DNA-binding domain of bZIPs (**Figure 7**). However, in contrast to the bZIPs transcription factor in which the interaction to the DNA groove is mediated through α -helix, the anti-parallel β -sheets are in contact with the DNA (Allen *et al.* 1998).

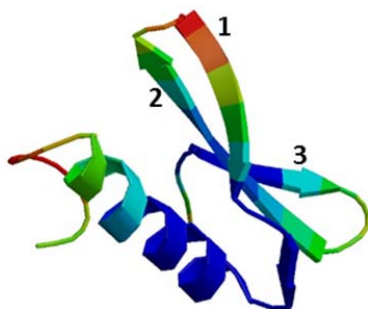


Figure 7: The three dimensional structure of AP2/ERF DNA binding domain modeled by SWISS-MODEL online program interface. Numbered 1, 2 and 3 are three-stranded anti-parallel β -sheet, the α -helix is packed approximately parallel to the β -sheet.

1.6.2 Functions for ERF1b transcription factor

It has been shown that genes in the AP2/ERF transcription factor family have vital functions in the transcriptional regulation of many physiological processes including growth, development and responses to biotic and abiotic stresses. Some of the functions by the transcription factor family include, dehydration and cold stress (Agarwal *et al.* 2006, Sakuma *et al.* 2002), disease resistance (Gutterson and Reuber 2004, Lorenzo *et al.* 2003) and flower development (Elliott *et al.* 1996). ERF1b (Rap2.4) transcription factor family is not well

characterized in terms of function. Among the reports on the ERF1b transcription factor family, Rap2.4a was demonstrated to be involved in the redox regulation of nuclear encoded chloroplast antioxidant enzyme called 2-cysteine peroxiredoxin-A (2CPA) (Shaikhali *et al.* 2008). Rap2.4b, recently also named as Wound Induced Dedifferentiation 1 (WIND1), is highly expressed after wounding and activates the formation of callus thereby controlling cell dedifferentiation (Iwase *et al.* 2011). In addition Rap2.4b was shown to be involved in drought and salt stress tolerance (Lin *et al.* 2008b).

1.7 Chloroplast retrograde signaling and antioxidant enzymes

Plant cell maintains two semi-autonomously replicating organelles, chloroplast and mitochondria, that have their own genome. Nuclear gene expression controls organellar function. The number of proteins that are present in these organelles by far exceeds the coding capacity of the organelle genome (Richly and Leister 2004). Thus, to execute the desired function, chloroplast and mitochondria rely on the import of proteins that are encoded in the nuclear genome (Woodson and Chory 2008). It is believed that the prokaryotic ancestor of these organelles contained all the necessary genes to permit independent survival of the organisms (Dyall *et al.* 2004). After endosymbiosis of the free living ancestor, the genes presumed to be transferred to the host cell making it a master regulator of the association between the two. As plant cells maintain three separate genomes, efficient communication of signals and coordination of gene expression is required to support life. Given that the anabolism and catabolism processes which require high energy turnover are performed by chloroplast and mitochondria respectively, the two organelles have to be constantly protected from unwanted energy leak from the system. To this end, various environmental stresses were shown to disturb the process. Perturbation of the metabolic states of the chloroplast will be translated to changes in nuclear gene expression, that would help to restore the normal state again (Woodson and Chory 2008).

Chloroplasts are armed with a variety of antioxidant enzymes that are encoded in the nucleus and expressed in a response to chloroplast retrograde signals, whose mechanism is not yet well understood (**Figure 8**). Signaling cascades have been postulated sensing intersystem signals, ROS, antioxidant status, the acceptor availability at PSI and metabolites (Baier and Dietz 2005, Fey *et al.* 2005, Pfannschmidt *et al.* 2003). It has been shown that Rap2.4a, which belongs to the ERF1b transcription factor family, is a redox-sensor and transducer of redox information. It activates the expression of 2-cysteine peroxiredoxin-A (2CPA) which is a nuclear encoded chloroplast antioxidant enzyme (Shaikhali *et al.* 2008). In Rap2.4a knockdown lines, the transcript level of various other genes that encode chloroplast proteins, namely tAPx and sAPx, were misregulated (Shaikhali *et al.* 2008), indicating that Rap2.4a is involved in chloroplast retrograde signaling.

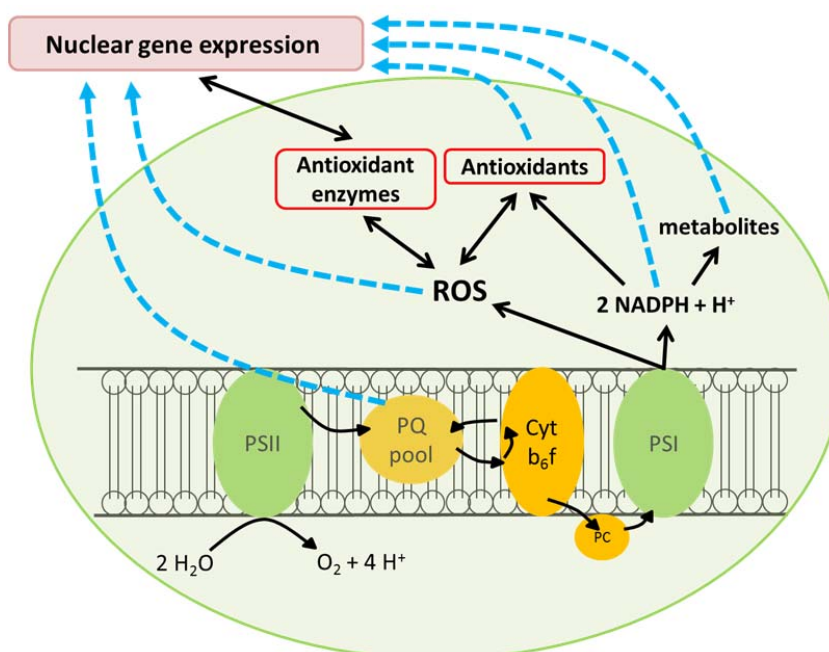


Figure 8: Schematics of chloroplast redox signaling. Various redox signals are produced in the electron transport chain as a result of excess electron pressure due high light or other environmental stress conditions. Under stressful condition, the plastoquinol (PQ) pool is mostly reduced and the energy liberated through non plastoquinol (NPQ) is inflated, that ultimately leads to the production of reactive oxygen species (ROS). The ROS signals through well uncharacterized mechanism communicate to the nucleus to induce nuclear gene expression, ultimately result in the production of chloroplast antioxidant enzymes.

1.7.1 Chloroplast antioxidant enzymes

Plants must cope up to continuously changing environmental conditions such as high light, drought heat and cold, which produce reactive oxygen species (ROS) that as a consequence results in oxidative damage (Asada 1999, Mittler 2002). ROS can even be produced under ideal growth conditions due to a wide array of cellular metabolic processes such as mitochondrial, chloroplastic and plasma membrane linked electron transport systems. Under normal conditions of photosynthesis, the amount of electrons excited are in surplus than the available electron acceptor such as NADP^+ , leading to the flow of excess electron into water-water cycle (Asada 1999). As a counteract mechanism, plants have developed a multitude of defense system against unregulated oxidation in cellular systems. One of the major source of ROS in chloroplast is the Mehler reaction (Mehler 1951), when the plants are exposed to light the molecular oxygen is reduced to super oxides (O_2^-) which has to be dismutated to H_2O_2 by superoxide dismutase which in turn is converted to water by ascorbate peroxidase (APx) in a general process called 'water-water' cycle (Asada 1999). The two variants of ascorbate peroxidase located in the chloroplast, namely sAPx and tAPx, are implicated in their involvement in the detoxification of hydrogen peroxide and their pivotal physiological role in photooxidative stress response (Ishikawa and Shigeoka 2008, Maruta *et al.* 2010). The two chloroplastic isoforms of APxs are encoded in the nucleus and have to be transported to the chloroplast which involves the communication of at least two distinct cellular compartments. The communication of the two cellular organelles implies that signal of some kind has to be transmitted to the nucleus to activate the expression of the antioxidant enzymes. In quest for the transcriptional regulator of the two antioxidant enzymes, the interaction of promoter of the tAPx and sAPx with that of ERF1b transcription factor family was investigated.

1.7.2 2-cysteine peroxiredoxin-A (2CPA)

Plants have four types of peroxiredoxin (Horling *et al.* 2002, Pitsch *et al.* 2010) which are peroxidases without heme group. Functionally active 2CPA exists in a homodimer form in which the catalytically active Cys residues are reduced (Konig *et al.* 2003). Upon oxidation, the two Cys residues form a disulphide linkage, which has to be reduced later to restore the catalytic function of the enzyme by an electron donor such as thioredoxin (Konig *et al.* 2003). 2-cysteine peroxiredoxin-A (2CPA) is one of the antioxidant enzymes in the chloroplast whose transcriptional regulation is well studied. The mRNA abundance reacts to the external application of low molecular weight antioxidants such as glutathione, cysteine and ascorbate as reviewed in Baier *et al.* (2010). On the other hand, the oxidative stressors that would normally lead to the increased transcript level of other antioxidant enzymes, hardly have effect on the transcript abundance of 2CPA (Baier and Dietz 1997, Baier *et al.* 2010). It has been shown that upon the deletion of chloroplastic ascorbate peroxidases, the expression of 2CPA is inflated (Kangasjarvi *et al.* 2008). The expression of 2CPA is high during early developmental stages of young tissue (Baier and Dietz 1996). The expression activity of 2CPA is proportional to the acceptor availability downstream of photosystem I in the electron transport chain of photosynthesis (Baier *et al.* 2004). In quest for the upstream regulator of 2CPA, a yeast-one-hybrid screen identified Rap2.4a as the modulator of the expression of 2CPA which itself is demonstrated to be redox sensitive. Upon mild stress condition, Rap2.4a binds to the redox-box of the CE3-like element (Baier *et al.* 2004, Shaikhali *et al.* 2008) in its dimeric form.

1.7.3 Chloroplast ascorbate peroxidase

Ascorbate peroxidases (APx) are nuclear encoded antioxidant enzymes whose activities are present in the cytosol, chloroplasts, mitochondria and peroxisomes. Continuous productions of H₂O₂ as a result of various internal and external perturbations of cellular processes are detoxified by various isoforms of APx. To this end, the transcript abundance of APx

demonstrated to be increased following the treatment of plants with high temperature, methyl viologen, high light and salt stress (Storozhenko *et al.* 1998). Comparisons of APx sequences in different species revealed that the organellar isoforms have the same ancestral origin with that of catalase-peroxidases (Zamocky *et al.* 2000). It is believed that the ancestral organellar APx was duplicated to give rise to the two isoforms present in the chloroplast, namely stromal ascorbate peroxidase (sAPx) and thylakoid-bound ascorbate peroxidase (tAPx) (Pitsch *et al.* 2010). *In vitro* analysis demonstrated that both chloroplastic APx isoforms can be catalytically nonfunctional by excess amount of H₂O₂ as compared to the cytosolic isoform, which could explain the cellular damage caused at severe stress condition (Hiner *et al.* 2000, Kitajima 2008). The increase in local concentration of H₂O₂ as a result of the inactive chloroplastic APx, which could be important for propagating ROS mediated signaling to control further damage, in which case the sensitivity of the isoforms make chloroplastic APx efficient H₂O₂ sensor (Kitajima 2008). Plants with double sAPx and tAPx knockdown did not show heightened stress symptoms, and the absence of the two isoforms were compensated by boosting the level of 2CPA (Kangasjarvi *et al.* 2008).

1.8 Motivation of the project

Given that the DNA binding domain of ERF1b transcription factors are highly conserved, the functional specificity or redundancy of the group members remains unknown. To delineate the functions of each transcription factor thereby determining the binding site and thus elucidate the downstream target genes of ERF1b transcription factors, several approaches were considered. One of the widely used approaches considered in defining the transcription factor binding site is CHIP-seq (Kaufmann *et al.* 2010). Nonetheless, there are inherent technical challenges associated to the routine application of the method. For instance, CHIP-seq grade antibodies have to be developed. For each transcription factor whose motif is to be elicited, one may need prior knowledge of the peak expression time and/or actively expressing tissue to predestinate the binding state of the transcription factor to its DNA target. Prior to target identification, the bound DNA sequences have to be

amplified and target libraries have to be developed and validated. Owing to the rapid evolution of sequencing technology, the cost of running ChIP-seq experiment has decreased, although it is still expensive.

Another alternative approach considered was protein binding microarray (PBM) which is an *in vitro* approach to define transcription factor binding sites (Mukherjee *et al.* 2004a). In addition to costs associated with instrumentation, sub-optimal buffer condition might affect protein folding and, consequently, the binding specificity of transcription factors. The results are not always reproducible *in vivo*, due to the typical pitfalls of an *in vitro* experiment.

A new approach has been developed for genome wide screening of transcription factor binding sites and to define binding motifs. Stringent inverse yeast-one-hybrid is designed as an easy and rapid as well as cheap method to screen genomes of any kind for transcription factor binding site, which extends on the advantages of previous yeast systems. Besides identification of the DNA binding preferences, it permits designation of target loci in a genome wide scale. In addition, based on the identified motifs, downstream target genes of the transcription factor was identified and functionally characterized.

2. Materials and Methods

2.1 Plant material

A. thaliana Col-0 and T-DNA insertion lines obtained from SALK-collection were used in the gene expression and transient gene expression analysis. SALK T-DNA insertion lines were obtained from Nottingham Arabidopsis Stock Centre (NASC, Loughborough, UK). Rap2.4c (At2g22200: SALK_064696 and SALK_108879) T-DNA insertion lines were genotyped by PCR (2.3.5) using CTACAGAGGCGTAAGGCAAAG, TATCCAATCTTGATGTTGGGC and GTGTATCGGTGAGGCTGAGAG, GTCCTCCTCCGGTAGTTTCAC as primer pairs, respectively, to select for homozygous lines before embarking with the actual experiment. Similarly, the Rap2.4d (At1g22190: SALK_139727) line was genotyped by using GTGTATCGGTGAGGCTGAGAG and GTCCTCCTCCGGTAGTTTCAC primer pair.

2.1.1 Seed sterilization

A. thaliana seeds were surface sterilized in 1.5 ml reaction tube initially by washing them with 300 µl of 70 % (v/v) ethanol for about 1 min followed by 300 µl of 20 % (v/v) household bleach (Glorix, Lever Farberge, The Netherlands) for 8 min. The seeds were thoroughly rinsed 5 times using sterile water.

2.1.2 Plant growth condition

Following surface sterilization, the seeds were vernalized in the dark at 4°C for two days to break the seed dormancy. The seeds were grown on petri plates containing sterile Murashige and Skoog (MS) (Duchefa, Haarlam The Netherlands) medium adjusted to pH 5.7

and supplemented with 2.5 g/L Phytigel (Sigma, Steinheim, Germany) and 0.5 % (w/v) sucrose. The seeds on the growth medium were incubated in a Percival (Percival Scientific Inc, Perry, USA) growth cabinet under short day conditions of 10 hrs of 120 $\mu\text{mol quanta m}^{-2}\text{s}^{-1}$ at 22°C and 14 hrs of dark at 20°C.

2.1.3 Mature Plant growth condition

Seeds vernalized for two days were sown on the nutrient rich soil substrate composed of 42.4 % (v/v) P-soil, T-soil 42.4 % (v/v), which were produced by Einheitserde® (Sinntal-Altengronau, Germany), and perlite 15.2 % (v/v) (Perligran G:Dortmund, Germany). The composition of each soil type is indicated in the **Table 1**. The seeds were grown in plant growth chamber under short day conditions of 10 hrs of 120 $\mu\text{mol quanta m}^{-2}\text{s}^{-1}$ at 20°C and 14 hrs of dark at 18°C.

Table1: Soil composition

Type	Ingredients	Org. sub.	pH	Salt KCl [g/l]	CaCl ₂ [mg/l]	P ₂ O ₅ [mg/l]	K ₂ O [mg/l]
P-soil	white peat, clay	75 %	5,8	1,5	150	150	210
T-soil	white peat, clay	75 %	5,8	2,5	310	300	420
Arabidopsis	white peat, clay,	50 %	5,8	1,7	194,5	189	267
Substrate	perligran G						

2.1.4 Cold treatment of plants

Plants grown for 4 weeks under short day conditions at 20°C in the growth chamber were cold treated for 24 hrs by transferring the plants to a 4°C growth chamber, while keeping all other growth parameters the same with the growth conditions prior to the cold treatment. Following the appropriate length of cold treatment, the leaves of the plant were harvested and frozen immediately in liquid nitrogen.

2.2 Gene expression analysis

2.2.1 RNA extraction

Leaves of at least three different plants (not more than 100 mg) grown under appropriate conditions were harvested and pooled into 2 ml reaction tube and immediately frozen in liquid nitrogen. The plant tissues were thoroughly ground to a fine powder using a tissue homogenizer (Retsch, Düsseldorf, Germany). RNA isolation was done using a Roboklon kit (Berlin, Germany) according to the manufacturer's instruction. To lyse the cells, the ground tissues were vigorously mixed by vortexing in 200 μ l LG and 100 μ l LR buffers, and the cellular debris were sedimented by spinning down at 14000 $\times g$ for 4 min. To further clarify the lysate and remove the DNA, after mixing with 200 μ l LR buffer, the supernatant was transferred to a homogenization spin column and centrifuged at 14000 $\times g$ for 2 min. The flow-through was mixed with 300 μ l 100% ethanol and applied to RNA binding spin-column and spun for a minute at 11000 $\times g$ to bind the RNA to the column matrix. The column was washed with DN1 buffer which further eliminated the remnants of DNA. After additional washing steps with 650 μ l and 350 μ l RBW buffer, the RNA was eluted with 40 μ l of RNase free water. The quality of the RNA was determined spectrophotometrically as well as by a running RNA gel.

2.2.2 First strand cDNA synthesis

High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Carlsbad, USA) was employed to synthesize cDNA. In the cDNA synthesis, 1 μ g RNA was used as a template in a total reaction volume of 10 μ l containing 1 \times RT Buffer, 4 mM dNTP, 1 \times RT random primers and 50 U MultiScribeTM Reverse Transcriptase assembled on ice. The reaction mixture was loaded to a thermocycler (Analytikjena, Jena, Germany) preprogrammed to 25°C for 10 min to extend the random primers, followed by 37°C for 120 min for cDNA synthesis and finally

heated to 85°C for 5 min to inactivate the enzyme. Prior to use for further analysis, the various cDNAs were equilibrated by diluting them with RNase-free water.

2.2.3 Quantitative real-time PCR

2.2.3.1 Primer Design

The primer design for quantitative Real-time PCR was performed using QuantPrime online program (<http://www.quantprime.de/>) (Arvidsson et al., 2008). The parameters considered in the program, when selecting the best primer pair alternatives, were melting temperature, length of primers, the GC content, proximity to the intron-exon border, the product length, and the possibility of dimer formation. The best selected 5′sense and 3′antisense primer pairs, as predicted in the program, were additionally checked manually for the formation of hair pin structures. The specificity of the primers was initially determined by examining the melting curve of each amplicon. Additionally, the efficiency of each primer pair was assessed by making the standard curve correlation coefficient of the reaction of a serially diluted known template with the primer pairs.

2.2.3.2 Fluorometry

The Roboklon SYBR® Green master mix (Berlin, Germany) contains Roboklon SYBR Green polymerase, dNTPs, optimized PCR buffer and SYBR Green. It was used for the real-time amplification reaction using the Bio-Rad CFX96 Real Time System (Hercules, USA). Specific transcript abundances were quantified with SYBR Green dye which intercalates with double stranded DNA as the amplicons continue to build up logarithmically. During the course of reaction, the fluorescence signal emitted at 585 nm when excited with 470 nm is measured in real time. The reaction mixtures were assembled in 10 µl reaction volumes and contained 1 µg cDNA template, 0.6 mM of 5′sense and 3′antisense primers, and 1 x

Materials and Methods

Roboklon SYBR® Green master mix. A three-step 40 cycle PCR protocol were employed as shown in the **Table 2**. All the reactions were performed in three technical replicates.

Table 2: General quantitative real-time PCR protocol

Reaction Step	Cycles	Temperature (°C)	Time
Activation of DNA polymerase	1	95	10 min
Denaturation	39	95	30 s
Primer annealing		60	30 s
Elongation		72	30 s
Dissociation curve	1	72	0.2°C/ sec

2.2.3.3 Standardization

For qRT-PCR normalization, Actin7 (At5g09810) and F-box (AT5G15710) were used as reference genes (An *et al.* 1996). They are constitutively expressed in a plant cells. The relative expression of the target genes compared to the reference genes were calculated according to $2^{-\Delta\Delta Ct}$ method using Bio-Rad CFX manager 2.1 program.

2.3 General methods

2.3.1 Spectrophotometric analysis of nucleic acids

The quality of the DNA / RNA was checked by analyzing the aliquots of nucleic acid preparation spectrophotometrically at 230, 260 and 280 nm using NanoDropND1000 (Nanodrop, Wilmington, USA). The peak absorbance of nucleic acids is at 260 nm whereas proteins, phenols and other contaminants have a peak around 280 nm and carbohydrate contaminants have a peak at 230. The 260/280 and 260/230 absorbance ratio of 1.8-2.0 and 1.8-2.2, respectively, were considered as a pure DNA/RNA preparation.

2.3.2 Ethanol precipitation of DNA

1/10 volume of sodium acetate (pH 5.2) was added to the DNA sample to a final concentration of 0.3 M and mixed well. Following the addition of 2.5 volumes of 100 % ethanol, the mixture was stored at -20°C for at least 30 min. The DNA was pelleted at 14000 x *g* for 15 min and afterwards washed with 1 ml 70 % ethanol. After removing the supernatant, the pellet was vacuum dried and resuspended in either water or Tris-EDTA buffer (pH 8.0).

2.3.3 Purification of DNA from enzyme reaction mixtures and PCR samples

DNA fragments from PCR and various enzymatic reactions were purified by using QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). The samples were mixed with Buffer PB in a 1: 5 volume ratio to promote the purification of DNA fragment greater than 100 bp. The mixture was loaded to a QIAquick spin column and centrifuged briefly for 30 sec and subsequently washed with 750 µl of Buffer PE. The DNA was eluted either by Buffer EB (10 mM Tris-Cl, pH 8.5) or double distilled water, both of which were pre heated to 65°C.

2.3.4 Extraction of DNA from agarose gels

DNA fragments separated by agarose gel electrophoresis were excised from the gel and purified using QIAquick Gel Extraction kit (Qiagen, Hilden, Germany). Each gel piece was dissolved in Buffer QG in 1: 3 volume (assuming that 1 mg ≈ 1 µl) at 50 °C in heating block with gentle agitation for 10 min. To enhance the binding of either less than 500 bp or greater 4 kb DNA fragments, one gel volume of isopropanol was added to the sample mixture, loaded to the QIAquick column and centrifuged briefly for 60 sec. The column was successively washed with 500 µl Buffer QC and 750 µl Buffer PE. After drying the column by additional centrifugation, the DNA was eluted with buffer EB (10 mM Tris-Cl, pH 8.5) or double distilled water, both of which were pre heated to 65°C.

2.3.5 Polymerase chain reaction (PCR)

10x PCR Buffer: 200 mM Tris/HCl (pH 8.4)
 500 mM KCl
 15 mM MgSO₄
 5 mM MgCl₂

Polymerase chain reaction (PCR) was routinely used to amplify a specific segment of DNA sequence using thermostable DNA polymerase. The general protocol for amplification is shown below in **Table 3**. Depending on the purpose of the product, the initial denaturation and final extension phase was varied between 1-7 min and 2-20 min, respectively. A typical PCR reaction in 20 μ l reaction volume contained 2 μ l (\approx 50-100 μ g) template DNA, 0.2 mM of each dNTP, 0.5 μ M of each primer (forward and reverse), and 1.25 U of *Taq* DNA polymerase in 1 x PCR buffer.

Table 3: General protocol for PCR

Reaction Step	Cycles	Temperature (°C)	Time (S)
Initial denaturation	1	94	180
Denaturation	39	94	15-45
Primer annealing		$T_m - 5$	15-45
Elongation		72	60 Kb ⁻¹
Final elongation	1	72	180

2.3.6 Agarose gel electrophoresis

5 x DNA loading buffer: 0.25 % (w/v) Bromophenolblue
 0.25 % (w/v) Xylenecyanol
 30 % (v/v) Glycerol

Materials and Methods

1 x TAE-buffer : 0.8 mM Tris-acetate pH 7.5
 0.02 mM EDTA

Various DNA fragments, such as generated by PCR or a result of enzymatic digestion, were analyzed electrophoretically on a TAE agarose gel. Depending on the fragment size to be separated, 0.8 - 3 % (w/v) gel was prepared by melting agarose in a TAE-buffer with subsequent addition of 0.5 µg/ml of Ethidium bromide. The DNA samples were mixed with 5 x DNA-loading buffer. Depending on the product size of interest, either 100 bp or 1 kb DNA-ladder (Fermentas, St. Leon-Rot, Germany) was loaded on the gel. The DNA samples were resolved at 80 - 130 V in 1 X TAE-buffer. For documentation, DNA fragments were visualized and digitally imaged in INTAS imaging system (Göttingen, Germany).

2.3.7 TA cloning

TOPO® TA Cloning® Kit (Invitrogen, Carlsbad, USA) was routinely used to clone the fragments generated by PCR (2.3.5). During the amplification procedure by PCR, the final extension time was extended to more than 20 min to produce an A overhang which would lead to ligation of the amplicon to the single 3'-thymidine overhangs provided by vectors in TOPO® TA Cloning® Kits. The Ligation was done according the proscriptio made by the manufacturer. 1-4 µl fresh PCR product, 1 µl TOPO® vector and 1 µl salt solution (1.2 M NaCl and 0.06 M MgCl₂) were assembled in 6 µl total reaction volume and incubated for 30 min at room temperature. The resulting ligation mixture of 2-4 µl was used to transform chemically competent DH5α or TOP10 *E.coli* cells. The insert orientation was checked by double digestion with enzymes that cut at the vector backbone and in the insert. To ensure that no mutation had been introduced, the inserts were further sequenced (2.3.11) with M13 forward primer (5'-GTAAAACGACGGCCAG-3') and M13 reverse primer (5'-CAGGAAACAGCTATGAC-3').

2.3.8 Competent cell preparation

CaCl₂-Medium: 60 mM CaCl₂
 15 % (v/v) Glycerin
 10 mM PIPES (pH 7.0 with KOH)

A freshly grown DH5 α *E. coli* colony was used to inoculate 5 ml LB media and allowed to grow overnight at 37°C while shaking at 250 rpm. The next morning, 2 ml of overnight culture was transferred to 200 ml pre-warmed LB media and grown for 2-3 hrs until the culture reached the logarithmic phase (OD₆₀₀ = 0.4-0.5). Following the incubation on ice for about 20 min to slow down the metabolic activity, the cells were sedimented at 2000 x *g* for 10 min at 4°C. The cells were washed by 10 ml ice-cold CaCl₂-medium by resuspending and subsequent sedimentation. For long treatment of the cells with CaCl₂-medium, the cells were resuspended again and incubated for 30 min on ice. After harvesting, the cells were resuspended with 2 ml CaCl₂-medium of which 80 μ l aliquots were frozen down in liquid nitrogen and transferred to -80°C for storage until later use.

2.3.9 Bacterial transformation

SOC Medium: 2 % Tryptone
 0.5 % Yeast extract
 10 mM NaCl
 25 mM KCl
 10 mM Mg Cl₂
 10 mM MgSO₄
 20 mM Glucose

For each clone to be transformed, one vial of either One Shot[®] TOP10, DH5 α or BL21 chemically competent *E. coli* cells were thawed on ice for 10 min. The cells were incubated

on ice for 30 min following the addition of 4 μ l of fresh ligation reaction to the cells and gentle mixing by flicking. The cells were heat-shocked for 30 sec at 42°C and placed back on ice for 2 min so that the cells can recover. The cells were suspended with 250 μ l pre warmed SOC media and incubated at 37°C for 1 hr while shaking. 80-100 μ l were spread on pre warmed LB agar containing the appropriate selection medium and incubated overnight at 37°C.

2.3.10 Quick DNA isolation from plant

Rapid extraction buffer (REB):	50 mM Tris-HCL pH 8.0
	25 mM EDTA
	250 mM NaCl
	0.5% SDS

One or two young Arabidopsis leaves were harvested and homogenized in 200 μ l REB buffer using a micropestle. For extraction of nucleic acid, an equal volumes of phenol : chloroform : Isoamyl alcohol (25:24:1) were applied and mixed thoroughly to the sample. Phase separation was done by spinning at 14000 x *g* for 5 min at room temperature. The aqueous phase was transferred to a fresh 1.5 ml reaction tube, mixed with an equal volume of isopropanol and subsequently stored at - 20°C for 1 hr to precipitate the DNA. The DNA was sedimented at top speed for 15 min, vacuum dried and resuspended either in 100 μ l sterile water or Tris-EDTA buffer (pH 8).

2.3.11 DNA sequencing

DNA sequencing was routinely performed to verify the identity of the sequences of the insert within clones. In the stringent inverse yeast-one-hybrid experiment, the interacting genomic fragments were identified from the growing colonies of the screening plate. Aliquots of purified plasmid DNA containing approximately 2 μ g DNA were sent for

sequencing to GATC Biotech (Konstanz, Germany) where they were sequenced with an ABI 3730xl system by chain terminator sequencing (Sanger sequencing).

2.4 Genomic library construction

2.4.1 Genomic DNA isolation

High molecular weight genomic DNA (> 30,000 bp) was isolated for construction of the genomic library, which serves as a prey in the stringent inverse yeast-one-hybrid screen. Genomic DNA was isolated using DNeasy Plant Maxi Kit (Qiagen, Hilden, Germany) from young leaves of *A. thaliana* (Col-0) that were grown in the green house under long day (2.1.2) conditions. Leaf tissue of about 1 g fresh weight was harvested and frozen in liquid nitrogen. It was then pulverized under liquid nitrogen to a fine powder using a precooled mortar and pestle. The powder samples were transferred to 50 ml tubes and mixed with 10 µl RNase (100mg/ml) and 5 ml of AP1 buffer that was preheated to 65°C. The mixtures were incubated for about 10 min at 65°C with gentle shaking to facilitate the lysis process. To precipitate the detergents, proteins and polysaccharides, 1.8 ml of AP2 buffer was added to the lysate and mixed. Preceded by 10 min incubation on ice, the lysate mixture was centrifuged at 5000 x *g* for 5 min at room temperature. To further clarify the lysate, the supernatant was transferred into the QIAshredder Maxi spin column and centrifuged for 5 min at 5000 x *g*. A volume of about 1.5 AP3/E was added and immediately mixed with the cleared lysate. To bind the genomic DNA to the column matrix, the mixture was transferred to DNeasy Maxi spin column and centrifuged for 5 min. The column was washed with 12 ml of AW buffer. The genomic DNA was eluted by washing twice with 1 ml of AE buffer. To determine the sample quality, aliquots of the isolated genomic DNA were analyzed spectrophotometrically (2.3.1) and by agarose gel electrophoresis (2.3.4).

2.4.2 Shearing of genomic DNA

In order to generate a random genomic library of *A. thaliana*, high molecular weight DNA was mechanically sheared to produce a fragment size of about 200-300 bp. For DNA fragmentation a S220 ultrasonicator (Covaris, Woburn, USA) was used.

3 µg of genomic DNA was suspended in 130 µl 10 mM Tris and 1 mM EDTA (pH 8.0) in Covaris Snap-Cap microTUBEs (Covaris, Woburn, USA) to prepare it for the shearing step. The ultrasonicator machine was configured at a duty factor of 10 %, intensity level 4, with 200 cycles per burst for 80 sec in the frequency sweeping mode. The water bath was filled with pure double distilled water, to a level of 12. As any water particulates interfere with the focusing of the acoustic beam, the water was purified by Covaris water conditioning system. The degas pump was started 30 minutes before the start and allowed to run continuously during sample treatment so that the acoustic coupling to the sample could be improved. To prevent sample heating and enhance the energetic acoustic activity inside the sample fluid during the sample running process, the initial water bath temperature was adjusted to 7°C. After configuring the machine, the empty microTUBE was mounted to the loading station in a vertical orientation. Genomic DNA sample was transferred into the tube with tapered pipette tip by penetrating the pre-split septa of the microTUBE alongside the inner wall to avoid the introduction of any bubbles. The formation of any air bubbles inside the microTUBE was removed by brief centrifugation. The microTUBE was then placed into the S220 apparatus together with its holder. The sample run was initiated according to the parameters already calibrated. After the sample run, the tube was removed from the holder and the sample was transferred to a 1.5 ml reaction tube.

2.4.3 Testing the quality of sheared DNA

The quality of the sheared DNA was analyzed in an Agilent 2100 Bioanalyzer (Agilent, Wilmington, USA) after preparation of the samples with Agilent DNA 1000 Kit (Agilent,

Wilmington, USA). The bioanalyzer is microfluidics-based platform to analyze and control the quality of DNA, RNA and protein samples in which the results could be produced in short period of time. Prior to sample preparation for the chip protocol, the bioanalyzer was set up such that the base plate of the priming station was attuned and the syringe at the chip priming station was replaced by a new one and its clip was adjusted. To prepare the gel-dye mix, the DNA gel matrix and the DNA dye were incubated at room temperature for 30 min. After DMSO was dissolved, the DNA dye concentrate was vortexed and centrifuged for 10 sec consecutively. 25 μ l of the DNA dye concentrate was transferred to the DNA gel matrix and vortexed to facilitate mixing of dye and gel. Gel-dye mix was then transferred to the top receptacle of a spin filter which in turn was spun down for 15 min at room temperature and 2240 $\times g$. The mixture was stored at 4°C in darkness until later use. 9 μ l gel-dye mix was pipetted in each well of the new DNA chip that was placed at chip priming station. The 12 sample wells were filled with 5 μ l of DNA marker and 1 μ l sample. Unused sample wells were filled with 5 μ l of DNA marker and 1 μ l deionized water. The chip was mounted to the vortex mixer and spun for 60 sec at 2400 rpm (Agilent, Wilmington, USA). The chip containing the samples was loaded on to the receptacle of the bioanalyzer and the sample run was started.

2.4.4 End repair of the sheared DNA fragment

NEBNext End Repair Module kit (BioLabs, Frankfurt am Main, Germany) contains a cocktail of enzymes that has been optimized for repairing the DNA fragments' end that were generated by acoustic shearing. After purification from the enzymatic admixtures, the DNA fragment treated by the enzymatic mixtures contained in the kit can be used for blunt end ligation. To facilitate the fragment end repair process a mixtures of three enzymes, namely T4 DNA polymerases, Klenow Fragment and T4 Polynucleotide Kinase, were used. The 5' \rightarrow 3' polymerase activity of T4 DNA polymerase and Klenow Fragment fills 3' recessed DNA ends while 3' \rightarrow 5' exonuclease activity removes 3' protruded DNA end. T4 Polynucleotide Kinase couples 5'-phosphates to ends of unphosphorylated sheared DNA fragments.

The fragment ends of 3 µg DNA were blunted with 15 U of T4 DNA polymerase and 50 U of T4 Polynucleotide Kinase provided in the kit as enzyme mix. In addition, the fragments were treated with 5 U of Klenow Fragment DNA Polymerase (BioLabs, Frankfurt am Main, Germany) in 10 µl 1x NEBNext End Repair Reaction Buffer (50 mM Tris-HCl, 10mM MgCl₂, 10mM DTT, 1mM ATP, 0.4 mM dATP, 0.4 mM dCTP, 0.4 mM dGTP, 0.4 mM dTTP, pH 7.5 at 25°C). For optimal results, the incubation time was set to 1 hour at 20°C. The end repaired DNA fragments were later cleaned from enzymes and buffer components by QIAquick Purification Kit (Qiagen, Hilden, Germany).

2.4.5 Prey plasmid construction

5 µg of pHIS2.2 vector were prepared for uptake of the genomic DNA fragments by digesting it with 5 µl FastDigest SmaI (Fermentas, St. Leon-Rot, Germany) with 5 µl 10 x FastDigest buffer in 50 µl reaction volume and subsequent blunting and dephosphorylation with NEBNext End Repair Module (BioLabs, Frankfurt am Main, Germany) and FastAP Thermosensitive Alkaline Phosphatase (Fermentas, St. Leon-Rot, Germany) respectively. Blunt end ligation of the random genomic library DNA into the vector was performed in a 1 to 10 molar ratio with T4 DNA Ligase (Fermentas, St. Leon-Rot, Germany). The ligation products were transformed into 20 aliquots of 100 µl ultra-competent Top10 *E. coli* cells (Invitrogen, Carlsbad, USA) by a heat shock and plated out on 50 150 mm size petri dishes containing Luria-Bertani (LB) medium supplemented with kanamycin (50 µg/ml).

2.4.6 Harvesting the genomic library

Following overnight growth of *A. thaliana* genomic library in *E. coli* at 37 °C, the colonies were washed off from each plate with 10 ml LB medium using 3 mm sterilized metal beads by 15 min gentle agitation on a shaker. To amplify the library, the washes from all plates were combined and used to start a 4 liter overnight culture supplemented with kanamycin (50 µg/ml). The cells were harvested by centrifugation and the plasmid library was prepared

using QIAGEN Plasmid Maxi Kit (Qiagen, Hilden, Germany) (2.4.7) according to the manufacturer's instruction.

2.4.7 Plasmid library preparation

Large scale library preparation was done using QIAGEN Plasmid Maxi Kit (Qiagen, Hilden, Germany). For plasmid library preparation, a 4 liter overnight culture of freshly harvested genomic library divided into 500 ml was used. For large scale single clone plasmid library preparation, freshly grown and well isolated single colony was used to start a 5 ml LB medium culture containing kanamycin (50 µg/ml). The colony was grown for about 8 hrs at 37°C. The growing culture was diluted into 500 ml media to start an overnight culture with vigorous shaking. The bacterial cells were harvested by spinning down at 6000 x *g* for 15 min at 4°C and subsequently resuspended in 20 ml P1 buffer containing 100 µg/ml RNase A. For alkaline lysis of the cells, 20 ml of P2 buffer was added, mixed gently but thoroughly and incubated for 5 min. The reaction was neutralized by adding 20 ml of pre-chilled P3 buffer, mixed gently and promptly poured to QIAfilter Cartridge to filter the fluffy precipitate. Following equilibration of QIAGEN-tip 500 with 10 ml QBT buffer, the cleared lysate was applied and allowed to enter the resin by gravity flow. The column was washed twice with 30 ml QC buffer and eluted with 15 ml QF buffer. To concentrate the plasmid DNA, it was ethanol precipitated and redissolved in 2 ml of Tris-EDTA buffer, pH 8.0.

2.5 Bait plasmid construction

Bait plasmid was developed by cloning the transcription factor of interest into the GAL4 activation domain of the pACT2 vector (Clontech, Mountain View, USA). Full-length coding sequence of the ERF1b transcription factors were used in the entire experiment. Since the genes do not have any intron to be spliced out, the primers indicated in the **appendix 6.1** were used to amplify genomic DNA. Cloning of coding sequences of the proteins into the vector was facilitated by the directional cloning using the restriction sites generated by the

PCR primers. The construct was sequenced to ensure that no mutation was introduced during the cloning procedures.

2.6 Yeast manipulation

2.6.1 Yeast strain

The auxotroph yeast (*Saccharomyces cerevisiae*) strain Y187 used in the study has the genotype of MAT α , ura3-52, his3-200, ade2-101, trp1-901, leu2-3, 112, gal4 Δ , met-, gal80 Δ , URA3::GAL1UASGAL1TATA-lacZ and MEL1 having trp1, leu2 and cyh2 as a transformation marker and his3 as an interaction marker.

2.6.2 Growth condition

YPD medium:	2% (w/v) Peptone
	2% (w/v) Glucose
	1% (w/v) Yeast extract
SD medium:	2.67 % (w/v) Minimal SD Base
	0.062 % (w/v) DO Supplement

Non transformed yeast cells were grown either on rich yeast extract-peptone-dextrose (YPD) or YPD supplemented with 80 mg/ml adenine (YPAD) medium at 30°C. Adenine was supplemented so that higher saturation densities could be achieved and to avoid the formation of pink coloration in ade1 and ade2 mutant strain. For selection and propagation of transformants, yeast cells were grown at 30°C for 5 days on minimal synthetic dropout (SD) medium (Clontech, Mountain View, and USA), in which all the nutrients but auxotrophic marker specific to the plasmid transformed was included.

2.6.3 Yeast transformation

The small scale yeast transformation strategy was adopted from Gietz and Schiestl (2007a). It is a lithium acetate/single-stranded carrier DNA/PEG method for transformation of *S. cerevisiae*. Although the transformation yield is lower, it is a method of choice for single and small scale transformation, and was used to transform the bait plasmid in the stringent inverse yeast-one-hybrid screen.

An overnight yeast culture was started by inoculating 5 ml of YPAD medium. To avoid the formation of cell clumps and ensure that all the cells were suspended, a 1.5 ml reaction tube containing 1 ml of the same medium was initially inoculated and vortexed for 30 sec. The content was then transferred to a 5 ml tube and incubated overnight at 30°C while shaking at 250 rpm. The next morning, about 2 ml of the overnight culture was transferred to 50 ml (to a final OD₆₀₀ of 0.1) of YPAD medium, and allowed to grow for 5-6 hours until the cells undergo at least two cell divisions. The cell culture was harvested at 3000 x *g* for 5 min at room temperature and washed with 25 ml of double distilled sterile water. The washed cells were treated with 1ml of 100 mM Lithium acetate (pH 8.4-8.9) for 5 min. Meanwhile, carrier DNA (salmon sperm DNA) of 2 mg/ml was denatured by heating it to 100°C for 10 min and cooled down immediately on ice. The cells were pelleted at 13000 x *g* for 10 sec and suspended in 500µl 100mM lithium acetate (pH 8.4-8.9). 80 ml aliquots of cell suspension were dispensed to 1.5 ml reaction tubes per single transformation. The transformation mixture was prepared by mixing 240µl PEG (3350 or 4000), 36µl 1M LiAc (pH= 8.4-8.9), 50µl denatured carrier-DNA, 500 ng of plasmid and additional sterile water to a final volume of 34 ml on ice. Following the removal of the lithium acetate, the transformation mixture was added to the pellet and vigorously agitated to avoid any cell clumps. The entire mixture was incubated initially at 30°C and then at 42°C each for 30 min. After pelleting the cells, the supernatant was removed, suspended in 200 µl water and plated on the appropriate selective plate.

2.6.4 Library scale yeast transformation

Library scale high-efficiency yeast transformation was done according to (Gietz and Schiestl 2007b). The yeast strain transformed with the bait construct was grown on SD/-Leu agar media. 50 ml overnight culture in SD/-Leu broth medium was started with a freshly grown colony. The next morning, the cell density of the culture was determined spectrophotometrically at 600 nm (**2.3.1**). Assuming that OD₆₀₀ of 0.1 carries 1 X 10⁶ yeast cells, the volume containing 6.25 x 10⁸ cells were harvested at 3000 x g for 5 min. The cells were resuspended in warm SD/-Leu to a final cell density of 5 x 10⁶ cells per ml, incubated for 4-5 hours and harvested by centrifugation. The cells were taken up in the transformation mixture (7.2 ml 50 % (w/v) PEG-3350, 1.08 ml 1M lithium acetate [pH 8.5], 1.5 ml denatured salmon sperm DNA [2mg/ml]), 60 µg plasmid DNA in a total volume of 1.02 ml). The transformation mixture was incubated for 120 min at 42°C, harvested by centrifugation and resuspended in 20 ml sterile water. The suspension was then plated on a 150 mm large plates on suitable selection media.

2.6.5 Yeast colony PCR

10 X Colony PCR buffer:	0.13 M Tris-HCl (pH 8.5)
	0.56 M KCl

A medium sized yeast colony was suspended in 30 µl of 0.2% SDS by vortexing for 15 sec, which was then incubated in a heating block at 90°C for 4 min. After cell lysis, the cellular debris were pelleted at 14000 x g and the supernatants were transferred to a new tube. The following components were combined in one tube on ice. The samples were tested by PCR in a thermocycler (Analytikjena, Jena, Germany) with the program as described in standard PCR (**Table 4**).

Materials and Methods

Table 4: General component mixture in yeast colony PCR

Components	Amount (μ l)
10x colony PCR buffer	5
μ l 50 mM MgCl ₂	1.5
10 mM mix of dNTPs	1
10 mM of each primer	1
25% Triton X-100	2
Taq Polymerase	0.3
H ₂ O	up to 49
crude DNA prep	1
Total Volume	50

2.6.6 Yeast plasmid isolation

Plasmid isolation from yeast cells was performed using Easy Yeast Plasmid Isolation Kit (Clontech, Mountain View, USA). Fresh yeast colonies growing on appropriate selective agar medium were used to streak 1cm x 1cm squares on a new plate, and incubated at 30°C for 2 days to grow the cell patches. Using a sterile inoculating loop, about 10 mg of cells were scooped and suspended in 500 μ l of 10 mM EDTA. The washed cells were pelleted at 11000 x *g* for 1.5 min. Following the resuspension of the pellet in 200 μ l zymolyase buffer (Clontech, Mountain View, USA) by vigorous vortexing, 20 μ l zymolyase was added and mixed thoroughly. For efficient disruption of cell walls, the mixture was incubated at 30°C for 1 hr with gentle agitation on a shaker. The spheroplasts were pelleted for 10 min at 2000 x *g*. For alkaline lysis of spheroplasts, the pellet was resuspended in 250 μ l of Y1 Buffer/RNase A solution, immediately followed by 250 μ l of Y2 lysis buffer and incubated for 5 min at room temperature. To neutralize the reaction, 300 μ l of Y3 buffer was added and gently mixed. The lysate was clarified by pelleting it two times at 11000 x *g* for 5 min. For binding reaction to the column matrix, the cleared lysate was passed through the Yeast Plasmid Spin Column by spinning down for 1 min at 11000 x *g*. The column was washed with

450 μ l of washing buffer. To ensure complete dryness, the column was centrifuged at 14000 $\times g$ for additional 3 min. The plasmid DNA was eluted using 30 μ l YE elution buffer. To rescue and further amplify the plasmid DNA, high efficiency competent *E. coli* cells were transformed with 10 μ l of the plasmid preparation (2.3.9).

2.6.7 Stringent inverse-yeast-one-hybrid experiment

The interaction assay was executed by plating out the double transformants on 150 mm plates with SD-dropout media lacking leucine, tryptophane and histidine (SD-/Leu/-Trp/-His) and supplemented with 0 mM, 60 mM, 100 mM, 160 mM, 200 mM, 250 mM and 300 mM 3-AT (3-Amino-1,2,4-triazole). The plates were incubated for 5 days at 30°C to grow colonies. As a transformation control, the double transformant was also plated on SD-dropout media lacking only leucine, tryptophane (SD-/Leu/-Trp).

2.7 Bioinformatic analysis

2.7.1 Global and multiple alignment

Following the global alignment of the interacting clone sequence with that of the vector sequence, the sequence of the interacting genomic fragments were identified after cutting the border vector sequences. For the identification of tagged genes, the genomic fragments interacting with the bait transcription factor were mapped to the genome using the BLAST algorithm with full genome sequence information available via National Center for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The BLAST was done with highly similar sequences (megablast) with *A. thaliana* (Col-0) reference genome 7.

2.7.2 Motif analysis

The interacting genomic DNA fragments were analyzed for common motifs using MEME (Multiple Em for Motif Elicitation) (<http://meme.nbcrc.net/meme/>) and SCOPE (Suite for Computational identification Of Promoter Elements) (<http://genie.dartmouth.edu/scope/>) online interfaces. MEME is a computational method to identify ungapped motifs and enables to choose the best width, number of occurrences, and description for each motif (Bailey and Elkan 1994). SCOPE is a parameter-free method to computationally determine a motif in sets of co-regulated genes and employs three component algorithms, each designed to identify different types of promoter motifs (Chakravarty *et al.* 2007). BEAM and PRISM search for non-degenerate motifs and degenerate motifs, respectively. In the SCOPE program, SPACER scans for bipartite motifs. During the analysis, SCOPE uses the genome sequence of an organism of interest in the background. In both MEME and SCOPE programs, FASTA sequence format was used as an input.

2.7.3 Public microarrays data

For *in silico* transcript abundance regulation and co-regulation analysis publicly available microarray information was analyzed using Genevestigator (<https://www.genevestigator.com/gv/plant.jsp>) and eFP Browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>).

2.8 Expression, purification and detection of recombinant proteins

2.8.1 Bacterial growth conditions

LB Medium: 1 % (w/v) Peptone
 0.5 % (w/v) Yeast extract
 1 % (w/v) NaCl
 1.5 % (w/v) Agar
 pH 7.2

Bacterial cultures were routinely grown either on solid or broth Luria Bertani (LB) media. TOP10 and DH5 α were used in the cloning procedures while BL21 was used for recombinant protein expression. The cultures were grown at 37°C except for low temperature expression of proteins in which case it was incubated at 25°C.

2.8.2 Cloning procedures

For the purpose of affinity purification of 6 x His tagged fusion proteins, the full length sequences of the genes of interest were cloned in pEXP5-NT/TOPO (Invitrogen, Carlsbad, USA). In addition to providing the N- terminal tag to the protein of interest, the vector uses a strong bacteriophage T7 promoter to drive the expression. The coding sequence was inserted into the TOPO TA cloning sites. Following purification of the PCR products, using QIAquick PCR Purification Kit (Hilden, Germany), 2 μ l freshly purified product, 1 μ l salt solution (1.2 M NaCl and 0.06 M MgCl₂) and 1 μ l TOPO[®] vector were combined in 6 μ l reaction volume. The reaction mixture was incubated for 30 min at room temperature prior to transformation of either competent TOP10 and DH5 α cells.

2.8.3 Heterologous expression of proteins in *E.coli*

A single colony of freshly transformed BL21 cells were used to start an overnight culture in LB broth media supplemented with 100 µg/ml ampicillin. The next morning, pre-warmed 500 ml LB media was inoculated with the overnight culture. The culture was allowed to grow for 1.5 to 2.5 hrs until the OD₆₀₀ reached 0.6-0.7. Protein expression was induced by applying Isopropyl-β-D-thiogalactoside (IPTG) (Duchefa, Haarlem, The Netherlands) to 0.5-1 mM final concentration of the culture medium. The cells were grown either at 19°C for overnight, 25°C for 5 hrs or 37°C for 3 hrs. The cells were harvested at 4000 x *g* at 4°C for 20 min and stored at -80°C until later use.

2.8.4 Purification of recombinant proteins

Buffer B: 100 mM NaH₂PO₄
10 mM Tris·Cl
8 M Urea
pH to 8.0 using NaOH

Buffer C: 100 mM NaH₂PO₄
10 mM Tris·Cl
8 M Urea
pH to 6.3 using NaOH

Buffer D: 100 mM NaH₂PO₄
10 mM Tris·Cl
8 M Urea
pH to 5.9 using NaOH

Materials and Methods

Buffer E: 100 mM NaH₂PO₄
 10 mM Tris·Cl
 8 M urea
 pH to 6.3 using NaOH

Frozen pellets from 500 ml cultures were thawed on ice and resuspended in 10 ml of buffer B. For efficient lyses, the cells were agitated at room temperature for 1 hr or until the solution became translucent. The crude cell lysate was centrifuged at 16000 x *g* for 20 min to sediment the cellular debris. The cleared lysate was either stored at -80°C or directly used in purification procedures. 4 ml of the cleared lysate was mixed to 1 ml of the 50 % Ni-NTA (Qiagen, Hilden, Germany) slurry. The mixture was gently agitated for 1 hr at room temperature. The lysate-resin mixture was loaded into empty column and the flow-through was collected. The column was washed twice with buffer C and finally, the bound proteins were eluted with 0.5 ml of Buffer D and Buffer E, each for four times.

2.8.5 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

6 % Stacking gel: 12.67 % (v/v) 1.5 M Tris-HCl, pH 6.8
 4.87 % (w/v) Acrylamide
 0.13 % (w/v) Bisacrylamide
 0.1 % (w/v) SDS
 0.1 % (w/v) APS
 0.3 % (v/v) TEMED

1 x Running buffer: 1.44 % (w/v) Glycine
 0.3 % (w/v) Tris-HCl
 0.1 % (w/v) SDS

Materials and Methods

12 % Resolving gel:	25.0 % (v/v) 1.5 M Tris-HCl, pH 6.8
	11.68 % (w/v) Acrylamide
	0.32 % (w/v) Bisacrylamide
	0.1 % (w/v) SDS
	0.1 % (w/v) APS
	0.012 % (v/v) TEMED
4 x Loading buffer:	500 mM Tris-HCl, pH 6.8
	8 % (w/v) SDS
	40 % (w/v) Glycerol
	20 % (w/v) Bromophenol blue

Proteins were separated approximately according to their molecular weight in an electric field. The polymerization of acrylamide with bis-acrylamide was activated by tertiary-amine (TEMED) and catalyzed by ammonium persulfate (APS). The resolving gel that exhibited small pores and higher pH was casted in the preassembled gel cassette and overlaid with isopropanol. Following 1 hr polymerization and removing the isopropanol, the stacking gel solution was poured on top of the resolving gel and polymerized for 15 min. Prior to loading the gel, the protein samples were mixed with one volume of protein-loading buffer and either directly loaded or heated to 95°C for 10 min for insoluble sample fraction. The protein samples were resolved at 30 mA constant current until the mark reached the lower edge.

2.8.6 Coomassie staining of protein gels

Staining solution:	50 % (v/v) Methanol
	0.1 % (w/v) Coomassie R 250
	2 % (v/v) Acetic acid

Materials and Methods

Destaining solution: 40 % (v/v) Methanol
 7 % (v/v) Acetic acid

Following electrophoresis, the gel was incubated for 60 min in the staining solution with gentle shaking. After discarding the staining solution, the unbound Coomassie brilliant blue was washed away by incubating the gel in destaining solution overnight.

2.8.7 Western blotting

Transfer buffer: 25 mM Tris
 150 mM Glycine
 10% (v/v) Methanol

Ponceau S: 0.2 % (w/v) Ponceau S
 3 % (v/v) Acetic acid

Succeeding the equilibration of the gel in TBS solution for 5 min, a semi-dry transfer apparatus was used to blot the protein from SDS-PAGE to the nitrocellulose membrane. Gel on top of the membrane was sandwiched together within pre-wetted Whatman paper with a transfer buffer. The transfer was accomplished for 45 min under the constant current of 2mA per cm² of the membrane. The effectiveness of the transfer was checked by staining the blotted membranes in Ponceau S solution for 2 min followed by 5 times washing with double distilled water.

2.8.7.1 His-Tag detection with a Ni-NTA-HPR conjugate

TBS Buffer: 10 mM Tris-Cl
 150 mM NaCl
 pH 7.5

Materials and Methods

Tris-saline:	0.9% (w/v) NaCl 0.1 M Tris-Cl pH 8.0
TBS-Tween buffer:	20 mM Tris-Cl 500 mM NaCl 0.05% Tween 20 pH 7.5
HPR staining solution (30ml):	0.6 % (w/v) 4-Chloro-1-naphthol 20 % (v/v) Methanol 80 % (v/v) ml Tris-saline 60µl 30 % H ₂ O ₂

After washing Western membrane twice with TBS, it was blocked in 3 % milk powder solution overnight at 4°C. Binding of Ni-NTA conjugate at 1: 1000 dilutions in TBS-Tween buffer was done succeeding the triple wash of the membrane with TBS-Tween each for 10 min. Detection was done after the three washing steps. Finally, the membrane was stained with horseradish peroxidase staining solution for 1-5 min.

2.8.7.2 Detection of proteins with specific antibody

Following the transfer of proteins on to a nitrocellulose membrane, the membrane was washed with TBS buffer for 5 min at room temperature. It was incubated twice for 30 min in 1 % milk powder dissolved in TBS for effective blocking. The membrane was incubated in a solution containing 1:5000 - 1:10,000 primary antibody in 1 % milk powder preceded by washing a second washing step by TBS for 10 min. The unbound antibody was removed by washing the membrane three times (3 min, 10 min and 15 min) in 1 % milk powder. A dilution of 1: 3000 secondary antibodies in 1 % milk powder solution was used to incubate

the membrane for 1 hr. After washing, the membrane was subjected for signal detection reaction with HRP staining solution for 1-5 min.

2.8.8 Determination of protein concentration

Protein quantification was performed according to Bradford (1976) by employing the BioRad protein assay kit (BioRad protein assay, Munich, Germany). The assay is based on the principle that Coomassie Brilliant Blue G-250 solution has its maximum absorbance at 465 nm and shifts its peak absorbance to 595 nm upon protein binding.

To make a standard curve, a range of BSA stock dilution (0 mg/ml – 1.5 mg/ml) was prepared. 10 µl of protein samples or standards were mixed to 790 µl double distilled water and 200 µl of BIORAD-reagent per sample and incubated for 10 min at room temperature. The absorbance of the various samples was measured spectrophotometrically at 595 nm. Determination of the protein concentration was done by comparing the absorbance of the samples to the standard curve for BSA.

2.8.9 Antibody development

Polyclonal antibodies that are produced by triggering various B-lymphocytes cells in an animal were developed by using purified Rap2.4d (At1g22190) gene. Highly purified 500 µg Rap2.4d protein was sent to Pineda (Berlin, Germany) to immunize the rabbit succeeding initial test of pre immunization serum by western blotting. The Rabbit was allowed to grow for 61 days after which the serum was collected.

2.9 Plant Protein extraction

Buffer E:	125 mM Tris-HCl pH 8.8
	1 % (w/v) SDS
	10 % (v/v) Glycerol
	50 mM Sodium metabisulfite ($\text{Na}_2\text{S}_2\text{O}_5$)

Total protein was extracted from about 100 mg of fresh plant material by grinding the tissue in 180 μl Buffer E in a 1.5 ml reaction tube. The extract was centrifuged at 13000 x *g* for 10 min to remove cellular debris. The supernatant was transferred to a fresh reaction tube.

2.10 Transient expression of genes in Arabidopsis seedlings

2.10.1 Site directed mutagenesis

The promoter region of Kin2 (AT5G15970) was mutagenized utilizing a site directed mutagenesis protocol developed by Montemartini *et al.* (1999). To introduce two different mutagenized variant of the Kin2 (COR6.6) promoter two overlapping primers containing the desired mutation were designed. To introduce the first mutation variant on short and long promoter variants, the first round of PCR was performed with Kin2FP1: Kin2MR1 and Kin2FP2: Kin2MR1 respectively in addition to Kin2RP1:Kin2MF1 (**Table 5**).

Table 5: Site directed mutagenesis primer

Primer	Sequence (5' →3')
Kin2FP1	CTTGTTAGTTGCAACCAACG
Kin2RP1	GGATTGAGAAGATATGTGGG
Kin2PF2	GAAAGAGTGTATCGGATGCG
Kin2PR2	AGTTTGGAGCCAGACATGTC
Kin2MF1	GAAAACACGGAGAAGAGAGGACCACACA
Kin2MR1	GTGTGTGGTCCTCTCTTCCGTGTTTTC
Kin2MF2	GAAAACACGGAAAAGAGAGGACCACACAC
Kin2MR2	GTGTGTGGTCCTCTCTTTCCGTGTTTTC

Agarose gel electrophoresis was run to check the size of the product. The product of appropriate size was excised from the gel and purified using QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany). Accordingly the second round of PCR was performed to join the mutagenized PCR product using Kin2FP1: Kin2RP1 and Kin2FP2: Kin2RP1 and to generate short and long promoter variants. The second mutation variant was introduced in the long and short promoter fragment in a similar fashion. The introduction of the mutation was checked by sequencing (2.3.11) using M13 forward primer.

2.10.2 Reporter plasmid construction

Various fragments of the Kin2 promoter (0-729 bp, 0-1508 bp and -729-1508 bp of the transcription initiation site) were generated by PCR (**Primers: Table 6**) and cloned in to the PCR 8/GW/TOPO (entry) vector provided within the TA Cloning Kit (Invitrogen, Carlsbad, USA). Similarly, the mutagenized promoters were ligated to the same vector and the introduction of the mutation was confirmed by sequencing (2.3.11) using M13 forward primer.

Materials and Methods

Table 6: Primer for Kin2 (COR6.6) amplification

Primer	Sequence (5' → 3')
Kin2FP1	CTGTAGTTGCAACCAACG
Kin2RP1	GGATTGAGAAGATATGTGGG
Kin2PF2	GAAAGAGTGTATCGGATGCG
Kin2PR2	AGTTTGGAGCCAGACATGTC

2.10.3 LR Reaction

The pHGWL7.0 destination vector allows the expression of luciferase under the control of an inserted promoter, in expression experiments. Since the entry vector (pCR8[®]/GW/TOPO[®]) has the same antibiotic selection marker as the destination vector, the spectinomycin cassette in the entry clone was removed by double digestion with XbaI and XhoI, followed by gel purification. Gateway[®] recombination cloning reactions were assembled by combining 100 ng of entry clone, 150 ng of digested destination vector and 2 µl LR Clonase[™]II enzyme mix (Invitrogen, Carlsbad, USA) in 10 µl total reaction volume followed by overnight incubation at room temperature. The reaction was terminated by addition of 1 µl of Proteinase K and incubating the samples for 10 hrs at 37°C.

2.10.4 Seedling Transfection

Transient expression of the reporter genes was accomplished according to Voinnet et al. (2003) which uses viral encoded p19 protein to suppress the gene silencing of the transiently expressed protein. Two weeks old seedlings grown on MS medium were infiltrated by GV3101 (pMMP90) agrobacteria harboring either the wild or mutagenized promoter construct in pHGWL7.0 and expressing luciferase upon interaction.

2.10.4.1 *Agrobacterium* suspension solution preparation

Activation buffer: 10 mM MES/KOH pH 5.6
 10 mM CaCl₂
 150 μM Acetosyringon

An overnight culture was started by inoculating 5 ml YEB medium with GV3101 (pMP90) strain of *agrobacterium* harboring mutagenized Kin2::pHGWL7.0 and p19::pBin61. The growing media was supplemented with rifampicin 100 μg ml⁻¹, gentamycin 25 μg ml⁻¹ and spectinomycin 100 μg ml⁻¹ for Kin2::pHGWL7.0 construct and rifampicin 100 μg ml⁻¹ and kanamycin 50 μg ml⁻¹ p19::pBin61. The *agrobacteria* were grown at 28°C while shaking at 200 rpm. The 5 ml overnight culture was used to start a 50-150 ml overnight culture and grown until the OD₆₀₀ of at least 0.5 attained. The *agrobacterium* harboring both the promoter construct and the p19 construct, were mixed in a proportion described by the following formula.

$$V_{\text{construct}} = n * V_{\text{final}} * 0.5 / \text{OD}_{600} \qquad V_{\text{p19}} = n * V_{\text{final}} * 0.3 / \text{OD}_{600}$$
$$n * V_{\text{final}} = 30 \text{ ml per plate} \qquad n = \text{number of plates}$$

The mixture of both *agrobacterium* clones were harvested at 3000 x *g* for 8 min at room temperature. The pellet was then resuspended in 30 ml per plate of the activation buffer and incubated for 4 hrs at room temperature.

2.10.4.2 Vacuum Infiltration

Infiltration of *Arabidopsis* seedlings was done using vacuum dessicator (Kartell, Noviglio, Ital). The vacuum was applied 6 times each for 90 seconds or until bubbles escaped as an indicator for intra tissue gas. The plants were washed with sterile water to remove the *Agrobacterium* on the surface and transferred to a fresh MS medium supplemented with 500 μg/ml of cefotaxime to kill the remaining *agrobacterium* which could possibly give false

positive signals. The plants were grown for two days before testing the reporter gene activity.

2.10.4.3 Luminescence measurement

Plants infiltrated with agrobacterium solution containing the luciferase as a transiently expressed reporter gene were sprayed with 1 mM D-luciferin solution and incubated in the dark for 5 min during which luciferase converts the D-luciferin substrate into oxyluciferin which would emit light as it turns to the ground state in its excited state. The luminescence was measured using NightSHADE LB 985 (BERTHOLD, Bad Wildbad, Germany) imaging system with the exposure time of 60 sec to 300 sec with low gain and slow readout.

3. Results

3.1 Genomic library development

3.1.1 High molecular weight DNA isolation

A modified yeast-one-hybrid approach has been developed to define the DNA binding sites of the various ERF1b transcription factors. Different yeast screening strategies such as Lex-A system (Fields and Sternglanz 1994) and Aureobasidin A resistance based ones, were considered. Finally, the GAL4 yeast-one-hybrid system with HIS3 DNA-protein interaction marker was chosen owing to the fact that GAL4 system is most versatile in respect of sensitivity and reproducibility. In addition, the system with HIS3 marker, can detect interactions over a wider range of interaction strengths as compared to the ADE2 based marker, and the stringency of the screen can be modulated by applying HIS3 inhibitors in various concentrations. Here the pACT2.2 vector encoding the GAL4 transcription activation domain and pHIS2 having HIS3 interaction marker were chosen to develop a yeast-one-hybrid protocol to identify transcription factors binding motifs. As opposed to the classical yeast-one-hybrid method, in which single or multiple copies of defined sequences of DNA are used as baits, stringent inverse yeast-one-hybrid method uses random genomic library fragments of an organism as prey (**Figure 9**). Target libraries are constructed by cloning short random genomic fragments in the yeast expression vector (pHIS2) upstream of the GAL4 promoter, and the auxotrophic reporter marker, HIS3.

To set up a library suited for screening putative transcription factor binding sites, high molecular weight genomic DNA (> 30, 000 bp) was isolated from young leaves of *A. thaliana* (Col-0). Considering that (1) some of the transcription factors in ERF1b are putatively targeted to the chloroplast (Schwacke *et al.* 2007), (2) that DNA binding sites in the

Results

chloroplasts are widely unknown and (3) that the techniques to screen for organular binding sites are hardly developed, the genomic DNA preparation included genomic DNA of chloroplast and mitochondria in addition to nuclei genome, to address a wider technical problem.

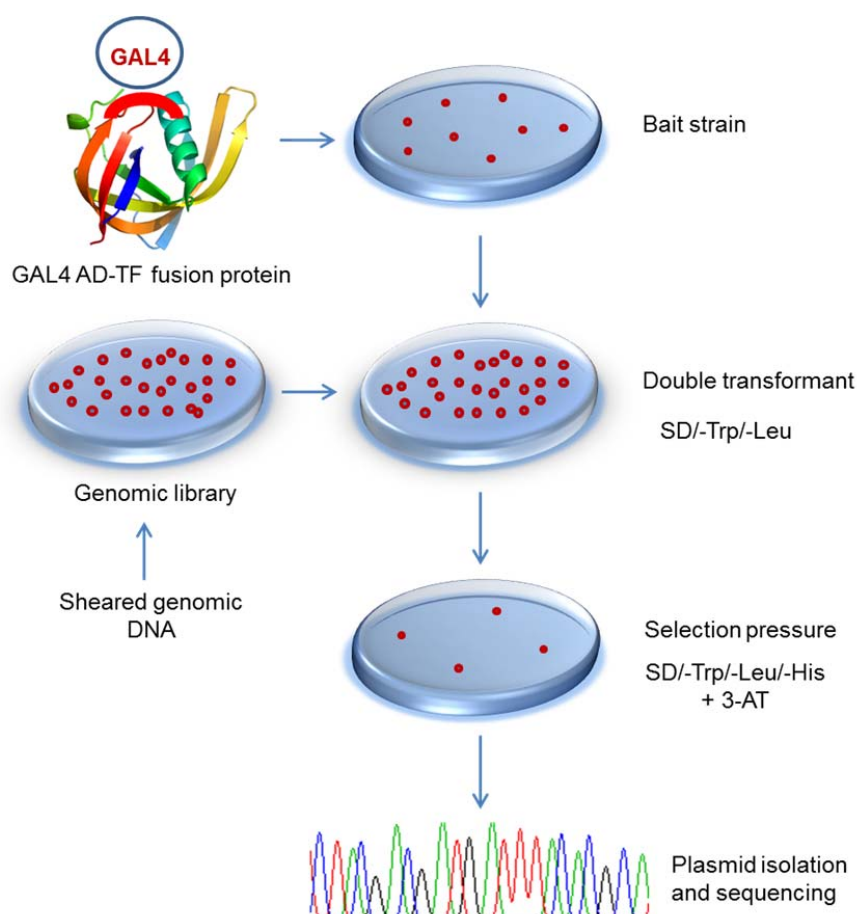


Figure 9: Schematic depiction of stringent inverse yeast-one-hybrid experiment. A plasmid genomic library constructed from random genomic fragments of *A. thaliana*, acts as prey.pHIS2 yeast reporter vector have histidine as an auxotrophic interaction marker. The plasmid library is used to optimize the library scale yeast transformation efficiency prior to the actual interaction assay. The cDNA of a transcription factor, as a bait, is cloned into the pACT2.2 yeast expression vector fused with GAL4 activation domain. Sequential yeast transformation strategy was adopted such that the prey plasmid library transformation is preceded by bait transformation. The doubly transformed yeasts are assayed on Synthetic Dextrose (SD) dropout medium lacking Leucine, tryptophan and histidine and supplemented with 3-AT. The reporter plasmid is isolated from the growing colonies after 5 days of incubation. The interacting plasmids are sequenced and analysed to determine the conserved motif.

3.1.2 Shearing genomic DNA

A protocol for genomic library preparation was developed (**Figure 10**). The DNA was isolated from *A. thaliana* and sheared to a fragment size of 200-300 bp. These lengths boost the resolution of the screen by keeping the distance of the transcription factor binding site and the GAL4 transcription initiation site short.

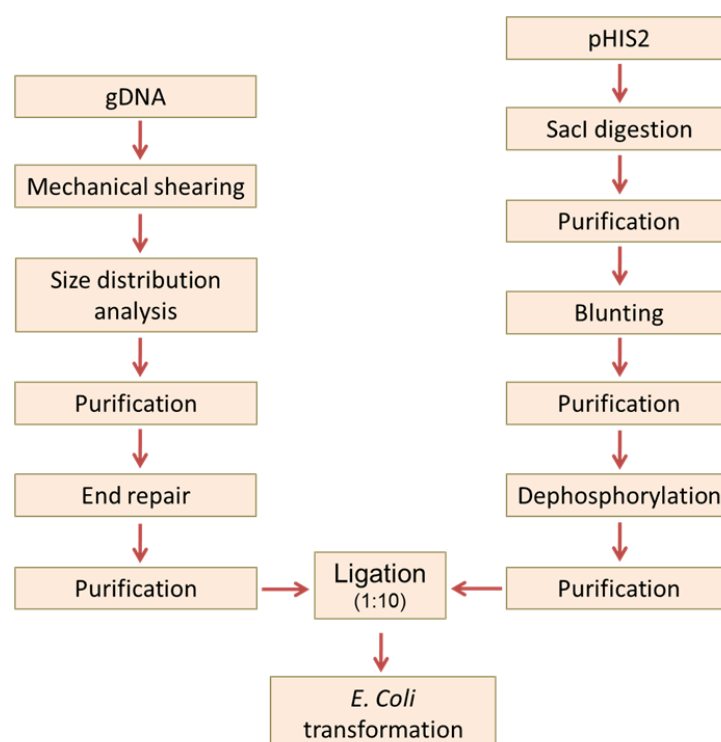


Figure 10: Flow chart of genomic library construction. High molecular weight DNA was isolated and subjected to mechanical shearing to produce average fragment size of about 300 bp. After quality control and purification, the ends of sheared fragments were blunted followed by second purification from enzymatic and buffer admixtures. The pHIS2 vector was cut by digestion with *SacI*. The ends were then sequentially blunted, preceded by purification. The dephosphorylated reporter vector was ligated to the genomic fragment in 1: 10 ratio. The ligation mix was used to transform high efficiency competent *E.coli* cells.

It has been shown that there is a strong bias in the functional binding of transcription factors toward transcription start site (Tabach *et al.* 2007), and thus with 200-300 bp there is a high probability that a regulating transcription factor activates the expression of interaction

Results

marker, HIS3. In addition, good coverage of Arabidopsis genome can be achieved while maintaining a manageable library size. With 200-300 bp library fragment length, it is expected that 5.4×10^5 clones would cover the *A. thaliana* genome once. For comparison, with 100 bp fragment size one would ideally already require 1.35×10^6 per one time coverage. With the higher library fragment size such as 500 bp, the distance to the transcription initiation site could be too large to activate the transcription of the auxotrophic marker.

Much effort had been placed to optimize the enzymatic digestion of the genomic DNA to get the fragments of desired range, given the simplicity of the procedure. Different four base cutter enzymes such as AluI and RsaI were tried to digest the genomic DNA. However, in addition to very low yield after purification of the DNA from enzymes and buffer admixtures, the fragment sizes were found to be widely distributed over large a range. In addition, some parts of the genome are more accessible for digestion than the others, which resulted in uneven digestion of the genome. In contrast, mechanical shearing is sequence independent. It uses physical forces and results in similar fragment sizes. Therefore, acoustic mechanical shearing device that is routinely used in other library preparation procedures, such in genome sequencing and ChIP-seq application, was chosen (Lundin *et al.* 2010, Meyer and Kircher 2010, Syed *et al.* 2009). The method has additional merit over the traditional enzymatic digestion in that there is no sample loss due to further sample purification and there is no contamination problem owing to the fact that there is no direct contact to the samples being processed. Furthermore, the method permits easy optimization to produce the fragments of desired range.

Different combinations of settings were initially tested by modulating the duty cycle between 2 % and 10 %. Different intensity levels (between 3 to 5) and exposure times (80-120 sec) were also tested (**Figure 11**). With only 2 % duty cycle and intensity level 3, the genomic DNA was insufficiently sheared. At the intensity level 4, the mean fragment size

Results

was around 350-400 bp. With higher duty cycle (10 %) and at intensity level 5, fragment size was in the desired range, with the mean size of about 250 bp.

Prior to proceeding to polish the sheared DNA fragments to make the genomic library ligation, the fragment distribution had to be determined. 1 μ l sheared DNA sample was directly used to analyse the size distribution of DNA fragments with the Agilent-bioanalyzer. The electropherogram (**Figure 11d**) demonstrates optimal size distribution of the sheared DNA fragments between 100 and 800 bp. Most fragments were around 200-300 bp and only very low amounts of fragments were larger than 500 bp. The right and the left sharp peaks in the electropherograms are 15 bp and 1500 bp marker loaded together with the sample. The proportion of the target fragment range was dependent upon the amount of starting material used in the shearing process.

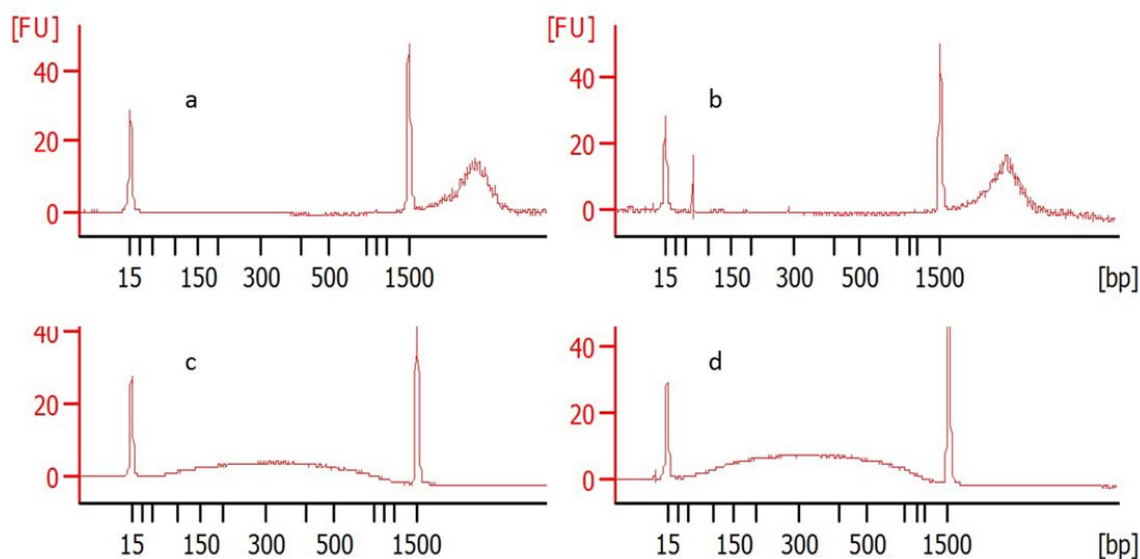


Figure 11: Optimization of DNA shearing with a S220 ultrasonicator (Covaris, Woburn, USA) as analyzed by Agilent 2100 Bioanalyzer (Agilent, Wilmington, USA) (a) Fragmentation with 2 % duty cycle and intensity level 3 for 80 sec (b) Fragmentation with 2 % duty cycle and intensity level 4 for 80 sec (c) Fragmentation with 10 % duty cycle and intensity level 4 for 80 sec (d) fragmentation with 10 % duty cycle and intensity level 5 for 120 sec.

Results

Ultimately, the configuration recommended by the manufactures was adopted with slight modification as shown in (**Figure 12**). Optimal DNA shearing was achieved by increasing the time of fragmentation and the intensity level. Fragments of approximately 200-300 bp length were generated by acoustic shearing of 3 μg (the maximum amount than can be loaded at a time) DNA suspended in 130 μl volume of 10 mM Tris and 1 mM EDTA (pH 8.0) using a S220 ultrasonicator (Covaris) at a duty factor of 10 % and an intensity level 5with 200 cycles per burst for 120 seconds. The machine was set up that the power mode was on Frequency Sweeping. The water bath was filled with pure double distilled water, to a level of 12.

Target Base Pair (Peak)	300
Duty Cycle	10%
Intensity	4
Cycles per Burst	200
Time (seconds)	80
Temperature (water bath)	6-8°C
Power mode	Frequency Sweeping
Degassing mode	Continuous
Volume	130 μl
Buffer	Tris EDTA, pH 8.0
DNA mass	<5 μg
Starting material	> 50kb
Water level (FILL/RUN)	S2 – level 12 E210 – level 6
AFA Intensifier	Yes

Figure 12: The S220 ultrasonicator (Covaris, Woburn, USA) parameter configurations for DNA fragmentation (adapted from Covaris protocol).

3.1.3 Optimization of ligation ratio

The random DNA fragments generated by acoustic/mechanical shearing were ligated into the pHIS2.2 reporter vector upstream of the HIS3 reporter element. Before embarking with

Results

the ligation procedure, 3 μg of mechanically sheared DNA were blunted with a cocktail of enzymes, such as 50 U T4 Polynucleotide Kinase, 15 U T4 DNA polymerases and 5 U Klenow Fragment for 1 hour to get the optimal result. In parallel, the pHIS2.2 vector that had been cut by *SacI* and dephosphorylated was treated by the same enzymes to prepare them for blunt end ligation. Different insert to vector ligation ratios (1:3 to 1:15) were tested to get the highest number of colonies following transformation of high efficiency competent *E. coli* cells (**Figure 13**). The ligation ratio (1:10) yielded the highest number of colonies compared to either extremely high or low ligation ratio. For library generation 150 μl ligation mixture were transformed into 20 aliquots of 100 μl ultra-competent cells and spread out on 50 150 mm sized plates. Following overnight growth, the colonies were washed off from each plate with 10 ml LB medium. To amplify the library, the washes from all plates were combined and used to start a 4 liter overnight culture supplemented with kanamycin.

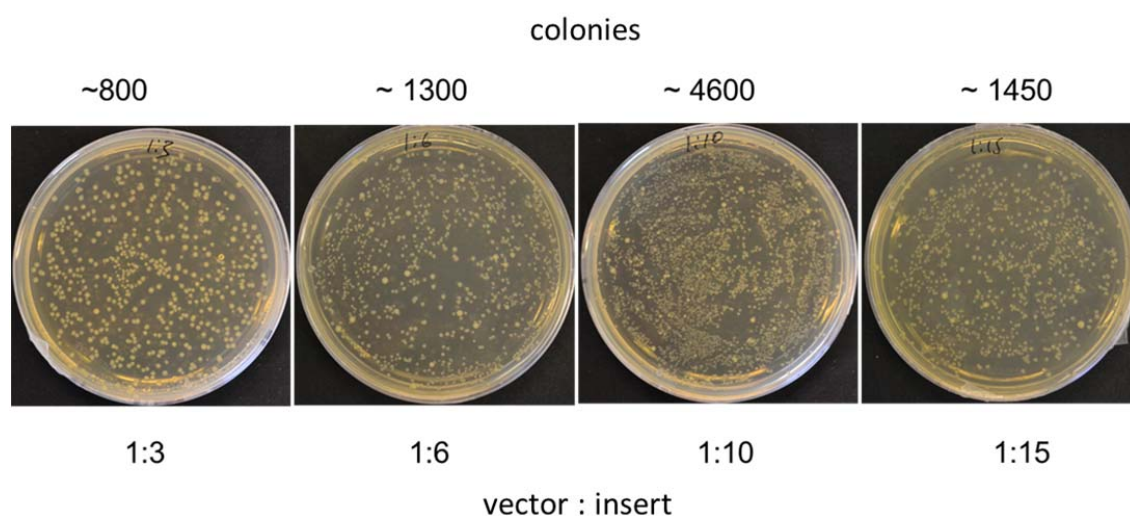


Figure 13: Ranges of vector to insert ligation ratio in Arabidopsis genomic library preparation. Blunt end ligation of the random genomic library DNA into the vector was performed in 1 to 10 molar ratios with T4 DNA Ligase. The ligation products were transformed into 100 μl ultra-competent Top10 *E. coli* cells by a heat shock and plated out on 150 mm size petri dishes containing Luria-Bertani (LB) medium supplemented with kanamycin (50 $\mu\text{g}/\text{ml}$). A range of insert to vector ratios (1:3 to 1:15) was tested.

3.1.4 Sequence analysis of genomic library

2.25×10^6 primary colonies that would cover the Arabidopsis genome more than four times, were used to prepare plasmid library. For further validation, 99 library clones were randomly selected, the plasmids were isolated and the inserts were sequenced. The sequencing result showed that, each colony represented a divergent clone and all colonies carried a plasmid harboring a fragment of genomic DNA. More than 75% of the inserts were in the desired fragment range of 200 – 300 bp with the average fragment length of 258 bp (Figure 14).

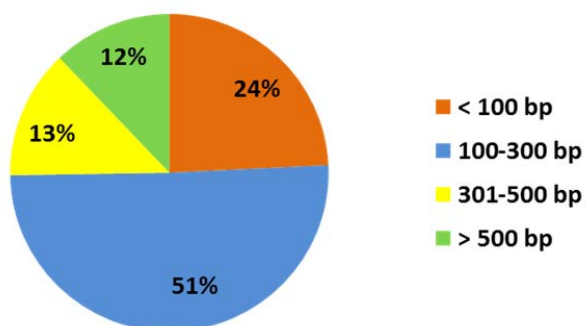


Figure 14: Insert size distribution of the genomic library as determined by randomly sequencing individual clones. Random sequencing of 99 clones gave 258 bp as a mean fragment size of the insert. Majority of the inserts are in the range of desired fragment length.

3.2 Stringent inverse yeast-one-hybrid

3.2.1 Optimization of library scale yeast transformation

For the actual screen it was important to achieve high yeast library transformation rates, but to avoid multiple uptake of various plasmids by the same yeast cell, which would otherwise perplex the succeeding analysis. Prior to the library screening, the yeast cells were transformed with 1-5 μg of the plasmid library per 100 μl cell preparation (1.7×10^8 cells). The scale-up factor for bulk library transformation was determined to optimize the

Results

efficiency and reliability of the screen. 2 μg was the lowest amount that gave comparably equal numbers of transformants in various transformation experiments (**Table 8**). For the actual library screen, a scale-up factor of 30 (60 μg plasmid library) was used for transformation.

In addition, the heat shock time was optimized for best transformation efficiency. Incubation time test at 42°C ranging from 30 min to 120 min demonstrated that longer heat shock time (120 min) yielded in a better transformation efficiency (**Figure 15**). After the optimization of all the parameters, the transformation technique resulted in 0.5 – 1 x 10⁶ double-transformed colonies per transformation and plate.

Table 8: The amount of plasmid (μg) transformed into 100 μl cell preparation (1.7×10^8 cells) and the resulting number colonies obtained after incubating the transformed cells for 3-4 days at 30°C.

Plasmid (μg) transformed	Number of colonies
2	1.01×10^4
3	1.15×10^4
4	1.04×10^4

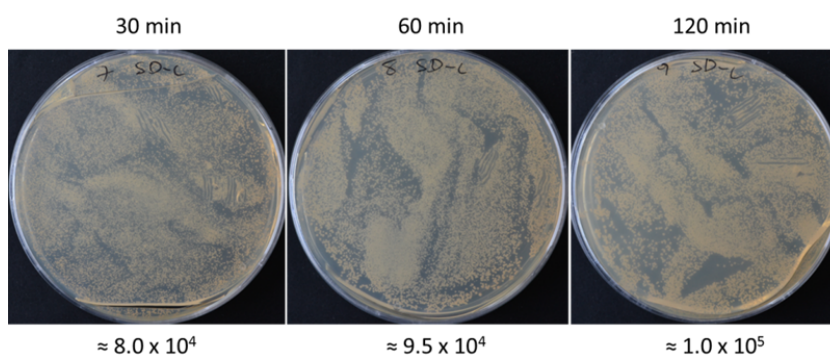


Figure 15: Various heat shock times in yeast transformation. Y187 yeast strain transformed with 2 μg plasmid and plated on SD/-Leu media (150 mm) after dilution and incubated at 30°C for 3-4 days.

3.2.2 Screening for promoter protein binding

Owing to the fact that HIS3, which served as the interaction marker, has a leaky expression in many yeast strains because of the affinity of yeast endogenous proteins to GAL4 promoter (Van Crielinge and Beyaert 1999), it was imperative to define the background suppression level. The yeast reporter strain, which was transformed only with prey constructs, was assayed on SD/-Trp/-His medium prior to the actual screening approach. The tryptophan auxotrophy served as a transformation efficiency control, while the histidine auxotrophy indicated activation of the GAL4 promoter. To overcome the leakiness of the GAL4 promoter and to minimize promoter activation by non-specific protein-DNA binding, such as by yeast's endogenous transcription factors, the interaction baseline was analyzed on 0 – 160 mM 3-amino-1,2,4-triazole (3-AT). 3-AT is a competitive inhibitor of imidazoleglycerol-phosphate dehydratase (encoded by the HIS3 gene) and acts as a histidine biosynthesis inhibitor. Thus, background suppression was optimized by transforming Y187 cells with aliquots of the genomic library and plating them on 3-AT gradients (0 mM to 160 mM) on (SD) agar lacking tryptophan and histidine (SD/-Trp/-His) (**Figure 16**). After 5 days at 30°C, the 3-AT concentration that gave no big and hardly any tiny colonies (160 mM) was chosen the optimal 3-AT concentration for background suppression.

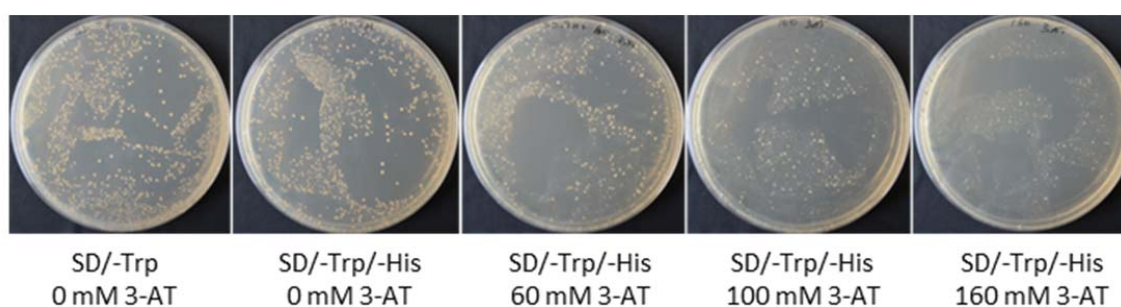


Figure 16: Yeast transformed only with prey plasmid (reporter plasmid construct) and plated on SD/-Trp and SD/-TH media supplemented with increasing amounts of 3-AT and allowed to grow for 3-4 days at 30°C.

Results

To increase the efficiency of prey and bait double transformations and to guarantee constant bait levels, a sequential yeast transformation scheme was chosen instead of the previously recommended mating of single transformed cells. Y187 yeast cells were first transformed with the Rap2.4d bait construct (Rap2.4d-cDNA in pACT2.2). Transformants were selected on SD/-Leu media. Subsequently, 5.0×10^9 fresh yeast cells harboring the bait construct were transformed with the prey library constructs. To monitor the transformation success independent from bait-prey-interaction, the cells were screened on SD/-Leu/-Trp. Markers on the bait (Leu2) and prey constructs (Tri1) complement the two auxotrophies of yeast Y187. The transformation technique applied here resulted in $0.5 - 1 \times 10^6$ double-transformed colonies per transformation and plate.

The interaction assays were executed on SD/-Leu/-Trp/-His plates. The media were supplemented with increasing concentrations of the histidine biosynthesis inhibitor 3-AT (0 mM, 60 mM, 100 mM, 160 mM, 200 mM, 250 mM, 300 mM) (**Figure 17**). For binding motif screening, only strongly interacting bait-prey combinations were selected on 160 mM and 200 mM 3-AT.

Table 9: The number of interacting colonies grown on selective media (SD/-leu/-Trp/-His supplemented with increasing amount of 3-Amino-1,2,4-triazole (3-AT)).

Selection pressure	Number of colonies
SD/-LT	1×10^6
SD/-LTH	5×10^5
SD/-LTH + 60 mM 3-AT	706
SD/-LTH + 100 mM 3-AT	442
SD/-LTH + 160 mM 3-AT	107
SD/-LTH + 200 mM 3-AT	54
SD/-LTH + 250mM 3-AT	16
SD/-LTH + 300 mM 3-AT	5

Results

From approximately 1.0×10^6 double transformed yeast colonies plated per plate, up to 100 and 50 colonies were vigorously growing at 160 mM and 200 mM 3-AT, respectively (**Table 9**). Almost no colonies grew on media supplemented with 300 mM 3-AT (**Figure 17**) demonstrating that this inhibitor concentration exceeds the HIS3 expression capacity of the vector system.

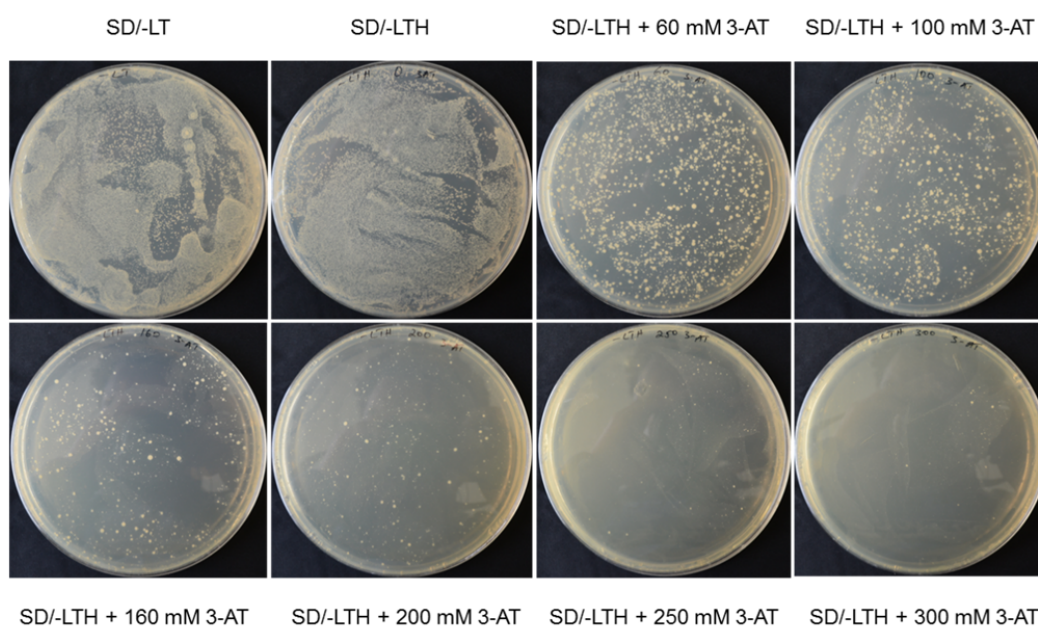


Figure 17: Stringent inverse yeast-one-hybrid assay. Y187 auxotrophic yeast strain sequentially transformed with both bait and prey plasmids and plated on SD dropout medium lacking leucine, tryptophan and histidine (SD/-Leu/-Trp/-His). The double transformed yeast strain was assayed on increasing amount 3-AT to delineate strongly and weakly interacting clones. For transformants assayed on 300 mM 3-AT, almost no colonies were observed even after 7 days of incubation. Only boldly growing colonies on selection media supplemented with 160 mM and 200 mM 3-AT were selected of further analysis of the prey sequence thereby defining a motif. The top left plate shows the same double transformant plated on SD dropout medium lacking leucine, tryptophan as a transformation control.

3.2.3 Conforming the interacting clones

For confirmation, the colonies that were found to be interacting in the screen were re-assayed for bait-prey interactions by streaking them out on fresh highly selective auxotrophic medium (SD/-Leu/-Trp/-His) supplemented with 160 and 200 mM 3-AT (**Figure 18**). Almost all colonies showed a profuse growth in their respective selective growth medium.

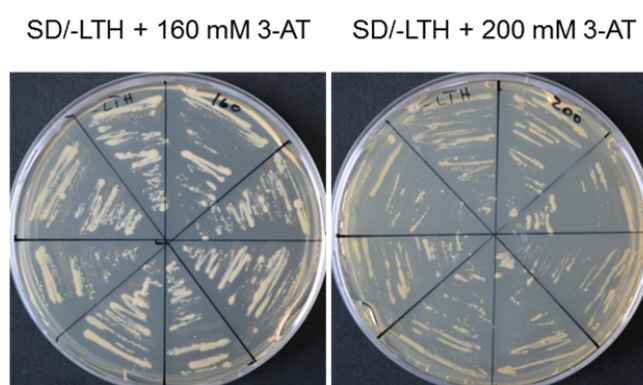


Figure 18: The figure shows examples of growing colonies when the strongly interacting clones were streaked on to the same selection media (SD/-Leu/-Trp/-His + 160 mM 3-AT and SD/-Leu/-Trp/-His + 200 mM 3-AT) for conformation.

3.2.4 Sequence analysis of Rap2.4a and Rap2.4d interacting clones

To determine the binding site of Rap2.4d, the plasmids of 96 strongly interacting clones were isolated and sequenced robotically with a primer binding to the genomic DNA flanking vector sequence. After removal of the vector sequences, the reads were analyzed for the presence of conserved motifs using MEME (Bailey *et al.* 2009) and SCOPE (Chakravarty *et al.* 2007) online programs.

To identify tagged genes, the genomic fragments interacting with the bait were mapped to the genome using the BLAST algorithm with full genome sequence information available via

National Center for Biotechnology Information (NCBI). All sequencing reads of the interacting clones could be mapped back to Arabidopsis DNA. The genomic library preparation used in the assay was a mixed library containing DNA fragments of nuclear, plastid and mitochondrial origin.

For Rap2.4d interacting clones, 84.4 % of the sequenced DNA-fragments mapped to the Arabidopsis genome, the rest to the plastome. No hit was obtained for mitochondrial DNA. Of sequences which mapped to nuclear DNA, 94.7 % mapped to promoter regions. The other fragments mapped to similar extents to 5'-UTRs and open reading frames.

Of the Rap2.4a interacting fragments, 24 % of the sequence reads mapped to the chloroplast genome while 12.1 % could not be mapped either to nuclear or chloroplast genome. When considering those that mapped to the nuclear genome, 88 % fell on the promoter regions while the remaining 12 % was found to be on the protein sequence.

3.3 Identification of Rap2.4 targets

For plasmid analysis of the interacting DNA-fragments, DH5 α *E.coli* cells were transformed with the plasmids that were isolated from the yeast colonies growing on SD/-Leu/-Trp/-His media and plated on kanamycin containing media. It selectively favored growth of cells harboring the prey plasmids. Following plasmid isolation and sequencing of inserts, the reads were mapped to the Arabidopsis genome using BLAST algorithm with *A. thaliana* genomic DNA on NCBI server to determine which locus of the genome contained the binding sites of the transcription factor. Owing to the fairly long fragment size (200 - 300 bp) of the DNA library elements in stringent inverse yeast-one-hybrid, the target sites within the genome were predicted with very low error probabilities. For example, in the present study with *A. thaliana*, the error probabilities for mapping back the fragment sequences to the genome were routinely below 10^{-50} . Transcription factor binding site determination and target gene identification were focused on Rap2.4a and Rap2.4d mainly because the two

transcription factors have high sequence variability as compared to the others, and thus would represent individual subgroups of ERF1b/Rap2.4 transcription factors.

3.4 Rap2.4a

3.4.1 Rap2.4a targets

Rap2.4a was identified as the transcription factor that binds to the 2CPA promoter regulating it a redox-dependent manner (Shaikhali *et al.* 2008). Motif analysis by EMSA with short variants of predicted target site delineated that C3 and G4 of the CGCG-core are vital for Rap2.4a binding. However the 2 bp (CG) is insufficient to define the specific binding targets. Here, the *A. thaliana* genome was screened for Rap2.4 a binding sites.

In due course, in an effort to identify the binding motif and more downstream target genes of Rap2.4a, 63 sequences that were tagged by stringent inverse yeast-one-hybrid were mapped to the genome and the neighboring gene was spotted. The genes that were unraveled displayed a diverse function ranging from early development to stress response. Online program interface (Usadel *et al.* 2006) was used to annotate and classify the tagged sequences. According to MapMan over representation analysis, the putative target genes of Rap2.4a classify into 11 groups with one group being unknown in its function (**Figure 19**).

Results

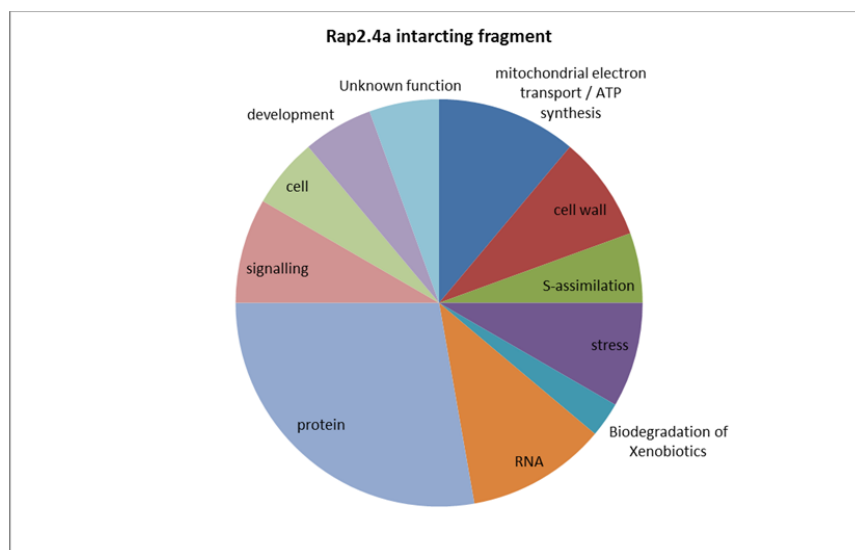


Figure 19: Putative Rap2.4a target genes classified into groups according to their function. The MapMan program was used to annotate and group the genes.

The most prominent groups of Rap2.4a targets, according to the program, were the protein (28 %), RNA (11 %), mitochondrial electron transport/ ATP synthesis (11 %) stress (8 %), and signaling (8 %).

A table with all identified hits can be found at the **appendix 6.2**. Among the tagged genes are Ethylene-responsive element binding factor 13 (ERF13, At2G44840) that contains AP2/ERF DNA binding domain. Its expression is modulated by ABA. ERF13 is a negative regulator of programmed cell death and implicated in the regulation of plant hypersensitive (HR) responses. Enhancing PRW8-mediated HR-like cell death (ERH1, At2G37940) was another gene among the targets of Rap2.4a. Similar in function to ERF13, ERH1 is involved in negative regulation of programmed cell death. Another likely target of Rap2.4a revealed was the WRKY DNA-binding protein 27 (WRKY27, At5G52830) which is involved in the biotic pressure response such as bacterial infection. In addition, the basic helix-loop-helix (bHLH) transcription factor (AT3G20640) that is located in the nucleus and in chloroplasts and is implicated to be involved in cell elongation and seed germination is also present in the list.

3.4.2 Rap2.4a motif analysis

3.4.2.1 MEME motif analysis of Rap2.4a

To determine the binding site of Rap2.4a, the insert sequences of clones that were found to interact with the transcription factor were analyzed for the presence of consensus sequence using MEME 4.9.0 (Multiple EM for Motif Elicitation)(Bailey *et al.* 2009, Bailey *et al.* 2006). The two motifs that stand out were CCG (A/G)C and (A/G)C(C/T)GAC with e-values of 5.8e+005 and 3.2e+005, respectively (Figure 20).

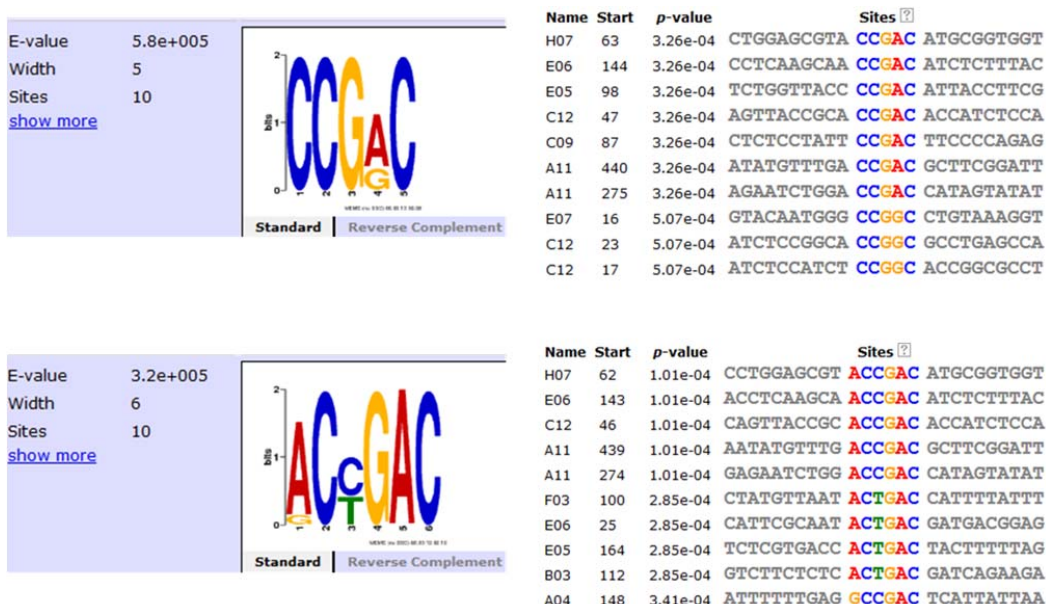


Figure 20: Motif of Rap2.4a and the corresponding sequence. Strongly interacting clones were chosen and sequenced from selection media supplemented with 160 mM and 200 mM 3-AT. The motif was elicited by analyzing using MEME, online *de-novo* motif discovery program to search ungapped consensus sequences. The sequences on the right side are strongly interacting parts of the fragment and are enriched by the eluded motif.

Taken together, the two motifs have resemble to the (A/G)CCGAC of the C-repeat (CRT)/Dehydration Responsive Element (DRE), which is present in many cold regulated (COR)

genes (Thomashow 1999, Yamaguchi-Shinozaki and Shinozaki 2006). The motif has a CG signature that was shown to be present in the Rap2.4a binding site (Shaikhali *et al.* 2008).

3.4.2.2 SCOPE motif analysis of Rap2.4a

To test the DNA fragments with an algorithm considering organism specific genome patterns and to search for bipartite motifs, the sequencing reads were further analyzed with SCOPE 2.1.0, which combines BEAM, PRISM and SPACER (Chakravarty *et al.* 2007). The sequences of the Rap2.4a interacting clones were additionally analyzed using the SCOPE 2.1.0 program which uses *A. thaliana* as a background sequence. The analysis gave a similar result as the first prediction with the MEME program but in a reverse orientation (**Figure 21**).

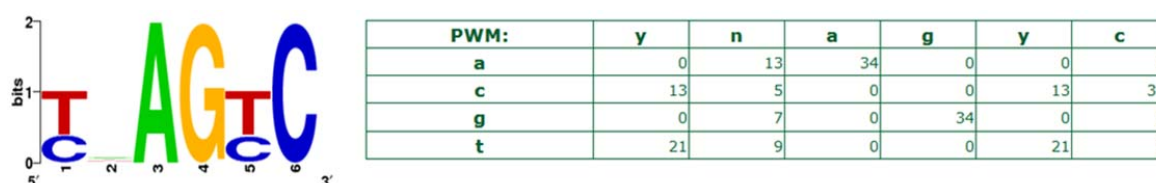


Figure 21: Rap2.4d motif as elicited by SCOPE 2.1.0 program. The table on the right indicates the occurrences of each nucleotide among the totally analyzed sequences at a given position, described by probability weight matrix (PWM).

3.5 Rap2.4d

3.5.1 Rap2.4d Targets

Similarly, the genes that were identified as putative downstream targets of Rap2.4d have a wide array of function ranging from redox control to heme transporter activity. According to the MapMan functional annotation (Usadel *et al.* 2006) (**Figure 22**), the tagged genes can be classified into 13 groups including one group whose function is unknown. Among the target genes of RAP2.4d, the major groups are involved in function related to protein (19 %), RNA

Results

(13%), transport (10 %), signaling (8 %), DNA (8 %) and lipid metabolism (8 %). A complete list of all identified hits as putative target genes is presented in the **appendix 6.3**. Among the candidate targets of Rap2.4d is Mitogen-activated protein kinase 3 (MPK3, AT3G45640) whose involvement include ABA mediated signaling pathway, negative regulation of programmed cell death, response to cold and response to oxidative stress. Sulfiredoxin (SRX, AT1G31170) that is located in the chloroplast and mitochondria, which respond to the oxidative stress, is another gene on the list.

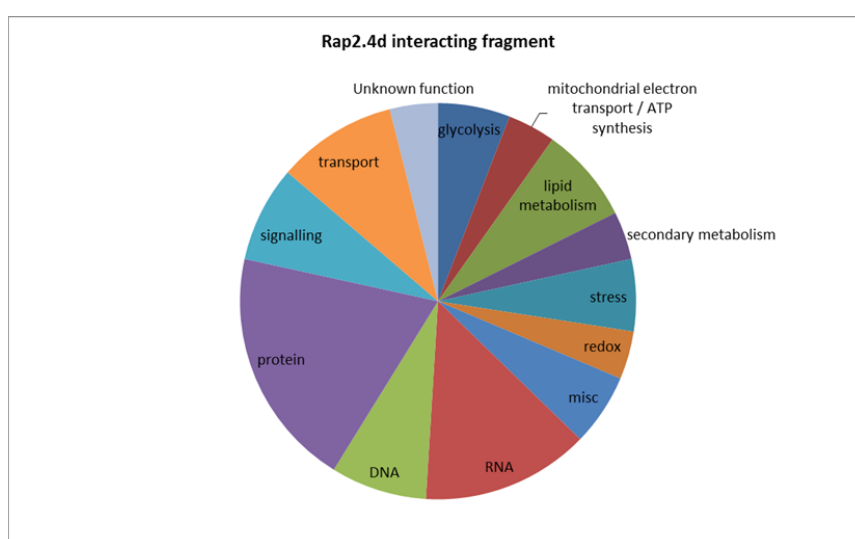


Figure 22: Putative Rap2.4a target genes classified into groups according to their function. MapMan program was used to annotate and group the genes.

3.5.2 Motif analysis of Rap2.4d

3.5.2.1 MEME motif analysis of Rap2.4d

The reads were analyzed for the presence of conserved motifs using MEME 4.9.0 (Multiple EM for Motif Elicitation). The motifs GHGGCG and AGGCS motifs were enriched within the interacting DNA fragments (**Figure 23**). The error probabilities for the motif prediction were $1.0e+003$ and $6.8e+001$, respectively, demonstrating that, despite some variations, the

Results

predicted motif signature is strong and sufficiently reliable to use it to screen the genome for additional putative DNA binding sites. MEME focuses on finding un-gapped non-degenerated and degenerated sequence patterns.

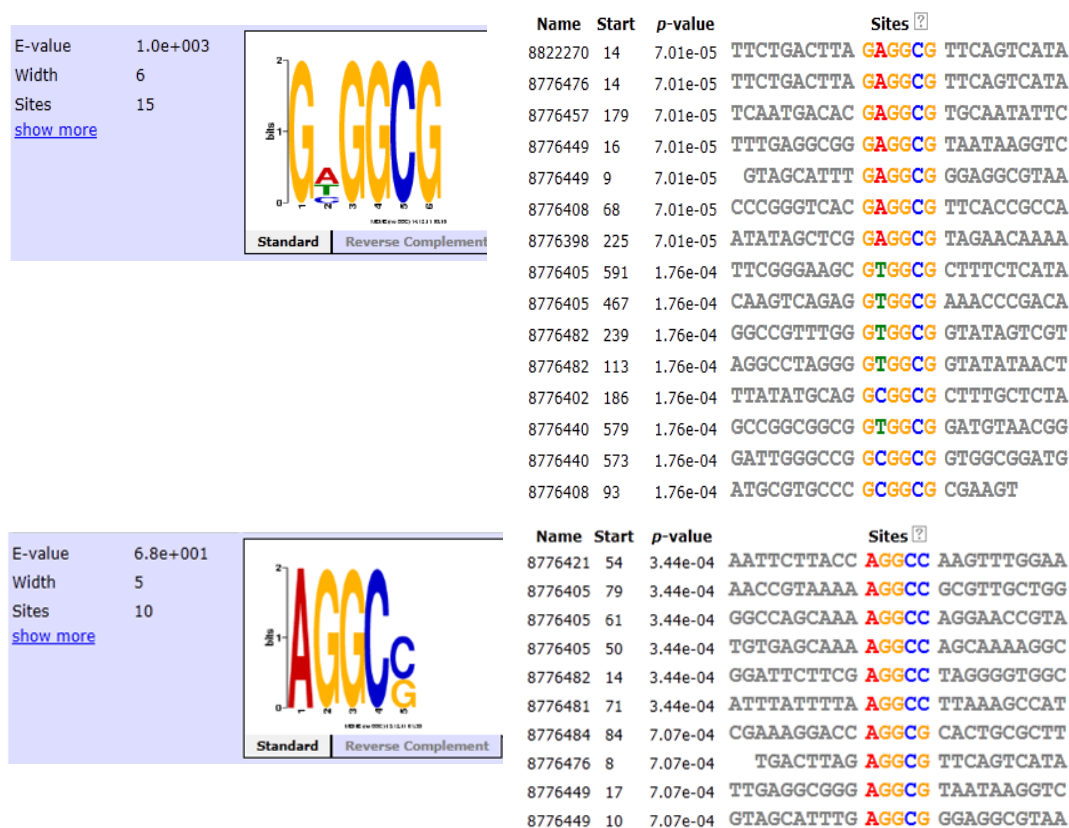


Figure 23: Determinations of Rap2.4d motif. Strongly interacting clones were chosen and sequenced from selection media supplemented with 160 mM and 200 mM 3-AT. The motif was elicited by analyzing using MEME, online *De novo* motif discovery program to search ungapped consensus sequences. The sequences on the right side are strongly interacting parts of the fragment and are enriched by the deduced motif.

3.5.2.2 SCOPE motif analysis of Rap2.4d

The sequences of the Rap2.4d interacting clones were additionally analyzed using the SCOPE 2.1.0 program which uses the *A. thaliana* genome as a background sequence. Searching for intergenic sequences of *A. thaliana* coupled with comparison of interacting DNA fragments,

Results

SCOPE analysis pointed out RGGCGTD (significance value 15.1, coverage 17.8 %) and GNCMBS (significance value 49.2, coverage 75.6 %) as putative Rap2.4d binding motifs (**Figure 24**). These motifs share the GGC core and match for the HGGCS consensus predicted by MEME analysis. The motif was found in more than 25 % of the analyzed DNA fragments indicating a high selectivity of the stringent inverse yeast-one-hybrid screen for the strongest bound DNA motifs. Another hexamer motif predicted by the program was CCGCC (significance value 54.2 and coverage 24.3 %). This motif resembles a GCC-box (GCCGCC) which is crucial element in responses to cold and dehydration (Shinozaki and Yamaguchi-Shinozaki 2000) has the affinity towards ERF1 transcription factor (Solano *et al.* 1998).

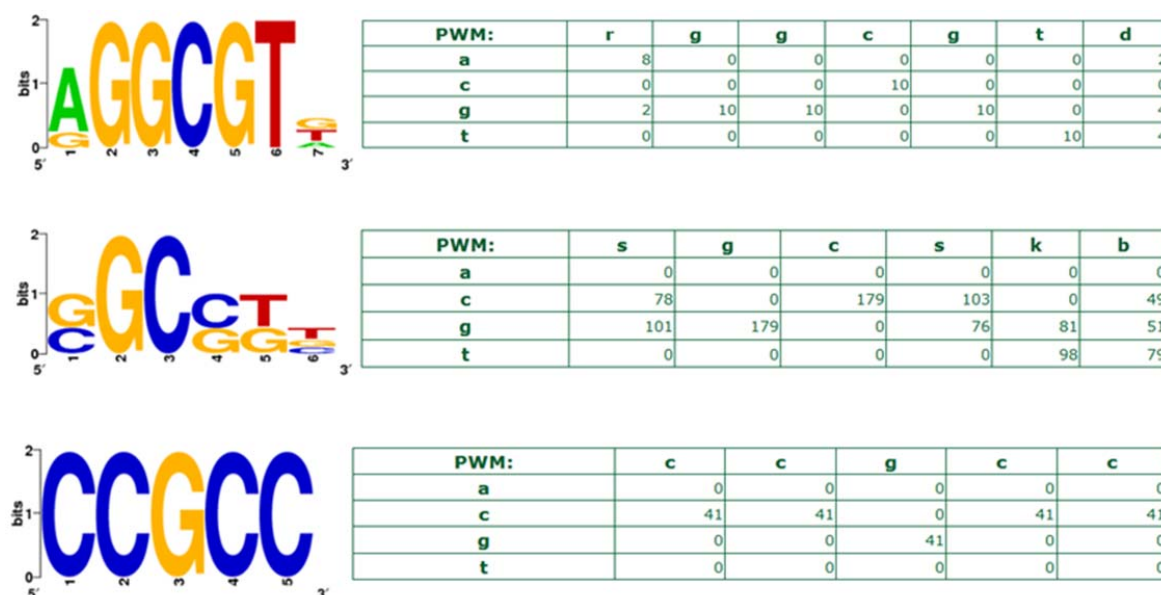


Figure 24: Rap2.4d motif as elicited by SCOPE 2.1.0 program. The table on the right indicates the occurrences of each nucleotide among the totally analyzed sequences at a given position.

3.6 Expression analysis of ERF1b genes

To evaluate the result obtained from the motif analysis to the other ERF1b family members and answer the central question of specificity versus redundancy of the transcription factor family, the expression profiles of all the group members were investigated from the publicly

available microarray dataset such as from eFP browser. Given that ERF1b transcription factor family has almost identical DNA binding domain as deduced from sequence comparison, the members with similar expression pattern in a variety of stress condition would likely target similar downstream genes and thus could have functional redundancy. Data analyzed from publicly available microarray showed that the expression pattern of Rap2.4d is similar to Rap2.4c, which has a highly similar (**Figure 6**) DNA binding domain, as compared to other ERF1b members suggesting that the two genes may have similar targets which entail redundancy in function. The expression patterns of both Rap2.4c and Rap2.4d (**Figure 25**) are closely related during cold treatment in which both genes are up regulated during the first few hours of cold treatment and get their peak expression after two hours of cold treatment. Upon further cold treatment (4°C) both genes are down regulated and after around 12 hours of cold treatment, the expression gets lower beyond the basal level. Similar to the time course response to cold, Rap2.4c and Rap2.4d show up regulation during the first few hours of drought treatment. Gradually, as duration of drought increases, both Rap2.4c and Rap2.4d are down regulated. Plants stressed with heat (30°C) gave a similar trend with that of cold and drought stress. In general, Rap2.4d responds faster than Rap2.4c in the cold than in a response to drought and heat stress. Taken together, the similarity of the expression pattern of both genes in different stress condition may indicate that they have functional similarity.

Results

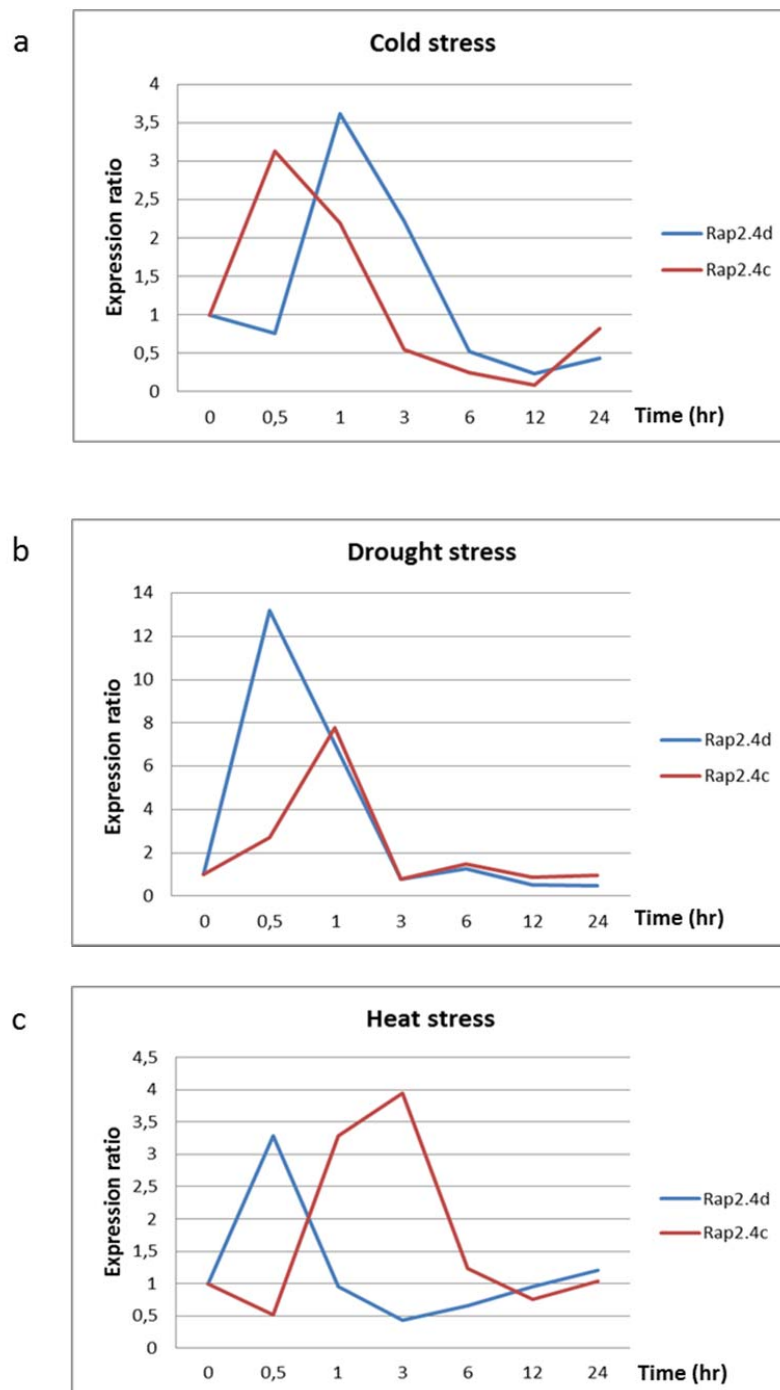


Figure 25: Expression pattern of Rap2.4c and Rap2.4d during a course of (a) cold (b) drought and (c) heat treatment. The expression values are expression ratio relative to Actin-7. The data was obtained from publicly available microarray via eFP browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>).

3.7 Promoter analysis of selected genes

All proteins in the ERF1b transcription factor family are characterized by a conserved AP2 domain, but non-conserved C- and N-termini (Nakano *et al.* 2006). It has been shown that Rap2.4a, which is a member of ERF1b transcription factor family, is a redox-sensor and a transducer of redox information and activates the expression of 2-Cys peroxiredoxin-A (2CPA) antioxidant enzyme (Shaikhali *et al.* 2008). It was, thus, imperative to analyze the promoters of selected similar genes using the determined motif of Rap2.4a and Rap2.4d. As the family members in ERF1b share a great deal of sequence similarity, the interaction of Rap2.4a and Rap2.4d transcription factors with promoters such as 2CPA, stromal ascorbate peroxidase (sAPx) and thylakoid membrane bound ascorbate peroxidase (tAPx) were investigated. Similar to 2CPA, sAPx and tAPx are nuclear encoded chloroplast antioxidant enzymes whose upstream regulator and signal transduction mechanism is yet to be investigated.

The promoter of 2CPA was analyzed using the determined Rap2.4a binding site and found to contain one version of the binding site (ACCGA) 170 bp upstream of the transcription start site (**Figure 26**) which mimics (A/G)CCGAC of the C-repeat (CRT)/Dehydration Responsive element (DRE). The CG signature of the core CE3-like element shown to be crucial for Rap2.4a binding to the redox box of 2CPA (Shaikhali *et al.* 2008). The identified binding site contained the same signature (CG) but downstream of the redox box and located fairly close to the transcription start site. Similarly, promoter analysis of tAPx unveiled that the second version of the Rap2.4a binding site (CCGAC) is present 504 bp upstream of the transcription start site (**Figure 26**) indicating that Rap2.4a could also involve in the regulation of tAPx. In contrary to tAPx, sAPx promoter does not have any of the two versions of the Rap2.4a binding sites. These observations are consistent with the observation that the transcript level of tAPx, but not sAPx was reduced by 40 % in Rap2.4a knockdown plants (in collaboration with Radoslaw Rudnik).

Results

Since the expressions of both COR6.6 and Kin1 are modulated by the DREB1 transcription factor (Liu *et al.* 1998) whose binding site is similar to that of Rap2.4a binding site, the promoters of COR6.6 and Kin1 were examined for the presence of Rap2.4a motifs. The promoter analysis divulge that Kin1 possesses both variants of the Rap2.4a binding sites at 85 and 683 bp upstream of the transcription start site (**figure 26**) hinting that Rap2.4a is involved involvement in cold stress regulation. Similarly, the promoter of COR6.6, which is a close homologue of Kin1, has both motif variants at 30 and 31 bp upstream of the transcription start site. The activity of Rap2.4a in cold regulation is further indicated by the fact that some of the identified target genes of Rap2.4a from library screen encode phosphatase 2C family protein (AT3G05640), 18Sr-RNA protein (AT2G01010), Leucine-rich repeat protein kinase family protein (AT2G28970) that are involved in cold and ABA signaling.

Based on *in silico* analysis, the promoter of tAPx (up to -700 bp upstream of the start codon) was found to contain two variants of the Rap2.4d binding motif (GAGGCG) located at -18 bp, and -308 bp, upstream of the transcription start site, while the sAPx promoter shows only the of variants (GAGGCG) very near (- 8 bp upstream) to the translation site. Consistent to promoter analysis, qRT-PCR analysis (in collaboration with Radoslaw Rudnik) indicated that Rap2.4c and Rap2.4d have differential regulatory impacts on the transcript level of chloroplast peroxidases as shown in the **figure 27**.

Results

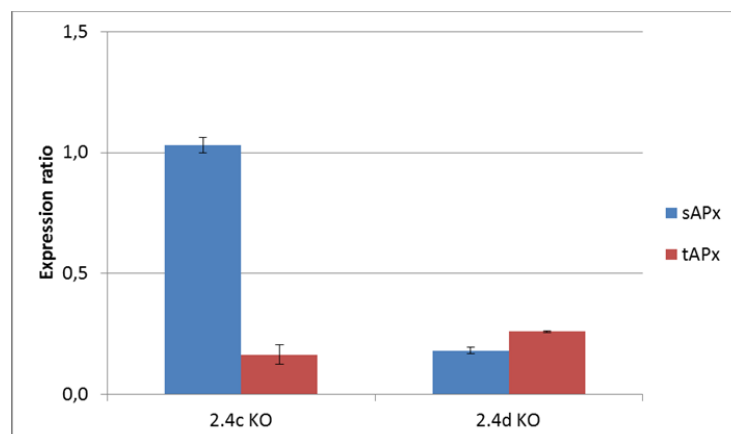


Figure 27: The relative transcript abundance of sAPx and tAPx as analyzed in Rap2.4C and Rap2.4d knock out lines (Collaboration with Radoslaw Rudnik). The expression ratio is normalized to Actin7 transcript level in Colombia wild type under standard conditions.

As for 2CPA, none of the Rap2.4d motif variants were present in the promoter region suggesting that Rap2.4d is not involved in the regulation of 2CPA and its function is different from that of Rap2.4a.

In silico promoter analysis also revealed that the full version of Rap 2.4d binding site is present in the two closely related cold regulated genes, namely Kin1 and COR 6.6 (Hughes and Dunn 1996). The Rap2.4d binding site (GAGGCG) is located -184 and -134 bp upstream of Kin1 and COR6.6 respectively, which is fairly close to the transcription start site. Since both COR6.6 and Kin1 are cold responsive genes (Shinozaki and Yamaguchi-Shinozaki 1997, Syed *et al.* 2009), it was hypothesized that Rap2.4d is involved in cold stress regulation.

Results



Figure 26: The promoter analysis of target genes by Rap2.4a and Rap2.4d transcription factors. The promoter sequences are to -700 bp upstream from the transcription start site. The sequence (GAGCCG) highlighted by the yellow is the Rap2.4d transcription factor binding site while the sequence written in red (ACCGA) and underlined by blue (CCGAC) show two putative Rap2.4a binding motifs.

3.7.1 Interaction of ERF1b with tAPx and sAPx promoter

Yeast-one-hybrid tests were employed to determine if there is a direct interaction between chloroplast APx promoters and the ERF1b transcription factors. The yeast strain Y187 was co-transformed with 200 ng of pHis2.2 plasmids harboring the tAPx promoter fragments of reporter plasmid and ERF1b transcription factor fusion constructs cloned in pACT2. For the interaction assay, the co-transformed yeast stains that were suspended in SD/-Leu/-Trp media to grow overnight were diluted to OD 0.01 with water **Figure 28**.

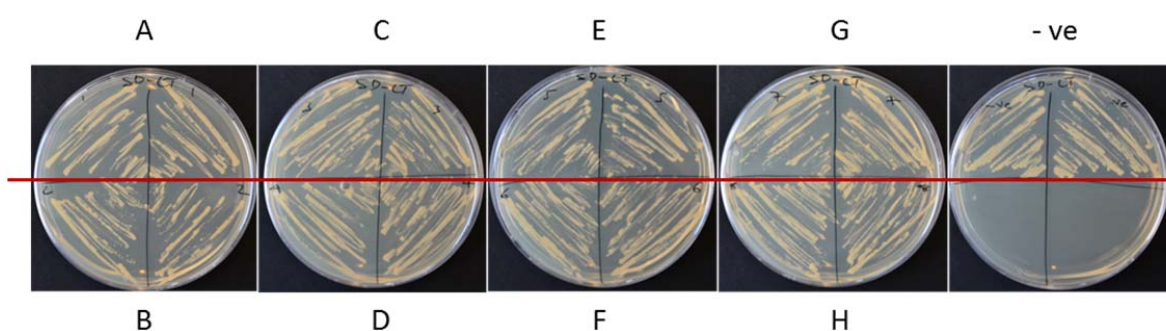


Figure 28: Transformation control, Y187 yeast strain co-transformed with both activation domain fusion protein construct and the reporter plasmid (tAPx-B2) plated on SD/-Leu/-Trp medium. The letter A to H indicates the eight ERF1b (Rap2.4a to Rap2.4 h) transcription factors. The pictures demonstrate equal dose transformation of both plasmids in all the case of transformation including the negative control. The negative control is a yeast transformed with normal reporter plasmid construct but empty activation domain vector.

Before proceeding to the interaction assay, the reporter vector was checked for the leakiness of the interaction marker, HIS3. For tAPx-B1 and tAPx-B2 promoter constructs, 5 mM and 70 mM supplement of 3-AT which is a competitive inhibitor of histidine biosynthesis into the SD/-Leu/-Trp/-His media, suppressed all the unspecific interactions caused by endogenous yeast transcription factors.

Results

Generally, the result indicated that there is a strong unspecific activation of tAPx-B1 by all ERF1b transcription factors whereas tAPx-B2 differentially interacted with the transcription factor family. Rap2.4a showed better interaction as compared to other 7 ERF1b transcription factors with the proximal fragment of the promoter (**Figure 29**).

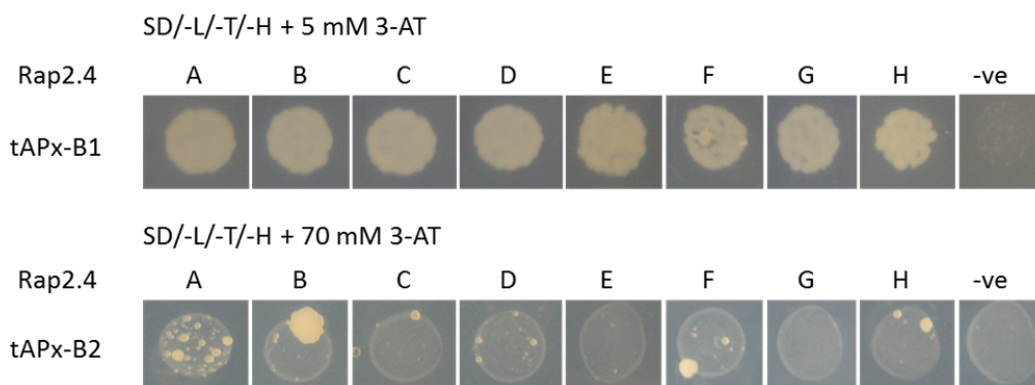


Figure 29: The interaction of ERF1b proteins with two (tAPx-B1 and tAPx-B2) overlapping fragments of the tAPx promoter. The promoter fragments and the transcription factors to be tested were cloned in pHIS2.2 and pACT2 vector respectively, and co-transformed into the yeast strain Y187. For the interaction assay the co-transformants of tAPx-B1 and tAPx-B2 were plated on SD/-Leu/-Trp/-His supplemented with 5 mM and 70 mM 3-AT respectively. Negative control (-ve) is a yeast transformed with respective pHIS2.2 promoter construct and with an empty expression vector (pACT2).

The promoter of sAPx was analyzed for its interaction with the ERF1b transcription factor. Before the actual screen, the yeast strain that was transformed only with the reporter plasmid was analyzed for the leakiness of the interaction marker by plating it on SD/-Leu/-Trp media supplemented with increasing amounts of 3-AT (0, 15, 50, 70, 90 mM) as shown in the **figure 30a**. The sAPx promoter shown to be leaky, demonstrated by the fact that even at 90 mM 3-AT expression of HIS3 was activated and there had been a growing colonies on histidine free media. In presence of Rap2.4a additional reporter gene activation was observed, whereas in presence of Rap2.4b, Rap2.4c and Rap2.4f the background activation was suppressed, indicating putative differential the interaction by ERF1b family.

Results

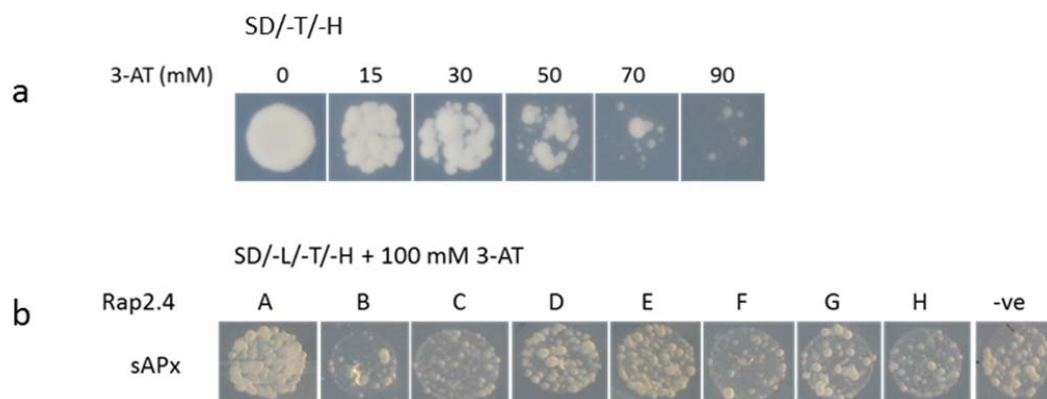


Figure 30: The interaction of sAPx promoter with ERF1b transcription factor. (a) The yeast strain Y187 transformed only with reporter construct containing the sAPx promoter. The construct was plated on increasing amounts of 3-AT on SD/-Trp/-His. (b) The yeast strain Y187 co-transformed with both promoter construct and fusion protein constructs and plated on SD/-Leu/-Trp/-His + 100 mM 3-AT.

3.7.2 Expression of tAPx and sAPx in the Rap2.4c knockdown and overexpression lines

To further characterize the functional relationship with Rap2.4c, whose DNA binding domain is identical to that of Rap2.4d, the expression of sAPx and tAPx was analyzed on Rap2.4c knockdown and overexpression lines. Rap2.4c knockdown line showed more than fourfold reduced expression level as compared to the wild type (Col-0) plants line, suggesting a vital regulatory role of Rap2.4c of tAPx (**Figure 31**). This observation is in agreement with the result that Rap2.4c knockdown line negatively affected the expression of tAPx but not sAPx (in collaboration with Radoslaw Rudnik). Overexpression of Rap2.4c led to a slight decrease in the transcript abundance of tAPx which might indicate that overexpression could result in a negative feedback loop. Consistent with yeast-one-hybrid data and the promoter motif analysis, the Rap 2.4c did not seem to regulate the expression of sAPx. Taken together, the results strongly suggest that Rap2.4c and Rap2.4d positively regulate the expression of tAPx directly.

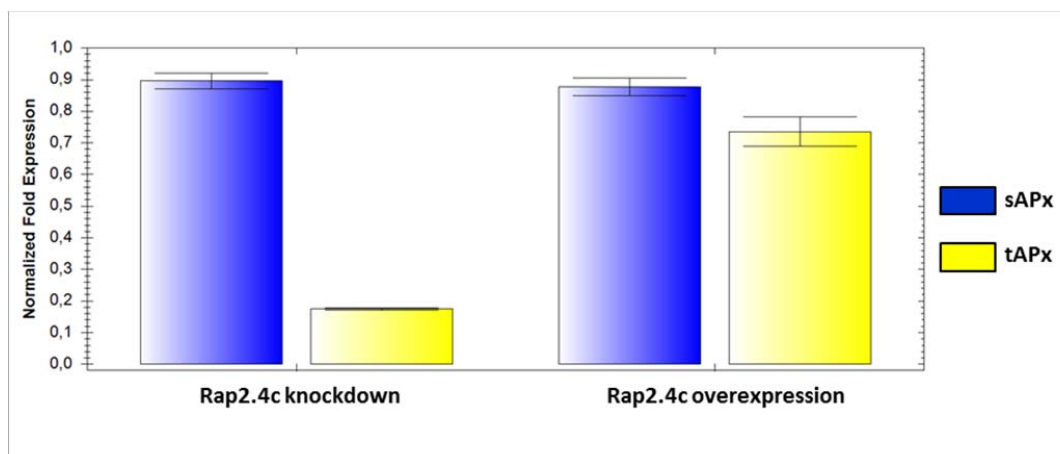


Figure 31: The relative transcript abundance of sAPx and tAPx analyzed on Rap2.4c knockdown and Rap2.4c overexpression line. The expression ratio is normalized to Actin7 transcript level in Colombia wild type under standard conditions.

3.7.3 Interaction of ERF1b with 2CPA promoter

2CPA belongs to the ascorbate independent chloroplast antioxidant defense system. ROS and ABA modulate the transcription activity of 2CPA antagonistically (Baier *et al.* 2004), where peroxides and photooxidative stress enhances its expression, and ABA decreases its expression. Shaikhali *et al.* (2008) for the first time demonstrated that Rap2.4a, which is a member of ERF1b family, is a transcriptional modulator of the 2CPA expression and that it regulates 2CPA in a redox dependent manner. The transcription factor dimerizes under moderately oxidizing condition leading to the binding to the 2CPA promoter. Under severely oxidizing conditions, the transcription factor forms oligomers, which are the inactive form and do not activate the expression of 2CPA. Given the high similarity of DNA binding domain of transcription factor family, the question of redundancy versus specificity of the group members had to be analyzed. Direct yeast-one-hybrid was employed to assess the interaction of the 2CPA promoter with the ERF1b transcription factor family.

The redox sensitive part of the 2CPA promoter (-315 bp to -706 bp) was cloned (Baier *et al.* 2004). As demonstrated by the reporter plasmid assay, yeast that was transformed with

Results

only 2CPA reporter construct showed no leaky effect when plated on SD/-Trp/-His media (**Figure 32a**). In conformity with the existing finding by Shaikhali et al. (2008), Rap2.4a interacts slightly with 2CPA promoter. In contrast, Rap2.4h much more strongly bound the 2CPA promoter as demonstrated by that fact that only Rap2.4h showed growth when streaked out on a selective plate containing 5 mM 3-AT (**Figure 32b and c**). In general, using pOne-1 reporter vector, in which the 2CPA promoter construct was built upon, shown to have advantage in that it hardly has leaky expression of interaction marker (**Figure 32a**).

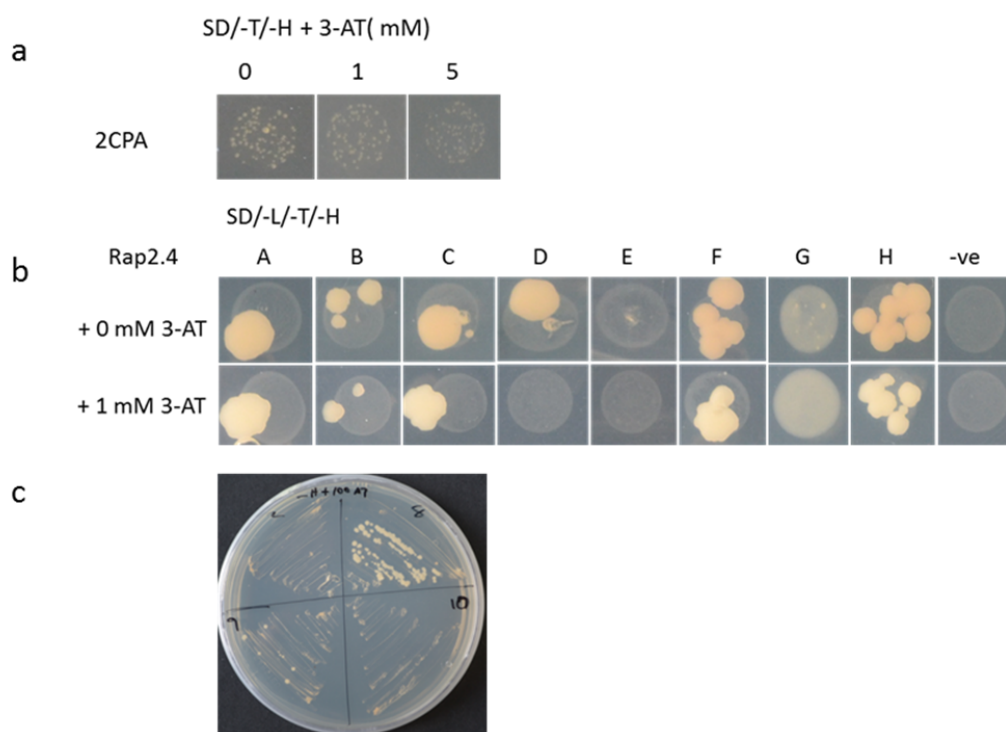


Figure 32: Interaction of 2CPA promoter with ERF1b transcription factor. (a) Yeast transformed with the reporter plasmid construct and plated on SD/-Trp/-His plus increasing amount of 3-AT. (b) Yeast co-transformed with promoter and transcription factor fusion construct and plated on SD/-Leu/-Trp/-His media with and without 3-AT supplement. (c) Yeast co-transformed with promoter and transcription factor fusion construct and streaked on SD/-Leu/-Trp/-His media supplemented with 5 mM 3-AT.

3.8 Rap2.4c and Rap2.4d in cold regulation

The identified Rap2.4d binding site was shown to be present close to the transcription start site on two closely related genes (Kin1 and COR6.6) implying that Rap2.4d is involved in cold stress regulation. It has been implicated that many genes that are regulated by the cold are also responsive to ABA, drought and osmotic stress (Shinozaki and Yamaguchi-Shinozaki 2007), which perplex the study of distinct cold regulation mechanism. The promoter activity analysis of COR6.6 (Kin2) in which the promoter was fused to β -glucuronidase (GUS) revealed that its transcription is induced by low temperature (Wang *et al.* 1995). Local cold treatment of transgenic *A. thaliana* (in the background of C24) with a luciferase marker fused to the promoter of COR6.6 demonstrated that, in warm grown plants COR6.6 reacts to a systemic signal emitted from remote but directly cold treated leaves (Gorsuch *et al.* 2010). In addition, the expression of COR6.6 is positively modulated by TINY, a dehydration-responsive element (DRE)-binding Protein-like transcription factor through direct binding to the promoter region, which itself is activated by cold, drought, ethylene, and slightly by methyl jasmonate (Sun *et al.* 2008). Cold induced genes such as Kin1, which is a close homologue of COR6.6, are thought to be regulated by CBF3 (Yamaguchi-Shinozaki and Shinozaki 2006). Genes, for instance Kin1, COR6.6 and COR47 do not necessarily require ABA and can be induced by drought (Shinozaki and Yamaguchi-Shinozaki 1997). Despite the fact that the COR6.6 response to cold is known for a long time, the exact regulator(s) that modulate the expression of the gene is/are subject of many studies. The expression pattern of Rap2.4d was analyzed from publicly available microarray data using eFP Browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>). Rap2.4d expression increases after 1 hour of cold treatment at 4°C, but gradually decreases when exposed to longer hours of cold as shown in the **figure 25**. Similarly, the transcript profile of Rap2.4c follows the same pattern as Rap2.4d (**Figure 33**), suggesting that Rap2.4c and Rap2.4d are cold responsive. In contrast, following 1 hour exposure of plants to cold leads to the slight decrease on COR6.6 expression, and upon prolonged cold treatment the expression value increases. The expression pattern of COR6.6 and Rap2.4d is therefore opposite and taken together with the

presence of Rap2.4d binding site on the promoter region of COR6.6, the results suggest that Rap2.4d is a direct but negative regulator of COR6.6 expression. The transcript abundance of BAP1 which is up regulated during cold seems to be negatively correlated with the expressions of Kin1 and COR6.6.

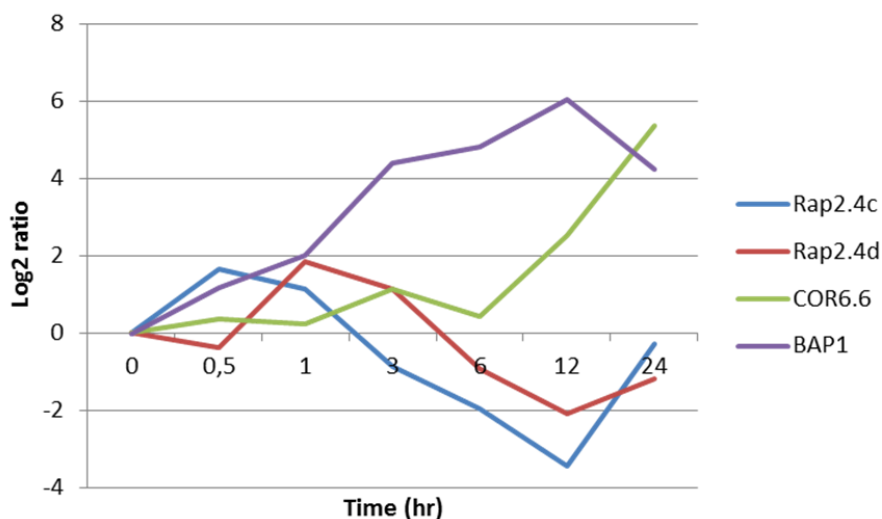


Figure 33: The relative expression level of Rap2.4c, Rap2.4d, COR6.6 and BAP1 during cold treatment and control. The data were obtained from Publicly available microarray eFP browser (<http://bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi>). The x-axis shows the length (in hours) of cold treatment.

3.8.1 Rap2.4d negatively regulates the expression of COR6.6

The Rap2.4d binding site was found to be present in close proximity to the transcription start site of COR6.6 promoter (**Figure 26**). Expression analysis from publicly available microarray data indicated that the expression pattern of the two genes is inversely regulated (**Figure 33**) indicating that Rap2.4d inhibits regulation of COR6.6 transcription. To determine the precise regulation pattern of COR6.6, a Rap2.4d T-DNA insertion line (overexpression line) was analyzed for the expression of COR6.6 and Rap2.4d. In addition, the transcript level of Rap2.4c, a close homologue of Rap2.4d, was analyzed to determine if there is a compensatory effect between the two close homologues. Furthermore, since the

expression pattern of Rap2.4c and Rap2.4d are similar when exposed to different stresses including cold (**Figure 25**), functional redundancy was expected. qRT-PCR was used to determine the expression level of Rap2.4c, Rap2.4d and COR6.6. The T-DNA insertion line of Rap2.4d, which was meant to knock out/down the gene, was found to be overexpressor line. This is plausible given the fact that the T-DNA insertion was not located in the reading frame but on the regulatory sequences, 3'UTR. It has been reported that there are many T-DNA insertion lines at the 3' or 5' side of the gene that lead to the activation of endogenous gene in its native location, because of the insertion in the regulatory sequences where enhancers or repressors could bind (Li *et al.* 2006, Ulker *et al.* 2008). To analyze the functional significance of Rap2.4d overexpression on COR6.6 which is induced by cold, both Rap2.4d T-DNA insertion line and Col-0 plants were treated with cold for 24 hrs at 4°C. Rap2.4d expression was up regulated by about 50 % on Rap2.4d overexpression line which led to the down regulation of COR6.6 by more than 20 % and Rap2.4c by about 70 % as compared to the wild type control plant in non-cold treated plants. In conformity to the publicly available microarray data, cold treatment led to decreased transcript level of Rap2.4d. However, the COR6.6 which should normally be up regulated during cold was not induced (**Figure 34**). This is assumed to be due to the slight overexpression of Rap2.4c which is presumed to have similar impact during cold and which possibly has a stronger negative regulation on COR6.6 than Rap2.4d.

Results

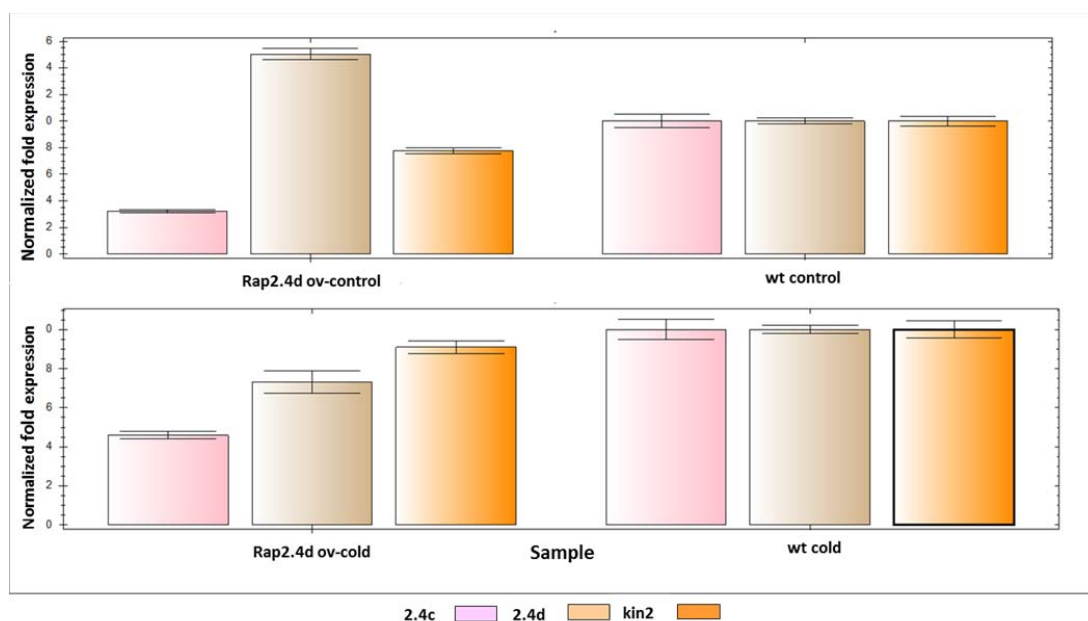


Figure 34: The relative transcript abundance of Rap2.4d T-DNA insertion line leading to overexpression and wild type (*Col-0*) control of cold treated and untreated samples analyzed by qRT-PCR. The plants were grown on nutrient rich soil substrate for 4 weeks at 20°C and transferred to 4°C for 24 hrs for cold treatment. The transcript abundance was determined relative to Actin-7 and Fbox.

3.8.2 Rap2.4c overexpression

To further establish the regulatory role of Rap2.4c on the expression of COR6.6 gene, the overexpression line of Rap2.4c was investigated. qRT-PCR was employed to analyze the expression pattern of the Rap2.4c in cold treated and non-treated plant. To check the effect of T-DNA insertion on the 3'UTR side of Rap2.4c, two primer pairs were used (**Figure 35**).

Results

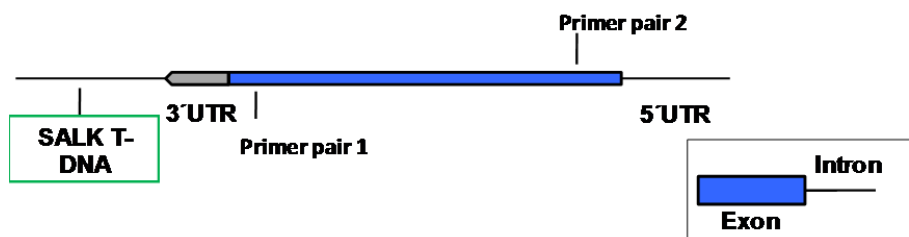


Figure 35: A schematic depiction of Rap2.4c overexpression line at the 3'UTR region. Primer pair 1 and 2 are the two different primer pairs that were used to check the transcript abundance of the Rap2.4c gene.

The T-DNA insertion on the 3'UTR side of the Rap2.4c led to the overexpression of Rap2.4c and the results were consistent when analyzed with both primer pairs (**Figure 36**). In addition, the overexpression pattern was the same after 24 hour the cold treatment at 4°C.

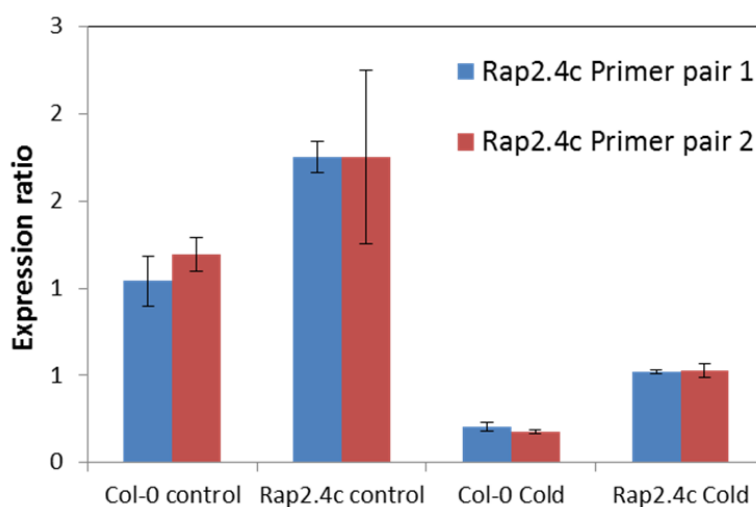


Figure 36: The relative transcript level of Rap2.4c in control and cold treated plants of Col-0 and Rap2.4c T-DNA insertion line analyzed by Rap2.4c primer pair1 and primer pair2. The plants were grown on nutrient rich soil substrate for 4 weeks at 20°C and transferred to 4°C for 24 hrs. The transcript abundance was determined relative to Actin-7.

Consistent with the publicly available microarray data, the expression of Rap2.4c was drastically reduced during cold treatment in both Col-0. In the Rap2.4c T-DNA insertion overexpression line, the Ra2.4c transcript level was also decreased, but still higher than in cold treated Col-0 plants.

3.8.3 Rap2.4c negatively regulates the expression of COR6.6 and BAP1

To determine the regulation of COR6.6 by Rap2.4c, the T-DNA overexpression line of Rap2.4c and wild type control (Col-0) were treated with and without cold and analyzed for the transcript levels of Rap2.4c, COR6.6 and BAP1. In line with the report that COR6.6 expression was high during cold treatment (Shinozaki and Yamaguchi-Shinozaki 1997, Syed *et al.* 2009), plants showed heightened expression in Col-0 and in the Rap2.4c insertion line after cold treatment. Contrary to the expression pattern of COR6.6, Rap2.4c expression was reduced after cold treatment suggesting that there is a negative regulation between the two genes. The basal expression of COR6.6 was higher as compared to that of Rap2.4c in Col-0 control plants (**Figure 37**).

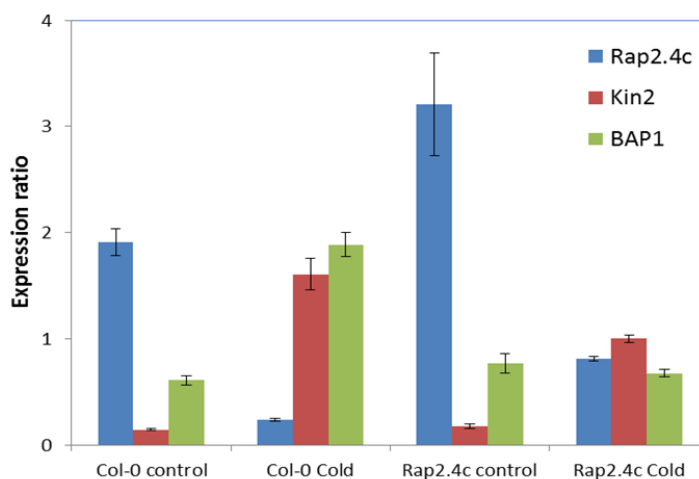


Figure 37: The relative transcript abundance of Rap2.4c, COR6.6 and BAP1 in control and cold treated plants of Col-0 and Rap2.4c T-DNA insertion line. The plants were grown on nutrient rich soil substrate for 4 weeks at 20°C and transferred to 4°C for 24 hrs. The transcript abundance was determined relative to Actin 7.

The transcript abundance of Rap2.4c was further boosted on Rap2.4c T-DNA insertion line owing to the T-DNA insertion. After 24 hrs of cold stress at 4°C, the transcript level of Rap2.4c in the T-DNA insertion line was not as low as compared to the expression in Col-0 cold treated plant due to the overexpression effect. However, the expression of COR6.6 in the overexpression line after cold stress was not as high as the expression in Col-0 cold plants accounting for the relatively higher presence of Rap2.4d which provides affirmation that Rap2.4c indeed negatively regulates COR6.6.

BON1-associated protein (BAP1) is a membrane associated C2 domain protein whose overexpression in plant abolishes reactive oxygen species (ROS) activated apoptosis (Yang *et al.* 2007). A wide range of stresses such as temperature fluctuation, biotic pressure, and mechanical stress could result in the induction of BAP1. It has been demonstrated that the activation of BAP1 by a moderate decrease in temperature is mediated by ICE1 as in the case of COR genes (Zhu *et al.* 2011). To check whether Rap2.4c regulates the expression pattern of BAP1, its transcript was also analyzed on the same cold treated plants of Col-0 and Rap2.4c insertion line. The expressional regulation pattern of BAP1 followed the same pattern as that of COR6.6, conforming that Rap2.4c is a negative regulator of cold regulons.

3.8.4 Rap2.4c as a general negative regulator of cold responsive genes

Rap2.4c whose expression profile across different stress condition is similar to that of Rap2.4d expression pattern demonstrated to be a strong negative regulator of COR6.6 and BAP1 genes which are normally induced by cold (**Figure 38**). Given that COR6.6 and BAP1 represent a wider gene regulation pathway, and that Rap2.4c regulates both, it was hypothesized that Rap2.4c is a general negative regulator of cold induced genes. To test this hypothesis, different cold responsive genes that belong to different gene families were tested for their expression level on the wild type control and Rap2.4c overexpression line. To investigate the functional regulatory role of Rap2.4c on cold responsive genes, the wild type control plants and the Rap2.4c overexpression plants were treated with cold for 24 hrs at

Results

4°C and the transcript level of COR15A (AT2G42540), COR47 (AT1G20440), COR6.6 (AT5G15970) and PLD1 (AT3G15730) were analyzed.

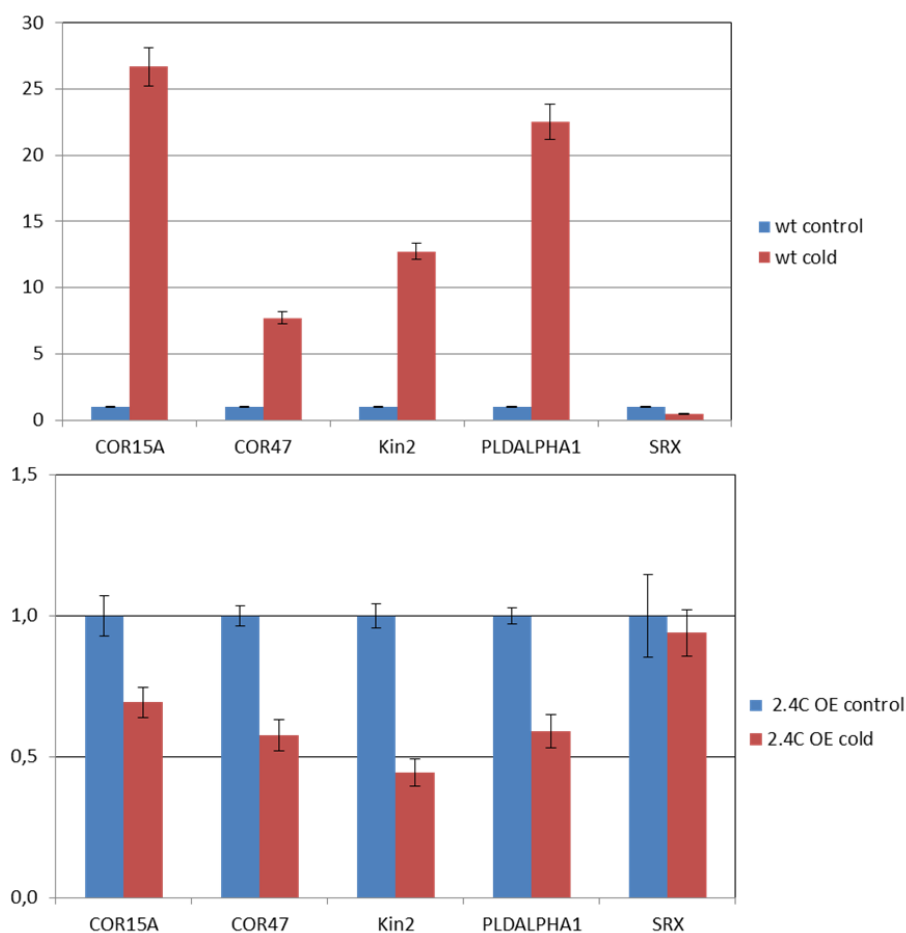


Figure 38: The relative transcript abundance of COR15A, COR47, COR6.6, PLDALPHA1 and SRX in control and cold treated plants of Col-0 and Rap2.4c T-DNA insertion line. The plants were grown on nutrient rich soil substrate for 4 weeks at 20°C and transferred to 4°C for 24 hrs for cold stress. The transcript abundance was determined relative to Actin-7.

Non-cold treated wild type plant showed low level of basal expression of the cold responsive genes. Upon cold treatment, the cold induced genes were up regulated as high as 27 times for COR15A, 8 times for COR47, 12.5 times for COR6.6 and 22 times for PLD1 fold expression. Upon cold treatment of Rap2.4c overexpression line, the transcript level of all of

the cold induced genes analyzed showed a dramatic decrease in the expression below the non-cold treated plant. Sulfiredoxin (SRX, AT1G31170), which was found to be one of the target genes of Rap2.4d in the stringent yeast one-hybrid screen, was also analyzed to see a functional redundancy of Rap2.4d and Rap2.4c apart from cold stress. Both wild type and Rap2.4c T-DNA insertion line plants did not show significant change during normal growth condition and cold treatment. This might indicate that Rap2.4c is a general negative regulator specific to cold stress. Rap2.4c and Rap2.4d may thus be important negative regulators that keep the cold induced genes under control at low level to avoid the unnecessary expenditure of energy.

3.8.5 Transient expression analysis of COR6.6 promoter

Rap2.4c was shown to negatively regulate the expression of COR6.6 and other cold regulated genes. To determine whether Rap2.4c modulates the expression of COR6.6 by direct binding to the promoter region, transient expression analysis was performed. To do the same, Arabidopsis wild type (Col-0) and Rap2.4c T-DNA insertion lines leading to the knockdown and overexpression of the Rap2.4c transcript was used. The result showed that luciferase reporter gene was activated in the wild type plant and overexpression line, but not in the knockdown line (Figure 39). Moreover, the luciferase activity is higher in the overexpression line than in the wild type plant, indicating that there is a direct interaction between Rap2.4c and COR6.6 promoter.

Results

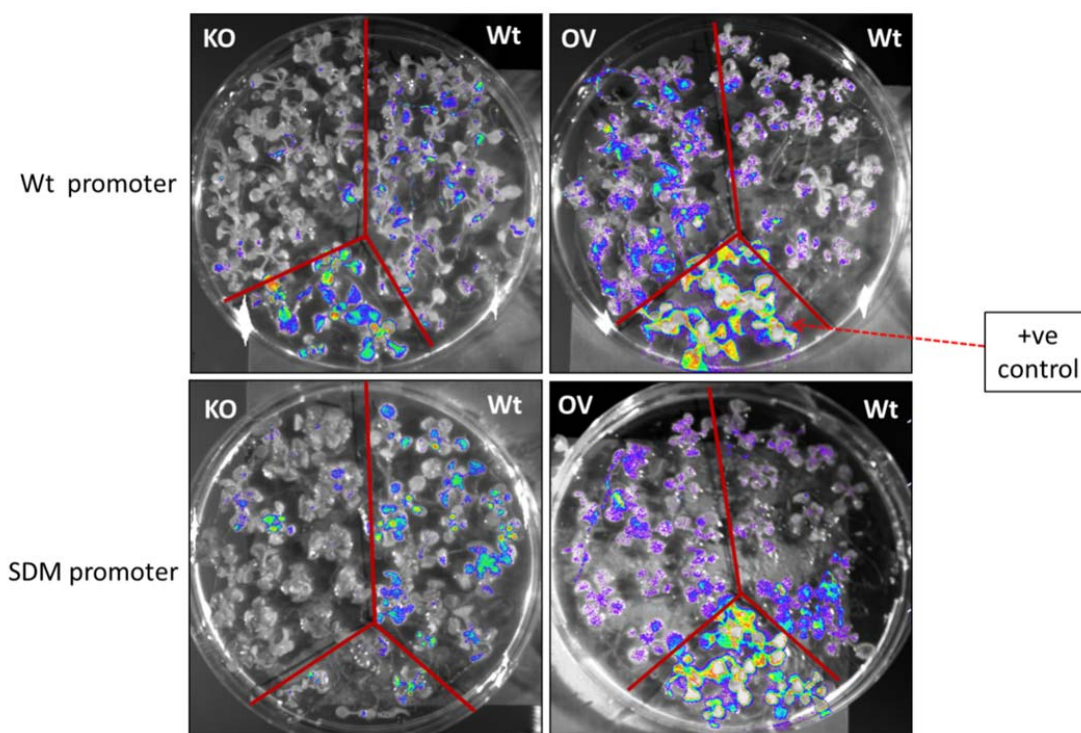


Figure 39: Transient reporter gene expression analysis of COR6.6 promoter. Agrobacterium transformed with wild type (wt) and site directed mutagenized (SDM) promoter construct in pHGWL vector were transfected to the T-DNA insertion knockdown (KO) and the overexpression (OV) line of *A. thaliana*. Before transfection, sterile seeds were grown under long day condition for 14 days in the growth chamber under standard growth conditions.

Rap2.4c is a close homologue of Rap2.4d whose binding motif was defined. To answer if the Rap2.4c binds to the same binding motif as Rap2.4d (GAGGCG), which is also present in the COR6.6 promoter, its binding site was mutated to GAAAAG in the COR6.6 promoter (0-729 bp). The result showed that, the Rap2.4c overexpression line had more luciferase activity than the wild type. In addition, similar to the case of wild type promoter, the knockdown line did not show any luciferase activity. Generally, the results indicate that Rap2.4c transcription factor has different binding site than Rap2.4d on 729 bp promoter length.

4. Discussion

4.1 Stringent inverse yeast-one-hybrid

ERF1b (Rap2.4) proteins are a family of transcription factors with highly conserved DNA binding domains (Nakano *et al.* 2006). Seven members are proposed to be nuclear targeted, one to be post-transcriptionally translocated into chloroplasts (Schwacke *et al.* 2007). The transcription factor family, contrary to the identical DNA binding domain, has highly variable C and N termini. To answer the functional redundancy or specificity of the transcription factors, their binding sites were determined. Here, a new method was developed to screen full genomes for transcription factor binding motifs and to identify transcription factor target genes (**3.2.2**), called stringent inverse yeast-one-hybrid. *A. thaliana* genomic library was set up to screen the genomes for the particular binding sites of Rap2.4a and Rap2.4d. With the average library fragment size of 258 bp (**Figure 14**), the genome of *A. thaliana* was covered more than four times, making the assay more versatile. In a recently developed method, a comparable genomic library used 300 bp for zebrafish genome and 700 bp for mouse genome, with the coverage of 4-7 and 3-4 times respectively (Zeng *et al.* 2008). With heterologously expressed Rap2.4a and Rap2.4d-GAL4 fusion proteins, strongly interacting genomic DNA fragments were isolated under highly stringent conditions (160 and 200 mM 3-AT). As a result of higher stringency set-ups, stringent inverse yeast-one-hybrid screen gave less variant motifs (**Figure 20, 23**) than the screens of mouse and zebrafish genomic DNA with Foxl1 (Zeng *et al.* 2008) and Gli (Milla *et al.* 2012), respectively, although the data were not pre filtered in the present study. In conclusion, as compared to the previously developed method by (Zeng *et al.* 2008), the sensitivity and the reliability of the stringent inverse yeast-one-hybrid is enhanced by HIS3 interaction marker and with optimal 3-AT.

4.2 Sequence analyses of Rap2.4a and Rap2.4d interacting clones

Rap2.4a and Rap2.4d transcription factor binding sites were determined using stringent inverse yeast-one-hybrid, by sequencing 96 strongly interacting colonies for each transcription factor (3.4.2, 3.5.2). The identity and the locus of each sequence were identified on the genome using the Blast algorithm. 84.4 % of the Rap2.4d interacting sequences were mapped to the nuclear genome, while the remaining sequences were identified on the chloroplast genome. As a consequence of a high number of chloroplasts in leaf cells (Koniger *et al.* 2008), a significant coverage of chloroplast genome fragments was expected. The low number of Rap2.4d hits on the chloroplast genome suggests that the transcription factor binds them specifically. Rap 2.4c, which is a close homologue of Rap2.4d, is putatively targeted to chloroplast (Schwacke *et al.* 2007) and, hence, expected that it has more number of targets on the chloroplast genome. Of the Rap2.4d hits detected on the nuclear genome, 94.7 % were found to be on the promoter region of a gene, demonstrating a high specificity for regulatory elements. The remaining Rap2.4d interacting fragments mapped to an open reading frame and to 5'-UTRs.

With regard to Rap2.4a, of the total hits that mapped to the nuclear genome, 88 % were identified on promoter regions of genes while the remaining were found to be mapped on the open reading frames. No hits in mitochondrial genome were identified for both Rap2.4a and Rap2.4d, indicating that the transcription factor has no activity in the coordination of mitochondrial gene expression. Taken collectively, the results indicate that stringent inverse yeast-one-hybrid, is sensitive enough to detect the *in vivo* regulatory regions of the genome.

In comparison, Zeng and coworkers (2008) have developed a yeast based method for genome wide screening of transcription factor binding sites. Despite the many advantages as compared to other techniques of the time, the method has hardly been applied. Milla *et al.* (2012) reported a need for extensive data curation to exclude repetitive DNA and fragments mapping to regions outside of 20 kb upstream of the transcriptional start site in

screens aiming at identification of *Gli* transcription factor binding sites in the zebrafish genome (Dahm *et al.* 2005). In contrast, it was not necessary to apply data curation in the present study. Since there was no experimental bias on non-coding regions, the high proportion of hits in upstream regions and the lack of mitochondrial hits indicates a specificity, which should guarantee a high data quality even with a much larger genome.

In conclusion, the low number of hits in plastid genome and the absence of hits in mitochondrial DNA indicate specificity. Stringent inverse yeast-one-hybrid may open insight into the function of organellar targeting of eukaryotic-type transcription factors into chloroplasts (Fujimoto *et al.* 2000).

4.3 Rap2.4a

4.3.1 Targets of Rap2.4a

Rap2.4a was shown to regulate the expression of the nuclear encoded chloroplast antioxidant enzyme called 2-cysteine peroxiredoxin-A (2CPA) in a redox dependent manner (Shaikhali *et al.* 2008). 2CPA is a peroxidase without heme group which detoxifies H₂O₂ that is produced as a result of various stresses. To identify further targets of Rap2.4a, cognate genes of the hits of the transcription factor, identified by stringent inverse yeast-one-hybrid, were analyzed. Based on MapMan over representation analysis program (3.4.1), the target genes of Rap2.4a were categorized into 11 groups (**Figure 19**). The function of the putative target genes of Rap2.4a varied from the signaling to cell development, suggesting that Rap2.4a, in addition to the regulation of antioxidant enzyme, has diverse biological functions that connect different signaling pathways.

One of the identified targets of Rap2.4a is ethylene-responsive element binding factor 13 (ERF13, At2G44840). ERF13 belongs to the same super-family as Rap2.4a which is characterized by having an AP2/ERF DNA binding domain (Nakano *et al.* 2006, Sakuma *et al.*

2002). The ERF13 transcription factor is involved in long distance signaling as demonstrated by the fact that the leaves of root wounded plants showed high levels of ERF13 gene expression (Sogabe *et al.* 2011). Abscisic acid (ABA) was shown to orchestrate its response by modulating downstream targets through ABA response elements (ABRE) and coupling elements (CE) that work in concert to effect the signaling (Kim 2006, Narusaka *et al.* 2003, Shen and Ho 1995). ERF13 was demonstrated to bind to CE1 element which is suggested to be a very important mediator in biotic and abiotic responses (Lee *et al.* 2010). *A. thaliana* plants infected with *Pseudomonas syringae* showed increased transcript abundance of ERF13 indicating that it plays a vital role in biotic stress response (Onate-Sanchez and Singh 2002).

Enhancing PRW8-mediated HR-like cell death (ERH1, At2G37940) that activates hypersensitive response (HR) through salicylic acid mediated signaling pathway is another putative downstream target gene of Rap2.4a identified by stringent inverse yeast-one-hybrid. ERH1 encodes inositolphosphorylceramide synthase (IPCS) which is an important enzyme in the sphingolipid metabolism that is involved in defense related programmed cell death (Wang *et al.* 2008).

In addition, Rap2.4a is found to be interacting with the WRKY27 (At5G52830) promoter in the screen for regulatory fragments. WRKY is a plant specific transcription factor family whose members are involved in plant specific processes such as trichome development and leaf senescence (Miao and Zentgraf 2007, Ulker and Somssich 2004). WRKY27 acts antagonistically in a response to pathogen pressure, as demonstrated by the fact that pathogen related (PR) genes were not expressed in the wrky27knockdown mutant (Mukhtar *et al.* 2008).

The NAN (At3G20640) transcription factor that belongs to the family of helix-loop-helix (bHLH) (Bailey *et al.* 2003) was isolated as putative downstream target of the Rap2.4a in the yeast based screen. The NAN protein is dually localized in the nucleus and in chloroplasts

(Kim *et al.* 2005). The NAN knockdown mutant has been characterized by having reduced hypocotyls and retarded seed germination, demonstrating that the protein is involved in germination of seeds and elongation of cells (Kim *et al.* 2005). The fact that Rap2.4a targets 2CPA and NAN, both of which are localized in chloroplasts, may show that it has an active function in regulating chloroplast activity.

Taken together, the target gene profile indicates that Rap2.4a is involved in mediating both biotic and abiotic stress responses and developmental processes. This conclusion is consistent with phenotype description of Rap2.4a knockdown lines (Shaikhali *et al.* 2008).

4.3.2 Motif analysis of Rap2.4a

The binding site of Rap2.4a was determined by aligning the sequences identified from strongly interacting clones. The MEME program identified two similar sequences, CCG(A/G)C and (A/G)C(C/T)GAC, as the binding sites of the Rap2.4a transcription factor (**Figure 20**). Taken the two binding motifs together, they are almost identical to the C-repeat (CRT)/Dehydration Responsive element (DRE), (A/G)CCGAC. DRE is a crucial *cis*-regulatory element that modulates the expression of RD29A in ABA mediated cold and drought stress responses (Yamaguchi-Shinozaki and Shinozaki 1994). Various cold and drought responsive genes harbor the DRE elements in their promoter regions indicating the importance of DRE in the abiotic stress response (Shinozaki and Yamaguchi-Shinozaki 2000, Thomashow 1999). Very similar to the DRE *cis*-acting element is CRT, (A/G)CCGA, that mediates cold stress response (Baker *et al.* 1994, Jiang *et al.* 1996a). C-repeat Binding Factor/DRE Binding protein 1 (CBF/DREB1) and DRE Binding protein 2 (DREB2) transcription factors, that belongs to the same superfamily as that of Rap2.4a which is characterized by single AP2/ERF DNA binding domain, interact with the CRT/DRE *cis*-acting element (Yamaguchi-Shinozaki and Shinozaki 2006). *A. thaliana* that overexpresses CBF1/DREB1 was shown to have heightened tolerance to freezing stress (Jaglo-Ottosen *et al.* 1998). The similarity of the Rap2.4a binding site with CRT/DRE suggests a possible involvement of Rap2.4a in cold stress regulation. The

downstream targets of CBF1/DREB1 include AP2/ERF-type and C2H2-type transcription factors, indicating that there is an elaborated regulatory network downstream of CBF1/DREB1 (Maruyama *et al.* 2004, Yamaguchi-Shinozaki and Shinozaki 2006).

4.4 Rap2.4d

4.4.1 Targets of Rap2.4d

Similar to Rap2.4a, the target genes of Rap2.4d were identified by analyzing the sequences of highly interacting clones in the stringent inverse yeast-one-hybrid assay. Using the MapMan functional annotation program, the identified putative targets of Rap2.4d were divided into 13 groups (**Figure 22**). Among the targets of Rap2.4d, a significant number of genes are involved in protein synthesis, DNA and RNA metabolism, signaling and transport functions.

Mitogen-activated protein kinase 3 (MPK3, At3G45640), one of the targets of Rap2.4d, is involved in biotic stress response. It is shown to mediate pattern-triggered immunity (PTI) responses that lead to programmed cell death and limit pathogen infection locally to protect uninfected areas from further infection (Igarashi *et al.* 2013). Programmed cell death is a highly orchestrated cascade of reactions that finally leads to the cell death, and proved to be crucial in the innate immunity response. Microbial associated molecular patterns (MAMPs) are sensed by plant cells via pattern recognition receptors (PRRs) to activate the PTI response (Boller and Felix 2009). MPK3 is demonstrated to play a pivotal role in mediating fungal pathogen response by phosphorylating Ethylene Response Factor (ERF6), which contains one AP2/ERF DNA binding domain. In addition MPK3 modulates stomatal closure as part of plant's generic defense system against pathogen challenge in ABA triggered signaling pathway (Melotto *et al.* 2006, Zeng *et al.* 2011). In addition to MPK3, two other genes that belong to Leucine-rich repeat protein kinase family (At1G64300 and At2G28990) are

identified targets of Rap2.4d with kinase activity, suggesting that Rap2.4d is positioned at higher place of many signaling cascades.

Sulfiredoxin (SRX, At1G31170) is another putative target of Rap2.4d that functions in the reduction of inactive cysteine residues of 2-Cys peroxiredoxins (2CPs) to restore the active form (Woo *et al.* 2003). SRX is dually targeted to chloroplasts and mitochondria. In chloroplasts, the activity of 2CPs are very important, as a consequence of high energy influx that leads to the generation of H₂O₂ (Baier and Dietz 1997). Upon extreme oxidative stress condition, the cysteine residues are overoxidized to form sulfinic acid (Cys-SO₂H) (Heiber *et al.* 2007), which inactivates 2CP (Yang *et al.* 2002). Although the inactivation of 2CP seems to be a disadvantage at a very high oxidative stress condition, H₂O₂ at such stress condition is required to activate other signaling cascades for initiating more comprehensive protection (Maruta *et al.* 2012). In addition to its activity as a redox dependent sulfinic acid reductase, SRX can also act as endonuclease in which either single or double stranded DNAs can be its substrates (Chi *et al.* 2012). To this end, another target gene of Rap2.4d, endoribonuclease L-PSP family protein (At3G20390), which is targeted to chloroplast and mitochondria, may also be involved in DNA and RNA metabolism (Elo *et al.* 2003). The fact that RAP2.4d putatively regulates proteins with the endonuclease activity, which are targeted to chloroplast and mitochondria, suggests that Rap2.4d may be involved in regulating chloroplast and mitochondrial metabolism of DNA and RNA.

In addition to SRX, three other proteins, copper amine oxidase family protein (At1G31710), oxidation-related zinc finger 1 (At2G19810) and Cinnamoyl COA:NADP oxidoreductase (At2G33590), that involve in the cellular redox reaction were found to be among the targets of Rap2.4d. This shows that there is a functional consistency among the targets of the Rap2.4d, which could suggest that Rap2.4d is involved in the redox regulation. It has already been shown that Rap2.4a, that is a close homologues of Rap2.4d, plays a redox-dependent regulatory function (Shaikhali *et al.* 2008).

Another three genes, Galacturonosyltransferase-like2 (At3G50760), Glucose 6-Phosphate/Phosphate Translocater1 (At5G54800) and Acyl-CoA thioesterase family protein (At1G01710), that interact with Rap2.4d are involved in carbohydrate transport activity. Galacturonosyltransferase-like2 (GATL2) mutant, for example, has a stunted growth phenotype upon exposure to low humidity condition (Lao *et al.* 2003). GATL2 is a glycosyltransferase that is presumed to be involved in the synthesis of components of cell wall (Lao *et al.* 2003). The involvement of Rap2.4d in stress responses could additionally be through the carbohydrate metabolism and cell wall synthesis, which affects the fitness of the plant.

4.4.2 Motif analyses of Rap2.4d

The DNA binding motif of Rap2.4d was determined by analyzing the sequences of strongly interacting fragments of the yeast based screen for consensus sequences (**Figure 23**). GHGGCG and AGGCS were identified as the binding site of Rap2.4d. It includes and extends the GGC element, which is well characterized as a target of many AP2-domain transcription factors (Fujimoto *et al.* 2000) and a proven binding site of Rap2.4b (Lin *et al.* 2008). The CG signature (HG at the beginning of the motif), was shown to be important for binding of Rap2.4a to the 2CPA promoter and adjusting expression of the nuclear encoded chloroplast peroxidases to chloroplast redox signals (Shaikhali *et al.* 2008).

4.5 Expressional analyses of ERF1b

A significant proportion of Rap2.4a and Rap2.4d target genes are involved in mediating abiotic stress response. All the ERF1b transcription factors have highly conserved or identical DNA binding domains. To find out if similar expression patterns exist among any of the ERF1b transcription factors and to address the question of functional redundancy or specificity, the expressions of ERF1b genes were analyzed under different abiotic stress conditions. The results indicated that Rap2.4c and Rap2.4d are closely related in terms of

the transcript abundance dynamics under different abiotic stress conditions (**Figure 25**), while the other family members did not show any correlation among each other. During cold stress conditions, the expression pattern of Rap2.4c and Rap2.4d is similar (**Figure 25a**). Both were upregulated during the first hour and gradually declined to reach the lowest level after 12 hours. Drought and heat stress also produced a similar pattern of gene expression as observed in cold stress for Rap2.4c and Rap2.4d (**Figure 25b, c**), but not for the other ERF1b transcription factor. Upon cold and drought stress Rap2.4d responds faster than Rap2.4c. There are two related groups of transcription factors, DREB1 and DREB2, that mediate the responses to drought and low cold, respectively (Liu *et al.* 1998). Taken together, the results suggest that there is an overlapping function between Rap2.4c and Rap2.4d, where the losses of one gene could be compensated by the other. It can be concluded that different abiotic stresses are perceived and modulate the expression of Rap2.4c and Rap2.4d in the same way. How Rap2.4c and Rap2.4d differentiate signals and activate target genes that correspond to a different stress remains to be identified.

4.6 Regulation of chloroplast antioxidant enzymes

4.6.1 2-Cys peroxiredoxin-A (2CPA) promoter analysis

The promoter of a prominent nuclear encoded chloroplast antioxidant enzyme called 2-Cys peroxiredoxin-A (2CPA) was analyzed for the predicted Rap2.4a and Rap2.4d binding sites, CCG(A/G)C and G(A/G)CCG, respectively. The result identified that the promoter of 2CPA has one version (ACCGA) of the Rap2.4a binding motif (**Figure 20**). It is located 170 bp upstream of the transcription initiation site, which in the proximal part of the promoter, makes a good candidate for functionally important binding site for the regulation 2CPA. Shaikhali *et al.* (2008) indicated that Rap2.4a binds to the redox sensitive part of 2CPA promoter (-706-315 up upstream of the transcription start site). Thus, it is presumed that, as 2CPA may have more than one Rap2.4a binding site in its promotor region. It has been already demonstrated that Rap2.4a modulates the expression of 2CPA (Shaikhali *et al.* 2008). To this

end, the expression of 2CPA is highly compromised in the Rap2.4a knockdown line (Radoslaw Rudnik, PhD thesis). In summary, the results indicated that there is a direct positive regulation of 2CPA by Rap2.4a via the direct binding of the ACCGA core element.

The binding site of Rap2.4a is very similar to that of the C-repeat (CRT)/Dehydration Responsive element (DRE) core element, (A/G)CCGAC. CRT/DRE element is known to mediate drought and cold stress responses (Sakuma *et al.* 2006, Yamaguchi-Shinozaki and Shinozaki 2006). It is already established that cold stress can lead to accumulation of reactive oxygen species (ROS) in chloroplasts. Since Rap2.4a regulates the expression of 2CPA that is located in the chloroplast in a redox dependent manner (Shaikhali *et al.* 2008), Rap2.4a senses redox signals emitted from chloroplast. Taken together, Rap2.4a responds to the accumulation of ROS produced as a result of cold stress in the chloroplast either by regulating the expression of 2CPA or by activating other cold responsive genes or both.

In silico promoter analysis revealed that the Rap2.4d transcription factor binding motif is not present in the promoter of 2CPA, indicating that Rap2.4d is not involved in the regulation of 2CPA expression, and Rap2.4a and Rap2.4d are not functionally redundant in this regard.

4.6.2 Interaction of ERF1b with 2CPA promoter

The DNA binding domains of the ERF1b transcription factors are almost identical. To answer the question of redundancy or specificity of the function of transcription factors, yeast-one-hybrid experiments were performed with all 3 ERF1b members. In line with the previous observation (Shaikhali *et al.* 2008), Rap2.4a was shown to bind the 2CPA promoter, but only weakly (**Figure 32**). In contrast, Rap2.4h, which is the previously uncharacterized member of the ERF1b, demonstrated to bind very strongly to the 2CPA promoter. In agreement to this, transient overexpression of Rap2.4h leads to low promoter activity of the 2CPA gene as compared to the wild type (in collaboration with Radoslaw Rudnik). In addition, we showed that Rap2.4h knockdown lines resulted in the up-regulation 2CPA transcript levels, indicating

that Rap2.4h negatively regulates the expression of 2CPA, as opposed to Rap2.4a. Collectively, the result suggests that Rap2.4a and Rap2.4h compete for the same binding site, and regulate the expression of 2CPA antagonistically. Rap2.4h is however, is a stronger competitor than Rap2.4a as demonstrated by vigorous binding of 2CPA promoter.

4.6.3 Chloroplast ascorbate peroxidases promoter analysis

The promoter analysis of chloroplast ascorbate peroxidases indicated that the second version (CCGAC) of the Rap2.4a binding site is present in the proximal promoter region (504 bp upstream of the transcription start site) in the thylakoid-bound ascorbate peroxidase (tAPx) (**Figure 26**). In contrast, none of the versions of Rap2.4a binding sites is present in the promoter of stromal ascorbate peroxidase (sAPx). In collaboration with Radoslaw Rudnik , it was shown that the expression of tAPx, but not sAPx is highly compromised in the Rap2.4a knockdown lines, which is in consistence with the promoter analysis. Taken together, the results indicated that Rap2.4a regulates the expression of another chloroplast localized antioxidant enzyme, tAPx, in addition to 2CPA. It has been indicated that silencing of tAPx leads to the suppression of CRT/DRE binding factors (CBF1) expression, which is a very important positive modulator of cold regulons (Maruta *et al.* 2012b). Similarly, a recent report (Balazadeh *et al.* 2012) indicated that the expression of RAP2.4a was positively regulated in a response to the H₂O₂ specific to the chloroplast. Similarly, the identified binding site of Rap2.4a on the tAPx promoter closely resembles the CRT/DRE element. It is therefore concluded that Rap2.4a connects the pathways of antioxidant enzyme regulation and cold stress responses.

2CPA and tAPx are known to be inactivated at severe oxidative stress conditions (Kitajima 2008, Yoshimura *et al.* 2000). As a consequence, the accumulation of H₂O₂ might be helpful for activating other signaling cascades to launch a general defense process (Hiner *et al.* 2000, Kitajima 2008, Yang *et al.* 2002). To this end, Rap2.4a may perceive redox signals from the chloroplast as a result of cold stress (and possibly other abiotic stresses) to either

activate the antioxidant enzymes (2CPA or tAPx) or cold responsive genes, in case of cold stress. Under relatively low redox imbalances Rap2.4a may activate expression of antioxidant enzymes such as 2CPA and tAPx, up to moderately high ROS signal. Rap 2.4a may furthermore activate the cold or other abiotic stress responsive genes upon more severe redox imbalances (**Figure 40**).

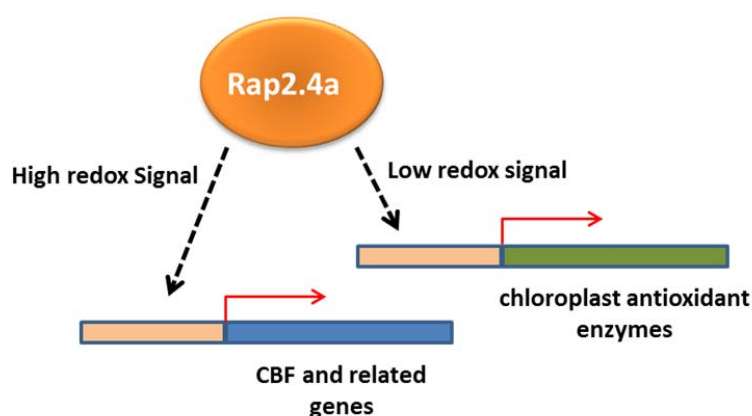


Figure 40: A model that depicts the regulation of chloroplast antioxidant enzymes and cold stress related genes by Rap2.4a as a result low and high levels of redox signal from the chloroplast.

Promoter analysis of tAPx revealed that the two versions of the Rap2.4d binding sites (G/AGGCG) are located close to the transcription start site (**Figure 26**), suggesting that Rap2.4d regulates the expression of tAPx. Similarly, the sAPx promoter was found to contain the binding site very close to the transcription start site. The transcript abundance analyses of sAPx and tAPx done in collaboration with Radoslaw Rudnik showed that Rap2.4a and Rap2.4d knockdown lines have reduced level of sAPx and tAPx expression and confirmed the regulatory effect predicted from promoter motif analysis and the direct yeast-one-hybrid.

4.6.4 Interaction of ERF1b with chloroplast ascorbate peroxidases promoter

Similarly, the interaction of ERF1b transcription factors with that of chloroplast ascorbate peroxidases was analyzed using yeast-one-hybrid-assay to test the possible direct regulation of the enzymes by ERF1b. The promoter region of tAPx was fragmented into tAPxB1 (distal part) and tAPxB2 (proximal part). The distal part of the tAPx promoter fragment showed unspecific interaction with all of the ERF1b transcription factors (**Figure 29**). In contrast, the proximal part of the promoter showed weak but specific interaction with ERF1b members. Rap2.4a, Rap2.4b, Rap2.4d and Rap2.4h showed weak but specific interaction. The promoter analysis of tAPx demonstrated that, both Rap2.4a and Rap2.4d transcription factor binding sites are present in the proximal part of the promoter, which is in conformity with the direct yeast-one-hybrid binding assay.

sAPx promoter was leaky in the yeast-one-hybrid assay in the sense that the reporter gene was activated by yeast intrinsic transcription factors to a significant extent. There seems to be a differential interaction pattern between the various ERF1b transcription factors and the sAPx promoter as compared to the negative control (**Figure 30**). Rap2.4a, Ra2.4d and Rap2.4e showed vigorous interaction while Rap2.4b, Rap2.4c and Rap2.4f showed suppressed interaction with sAPx promoter as compared to the control. This observation is consistent with the sAPx promoter analysis for Rap2.4d binding site, which is near to the transcription start site. However, this is not true for Rap2.4a whose binding site is not present on the promoter of sAPx, but shown to interact in the yeast-one-hybrid experiment.

In addition, the knockdown of Rap2.4a and Rap2.4d led to the reduced transcript abundance of tAPx (collaboration with Radoslaw Rudnik). In sum, the results indicate that the proximal part of the promoter has the functionally important Rap2.4a and Rap2.4d binding element that gives specificity to the ERF1b family in the regulation of tAPx. As in the case of 2CPA, it is proposed that, Rap2.4a and Rap2.4d promote the expression of tAPx upon moderate stress conditions.

4.6.5 Regulation of tAPx by Rap2.4c

In silico promoter analysis showed that the Rap2.4d binding site is present in the proximal part of the tAPx promoter (**Figure 26**). In addition yeast-one-hybrid assays demonstrated that there is a direct binding between the proximal part of the tAPx promoter and the Rap2.4d transcription factor. Since the expressional analysis of the ERF1b transcription factor family revealed that Rap2.4c and Rap2.4d are regulated in the same manner in different abiotic stress condition, it was hypothesized that Rap2.4c could also regulate the expression of tAPx. To test the hypothesis, that would answer the question of functional redundancy verses specificity the two transcription factors, the transcript level of tAPx was analyzed in the Rap2.4c knockdown line. Consistent with the function of Rap2.4d, Rap2.4c knockdown mutant has reduced levels of tAPx (**Figure 27, 31**). Overexpression of Rap2.4c restores the expression of tAPx but with a slightly reduced level of expression as compared to wild type, which is presumed to be a result of negative feedback loop regulation. It is widely established that plants can regulate the expression of a gene in the context of the whole system requirement by negative feedback loop (Leivar *et al.* 2012, Pokhilko *et al.* 2012). However, in contrary to the Rap2.4d knockdown line where the expression of sAPx is reduced, the transcript level of sAPx was not affected by either knockdown or overexpression of Rap2.4c. In general, it is concluded that Rap2.4c and Rap2.4d share functional similarity for most parts, but are different in the sAPx regulation.

4.7 Rap2.4c and Rap2.4d in cold regulation

4.7.1 Promoter analysis of COR6.6 and Kin 1

Since Rap2.4a has been linked to the regulation of cold stress response, its binding site was used to analyze the promoter regions of cold responsive genes. The results indicated that both variants of Rap2.4a binding site are present on the proximal promoter of Kin1 and Kin2 (Cold regulated 6.6, COR6.6) (**Figure 37**). Kin1 and COR6.6 are closely related homologues

genes that mediate cold stress response (Hughes and Dunn 1996). They are induced by cold and drought stress. Although the expression of Kin1 and COR6.6 is induced by ABA, ABA-deficient mutant (*abi*) treated with cold does express both genes indicating that the genes can directly respond to cold and drought without activating ABA signaling pathway (Nordin *et al.* 1991, Shinozaki and Yamaguchi-Shinozaki 1996). It has been demonstrated that overexpression of CBF3 leads to changes in the gene expression pattern that mimics the response to cold stress, in which cold regulated genes such as Kin1 and COR6.6 are up-regulated (Gilmour *et al.* 2000).

Stringent inverse yeast-one-hybrid screen for Rap2.4a interacting fragment identified at least two genes, protein phosphatase 2C (AT3G05640) and leucine-rich repeat protein kinase family protein (AT2G28970), that are involved in cold stress response and ABA signaling, as putative targets of Rap2.4a. All in all, the results suggest that Rap2.4a is involved in cold stress response. Promoter analyses of Kin1 and COR6.6 additionally unveiled that Rap2.4d binding motif (GAGGCG) is present in the proximal part of the promoter suggesting that Rap2.4d is also involved in the cold stress response (**Figure 26**).

4.7.2 Negative regulation of COR6.6 by Rap2.4d

The promoter analysis of cold responsive genes such as Kin1 and COR6.6 indicated that the binding site of the Rap2.4d transcription factor is present in the proximal part of their promoter. The expressional analysis of Rap2.4d and COR6.6 from the publicly available microarray database identified that the expression pattern of the two genes are opposite, hinting that Rap2.4d regulates the expression of COR6.6 negatively. Consequently, Rap2.4d would antagonize the regulation induced by cold.

COR6.6, a homologue of kin 1, was identified as cold regulated gene (Kurkela and Borg-Franck 1992, Wang *et al.* 1995). Cloning and characterization of the COR6.6 gene revealed that it is similar to type I fish antifreeze protein (Kurkela and Borg-Franck 1992, Kurkela and

Franck 1990). Both Kin1 and COR6.6 were found to be induced by cold, dehydration, osmotic pressure and ABA (Wang *et al.* 1995). The promoters of Kin1 and COR6.6 are active in young developing seeds, guard cells and pollen (Wang and Cutler 1995). Kin1 and COR6.6 genes code for 6.5 and 6.6 kD polypeptides both being located in the cytosol (Gilmour *et al.* 1996, Wang *et al.* 1994). A study to investigate the effects of COR6.6 protein on the destabilization of cell membrane, as a result of cellular dehydration caused by freezing, indicated that it is not involved in particular phospholipid-protein interaction that changes the phospholipid vesicles formation (Webb *et al.* 1996). ABA mediated stress response modulates the expression of genes such as COR6.6 via phospholipase C1 (Sanchez and Chua 2001). In the early ABA signaling pathway, Cyclic ADP-ribose (cADPR) activates a subset of cold responsive genes including COR6.6 (Sanchez *et al.* 2004). CBF3 gene is the upstream regulator that modulates the expression of cold responsive genes including COR6.6 and Kin1 (Yamaguchi-Shinozaki and Shinozaki 2006).

To further characterize the regulation of COR6.6 by Rap2.4d, the expression was compared in a Rap2.4d and Rap2.4c overexpression line with the wild type plant (Col-0) in their response to cold. In agreement with the expression analysis of the microarray data, the transcript abundance of Rap2.4d was reduced after cold treatment (**Figure 34**). In non-cold treated plants of the Rap2.4d T-DNA insertion overexpression line, the transcript level of COR6.6 was reduced showing that Rap2.4d and COR6.6 are inversely regulated. In the cold, the transcript level of Rap2.4d was declined in the Rap2.4d overexpression line. However in the same line, the expression of COR6.6 did not show a significant increase in its transcript abundance as expected for an inversely regulated gene. To explain the discrepancy, the expression of Rap2.4d close homologue, Rap2.4c, was analyzed (**Figure 34**). The result showed a slightly higher Rap2.4c transcript level in cold treated Rap2.4d overexpression plants. Functional redundancy was assumed to be present between Rap2.4c and Rap2.4d after analyzing their expression data (**Figure 25**). The fact that the expression of COR6.6 did not show a significant increase in cold treated plants of the Rap2.4d overexpression line is

explained by the fact that the Rap2.4c transcript level was also increased and compensated for reduced level of Rap2.4d in cold.

When considered collectively, the results indicate that Rap2.4d regulates the expression of COR6.6 negatively and that the Rap2.4c and Rap2.4d show functional redundancy in the cold stress response.

4.7.3 Negative regulation of COR6.6 and BAP1 by Rap2.4c

A small increase in the transcript abundance of Rap2.4c significantly restricted the increase in the expression of COR6.6 as expected from cold treated plant. It was, thus, presumed that Rap2.4c is a stronger negative regulator of COR6.6 than Rap2.4d. To test this hypothesis, the Rap2.4c T-DNA insertion overexpression line was analyzed for the expression of COR6.6 in both cold treated and non-cold treated plant. Consistent with the expressional analysis of microarray data (**Figure 37**), the transcript level of Rap2.4c, whose basal expression is high under normal conditions in the wild type and in the Rap2.4c overexpression line, was decreased in the cold. Due to the low level of Rap2.4c transcript abundance in the wild type control plants, the expression of COR6.6 is highly induced upon cold treatment. However, the induction of COR6.6 in the cold treated Rap2.4c overexpression line was not as high as it should be when compared to the cold treated wild type control. This demonstrates that higher amount of Rap2.4c in the cold treated overexpression line, as compared to the cold control, antagonizes the induction of COR6.6. In summary, the results indicate that Rap2.4c is a strong negative regulator of COR6.6.

In addition to the COR6.6, another cold responsive gene, BAP1, was analyzed to assess a possible regulation by Rap2.4c. BAP1 was shown to be involved in the regulation of programmed cell death demonstrated by the fact that its loss leads to the heightened hypersensitive response (Yang *et al.* 2007). The C2-like domain of BAP1 binds to phospholipids, whose interaction is calcium dependent, thereby affecting the defense

responses negatively (Yang *et al.* 2006). BAP1 expression is modulated by temperature and its expression affects cell expansion and cell division at low temperature (Hua *et al.* 2001).

Surprisingly, the expression of BAP1 followed a similar pattern as that of COR6.6 in the Rap2.4c overexpression line with and without cold treatment (**Figure 37**). The expression of BAP1 was up regulated in cold treated wild type plants as expected. However, the expression of BAP1 did not show any increase in the cold treated Rap2.4c overexpression line, indicating that Rap2.4c is also negatively controlling the expression BAP1. All results considered, it is hypothesized that Rap2.4c is a general negative regulator of cold induced genes.

4.7.4 Rap2.4c as a general negative regulator of cold responsive genes

Rap2.4c, a close homologue of Rap2.4d, negatively regulated the expression of different cold induced genes, COR6.6 and BAP1 (**Figure 38**), that belong to different gene families. Hence, it was assumed that Rap2.4c is a general negative regulator of cold responsive genes. To test the hypothesis whether Rap2.4c is a general negative regulator of cold induced genes, the expression of additional well-characterized cold responsive genes (COR15A, COR47, and PLD1) were analyzed in the Rap2.4c overexpression line.

COR15A encodes a boiling-stable hydrophilic 14.7 kD polypeptide with a N-terminal transit peptide for chloroplast targeting (Lin and Thomashow 1992b). The protein is modified posttranslationally to 9.4 kD during import into chloroplasts (Gilmour *et al.* 1996). Its ability to protect proteins against freezing damage was demonstrated by the fact that COR15A abolishes the inactivation of cold-labile enzymes against freezing *in vitro*. At freezing temperatures, it is present in the soluble fraction of chloroplast, stroma (Lin and Thomashow 1992a). It has been indicated that increased freezing tolerance of chloroplasts in plant leaves and isolated protoplasts could be achieved by overexpressing COR15A (Artus *et al.* 1996, Steponkus *et al.* 1998). The increase in the freezing tolerance is a result of lowering

the temperature of freeze induced lamellar- to-hexagonal II phase transitions, which happens due to cellular dehydration because of freezing temperature (Steponkus *et al.* 1998). It has no chaperon activity, i.e. it is not involved in the folding of denatured proteins. It protects plants by mitigating aggregation of proteins during cold (Nakayama *et al.* 2008) by forming oligomers around its substrates (Nakayama *et al.* 2007). COR15A has a close homologue called COR15b, both of which respond to cold and ABA, but not to drought (Wilhelm and Thomashow 1993). The expression of COR15A in response to low temperature is mediated through CRT/DRE element, and CBF1 and CBF4 are involved in the induction COR15 as a result of cool temperature (Wang and Hua 2009).

COR47 encodes a boiling-stable hydrophilic polypeptide, as in the case of other cold regulated genes. It has a size of approximately 47 kD (Lin *et al.* 1990) and its expression is modulated by ABA, cold, dehydration and salt stress (Nylander *et al.* 2001, Welin *et al.* 1995). Heat stress results in the induction of DREB2 (dehydration responsive element-binding factor 2) and related cold responsive genes such as COR47, which indicates that there is a cross talk between heat, cold and dehydration signaling pathways (Lim *et al.* 2006). The HOS1 (High Expression of Osmotically Responsive genes1) locus negatively controls the expression of cold responsive genes such as COR47 and COR15A (Ishitani *et al.* 1998). In addition, ATAF1, which is a member of plant specific transcription factor family, negatively modulates the expression of COR47 upon drought stress (Lu *et al.* 2007). CBF1 and CBF3 work in concert to positively regulate the expression of COR47 (Novillo *et al.* 2007).

In the Rap2.4c regulation of cold induced genes, Phospholipase D1 (PLDALPHA1, PLD1) was additionally analyzed. PLD1 is important in hydrolyzing glycerophospholipids to generate phosphatidic acid, that serves as secondary messenger in the transmembrane signaling, vesicular trafficking and many other signaling pathways (Ktistakis *et al.* 1996, Pappan *et al.* 1997a). The activity of PLD1 is regulated by the presence of phosphatidylinositol 4, 5-bisphosphate and nanomolar concentrations of calcium changes, which is perceived thought

the binding to its K/R rich motif and C2 domain, respectively (Pappan and Wang 1999, Pappan *et al.* 1997b). Down regulation of PLD1 slows down the process of senescence mediated by ethylene and ABA signaling pathway (Fan *et al.* 1997). Cell membrane distortion as a result of altered lipid composition results in the loss of membrane integrity, which facilitates the process of senescence. PLD1 was found to be present in the plasma membrane, mitochondria, chloroplast and other intracellular membranes (Fan *et al.* 1999). It was shown that cold treatment of plants leads to the accumulation of phosphatidic acid which is the product of the PLD1 activity as determined by radioactive labeling (Ruelland *et al.* 2002). The accumulation of phosphatidic acid during cold is presumed to be a signal to activate cold responsive genes (Ruelland *et al.* 2002).

The expressions of COR15A and COR47 that are modulated by different sets of CBF genes were up regulated 27 and 8 times, respectively, during cold treatment of the wild type plants, showing their importance in cold stress response (**Figure 38**). Similarly, the expression of COR6.6 and PLD1, whose upstream regulation is totally different, were induced by a factor of 12.5 and 22 times, respectively in cold treatment of wild type plants. In the Rap2.4c overexpression line, however, the induction of all the cold responsive genes was abolished. Taken together, the results indicate that Rap2.4c is a general negative regulator of cold induced genes.

There are numerous biochemical and molecular changes that take place in plant cells during cold acclimation. Gene expression regulation that ultimately results in cold stress responses is very important for survival of the plants (Chinnusamy *et al.* 2010). To this end, the upstream transcriptional activator of well-established regulator of cold induced gene (CBF), called Inducer of CBF expression-1 (ICE1) was identified (Chinnusamy *et al.* 2003). It has been shown that the expression of many cold regulated genes was compromised in the ice mutant (Lee *et al.* 2005). ICE1 encodes a bHLH transcription factors that interacts with the promoter of CBF3. CBF3, in turn, modulates the expression of cold responsive genes by binding to CBF/DREB1 elements. The present work demonstrates that Rap2.4c is a general

Discussion

cold regulated gene that negatively affects the expression of cold responsive genes, antagonistically to that of CBF3. The promoter analysis of the cold responsive genes (Kin1 and COR6.6) suggested that Rap2.4d, a close homologue of Rap2.4c, directly interacts with cold responsive genes. In conclusion, we reviewed the present model (Yamaguchi-Shinozaki and Shinozaki 2006) of cold regulation by adding Rap2.4c as a negative regulator of cold regulons. Based on the high similarity of Rap2.4a binding motif to that of CRT/DRE element which is a binding site of CBF3/DREB1A genes, it is hypothesized that Rap2.4a positively regulates the cold responsive genes.

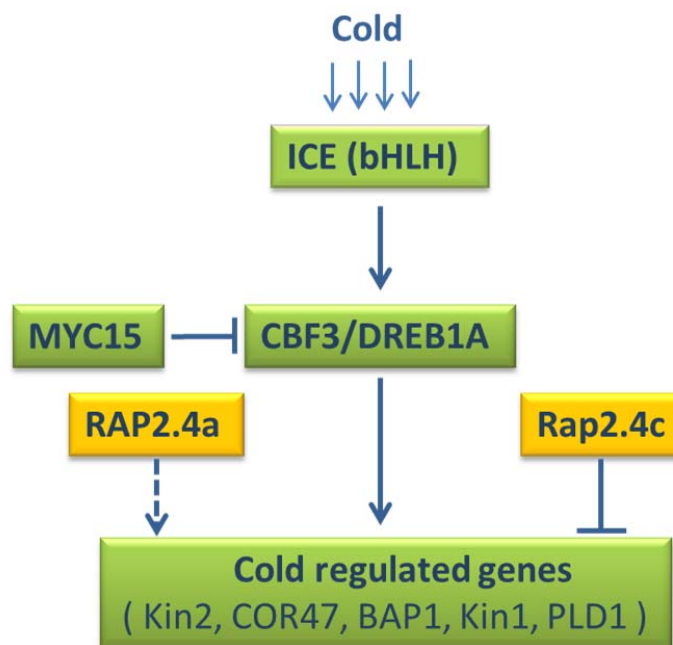


Figure 41: A model that shows the regulation of genes in a response to cold stress. Cold stress sensed by the ICE (inducer of CBF expression) gene is communicated downstream to modulate the cold responsive genes. While Rap2.4c negatively regulates cold responsive genes, Rap2.4a has a putative positive influence on their expression.

4.7.5 Direct interaction of Rap2.4c and COR6.6 promoter

Gene expression analysis of Rap2.4c and Rap2.4d showed that both genes respond to cold and other biotic stresses in a similar expression pattern (**Figure 25**). Analysis of the two genes in the T-DNA insertion line indicated that both genes negatively regulate the expression of COR6.6 in cold. Taken together, the results indicated that Rap2.4c and Rap2.4d genes are redundant in function. Since the binding site of Rap2.4c was not defined, to determine if the regulation of COR6.6 is by direct binding of Rap2.4c, transient reporter gene expression analysis was done (**Figure 39**). The result indicated that there is a direct interaction between the promoter of COR6.6 and Rap2.4c, as determined by the active reporter gene expression in the Rap2.4c overexpression line. In conclusion, it was demonstrated that Rap2.4c negatively regulates cold responsive genes by direct binding. This conclusion is in conformity with the proposed model of regulation of cold responsive genes.

To answer the if Rap2.4c and Rap2.4d exert their regulatory function by binding to the same binding motif, the binding site of Rap2.4d was mutated in the COR6.6 promoter. It was shown that Rap2.4c equally activates the reporter gene activity of the wild type and mutagenized promoter constructs of the *Agrobacterium* infiltrated seedlings. The result demonstrated that the two transcription factors do not compete for the same binding motif (**Figure 39**). It is, thus, assumed that their regulation of cold responsive genes could be additive and the overexpression or suppression of both genes could result in dramatic decrease or increase response of plant to cold stress.

5. References

- Agarwal, P.K., Agarwal, P., Reddy, M.K. and Sopory, S.K. (2006) Role of DREB transcription factors in abiotic and biotic stress tolerance in plants. *Plant Cell Rep*, **25**, 1263-1274.**
- Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M. and Suzuki, M. (1998) A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *Embo Journal*, **17**, 5484-5496.**
- An, Y.Q., McDowell, J.M., Huang, S.R., McKinney, E.C., Chambliss, S. and Meagher, R.B. (1996) Strong, constitutive expression of the Arabidopsis ACT2/ACT8 actin subclass in vegetative tissues. *Plant J*, **10**, 107-121.**
- Artus, N.N., Uemura, M., Steponkus, P.L., Gilmour, S.J., Lin, C.T. and Thomashow, M.F. (1996) Constitutive expression of the cold-regulated Arabidopsis thaliana COR15a gene affects both chloroplast and protoplast freezing tolerance. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 13404-13409.**
- Asada, K. (1999) The water-water cycle in chloroplasts: Scavenging of Active Oxygens and Dissipation of Excess Photons. *Annual review of plant physiology and plant molecular biology*, **50**, 601-639.**
- Bagga, R., Michalowski, S., Sabnis, R., Griffith, J.D. and Emerson, B.M. (2000) HMG I/Y regulates long-range enhancer-dependent transcription on DNA and chromatin by changes in DNA topology. *Nucleic acids research*, **28**, 2541-2550.**
- Baier, M. and Dietz, K.J. (1996) Primary structure and expression of plant homologues of animal and fungal thioredoxin-dependent peroxide reductases and bacterial alkyl hydroperoxide reductases. *Plant Mol Biol*, **31**, 553-564.**

References

- Baier, M. and Dietz, K.J.** (1997) The plant 2-Cys peroxiredoxin BAS1 is a nuclear-encoded chloroplast protein: its expressional regulation, phylogenetic origin, and implications for its specific physiological function in plants. *Plant J*, **12**, 179-190.
- Baier, M. and Dietz, K.J.** (2005) Chloroplasts as source and target of cellular redox regulation: a discussion on chloroplast redox signals in the context of plant physiology. *J Exp Bot*, **56**, 1449-1462.
- Baier, M., Pitsch, N.T., Mellenthin, M. and Guo, W.** (2010) Regulation of Genes Encoding Chloroplast Antioxidant Enzymes in Comparison to Regulation of the Extra-plastidic Antioxidant Defense System. *Ascorbate-Glutathione Pathway and Stress Tolerance in Plants*, 337-386.
- Baier, M., Stroher, E. and Dietz, K.J.** (2004) The acceptor availability at photosystem I and ABA control nuclear expression of 2-cys peroxiredoxin-alpha in Arabidopsis thaliana. *Plant and Cell Physiology*, **45**, 997-1006.
- Bailey, P.C., Martin, C., Toledo-Ortiz, G., Quail, P.H., Huq, E., Heim, M.A., Jakoby, M., Werber, M. and Weisshaar, B.** (2003) Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana. *Plant Cell*, **15**, 2497-2501.
- Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J.Y., Li, W.W. and Noble, W.S.** (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic acids research*, **37**, W202-W208.
- Bailey, T.L. and Elkan, C.** (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proceedings / International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, **2**, 28-36.
- Bailey, T.L., Williams, N., Misleh, C. and Li, W.W.** (2006) MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic acids research*, **34**, W369-W373.
- Baker, S.S., Wilhelm, K.S. and Thomashow, M.F.** (1994) The 5'-Region of Arabidopsis-Thaliana Cor15a Has Cis-Acting Elements That Confer Cold-Regulated, Drought-Regulated and Aba-Regulated Gene-Expression. *Plant Mol Biol*, **24**, 701-713.

References

- Balazadeh, S., Jaspert, N., Arif, M., Mueller-Roeber, B. and Maurino, V.G.** (2012) Expression of ROS-responsive genes and transcription factors after metabolic formation of H₂O₂ in chloroplasts. *Frontiers in plant science*, **3**, 234.
- Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K.** (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823-837.
- Berger, M.F., Philippakis, A.A., Qureshi, A.M., He, F.S., Estep, P.W., III and Bulyk, M.L.** (2006) Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities. *Nature Biotechnology*, **24**, 1429-1435.
- Boller, T. and Felix, G.** (2009) A renaissance of elicitors: perception of microbe-associated molecular patterns and danger signals by pattern-recognition receptors. *Annual review of plant biology*, **60**, 379-406.
- Bortz, P.D.S. and Wamhoff, B.R.** (2011) Chromatin Immunoprecipitation (ChIP): Revisiting the Efficacy of Sample Preparation, Sonication, Quantification of Sheared DNA, and Analysis via PCR. *PloS one*, **6**.
- Bradford, M.M.** (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*, **72**, 248-254.
- Brent, R. and Ptashne, M.** (1985) A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell*, **43**, 729-736.
- Bulyk, M.L.** (2004) Computational prediction of transcription-factor binding site locations. *Genome Biol*, **5**.
- Bulyk, M.L.** (2007) Protein binding microarrays for the characterization of DNA-protein interactions. *Advances in biochemical engineering/biotechnology*, **104**, 65-85.
- Bulyk, M.L., Johnson, P.L.F. and Church, G.M.** (2002) Nucleotides of transcription factor binding sites exert interdependent effects on the binding affinities of transcription factors. *Nucleic acids research*, **30**, 1255-1261.
- Calkhoven, C.F. and Ab, G.** (1996) Multiple steps in the regulation of transcription-factor level and activity. *Biochem J*, **317**, 329-342.

References

- Chai, C.L., Xie, Z.D. and Grotewold, E.** (2011) SELEX (Systematic Evolution of Ligands by EXponential Enrichment), as a Powerful Tool for Deciphering the Protein-DNA Interaction Space. *Plant Transcription Factors: Methods and Protocols*, **754**, 249-258.
- Chakravarty, A., Carlson, J.M., Khetani, R.S. and Gross, R.H.** (2007) A novel ensemble learning method for de novo computational identification of DNA binding sites. *Bmc Bioinformatics*, **8**.
- Chi, Y.H., Kim, S.Y., Jung, I.J., Shin, M.R., Jung, Y.J., Park, J.H., Lee, E.S., Maibam, P., Kim, K.S., Park, J.H., Kim, M.J., Hwang, G.Y. and Lee, S.Y.** (2012) Dual functions of Arabidopsis sulfiredoxin: Acting as a redox-dependent sulfinic acid reductase and as a redox-independent nuclease enzyme. *FEBS letters*, **586**, 3493-3499.
- Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B.H., Hong, X.H., Agarwal, M. and Zhu, J.K.** (2003) ICE1: a regulator of cold-induced transcriptome and freezing tolerance in Arabidopsis. *Gene Dev*, **17**, 1043-1054.
- Chinnusamy, V., Zhu, J.K. and Sunkar, R.** (2010) Gene Regulation During Cold Stress Acclimation in Plants. *Plant Stress Tolerance: Methods and Protocols*, **639**, 39-55.
- Dahm, R., Geisler, R. and Nüsslein-Volhard, C.** (2005) *Zebrafish (Danio rerio) Genome and Genetics*. 2 edn.: Wiley-VCH, Weinheim.
- Deplancke, B., Dupuy, D., Vidal, M. and Walhout, A.J.M.** (2004) A gateway-compatible yeast one-hybrid system. *Genome Research*, **14**, 2093-2101.
- Dingwall, C. and Laskey, R.A.** (1991) Nuclear Targeting Sequences - a Consensus. *Trends in biochemical sciences*, **16**, 478-481.
- Djordjevic, M.** (2007) SELEX experiments: New prospects, applications and data analysis in inferring regulatory pathways. *Biomol Eng*, **24**, 179-189.
- Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C., Doyle, F., Epstein, C.B., Fietze, S., Harrow, J., Kaul, R., Khatun, J., Lajoie, B.R., Landt, S.G., Lee, B.K., Pauli, F., Rosenbloom, K.R., Sabo, P., Safi, A., Sanyal, A., Shores, N., Simon, J.M., Song, L., Trinklein, N.D., Altshuler, R.C., Birney, E., Brown, J.B., Cheng, C., Djebali, S., Dong, X.J., Dunham, I., Ernst, J., Furey, T.S., Gerstein, M., Giardine, B., Greven, M., Hardison, R.C., Harris, R.S., Herrero, J., Hoffman, M.M., Iyer, S., Kellis, M., Khatun, J.,**

References

Kheradpour, P., Kundaje, A., Lassmann, T., Li, Q.H., Lin, X., Marinov, G.K., Merkel, A., Mortazavi, A., Parker, S.C.J., Reddy, T.E., Rozowsky, J., Schlesinger, F., Thurman, R.E., Wang, J., Ward, L.D., Whitfield, T.W., Wilder, S.P., Wu, W., Xi, H.L.S., Yip, K.Y., Zhuang, J.L., Bernstein, B.E., Birney, E., Dunham, I., Green, E.D., Gunter, C., Snyder, M., Pazin, M.J., Lowdon, R.F., Dillon, L.A.L., Adams, L.B., Kelly, C.J., Zhang, J., Wexler, J.R., Green, E.D., Good, P.J., Feingold, E.A., Bernstein, B.E., Birney, E., Crawford, G.E., Dekker, J., Elnitski, L., Farnham, P.J., Gerstein, M., Giddings, M.C., Gingeras, T.R., Green, E.D., Guigo, R., Hardison, R.C., Hubbard, T.J., Kellis, M., Kent, W.J., Lieb, J.D., Margulies, E.H., Myers, R.M., Snyder, M., Stamatoyannopoulos, J.A., Tenenbaum, S.A., Weng, Z.P., White, K.P., Wold, B., Khatun, J., Yu, Y., Wrobel, J., Risk, B.A., Gunawardena, H.P., Kuiper, H.C., Maier, C.W., Xie, L., Chen, X., Giddings, M.C., Bernstein, B.E., Epstein, C.B., Shores, N., Ernst, J., Kheradpour, P., Mikkelsen, T.S., Gillespie, S., Goren, A., Ram, O., Zhang, X.L., Wang, L., Issner, R., Coyne, M.J., Durham, T., Ku, M., Truong, T., Ward, L.D., Altshuler, R.C., Eaton, M.L., Kellis, M., Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C.H., Marinov, G.K., Khatun, J., Williams, B.A., Zaleski, C., Rozowsky, J., Roeder, M., Kokocinski, F., Abdelhamid, R.F., Alioto, T., Antoshechkin, I., Baer, M.T., Batut, P., Bell, I., Bell, K., Chakraborty, S., Chen, X., Chrast, J., Curado, J., Derrien, T., Drenkow, J., Dumais, E., Dumais, J., Duttagupta, R., Fastuca, M., Fejes-Toth, K., Ferreira, P., Foissac, S., Fullwood, M.J., Gao, H., Gonzalez, D., Gordon, A., Gunawardena, H.P., Howald, C., Jha, S., Johnson, R., Kapranov, P., King, B., Kingswood, C., Li, G.L., Luo, O.J., Park, E., Preall, J.B., Presaud, K., Ribeca, P., Risk, B.A., Robyr, D., Ruan, X.A., Sammeth, M., Sandhu, K.S., Schaeffer, L., See, L.H., Shahab, A., Skancke, J., Suzuki, A.M., Takahashi, H., Tilgner, H., Trout, D., Walters, N., Wang, H.E., Wrobel, J., Yu, Y.B., Hayashizaki, Y., Harrow, J., Gerstein, M., Hubbard, T.J., Reymond, A., Antonarakis, S.E., Hannon, G.J., Giddings, M.C., Ruan, Y.J., Wold, B., Carninci, P., Guigo, R., Gingeras, T.R., Rosenbloom, K.R., Sloan, C.A., Learned, K., Malladi, V.S., Wong, M.C., Barber, G., Cline, M.S., Dreszer, T.R., Heitner, S.G., Karolchik, D., Kent, W.J., Kirkup, V.M.,

References

Meyer, L.R., Long, J.C., Maddren, M., Raney, B.J., Furey, T.S., Song, L.Y., Grasfeder, L.L., Giresi, P.G., Lee, B.K., Battenhouse, A., Sheffield, N.C., Simon, J.M., Showers, K.A., Safi, A., London, D., Bhinge, A.A., Shestak, C., Schaner, M.R., Kim, S.K., Zhang, Z.Z.Z., Mieczkowski, P.A., Mieczkowska, J.O., Liu, Z., McDaniell, R.M., Ni, Y.Y., Rashid, N.U., Kim, M.J., Adar, S., Zhang, Z.C., Wang, T.Y., Winter, D., Keefe, D., Birney, E., Iyer, V.R., Lieb, J.D., Crawford, G.E., Li, G.L., Sandhu, K.S., Zheng, M.Z., Wang, P., Luo, O.J., Shahab, A., Fullwood, M.J., Ruan, X.A., Ruan, Y.J., Myers, R.M., Pauli, F., Williams, B.A., Gertz, J., Marinov, G.K., Reddy, T.E., Vielmetter, J., Partridge, E.C., Trout, D., Varley, K.E., Gasper, C., Bansal, A., Pepke, S., Jain, P., Amrhein, H., Bowling, K.M., Anaya, M., Cross, M.K., King, B., Muratet, M.A., Antoshechkin, I., Newberry, K.M., Mccue, K., Nesmith, A.S., Fisher-Aylor, K.I., Pusey, B., DeSalvo, G., Parker, S.L., Balasubramanian, S., Davis, N.S., Meadows, S.K., Eggleston, T., Gunter, C., Newberry, J.S., Levy, S.E., Absher, D.M., Mortazavi, A., Wong, W.H., Wold, B., Blow, M.J., Visel, A., Pennachio, L.A., Elnitski, L., Margulies, E.H., Parker, S.C.J., Petrykowska, H.M., Abyzov, A., Aken, B., Barrell, D., Barson, G., Berry, A., Bignell, A., Boychenko, V., Bussotti, G., Chrast, J., Davidson, C., Derrien, T., Despacio-Reyes, G., Diekhans, M., Ezkurdia, I., Frankish, A., Gilbert, J., Gonzalez, J.M., Griffiths, E., Harte, R., Hendrix, D.A., Howald, C., Hunt, T., Jungreis, I., Kay, M., Khurana, E., Kokocinski, F., Leng, J., Lin, M.F., Loveland, J., Lu, Z., Manthravadi, D., Mariotti, M., Mudge, J., Mukherjee, G., Notredame, C., Pei, B.K., Rodriguez, J.M., Saunders, G., Sboner, A., Searle, S., Sisu, C., Snow, C., Steward, C., Tanzer, A., Tapanari, E., Tress, M.L., van Baren, M.J., Walters, N., Washietl, S., Wilming, L., Zadissa, A., Zhang, Z.D., Brent, M., Haussler, D., Kellis, M., Valencia, A., Gerstein, M., Reymond, A., Guigo, R., Harrow, J., Hubbard, T.J., Landt, S.G., Fietze, S., Abyzov, A., Addleman, N., Alexander, R.P., Auerbach, R.K., Balasubramanian, S., Bettinger, K., Bhardwaj, N., Boyle, A.P., Cao, A.R., Cayting, P., Charos, A., Cheng, Y., Cheng, C., Eastman, C., Euskirchen, G., Fleming, J.D., Grubert, F., Habegger, L., Hariharan, M., Harmanci, A., Iyengar, S., Jin, V.X., Karczewski, K.J., Kasowski, M., Lacroute, P., Lam, H., Lamarre-Vincent, N., Leng, J., Lian, J., Lindahl-Allen, M., Min, R.Q., Miotto, B.,

References

Monahan, H., Moqtaderi, Z., Mu, X.M.J., O'Geen, H., Ouyang, Z.Q., Patacsil, D., Pei, B.K., Raha, D., Ramirez, L., Reed, B., Rozowsky, J., Sboner, A., Shi, M.Y., Sisu, C., Slifer, T., Witt, H., Wu, L.F., Xu, X.Q., Yan, K.K., Yang, X.Q., Yip, K.Y., Zhang, Z.D., Struhl, K., Weissman, S.M., Gerstein, M., Farnham, P.J., Snyder, M., Tenenbaum, S.A., Penalva, L.O., Doyle, F., Karmakar, S., Landt, S.G., Bhanvadia, R.R., Choudhury, A., Domanus, M., Ma, L.J., Moran, J., Patacsil, D., Slifer, T., Victorsen, A., Yang, X.Q., Snyder, M., White, K.P., Auer, T., Centanin, L., Eichenlaub, M., Gruhl, F., Heermann, S., Hoeckendorf, B., Inoue, D., Kellner, T., Kirchmaier, S., Mueller, C., Reinhardt, R., Schertel, L., Schneider, S., Sinn, R., Wittbrodt, B., Wittbrodt, J., Weng, Z.P., Whitfield, T.W., Wang, J., Collins, P.J., Aldred, S.F., Trinklein, N.D., Partridge, E.C., Myers, R.M., Dekker, J., Jain, G., Lajoie, B.R., Sanyal, A., Balasundaram, G., Bates, D.L., Byron, R., Canfield, T.K., Diegel, M.J., Dunn, D., Ebersol, A.K., Frum, T., Garg, K., Gist, E., Hansen, R.S., Boatman, L., Haugen, E., Humbert, R., Jain, G., Johnson, A.K., Johnson, E.M., Kutuyavin, T.V., Lajoie, B.R., Lee, K., Lotakis, D., Maurano, M.T., Neph, S.J., Neri, F.V., Nguyen, E.D., Qu, H.Z., Reynolds, A.P., Roach, V., Rynes, E., Sabo, P., Sanchez, M.E., Sandstrom, R.S., Sanyal, A., Shafer, A.O., Stergachis, A.B., Thomas, S., Thurman, R.E., Vernot, B., Vierstra, J., Vong, S., Wang, H., Weaver, M.A., Yan, Y.Q., Zhang, M.H., Akey, J.M., Bender, M., Dorschner, M.O., Groudine, M., MacCoss, M.J., Navas, P., Stamatoyannopoulos, G., Kaul, R., Dekker, J., Stamatoyannopoulos, J.A., Dunham, I., Beal, K., Brazma, A., Flicek, P., Herrero, J., Johnson, N., Keefe, D., Lusk, M., Luscombe, N.M., Sobral, D., Vaquerizas, J.M., Wilder, S.P., Batzoglou, S., Sidow, A., Hussami, N., Kyriazopoulou-Panagiotopoulou, S., Libbrecht, M.W., Schaub, M.A., Kundaje, A., Hardison, R.C., Miller, W., Giardine, B., Harris, R.S., Wu, W., Bickel, P.J., Banfai, B., Boley, N.P., Brown, J.B., Huang, H.Y., Li, Q.H., Li, J.J., Noble, W.S., Bilmes, J.A., Buske, O.J., Hoffman, M.M., Sahu, A.D., Kharchenko, P.V., Park, P.J., Baker, D., Taylor, J., Weng, Z.P., Iyer, S., Dong, X.J., Greven, M., Lin, X.Y., Wang, J., Xi, H.L.S., Zhuang, J.L., Gerstein, M., Alexander, R.P., Balasubramanian, S., Cheng, C., Harmanci, A., Lochovsky, L., Min, R., Mu, X.M.J., Rozowsky, J., Yan, K.K., Yip, K.Y.,

References

- Birney, E. and Consortium, E.P.** (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*, **489**, 57-74.
- Duret, L. and Bucher, P.** (1997) Searching for regulatory elements in human noncoding sequences. *Current opinion in structural biology*, **7**, 399-406.
- Dyall, S.D., Brown, M.T. and Johnson, P.J.** (2004) Ancient invasions: From endosymbionts to organelles. *Science*, **304**, 253-257.
- Ellington, A.D. and Szostak, J.W.** (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature*, **346**, 818-822.
- Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q.J., Gerentes, D., Perez, P. and Smyth, D.R.** (1996) AINTEGUMENTA, an APETALA2-like gene of arabidopsis with pleiotropic roles in ovule development and floral organ growth. *Plant Cell*, **8**, 155-168.
- Elo, A., Lyznik, A., Gonzalez, D.O., Kachman, S.D. and Mackenzie, S.A.** (2003) Nuclear genes that encode mitochondrial proteins for DNA and RNA metabolism are clustered in the Arabidopsis genome. *Plant Cell*, **15**, 1619-1631.
- Fan, L., Zheng, S., Cui, D. and Wang, X.** (1999) Subcellular distribution and tissue expression of phospholipase Dalpha, Dbeta, and Dgamma in Arabidopsis. *Plant physiology*, **119**, 1371-1378.
- Fan, L., Zheng, S. and Wang, X.** (1997) Antisense suppression of phospholipase D alpha retards abscisic acid- and ethylene-promoted senescence of postharvest Arabidopsis leaves. *Plant Cell*, **9**, 2183-2196.
- Featherstone, M.** (2002) Coactivators in transcription initiation: here are your orders. *Curr Opin Genet Dev*, **12**, 149-155.
- Fessele, S., Maier, H., Zischek, C., Nelson, P.J. and Werner, T.** (2002) Regulatory context is a crucial part of gene function. *Trends in Genetics*, **18**, 60-63.
- Fey, V., Wagner, R., Brautigam, K. and Pfanschmidt, T.** (2005) Photosynthetic redox control of nuclear gene expression. *J Exp Bot*, **56**, 1491-1498.
- Fields, S. and Sternglanz, R.** (1994) The two-hybrid system: an assay for protein-protein interactions. *Trends in genetics : TIG*, **10**, 286-292.

References

- Fox, K.R., Yan, Y.F. and Gong, B.** (1999) DNA sequence recognition by a novel series of minor groove-binding ligands. *Anti-Cancer Drug Des*, **14**, 219-230.
- Fried, M.G.** (1989) Measurement of protein-DNA interaction parameters by electrophoresis mobility shift assay. *Electrophoresis*, **10**, 366-376.
- Fried, M.G. and Bromberg, J.L.** (1997) Factors that affect the stability of protein-DNA complexes during gel electrophoresis. *Electrophoresis*, **18**, 6-11.
- Frith, M.C., Fu, Y.T., Yu, L.Q., Chen, J.F., Hansen, U. and Weng, Z.P.** (2004) Detection of functional DNA motifs via statistical over-representation. *Nucleic acids research*, **32**, 1372-1381.
- Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H. and Ohme-Takagi, M.** (2000) Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *The Plant cell*, **12**, 393-404.
- Garner, M.M. and Revzin, A.** (1981) A Gel-Electrophoresis Method for Quantifying the Binding of Proteins to Specific DNA Regions - Application to Components of the Escherichia-Coli Lactose Operon Regulatory System. *Nucleic Acids Res*, **9**, 3047-3060.
- Garner, M.M. and Revzin, A.** (1986) The Use of Gel-Electrophoresis to Detect and Study Nucleic-Acid Protein Interactions. *Trends in biochemical sciences*, **11**, 395-396.
- Gietz, R.D. and Schiestl, R.H.** (2007a) High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, **2**, 31-34.
- Gietz, R.D. and Schiestl, R.H.** (2007b) Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, **2**, 38-41.
- Gilmour, S.J., Lin, C.T. and Thomashow, M.F.** (1996) Purification and properties of Arabidopsis thaliana COR (cold-regulated) gene polypeptides COR15am and COR6.6 expressed in Escherichia coli. *Plant physiology*, **111**, 293-299.
- Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D. and Thomashow, M.F.** (2000) Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation. *Plant physiology*, **124**, 1854-1865.

References

- Gorsuch, P.A., Sargeant, A.W., Penfield, S.D., Quick, W.P. and Atkin, O.K.** (2010) Systemic low temperature signaling in Arabidopsis. *Plant & cell physiology*, **51**, 1488-1498.
- Gregory, T.R.** (2001) Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. *Biological reviews of the Cambridge Philosophical Society*, **76**, 65-101.
- Gregory, T.R.** (2005) The C-value enigma in plants and animals: a review of parallels and an appeal for partnership. *Annals of botany*, **95**, 133-146.
- Guo, A.Y., He, K., Liu, D., Bai, S.N., Gu, X.C., Wei, L.P. and Luo, J.C.** (2005) DATF: a database of Arabidopsis transcription factors. *Bioinformatics*, **21**, 2568-2569.
- Gutterson, N. and Reuber, T.L.** (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. *Curr Opin Plant Biol*, **7**, 465-471.
- Hampshire, A.J., Rusling, D.A., Broughton-Head, V.J. and Fox, K.R.** (2007) Footprinting: A method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods*, **42**, 128-140.
- Hannenhalli, S.** (2008) Eukaryotic transcription factor binding sites - modeling and integrative search methods. *Bioinformatics*, **24**, 1325-1331.
- Hawkins, J., Grant, C., Noble, W.S. and Bailey, T.L.** (2009) Assessing phylogenetic motif models for predicting transcription factor binding sites. *Bioinformatics*, **25**, 1339-1347.
- Heiber, I., Stroher, E., Raatz, B., Busse, I., Kahmann, U., Bevan, M.W., Dietz, K.J. and Baier, M.** (2007) The redox imbalanced mutants of arabidopsis differentiate signaling pathways for redox regulation of chloroplast antioxidant enzymes. *Plant physiology*, **143**, 1774-1788.
- Hellman, L.M. and Fried, M.G.** (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc*, **2**, 1849-1861.
- Higo, K., Ugawa, Y., Iwamoto, M. and Higo, H.** (1998) PLACE: a database of plant cis-acting regulatory DNA elements. *Nucleic acids research*, **26**, 358-359.
- Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T.** (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic acids research*, **27**, 297-300.

References

- Hiner, A.N.P., Rodriguez-Lopez, J.N., Arnao, M.B., Raven, E.L., Garcia-Canovas, F. and Acosta, M.** (2000) Kinetic study of the inactivation of ascorbate peroxidase by hydrogen peroxide. *Biochem J*, **348**, 321-328.
- Ho, J.W.K., Bishop, E., Karchenko, P.V., Negre, N., White, K.P. and Park, P.J.** (2011) ChIP-chip versus ChIP-seq: Lessons for experimental design and data analysis. *Bmc Genomics*, **12**.
- Hope, I.A. and Struhl, K.** (1986) Functional Dissection of a Eukaryotic Transcriptional Activator Protein, Gcn4 of Yeast. *Cell*, **46**, 885-894.
- Horling, F., Konig, J. and Dietz, K.J.** (2002) Type II peroxiredoxin C, a member of the peroxiredoxin family of Arabidopsis thaliana: its expression and activity in comparison with other peroxiredoxins. *Plant Physiol Bioch*, **40**, 491-499.
- Hua, J., Grisafi, P., Cheng, S.H. and Fink, G.R.** (2001) Plant growth homeostasis is controlled by the Arabidopsis BON1 and BAP1 genes. *Gene Dev*, **15**, 2263-2272.
- Huang, H., Tudor, M., Su, T., Zhang, Y., Hu, Y. and Ma, H.** (1996) DNA binding properties of two arabidopsis MADS domain proteins: Binding consensus and dimer formation. *Plant Cell*, **8**, 81-94.
- Hudson, M.E. and Quail, P.H.** (2003) Identification of promoter motifs involved in the network of phytochrome A-regulated gene expression by combined analysis of genomic sequence and microarray data. *Plant physiology*, **133**, 1605-1616.
- Hughes, M.A. and Dunn, M.A.** (1996) The molecular biology of plant acclimation to low temperature. *J Exp Bot*, **47**, 291-305.
- Igarashi, D., Bethke, G., Xu, Y., Tsuda, K., Glazebrook, J. and Katagiri, F.** (2013) Pattern-Triggered Immunity Suppresses Programmed Cell Death Triggered by Fumonisin B1. *PLoS one*, **8**.
- Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, K., Toyoda, T., Konagaya, A. and Shinozaki, K.** (2005) RARTE: Database and tools for complete sets of Arabidopsis transcription factors. *DNA Res*, **12**, 247-256.
- Ioshikhes, I., Trifonov, E.N. and Zhang, M.Q.** (1999) Periodical distribution of transcription factor sites in promoter regions and connection with chromatin structure.

References

- Proceedings of the National Academy of Sciences of the United States of America*, **96**, 2891-2895.
- Ishikawa, T. and Shigeoka, S.** (2008) Recent advances in ascorbate biosynthesis and the physiological significance of ascorbate peroxidase in photosynthesizing organisms. *Bioscience, biotechnology, and biochemistry*, **72**, 1143-1154.
- Ishitani, M., Xiong, L., Lee, H., Stevenson, B. and Zhu, J.K.** (1998) HOS1, a genetic locus involved in cold-responsive gene expression in arabidopsis. *Plant Cell*, **10**, 1151-1161.
- Iwase, A., Mitsuda, N., Koyama, T., Hiratsu, K., Kojima, M., Arai, T., Inoue, Y., Seki, M., Sakakibara, H., Sugimoto, K. and Ohme-Takagi, M.** (2011) The AP2/ERF Transcription Factor WIND1 Controls Cell Dedifferentiation in Arabidopsis. *Current Biology*, **21**, 508-514.
- Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O. and Thomashow, M.F.** (1998) Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance. *Science*, **280**, 104-106.
- Jiang, C., lu, B. and Singh, J.** (1996a) Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter Brassica napus. *Plant Mol Biol*, **30**, 679-684.
- Jiang, Y., Smale, S.T. and Gralla, J.D.** (1993) A Common Atp Requirement for Open Complex-Formation and Transcription at Promoters Containing Initiator or Tata Elements. *Journal of Biological Chemistry*, **268**, 6535-6540.
- Jiang, Y., Yan, M. and Gralla, J.D.** (1996b) A three-step pathway of transcription initiation leading to promoter clearance at an activated RNA polymerase II promoter. *Molecular and cellular biology*, **16**, 1614-1621.
- Jofuku, K.D., Denboer, B.G.W., Vanmontagu, M. and Okamoto, J.K.** (1994) Control of Arabidopsis Flower and Seed Development by the Homeotic Gene *Apetala2*. *Plant Cell*, **6**, 1211-1225.
- Johnston, S.A., Zavortink, M.J., Debouck, C. and Hopper, J.E.** (1986) Functional Domains of the Yeast Regulatory Protein Gal4. *Proceedings of the National Academy of Sciences of the United States of America*, **83**, 6553-6557.

References

- Jost, J.P., Munch, O. and Andersson, T.** (1991) Study of protein-DNA interactions by surface plasmon resonance (real time kinetics). *Nucleic Acids Res*, **19**, 2788.
- Kadonaga, J.T.** (2004) Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell*, **116**, 247-257.
- Kangasjarvi, S., Lepisto, A., Hannikainen, K., Piippo, M., Luomala, E.M., Aro, E.M. and Rintamaki, E.** (2008) Diverse roles for chloroplast stromal and thylakoid-bound ascorbate peroxidases in plant stress responses. *Biochem J*, **412**, 275-285.
- Kaufmann, K., Muino, J.M., Osteras, M., Farinelli, L., Krajewski, P. and Angenent, G.C.** (2010) Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nature Protocols*, **5**, 457-472.
- Keegan, L., Gill, G. and Ptashne, M.** (1986) Separation of DNA-Binding from the Transcription-Activating Function of a Eukaryotic Regulatory Protein. *Science*, **231**, 699-704.
- Kim, J.A., Yun, J., Lee, M., Kim, Y.S., Woo, J.C. and Park, C.M.** (2005) A basic helix-loop-helix transcription factor regulates cell elongation and seed germination. *Mol Cells*, **19**, 334-341.
- Kim, S., Shi, H., Lee, D. and Lis, J.T.** (2003) Specific SR protein-dependent splicing substrates identified through genomic SELEX. *Nucleic acids research*, **31**, 1955-1961.
- Kim, S., Soltis, P.S., Wall, K. and Soltis, D.E.** (2006) Phylogeny and domain evolution in the APETALA2-like gene family. *Mol Biol Evol*, **23**, 107-120.
- Kim, S.Y.** (2006) The role of ABF family bZIP class transcription factors in stress response. *Physiol Plantarum*, **126**, 519-527.
- Kitajima, S.** (2008) Hydrogen Peroxide-mediated Inactivation of Two Chloroplastic Peroxidases, Ascorbate Peroxidase and 2-Cys Peroxiredoxin. *Photochem Photobiol*, **84**, 1404-1409.
- Konig, J., Lotte, K., Plessow, R., Brockhinke, A., Baier, M. and Dietz, K.J.** (2003) Reaction mechanism of plant 2-Cys peroxiredoxin - Role of the C terminus and the quaternary structure. *Journal of Biological Chemistry*, **278**, 24409-24420.

References

- Koniger, M., Delamaide, J.A., Marlow, E.D. and Harris, G.C.** (2008) Arabidopsis thaliana leaves with altered chloroplast numbers and chloroplast movement exhibit impaired adjustments to both low and high light. *J Exp Bot*, **59**, 2285-2297.
- Koop, B.F.** (1995) Human and Rodent DNA-Sequence Comparisons - a Mosaic Model of Genomic Evolution. *Trends in Genetics*, **11**, 367-371.
- Kornberg, R.D. and Lorch, Y.** (2002) Chromatin and transcription: where do we go from here? Commentary. *Curr Opin Genet Dev*, **12**, 249-251.
- Ktistakis, N.T., Brown, H.A., Waters, M.G., Sternweis, P.C. and Roth, M.G.** (1996) Evidence that phospholipase D mediates ADP ribosylation factor-dependent formation golgi coated vesicles. *J Cell Biol*, **134**, 295-306.
- Kurkela, S. and Borg-Franck, M.** (1992) Structure and expression of kin2, one of two cold- and ABA-induced genes of Arabidopsis thaliana. *Plant Mol Biol*, **19**, 689-692.
- Kurkela, S. and Franck, M.** (1990) Cloning and characterization of a cold- and ABA-inducible Arabidopsis gene. *Plant Mol Biol*, **15**, 137-144.
- Lao, N.T., Long, D., Kiang, S., Coupland, G., Shoue, D.A., Carpita, N.C. and Kavanagh, T.A.** (2003) Mutation of a family 8 glycosyltransferase gene alters cell wall carbohydrate composition and causes a humidity-sensitive semi-sterile dwarf phenotype in Arabidopsis. *Plant Mol Biol*, **53**, 687-701.
- Laskey, R.A. and Dingwall, C.** (1993) Nuclear Shuttling - the Default Pathway for Nuclear Proteins. *Cell*, **74**, 585-586.
- Lee, B.H., Henderson, D.A. and Zhu, J.K.** (2005) The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. *Plant Cell*, **17**, 3155-3175.
- Lee, S.J., Park, J.H., Lee, M.H., Yu, J.H. and Kim, S.Y.** (2010) Isolation and functional characterization of CE1 binding proteins. *Bmc Plant Biol*, **10**.
- Leivar, P., Monte, E., Cohn, M.M. and Quail, P.H.** (2012) Phytochrome Signaling in Green Arabidopsis Seedlings: Impact Assessment of a Mutually Negative phyB-PIF Feedback Loop. *Mol Plant*, **5**, 734-749.
- Lemon, B. and Tjian, R.** (2000) Orchestrated response: a symphony of transcription factors for gene control. *Gene Dev*, **14**, 2551-2569.

References

- Lescot, M., Dehais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouze, P. and Rombauts, S.** (2002) PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic acids research*, **30**, 325-327.
- Levine, M. and Tjian, R.** (2003) Transcription regulation and animal diversity. *Nature*, **424**, 147-151.
- Li, J.J. and Herskowitz, I.** (1993) Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. *Science*, **262**, 1870-1874.
- Li, Q., Li, N., Hu, X., Li, J., Du, Z., Chen, L., Yin, G., Duan, J., Zhang, H., Zhao, Y. and Wang, J.** (2011) Genome-wide mapping of DNA methylation in chicken. *Plos One*, **6**, e19428.
- Li, Y., Rosso, M.G., Ulker, B. and Weisshaar, B.** (2006) Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions. *Genomics*, **87**, 645-652.
- Lim, C.J., Yang, K.A., Hong, J.K., Choi, J.S., Yun, D.J., Hong, J.C., Chung, W.S., Lee, S.Y., Cho, M.J. and Lim, C.O.** (2006) Gene expression profiles during heat acclimation in *Arabidopsis thaliana* suspension-culture cells. *Journal of plant research*, **119**, 373-383.
- Lin, C.T. and Thomashow, M.F.** (1992a) A Cold-Regulated *Arabidopsis* Gene Encodes a Polypeptide Having Potent Cryoprotective Activity. *Biochem Bioph Res Co*, **183**, 1103-1108.
- Lin, C.T. and Thomashow, M.F.** (1992b) DNA-Sequence Analysis of a Complementary-DNA for Cold-Regulated *Arabidopsis* Gene Cor15 and Characterization of the Cor-15 Polypeptide. *Plant physiology*, **99**, 519-525.
- Lin, C.T., Wei, W.G., Everson, E. and Thomashow, M.F.** (1990) Cold-Acclimation in *Arabidopsis* and Wheat - a Response Associated with Expression of Related Genes Encoding Boiling-Stable Polypeptides. *Plant physiology*, **94**, 1078-1083.
- Lin, R.C., Park, H.J. and Wang, H.Y.** (2008) Role of *Arabidopsis* RAP2.4 in regulating light- and ethylene-mediated developmental processes and drought stress tolerance. *Molecular plant*, **1**, 42-57.

References

- Liu, J. and Stormo, G.D.** (2005) Combining SELEX with quantitative assays to rapidly obtain accurate models of protein-DNA interactions. *Nucleic acids research*, **33**, e141.
- Liu, L.S., White, M.J. and MacRae, T.H.** (1999) Transcription factors and their genes in higher plants - Functional domains, evolution and regulation. *Eur J Biochem*, **262**, 247-257.
- Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K.** (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. *Plant Cell*, **10**, 1391-1406.
- Lockhart, D.J. and Winzeler, E.A.** (2000) Genomics, gene expression and DNA arrays. *Nature*, **405**, 827-836.
- Loots, G.G., Locksley, R.M., Blakespoor, C.M., Wang, Z.E., Miller, W., Rubin, E.M. and Frazer, K.A.** (2000) Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science*, **288**, 136-140.
- Lorenz, C., von Pelchrzim, F. and Schroeder, R.** (2006) Genomic systematic evolution of ligands by exponential enrichment (Genomic SELEX) for the identification of protein-binding RNAs independent of their expression levels. *Nature Protocols*, **1**, 2204-2212.
- Lorenzo, O., Piqueras, R., Sanchez-Serrano, J.J. and Solano, R.** (2003) ETHYLENE RESPONSE FACTOR1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell*, **15**, 165-178.
- Lu, P.L., Chen, N.Z., An, R., Su, Z., Qi, B.S., Ren, F., Chen, J. and Wang, X.C.** (2007) A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in Arabidopsis. *Plant Mol Biol*, **63**, 289-305.
- Lundin, S., Stranneheim, H., Pettersson, E., Klevebring, D. and Lundeberg, J.** (2010) Increased throughput by parallelization of library preparation for massive sequencing. *PloS one*, **5**, e10029.
- Magnani, E., Sjolander, K. and Hake, S.** (2004) From endonucleases to transcription factors: Evolution of the AP2 DNA binding domain in plants. *Plant Cell*, **16**, 2265-2277.

References

- Majka, J. and Speck, C.** (2007) Analysis of protein-DNA interactions using surface plasmon resonance. *Adv Biochem Eng Biot*, **104**, 13-36.
- Man, T.K. and Stormo, G.D.** (2001) Non-independence of Mnt repressor-operator interaction determined by a new quantitative multiple fluorescence relative affinity (QuMFRA) assay. *Nucleic acids research*, **29**, 2471-2478.
- Maniatis, T. and Reed, R.** (2002) An extensive network of coupling among gene expression machines. *Nature*, **416**, 499-506.
- Maruta, T., Noshi, M., Tanouchi, A., Tamoi, M., Yabuta, Y., Yoshimura, K., Ishikawa, T. and Shigeoka, S.** (2012) H₂O₂-triggered retrograde signaling from chloroplasts to nucleus plays specific role in response to stress. *The Journal of biological chemistry*, **287**, 11717-11729.
- Maruta, T., Tanouchi, A., Tamoi, M., Yabuta, Y., Yoshimura, K., Ishikawa, T. and Shigeoka, S.** (2010) Arabidopsis chloroplastic ascorbate peroxidase isoenzymes play a dual role in photoprotection and gene regulation under photooxidative stress. *Plant & cell physiology*, **51**, 190-200.
- Maruyama, K., Sakuma, Y., Kasuga, M., Ito, Y., Seki, M., Goda, H., Shimada, Y., Yoshida, S., Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2004) Identification of cold-inducible downstream genes of the Arabidopsis DREB1A/CBF3 transcriptional factor using two microarray systems. *Plant J*, **38**, 982-993.
- McKenna, N.J. and O'Malley, B.W.** (2002) Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell*, **108**, 465-474.
- Mehler, A.H.** (1951) Studies on Reactions of Illuminated Chloroplasts .2. Stimulation and Inhibition of the Reaction with Molecular Oxygen. *Arch Biochem Biophys*, **34**, 339-351.
- Melotto, M., Underwood, W., Koczan, J., Nomura, K. and He, S.Y.** (2006) Plant stomata function in innate immunity against bacterial invasion. *Cell*, **126**, 969-980.
- Meyer, M. and Kircher, M.** (2010) Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harbor protocols*, **2010**, pdb prot5448.

References

- Meyer, P.** (2000) Transcriptional transgene silencing and chromatin components. *Plant Mol Biol*, **43**, 221-234.
- Miao, Y. and Zentgraf, U.** (2007) The antagonist function of Arabidopsis WRKY53 and ESR/ESP in leaf senescence is modulated by the jasmonic and salicylic acid equilibrium. *Plant Cell*, **19**, 819-830.
- Milla, L.A., Cortes, C.R., Christian, Q., Onate, M.G., Cambiazo, V., Burgess, S.M. and Palma, V.** (2012) Yeast-based assay identifies novel Shh/Gli target genes in vertebrate development. *BMC genomics*, **13**.
- Mindell, D.P. and Meyer, A.** (2001) Homology evolving. *Trends Ecol Evol*, **16**, 434-440.
- Mitsuda, N. and Ohme-Takagi, M.** (2009) Functional Analysis of Transcription Factors in Arabidopsis. *Plant and Cell Physiology*, **50**, 1232-1248.
- Mittler, R.** (2002) Oxidative stress, antioxidants and stress tolerance. *Trends in plant science*, **7**, 405-410.
- Mokry, M., Hatzis, P., de Bruijn, E., Koster, J., Versteeg, R., Schuijers, J., van de Wetering, M., Guryev, V., Clevers, H. and Cuppen, E.** (2010) Efficient Double Fragmentation ChIP-seq Provides Nucleotide Resolution Protein-DNA Binding Profiles. *Plos One*, **5**.
- Montemartini, M., Kalisz, H.M., Hecht, H.J., Steinert, P. and Flohe, L.** (1999) Activation of active-site cysteine residues in the peroxiredoxin-type trypanoxin peroxidase of *Crithidia fasciculata*. *Eur J Biochem*, **264**, 516-524.
- Mukherjee, S., Berger, M.F., Jona, G., Wang, X.S., Muzzey, D., Snyder, M., Young, R.A. and Bulyk, M.L.** (2004a) Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nature Genetics*, **36**, 1331-1339.
- Mukherjee, S., Berger, M.F., Jona, G., Wang, X.S., Muzzey, D., Snyder, M., Young, R.A. and Bulyk, M.L.** (2004b) Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. *Nat Genet*, **36**, 1331-1339.
- Mukhtar, M.S., Deslandes, L., Auriac, M.C., Marco, Y. and Somssich, I.E.** (2008) The Arabidopsis transcription factor WRKY27 influences wilt disease symptom development caused by *Ralstonia solanacearum*. *Plant J*, **56**, 935-947.

References

- Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H.** (2006) Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol*, **140**, 411-432.
- Nakayama, K., Okawa, K., Kakizaki, T., Honma, T., Itoh, H. and Inaba, T.** (2007) Arabidopsis Cor15am is a chloroplast stromal protein that has cryoprotective activity and forms oligomers. *Plant physiology*, **144**, 513-523.
- Nakayama, K., Okawa, K., Kakizaki, T. and Inaba, T.** (2008) Evaluation of the protective activities of a late embryogenesis abundant (LEA) related protein, Cor15am, during various stresses in vitro. *Biosci Biotech Bioch*, **72**, 1642-1645.
- Narlikar, G.J., Fan, H.Y. and Kingston, R.E.** (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, **108**, 475-487.
- Narusaka, Y., Nakashima, K., Shinwari, Z.K., Sakuma, Y., Furihata, T., Abe, H., Narusaka, M., Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2003) Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. *Plant J*, **34**, 137-148.
- Nikolov, D.B. and Burley, S.K.** (1997) RNA polymerase II transcription initiation: A structural view. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 15-22.
- Nordin, K., Heino, P. and Palva, E.T.** (1991) Separate signal pathways regulate the expression of a low-temperature-induced gene in Arabidopsis thaliana (L.) Heynh. *Plant Mol Biol*, **16**, 1061-1071.
- Novillo, F., Medina, J. and Salinas, J.** (2007) Arabidopsis CBF1 and CBF3 have a different function than CBF2 in cold acclimation and define different gene classes in the CBF regulon. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 21002-21007.
- Nylander, M., Svensson, J., Palva, E.T. and Welin, B.V.** (2001) Stress-induced accumulation and tissue-specific localization of dehydrins in Arabidopsis thaliana. *Plant Mol Biol*, **45**, 263-279.
- Ohmetakagi, M. and Shinshi, H.** (1995) Ethylene-Inducible DNA-Binding Proteins That Interact with an Ethylene-Responsive Element. *Plant Cell*, **7**, 173-182.

References

- Onate-Sanchez, L. and Singh, K.B.** (2002) Identification of Arabidopsis ethylene-responsive element binding factors with distinct induction kinetics after pathogen infection. *Plant physiology*, **128**, 1313-1322.
- Orphanides, G. and Reinberg, D.** (2002) A unified theory of gene expression. *Cell*, **108**, 439-451.
- Ouwerkerk, P.B. and Meijer, A.H.** (2001) Yeast one-hybrid screening for DNA-protein interactions. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, **Chapter 12**, Unit 12 12.
- Pappan, K., Qin, W., Dyer, J.H., Zheng, L. and Wang, X.** (1997a) Molecular cloning and functional analysis of polyphosphoinositide-dependent phospholipase D, PLDbeta, from Arabidopsis. *The Journal of biological chemistry*, **272**, 7055-7061.
- Pappan, K. and Wang, X.** (1999) Molecular and biochemical properties and physiological roles of plant phospholipase D. *Biochimica et biophysica acta*, **1439**, 151-166.
- Pappan, K., Zheng, S. and Wang, X.** (1997b) Identification and characterization of a novel plant phospholipase D that requires polyphosphoinositides and submicromolar calcium for activity in Arabidopsis. *The Journal of biological chemistry*, **272**, 7048-7054.
- Park, P.J.** (2009) ChIP-seq: advantages and challenges of a maturing technology. *Nature reviews. Genetics*, **10**, 669-680.
- Pennacchio, L.A. and Rubin, E.M.** (2001) Genomic strategies to identify mammalian regulatory sequences. *Nature Reviews Genetics*, **2**, 100-109.
- Pfanschmidt, T., Schutze, K., Fey, V., Sherameti, I. and Oelmuller, R.** (2003) Chloroplast redox control of nuclear gene expression--a new class of plastid signals in interorganellar communication. *Antioxidants & redox signaling*, **5**, 95-101.
- Philippakis, A.A., Qureshi, A.M., Berger, M.F. and Bulyk, M.L.** (2008) Design of compact, universal DNA microarrays for protein binding microarray experiments. *J Comput Biol*, **15**, 655-665.

References

- Pillai, S. and Chellappan, S.P.** (2009) CHIP on chip assays: genome-wide analysis of transcription factor binding and histone modifications. *Methods Mol Biol*, **523**, 341-366.
- Pitsch, N.T., Witsch, B. and Baier, M.** (2010) Comparison of the chloroplast peroxidase system in the chlorophyte *Chlamydomonas reinhardtii*, the bryophyte *Physcomitrella patens*, the lycophyte *Selaginella moellendorffii* and the seed plant *Arabidopsis thaliana*. *Bmc Plant Biol*, **10**.
- Pokhilko, A., Fernandez, A.P., Edwards, K.D., Southern, M.M., Halliday, K.J. and Millar, A.J.** (2012) The clock gene circuit in *Arabidopsis* includes a repressilator with additional feedback loops. *Mol Syst Biol*, **8**.
- Priest, H.D., Filichkin, S.A. and Mockler, T.C.** (2009) cis-Regulatory elements in plant cell signaling. *Curr Opin Plant Biol*, **12**, 643-649.
- Proudfoot, N.** (2000) Connecting transcription to messenger RNA processing. *Trends in biochemical sciences*, **25**, 290-293.
- Raha, D., Hong, M. and Snyder, M.** (2010) ChIP-Seq: a method for global identification of regulatory elements in the genome. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, **Chapter 21**, Unit 21 19 21-14.
- Raikhel, N.** (1992) Nuclear Targeting in Plants. *Plant physiology*, **100**, 1627-1632.
- Reece-Hoyes, J.S., Barutcu, A.R., McCord, R.P., Jeong, J.S., Jiang, L., MacWilliams, A., Yang, X., Salehi-Ashtiani, K., Hill, D.E., Blackshaw, S., Zhu, H., Dekker, J. and Walhout, A.J.M.** (2011) Yeast one-hybrid assays for gene-centered human gene regulatory network mapping. *Nature Methods*, **8**, 1050-+.
- Riano-Pachon, D.M., Ruzicic, S., Dreyer, I. and Mueller-Roeber, B.** (2007) PlnTFDB: an integrative plant transcription factor database. *Bmc Bioinformatics*, **8**.
- Richly, E. and Leister, D.** (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. *Gene*, **329**, 11-16.
- Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z.,**

References

- Ghandehari, D., Sherman, B.K. and Yu, C.L.** (2000) Arabidopsis transcription factors: Genome-wide comparative analysis among eukaryotes. *Science*, **290**, 2105-2110.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y.J., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O.L., He, A., Marra, M., Snyder, M. and Jones, S.** (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, **4**, 651-657.
- Roeder, R.G.** (1991) The Complexities of Eukaryotic Transcription Initiation - Regulation of Preinitiation Complex Assembly. *Trends in biochemical sciences*, **16**, 402-408.
- Rombauts, S., Florquin, K., Lescot, M., Marchal, K., Rouze, P. and Van de Peer, Y.** (2003) Computational approaches to identify promoters and cis-regulatory elements in plant genomes. *Plant physiology*, **132**, 1162-1176.
- Ruelland, E., Cantrel, C., Gawer, M., Kader, J.C. and Zachowski, A.** (2002) Activation of phospholipases C and D is an early response to a cold exposure in Arabidopsis suspension cells. *Plant physiology*, **130**, 999-1007.
- Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Commun*, **290**, 998-1009.
- Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2006) Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell*, **18**, 1292-1309.
- Sanchez, J.P. and Chua, N.H.** (2001) Arabidopsis PLC1 is required for secondary responses to abscisic acid signals. *Plant Cell*, **13**, 1143-1154.
- Sanchez, J.P., Duque, P. and Chua, N.H.** (2004) ABA activates ADPR cyclase and cADPR induces a subset of ABA-responsive genes in Arabidopsis. *Plant J*, **38**, 381-395.
- Schneider, R. and Grosschedl, R.** (2007) Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Gene Dev*, **21**, 3027-3043.

References

- Schwacke, R., Fischer, K., Ketelsen, B., Krupinska, K. and Krause, K.** (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. *Mol Genet Genomics*, **277**, 631-646.
- Schwechheimer, C., Zourelidou, M. and Bevan, M.W.** (1998) Plant transcription factor studies. *Annual review of plant physiology and plant molecular biology*, **49**, 127-150.
- Seoighe, C. and Gehring, C.** (2004) Genome duplication led to highly selective expansion of the Arabidopsis thaliana proteome. *Trends in Genetics*, **20**, 461-464.
- Shaikhali, J., Heiber, I., Seidel, T., Stroher, E., Hiltcher, H., Birkmann, S., Dietz, K.J. and Baier, M.** (2008) The redox-sensitive transcription factor Rap2.4a controls nuclear expression of 2-Cys peroxiredoxin A and other chloroplast antioxidant enzymes. *BMC plant biology*, **8**, 48.
- Shen, Q.X. and Ho, T.H.D.** (1995) Functional Dissection of an Abscisic-Acid (Aba)-Inducible Gene Reveals 2 Independent Aba-Responsive Complexes Each Containing a G-Box and a Novel Cis-Acting Element. *Plant Cell*, **7**, 295-307.
- Shimada, T., Fujita, N., Maeda, M. and Ishihama, A.** (2005) Systematic search for the Cra-binding promoters using genomic SELEX system. *Genes Cells*, **10**, 907-918.
- Shinozaki, K. and Yamaguchi-Shinozaki, K.** (1996) Molecular responses to drought and cold stress. *Curr Opin Biotech*, **7**, 161-167.
- Shinozaki, K. and Yamaguchi-Shinozaki, K.** (1997) Gene Expression and Signal Transduction in Water-Stress Response. *Plant physiology*, **115**, 327-334.
- Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2000) Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. *Curr Opin Plant Biol*, **3**, 217-223.
- Shinozaki, K. and Yamaguchi-Shinozaki, K.** (2007) Gene networks involved in drought stress response and tolerance. *J Exp Bot*, **58**, 221-227.
- Sogabe, Y., Nakamura, H., Nakagawa, T., Hasegawa, S., Asano, T., Ohta, H., Yamaguchi, K., Mueller, M.J., Kodama, H. and Nishiuchi, T.** (2011) Visualization of wounding-induced root-to-shoot communication in Arabidopsis. *Plant signaling & behavior*, **6**, 1037-1039.

References

- Solano, R., Stepanova, A., Chao, Q.M. and Ecker, J.R.** (1998) Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Gene Dev*, **12**, 3703-3714.
- Steponkus, P.L., Uemura, M., Joseph, R.A., Gilmour, S.J. and Thomashow, M.F.** (1998) Mode of action of the COR15a gene on the freezing tolerance of Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 14570-14575.
- Stormo, G.D.** (2000) DNA binding sites: representation and discovery. *Bioinformatics*, **16**, 16-23.
- Storozhenko, S., De Pauw, P., Van Montagu, M., Inze, D. and Kushnir, S.** (1998) The heat-shock element is a functional component of the Arabidopsis APX1 gene promoter. *Plant physiology*, **118**, 1005-1014.
- Struhl, K.** (2001) Gene regulation. A paradigm for precision. *Science*, **293**, 1054-1055.
- Struhl, K., Kadosh, D., Keaveney, M., Kuras, L. and Moqtaderi, Z.** (1998) Activation and repression mechanisms in yeast. *Cold Spring Harbor symposia on quantitative biology*, **63**, 413-421.
- Sun, S., Yu, J.P., Chen, F., Zhao, T.J., Fang, X.H., Li, Y.Q. and Sui, S.F.** (2008) TINY, a dehydration-responsive element (DRE)-binding protein-like transcription factor connecting the DRE- and ethylene-responsive element-mediated signaling pathways in Arabidopsis. *The Journal of biological chemistry*, **283**, 6261-6271.
- Suzuki, M., Kao, C.Y. and McCarty, D.R.** (1997) The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *Plant Cell*, **9**, 799-807.
- Suzuki, M., Ketterling, M.G. and McCarty, D.R.** (2005) Quantitative statistical analysis of cis-regulatory sequences in ABA/VP1-and CBF/DREB1-regulated genes of Arabidopsis. *Plant physiology*, **139**, 437-447.
- Syed, F., Grunenwald, H. and Caruccio, N.** (2009) Optimized library preparation method for next-generation sequencing. *Nature Methods*, **6**, 1-11.

References

- Tabach, Y., Brosh, R., Buganim, Y., Reiner, A., Zuk, O., Yitzhaky, A., Koudritsky, M., Rotter, V. and Domany, E.** (2007) Wide-Scale Analysis of Human Functional Transcription Factor Binding Reveals a Strong Bias towards the Transcription Start Site. *PLoS one*, **2**.
- Thijs, G., Lescot, M., Marchal, K., Rombauts, S., De Moor, B., Rouze, P. and Moreau, Y.** (2001) A higher-order background model improves the detection of promoter regulatory elements by Gibbs sampling. *Bioinformatics*, **17**, 1113-1122.
- Thomashow, M.F.** (1999) Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annual review of plant physiology and plant molecular biology*, **50**, 571-599.
- Tjian, R. and Maniatis, T.** (1994) Transcriptional Activation - a Complex Puzzle with Few Easy Pieces. *Cell*, **77**, 5-8.
- Tompa, M., Li, N., Bailey, T.L., Church, G.M., De Moor, B., Eskin, E., Favorov, A.V., Frith, M.C., Fu, Y.T., Kent, W.J., Makeev, V.J., Mironov, A.A., Noble, W.S., Pavese, G., Pesole, G., Regnier, M., Simonis, N., Sinha, S., Thijs, G., van Helden, J., Vandenbogaert, M., Weng, Z.P., Workman, C., Ye, C. and Zhu, Z.** (2005) Assessing computational tools for the discovery of transcription factor binding sites. *Nature Biotechnology*, **23**, 137-144.
- Tuerk, C. and Gold, L.** (1990) Systematic Evolution of Ligands by Exponential Enrichment - Rna Ligands to Bacteriophage-T4 DNA-Polymerase. *Science*, **249**, 505-510.
- Ulker, B., Peiter, E., Dixon, D.P., Moffat, C., Capper, R., Bouche, N., Edwards, R., Sanders, D., Knight, H. and Knight, M.R.** (2008) Getting the most out of publicly available T-DNA insertion lines. *Plant J*, **56**, 665-677.
- Ulker, B. and Somssich, I.E.** (2004) WRKY transcription factors: from DNA binding towards biological function. *Curr Opin Plant Biol*, **7**, 491-498.
- Usadel, B., Nagel, A., Steinhauser, D., Gibon, Y., Blasing, O.E., Redestig, H., Sreenivasulu, N., Krall, L., Hannah, M.A., Poree, F., Fernie, A.R. and Stitt, M.** (2006) PageMan: An interactive ontology tool to generate, display, and annotate overview graphs for profiling experiments. *Bmc Bioinformatics*, **7**.

References

- Van Bortle, K. and Corces, V.G.** (2012) Nuclear Organization and Genome Function. *Annu Rev Cell Dev Bi*, **28**, 163-187.
- Van Criekinge, W. and Beyaert, R.** (1999) Yeast Two-Hybrid: State of the Art. *Biological procedures online*, **2**, 1-38.
- Venters, B.J. and Pugh, B.F.** (2009) How eukaryotic genes are transcribed. *Crit Rev Biochem Mol*, **44**, 117-141.
- Vidal, M. and Legrain, P.** (1999) Yeast forward and reverse 'n'-hybrid systems. *Nucleic acids research*, **27**, 919-929.
- Voinnet O., Rivas S, Mestre P., Baulcombe D. (2003)** An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J* **33**, 949-956
- Wang, H. and Cutler, A.J.** (1995) Promoters from kin1 and cor6.6, two Arabidopsis thaliana low-temperature- and ABA-inducible genes, direct strong beta-glucuronidase expression in guard cells, pollen and young developing seeds. *Plant Mol Biol*, **28**, 619-634.
- Wang, H., Datla, R., Georges, F., Loewen, M. and Cutler, A.J.** (1995) Promoters from kin1 and cor6.6, two homologous Arabidopsis thaliana genes: transcriptional regulation and gene expression induced by low temperature, ABA, osmoticum and dehydration. *Plant Mol Biol*, **28**, 605-617.
- Wang, H., Georges, F., Pelcher, L.E., Saleem, M. and Cutler, A.J.** (1994) A 5.3-Kilobase Genomic Fragment from Arabidopsis-Thaliana Containing Kin1 and Cor6.6. *Plant physiology*, **104**, 291-292.
- Wang, W.M., Yang, X.H., Tangchaiburana, S., Ndeh, R., Markham, J.E., Tsegaye, Y., Dunn, T.M., Wang, G.L., Bellizzi, M., Parsons, J.F., Morrissey, D., Bravo, J.E., Lynch, D.V. and Xiao, S.Y.** (2008) An Inositolphosphorylceramide Synthase Is Involved in Regulation of Plant Programmed Cell Death Associated with Defense in Arabidopsis. *Plant Cell*, **20**, 3163-3179.

References

- Wang, Y. and Hua, J.** (2009) A moderate decrease in temperature induces COR15a expression through the CBF signaling cascade and enhances freezing tolerance. *Plant J*, **60**, 340-349.
- Wasserman, W.W. and Fickett, J.W.** (1998) Identification of regulatory regions which confer muscle-specific gene expression. *Journal of molecular biology*, **278**, 167-181.
- Wasserman, W.W., Palumbo, M., Thompson, W., Fickett, J.W. and Lawrence, C.E.** (2000) Human-mouse genome comparisons to locate regulatory sites. *Nature Genetics*, **26**, 225-228.
- Webb, M.S., Gilmour, S.J., Thomashow, M.F. and Steponkus, P.L.** (1996) Effects of COR6.6 and COR15am polypeptides encoded by COR (cold-regulated) genes of *Arabidopsis thaliana* on dehydration-induced phase transitions of phospholipid membranes. *Plant physiology*, **111**, 301-312.
- Welin, B.V., Olson, A. and Palva, E.T.** (1995) Structure and organization of two closely related low-temperature-induced *dhn/lea/rab*-like genes in *Arabidopsis thaliana* L. Heynh. *Plant Mol Biol*, **29**, 391-395.
- Wilhelm, K.S. and Thomashow, M.F.** (1993) *Arabidopsis-Thaliana Cor15b*, an Apparent Homolog of *Cor15a*, Is Strongly Responsive to Cold and Aba, but Not Drought. *Plant Mol Biol*, **23**, 1073-1077.
- Woo, H.A., Chae, H.Z., Hwang, S.C., Yang, K.S., Kang, S.W., Kim, K. and Rhee, S.G.** (2003) Reversing the inactivation of peroxiredoxins caused by cysteine sulfinic acid formation. *Science*, **300**, 653-656.
- Woodson, J.D. and Chory, J.** (2008) Coordination of gene expression between organellar and nuclear genomes. *Nature Reviews Genetics*, **9**, 383-395.
- Wray, G.A., Hahn, M.W., Abouheif, E., Balhoff, J.P., Pizer, M., Rockman, M.V. and Romano, L.A.** (2003) The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol*, **20**, 1377-1419.
- Xiong, Y.Q., Liu, T.Y., Tian, C.G., Sun, S.H., Li, J.Y. and Chen, M.S.** (2005) Transcription factors in rice: A genome-wide comparative analysis between monocots and eudicots. *Plant Mol Biol*, **59**, 191-203.

References

- Yamaguchi-Shinozaki, K. and Shinozaki, K.** (1994) A Novel Cis-Acting Element in an Arabidopsis Gene Is Involved in Responsiveness to Drought, Low-Temperature, or High-Salt Stress. *Plant Cell*, **6**, 251-264.
- Yamaguchi-Shinozaki, K. and Shinozaki, K.** (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annual review of plant biology*, **57**, 781-803.
- Yamasaki, K., Kigawa, T., Seki, M., Shinozaki, K. and Yokoyama, S.** (2013) DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. *Trends in plant science*, **18**, 267-276.
- Yang, H.J., Li, Y.Q. and Hua, J.** (2006) The C2 domain protein BAP1 negatively regulates defense responses in Arabidopsis. *Plant J*, **48**, 238-248.
- Yang, H.J., Yang, S.H., Li, Y.Q. and Hua, J.** (2007) The Arabidopsis BAP1 and BAP2 genes are general inhibitors of programmed cell death. *Plant physiology*, **145**, 135-146.
- Yang, K.S., Kang, S.W., Woo, H.A., Hwang, S.C., Chae, H.Z., Kim, K. and Rhee, S.G.** (2002) Inactivation of human peroxiredoxin I during catalysis as the result of the oxidation of the catalytic site cysteine to cysteine-sulfinic acid. *Journal of Biological Chemistry*, **277**, 38029-38036.
- Yoshimura, K., Yabuta, Y., Ishikawa, T. and Shigeoka, S.** (2000) Expression of spinach ascorbate peroxidase isoenzymes in response to oxidative stresses. *Plant physiology*, **123**, 223-233.
- Zamocky, M., Janecek, S. and Koller, F.** (2000) Common phylogeny of catalase-peroxidases and ascorbate peroxidases. *Gene*, **256**, 169-182.
- Zeng, J., Yan, J., Wang, T., Mosbrook-Davis, D., Dolan, K.T., Christensen, R., Stormo, G.D., Haussler, D., Lathrop, R.H., Brachmann, R.K. and Burgess, S.M.** (2008) Genome wide screens in yeast to identify potential binding sites and target genes of DNA-binding proteins. *Nucleic acids research*, **36**.
- Zeng, W., Brutus, A., Kremer, J.M., Withers, J.C., Gao, X., Jones, A.D. and He, S.Y.** (2011) A genetic screen reveals Arabidopsis stomatal and/or apoplastic defenses against *Pseudomonas syringae* pv. tomato DC3000. *PLoS pathogens*, **7**, e1002291.

References

- Zhu, Y., Yang, H.J., Mang, H.G. and Hua, J.A.** (2011) Induction of BAP1 by a Moderate Decrease in Temperature Is Mediated by ICE1 in Arabidopsis. *Plant physiology*, **155**, 580-588.

6. Appendix

6.1 ERF1b primers

AT1G36060-EcoR1- F	TTTTT GAATTC GAGCGGATCTCTTCGGTGGT
AT1G36060-Xho1- R	TTTTT CTCGAGT CACGATAAAATTGAAGC
AT1G78080-BamH1-F	TTTTT GGATCC GTGCAGCTGCTATGAATTTG
AT1G78080-Xho1-R	TTTTT CTCGAGT CTAAGCTAGAATCGAATC
AT2G22200-BamH1-F	TTTTT GGATCC GAGAACTGCTTCTCTTTCTTTC
AT2G22200-EcoR1- R	TTTTT GAATTC TTAAGAATTGGCCAGTTTAC
AT1G22190-BamH1-F	TTTTT GGATCC GAACAACCTTCTATGGATTTTACAG
AT1G22190-EcoR1- R	TTTTT GAATTC CCTAATTTACAAGACTCGAACAC
AT5G65130-BamH1-F	TTTTT GGATCC GTGCTTTAAACATGAATGCTTAC
AT5G65130-Xho1-R	TTTTT CTCGAGT CAGAAGAGTTTCTCTAT
AT4G39780-BamH1-F	TTTTT GGATCC GTGCAGCCATAGATATGTTCAATAG
AT4G39780-Xho1-R	TTTTT CTCGAGT TTAAGATTCGGACAATTTG
AT1G64380-BamH1-F	TTTTT GGATCC GAGAAGAAAGCAATGATATTTTTC
AT1G64380-EcoR1- R	TTTTT GAATTC TTAATTGGCAAGAACTTCC
AT4G13620-BamH1-F	TTTTT GGATCC GTATCACACCAATACACACACAAC
AT4G13620-Xho1-R	TTTTT CTCGAGT CAAGAAGAATGAGGAAA

6.2 Rap2.4a targets

Gene	Function
at5g52840	mitochondrial electron transport / ATP synthesis.NADH-DH.localisation not clear
at2g07771	mitochondrial electron transport / ATP synthesis.cytochrome c
at2g45470	cell wall.cell wall proteins.AGPs.AGP
at4g39940	S-assimilation.AKN
at3g61177	stress.biotic.PR-proteins
at3g57620	Biodegradation of Xenobiotics
at5g05090	RNA.regulation of transcription.G2-like transcription factor family, GARP
at5g52830	RNA.regulation of transcription.WRKY domain transcription factor family
at2g01010	protein.synthesis.ribosomal RNA
at2g43930	protein.postranslational modification
at3g05640	protein.postranslational modification
at5g60900	protein.postranslational modification
at3g49030	protein.degradation.ubiquitin.E3.SCF.FBOX
at3g49040	protein.degradation.ubiquitin.E3.SCF.FBOX
at5g46210	protein.degradation.ubiquitin.E3.SCF.cullin
at2g28970	signalling.receptor kinases.misc
at2g28990	signalling.receptor kinases.misc
at5g54620	cell.organisation
at1g70895	development.unspecified
at5g29560	development.unspecified
at5g60910	development.unspecified
at2g44578	not assigned.no ontology
at1g70900	not assigned.unknown
at2g01021	not assigned.unknown
at2g37940	not assigned.unknown
at2g45250	not assigned.unknown
at4g06479	not assigned.unknown
at5g05100	not assigned.unknown
at5g11840	not assigned.unknown
at5g46220	not assigned.unknown

6.3 Rap2.4d targets

Gene	Function
at1g09935	glycolysis.unclear/dually targeted.phosphoglycerate mutase
at2g07771	mitochondrial electron transport / ATP synthesis.cytochrome c
at1g01710	lipid metabolism.lipid degradation.beta-oxidation.acyl-CoA thioesterase
at2g33590	secondary metabolism.phenylpropanoids
at1g18980	stress.abiotic.unspecified
at1g31170	redox.misc
at3g50760	misc.UDP glucosyl and glucuronyl transferases
at1g31710	misc.oxidases - copper, flavone etc
at3g20390	RNA.processing.ribonucleases
at5g06070	RNA.regulation of transcription.C2H2 zinc finger family
at2g19810	RNA.regulation of transcription.C3H zinc finger family
at3g30210	RNA.regulation of transcription.MYB domain transcription factor family
at5g33406	DNA.synthesis/chromatin structure.retrotransposon/transposase.hat-like transposase
at2g01010	protein.synthesis.ribosomal RNA
at1g64300	protein.postranslational modification
at2g12550	protein.degradation.ubiquitin
at1g10650	protein.degradation.ubiquitin.E3.RING
at4g05620	protein.degradation.ubiquitin.E3.SCF.FBOX
at2g28990	signalling.receptor kinases.misc
at3g45640	signalling.MAP kinases
at5g54800	transport.metabolite transporters at the envelope membrane
at4g08700	transport.nucleotides
at1g01620	transport.Major Intrinsic Proteins.PIP
at1g30135	not assigned.unknown
at1g49930	not assigned.unknown
at4g06479	not assigned.unknown

Poster presentation on the conferences

- Cell free protein synthesis, Potsdam, 31.01.13
- Havel-Spree-Colloquium 2012, Potsdam, Germany (Selected for oral presentation)
- GRC: Mitochondria and Chloroplast 2012, Smithfield, USA
- Molecular Interactions Workshop 2011, Berlin, Germany
- Botanikertagung (German Botanical Society meeting) 2011, Berlin, Germany
- Signals, Sensing and Plant Primary Metabolism, Potsdam, Germany
- 19th Photosynthesis workshop, Frankfurt, Germany

Curriculum vitae

For reasons of data protection, the curriculum vitae is not published in the electronic version

Acknowledgement

Above all, I thank GOD who not only preserved me alive but also protected and supported me during tough times of my life. First and foremost, I offer my sincerest gratitude to my supervisor, Prof. Dr. Margarete Baier, who has supported me throughout my PhD work with her immense knowledge whilst allowing me the room to think and work in my own way. This work would not have been possible without the constant support, patience, and encouragement of my advisor. I would also like to thank her for accepting me as her student while funding my stipend from her own project. The joy and enthusiasm she has for her research was contagious and motivational for me, even during tough times of my PhD work. One simply could not wish for a better supervisor.

I extend my thanks to the present and former group members of Prof. Dr. Baier who have helped me in one way or the other during my stay in the laboratory. The group has been a source of friendships as well as good advice and collaboration. Importantly, I am grateful to Wei Guo for his friendly help during my studies and thesis write up. Wei, is someone you will instantly love and never forget once you meet him. My earnest appreciation also goes to Radoslaw Rudnik not only for friendly and insightful discussions about our research topic but also for his kind and enormous help as my 'German language translator'. I would like to thank my thesis examination committee member and second supervisor, Prof. Dr. Rupert Mutzel, especially for his time in reviewing my thesis with in a very short time.

I would also like to thank my family for the support they have provided me through my entire life. In particular, I must acknowledge my mother, Yeshareg Abate, without her love and encouragement I would not be where I am now. Last but on the least, I would like to thank all my friends in Berlin who made my life warm and enjoyable in Germany.

I hereby declare that I have prepared and written the presented doctoral thesis myself using only the presented methods and sources. All sources from literature are marked as such and are properly cited.

Berlin, June, 2013

Jote Bulcha