PROSPECTING CANDIDATE GENES FOR DROUGHT TOLERANCE FROM FINGER MILLET (*Eleusine coracana* (L.) Gaertn.)

PARVATHI M SREEKUMAR
PAK 8157

DEPARTMENT OF CROP PHYSIOLOGY
UNIVERSITY OF AGRICULTURAL SCIENCES
BENGALURU

2010

PROSPECTING CANDIDATE GENES FOR DROUGHT TOLERANCE FROM FINGER MILLET (*Eleusine coracana* (L.) Gaertn.)

PARVATHI M SREEKUMAR

PAK 8157

THESIS SUBMITTED TO THE
UNIVERSITY OF AGRICULTURAL SCIENCES, BENGALURU
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE

***MASTER OF SCIENCE* (AGRICULTURE)**
IN
CROP PHYSIOLOGY

BENGALURU JULY 2010

Dedicated to
My Beloved Family
At home as well as in the department

Acknowledgement

As goes the adage, all that happens is for good……, it has been surely one of God's whims that I joined this wonderful department: Department of Crop Physiology, UAS, GKVK, Bengaluru, which is always engraved in gold on my heart. At this moment and forever, I fondly remember all the people who have helped me sculpt my career. I take this opportunity to convey my love and appreciation to all those who have been a part of my life during this phase of my life.

It is with immense pleasure that I express my profound gratitude to ***Dr. Nataraja Karaba. N.****, the Chairman of my Advisory Committee. He was more a mentor and friend than just a guide, who was a strong pillar of support in all my attempts with adequate appreciation as well as criticisms. He helped me to inculcate a logical way of approaching science with accuracy and I extend my warm regards for him for being whatever he has been to me.*

The relentless hardwork of ***Professor M. Udayakumar,*** *to extend his scientific knowledge has been a sheer source of inspiration. I am so grateful to him for the priceless guidance that he has given, not only as an Advisory Committee member but also as a terrific teacher, which helped me to excel in the field of science.* ***Dr. V. R. Sashidhar*** *surely has infused an insight into me to think, speak and write with a good vocabulary. As an Advisory Committee member, it was so kind of him to give his valuable and frank remarks which helped me interpret my research well. It is with great love and respect that I acknowledge all the encouragement and assistance rendered by* ***Dr. Anita Peter****, my Advisory Committee member, during my educational carrier at various stages of my study and investigation.*

All my ***Gurus*** *of the Department of Crop Physiology require a special mention as they are so unique in their own way. I have no words to explain my respect and regard for the ever caring* ***T. G. Prasad sir*** *who has a very strong role in the accomplishment of this achievement in my life. They have all instilled in me a new vision towards science which has made me to decide to tirelessly contribute for the advancement of science. At this juncture, I warmly remember* ***Dr. Roy Stephen*** *and* ***Dr. K. D. Prathapan****, my undergraduate teachers, for their adequate guidance at the right time. I also thank all my course teachers who has imparted me knowledge and wisdom during the course of my study.*

With all respect, love and affection, I tenderly remember all the seniors of our department who have been there to clarify any doubt and assist me during the period of my research. All my labmates were so special for me for having made me feel at home, especially Mamrutha Di and Yashoda Di, who held my hands as a toddler into the field of molecular biology. Pruthvi Di, Yamuna akka, Nagu akka, Babitha Di, Basu.., Vasantha akka, Nethra Di, Pallavi, Nagarjun Sir, Jali Sir, Rame Sir, Rahul Sir and Sasi really helped me keep my tempo high during my research period. The threesome- Shailesh, Patil Mahesh and Raju were always there to cheer me up. Lakshmi Madam was more like a mother to me with all her warmth and care.

I fondly remember Jasee for her lovely friendship and can never forget all those lunch boxes from home. I cherish and value all those discussions and heated arguments on science that I shared with Sajeevan, Ashwini, Reddy Sir and Mahesh Sir, which helped each of us a lot. This helped me to chisel out skills required for a scientific career. Their encouragement and care meant a lot to me and have helped me keep up my enthusiasm. I fondly remember the helping hands of all my other senior and junior colleagues without mentioning anybody in particular as it would be unfair if I leave out anybody.

To all other supporting staff of Crop Physiology Department especially Devatha Madam for providing us with all the material requirements for smooth execution of my research work, ever caring Prema akka, all members of the Department garden and all other non teaching staff.

It has been a worthy privilege for me to have a wonderful bunch of classmates. Pushpa, the two Pavis – CB and HV, Soumya, Smitha, Jaya, Shankar, Manoa, Tashi, Dhanraj, Nagaraj and Sunil – thanks for being there for me whenever I was in need. Thanks buddies. All my juniors made the days of research chores alone more alive with their apt interactions and thanks for their support and encouragement. Those whom I deeply missed during my stay here were my dear friends Maria, Lisa, Chinnu and Anjitha and Deepoo, my little brother, who always held me up with their prayers and comforting words during my difficult times.

I fall short of words to express my love beyond limits for my two room mates, Pushpa and Nandini, for their bountiful affection and for all that we shared during our good as well as bad times. I also realise the affection of my good natured hostel mates who made my stay at GKVK enjoyable and memorable.

I sincerely acknowledge ICAR, New Delhi for providing me financial assistance with its Junior Research Fellowship during the entire course of my study.

And last but not the least, my ***Family*** *….. It would have been impossible for me to successfully realise this venture in life without the unconditional love, warmth and encouragement of my dearest Achan and Amma in all my endeavours and laying their belief on my capabilities. I could literally feel the strength of the prayers of my grandmother, all other family members and all kith and kin who wished for my well being. And my bro, Vichu, for all his unexpressed love and affection, I knew you would be there for me always.*

Finally, zillions of thanks from my teeming heart to all those who have been involved in some way or the other in the completion of this thesis.

Bengaluru *Parvathi M Sreekumar*

July 2010

Prospecting candidate genes for drought tolerance from finger millet (*Eleusine coracana* (L.) Gaertn.)

ABSTRACT

Drought, in conjugation with high temperature and radiation, is the most critical environmental constraint on a global basis for plant survival and productivity. Better understanding of the effects of drought on plants and the recuperative mechanisms that grant them cellular tolerance (CT) is vital for crop improvement. Prospecting genes from stress adapted species than model systems is advantageous as the adaptations would have resulted from a novel biosynthetic pathway or novel protein. The major objective of the study was to identify and characterise drought stress responsive genes from a drought adapted crop, finger millet (*Eleusine coracana* (L.) Gaertn.).

Drought specific cDNA phage library developed in the Department of Crop Physiology, from finger millet leaf tissue, was screened and evaluated. Over 1100 clones were screened by plaque lifting and nearly 600 clones were short listed based on fragment length for sequencing. Annotation of the sequenced clones revealed that stress inducible genes belonged to diverse classes like cytoskeletal structures (10%), protective proteins (10%), catalytic genes (9%) etc., which are thought to maintain and enhance CT. Functional validation of selected genes by different approaches like e-northern, semi-quantitative and quantitative RT-PCR indicated the drought responsive nature. A few short listed stress responsive genes include metallothionein, sinapyl alcohol dehydrogenase, farnesylated protein, protein phosphatase, farnesyl pyrophosphate synthase, TAF6 and signal recognition particle receptor (SR). Two novel upstream regulatory genes TAF6 and SR were also found to be responsive to other abiotic stresses like salt, osmotic and oxidative stresses. These genes are hence proposed to be the prospected candidate genes which could be useful in manipulating cellular tolerance.

Signature of Student — Signature of Chairman

(Parvathi M Sreekumar) — (Nataraja Karaba N)

CONTENTS

CHAPTER	TITLE	PAGE NUMBERS
I	INTRODUCTION	1- 2
II	REVIEW OF LITERATURE	3- 21
III	MATERIALS AND METHODS	22- 36
IV	RESULTS	37- 62
V	DISCUSSION	63- 70
VI	SUMMARY	71- 73
VII	REFERENCES	74- 91
	APPENDICES	92- 132

LIST OF TABLES

Sl. No.	Title	Page No.
1	Typical transcriptional elements related to abiotic stresses in plants and crops	14
2	Potential genes mediating resistance to soil water stress and related abiotic stress in plants	15-18
3	Some examples of osmoregulatory genes associated with abiotic stress resistance	19
4	List of annotated genes cloned from drought stress specific cDNA library developed from finger millet leaf tissue	40-50
5	Details of the clone identities of selected clones	51
6	AGI codes of the homologous genes of *Arabidopsis thaliana* shortlisted for e-Northern analysis	53
7	Soil water potential of pots maintained at different field capacities to impose drought on finger millet	56
8	Leaf solute potential and net osmotic adjustment of finger millet plants maintained at various field capacities	56
9	List of stress responsive genes and primers used for expression analysis	59

LIST OF FLOWCHARTS

Sl. No.	Title	Between Pages
1	Schematic representation of RACE protocol used to clone full length genes (Invitrogen; USA)	36-37

LIST OF FIGURES

Sl. No.	Title	Between Pages
A	Stresses and stress related strains in plant tissues	5-7
B	General stress concept: emphasis on the incidence of a specific primary and a less specific secondary strain by a specific stressor	5-7
C	Basic plant gene regulatory network system	8-10
D	Model showing events associated with gene discovery using specific traits	19-21
1	Classification of annotated ESTs developed from finger millet leaf tissue collected from plants experiencing drought stress.	38-39
2	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under osmotic stress	54-55
3	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under drought stress	54-55
4	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under cold stress	54-55
5	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under salt stress	54-55
6	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under radiation stress	54-55
7	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under genotoxic stress	54-55
8	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under oxidative stress	54-55
9	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under wounding stress	54-55

Sl. No.	Title	Between Pages
10	E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues tissues under heat stress	54-55
11	Imposition of drought stress in finger millet by gravimetric approach	55-56
12	Phenotype of finger millet plants experiencing drought stress at whole plant level	55-56
13	Relative water content (RWC) of finger millet leaf tissue experiencing different levels of drought stress	56-57
14	Loss of membrane integrity of finger millet leaf tissue experiencing different levels of drought stress	56-57
15	Reduction in total chlorophyll content (R %) in finger millet leaf tissue experiencing different levels of drought stress	57-58
16	Reduction in cell viability in finger millet leaf tissue experiencing different levels of drought stress	57-58
17	Seedling assays: Reduction in growth on stress imposition on finger millet seedlings using NaCl (300 mM), PEG (- 8 bars) and Methyl Viologen (10 μM)	57-58
18	Quantitative RT PCR expression levels of (a)TAF6 and (b) signal recognition particle receptor in seedling (shoot) system	61-62

LIST OF PLATES

SL. No.	TITLE	Between Pages
1	Ethidium bromide stained agarose gel showing PCR amplified products lifted plaques	38-39
2	Ethidium bromide stained agarose gel showing purified PCR amplified products	38-39
3	Ethidium bromide stained agarose gel showing PCR amplified products of converted phagemids after *in vivo* excision.	52-53
4	Ethidium bromide stained agarose gel showing PCR amplified products of converted phagemids after *in vivo* excision	52-53
5	Ethidium bromide stained agarose gel showing PCR amplified products of the selected genes in converted phagemids after *in vivo* excision.	52-53
6	Ethidium bromide stained agarose gel showing PCR amplified products of the isolated plasmids	52-53
7	Effect of stress imposition on finger millet seedlings using NaCl (300 mM), PEG (- 8 bars) and Methyl Viologen (10 μM)	57-58
8	Semi quantitative RT PCR of known drought responsive gene, early light inducible protein (ELIP).	58-59
9	Semi quantitative RT PCR of signal recognition particle receptor at 15, 20, 23,25 and 30 cycles	58-59
10	Semi quantitative RT PCR to study the pattern of expression of stress responsive genes cloned from finger millet	60-61
11	Quantitative RT PCR to study the pattern of expression of metallotheionin and farnesylated protein cloned from finger millet	60-61
12	Quantitative RT PCR to study the pattern of expression of protein phosphatase and RIS BZ4 cloned from finger millet	60-61
13	Quantitative RT PCR to study the pattern of expression of sinapyl alcohol dehydrogenase and transcriptional regulator cloned from finger millet	60-61

SL. No.	TITLE	Between Pages
14	Quantitative RT PCR to study the pattern of expression of TAF6 and Signal recognition particle receptor cloned from finger millet	60-61
15	Quantitative RT PCR to study the pattern of expression of Light inducible protein cloned from finger millet	60-61
16	Semi quantitative RT PCR of TAF6 and SR in seedling (shoot) system	61-62
17	Partial sequence information and predicted structural features of TAF6	61-62
18	Partial sequence information and predicted structural features of Signal recognition particle receptor	61-62
19	Cloning of full length genes of TAF6 and signal recognition particle receptor through RACE approach	62-63

INTRODUCTION

I. INTRODUCTION

Drought along with high temperature is the most important abiotic factor affecting plant survival and productivity globally (Chaves *et al*., 2003). A better understanding of the plant processes affected by drought and adaptive mechanisms associated with tolerance is necessary for developing crop improvement strategies and improving management practices. Plant's response to stress is a complex quantitative trait and is controlled by many genes. There is a gap in our understanding on the molecular- to- physiological mechanism underlying plant responses (Mittler *et al*., 2001; Cushman and Bohnert, 2000). In order to understand these responses to stress at molecular level, it would be ideal to study the expression profiles of all the genes under stressful condition using stress adapted plant system.

A wide array of functional genomics tools have helped to evaluate stress signal perception and transduction pathways, and of the associated molecular regulatory network (Bray *et al*; 2004; Shinozaki *et* al; 2007; Shao *et al*; 2008). Different approaches like subtractive hybridization, suppressive subtractive hybridization (Diatchenko *et al*., 1996), differential display (Liang and Pardee, 1992; Cho *et al*., 2001), cDNA-AFLP (Kivioja *et al.,* 2005), microarray technology and other means (Seki *et al*., 2002; Shinozaki and Yamaguchi-Shinozaki, 2007) have been employed in the identification of the differentially expressed genes. Recent progress in molecular biology and bioinformatics (especially, DNA microarray technology), genomics, proteomics, metabolomics and transcriptomics) has provided insight into plant gene regulatory network system, which is mainly composed of inducible-genes (environmental factors and developmental cues), their expression programming and regulatory elements (cis-element and trans-element), corresponding biochemical pathways and diverse signal factors (Shinozaki *et al*; 2003).

Expressed Sequence Tags (ESTs) is an effective genomic approach for rapid identification of expressed genes under a particular condition and has been widely used in genome-wide gene expression studies in various tissues, developmental stages or under different environmental conditions. Generation and characterization of stress specific ESTs can be a viable approach when genome information is limited. Such genomic resources are considered as important tools for plant functional genomics, which forms

the basis to combine relevant traits of interest into a desired genetic background. Expresssed sequence tags and gain-in-function mutants are powerful tools in this regard. In many studies, the information generated by ESTs is being used to understand the adaptive responses. The BLAST result of groundnut Expresssed sequence tags (ESTs) revealed that significant number of unique groundnut seed ESTs match soybean (396), *Arabidopsis* (2952), rice (682), and other plant species (Guo *et al*., 2008). In plants, cDNA libraries are widely being used as the choice material for transcriptome profiling (Seki *et al*., 2002). Expression profiling of stress specific ESTs and full-length cDNA clones has lead to the identification of key regulators in plant signal transduction cascades. In light of these facts, one can authentically proclaim the importance of identifying Expressed Sequence Tags (ESTs).

Various experiments conducted in the Department of Crop Physiology have proved that certain species like groundnut, finger millet and horse gram are better adapted to drought (Gopalakrishna, 2001; Geetha-Govind, 2006). Hence, identification of highly stress responsive novel genes from finger millet, a drought adapted crop, can be used as candidate genes for improving drought susceptible species. With this broad objective in mind, a drought stress specific cDNA library was developed from finger millet plants maintained at different levels of drought stress (Ramegowda, 2008). It was proposed to use the library for prospecting novel stress responsive genes and validate their relevance by gene expression studies. Thus, this study was aimed at discovering stress responsive genes from stress adapted species finger millet.

From this context, the major objective of the study was to prospect a few candidate genes for imparting drought tolerance. The title of the investigation was framed as **"Prospecting candidate genes for drought tolerance from finger millet [*Eleusine coracana* (L.) Gaertn.] ".** The objectives set up for achieving this broad target were:

Major objectives:

1. Screening and evaluation of finger millet stress cDNA library
2. Identification of novel genes by sequencing and annotation using bioinformatics tools
3. Expression analysis of selected genes by
    (i) e-northern using database information
    (ii) examining transcript levels in the tissues exposed to drought stress

REVIEW OF LITERATURE

II. REVIEW OF LITERATURE

2.1 Introduction

Indian agriculture is much affected by climatic vagaries at the same time being prone to suffer under the clutches of various environmental stresses. Crop plants are exposed to a variety of abiotic stresses such as drought, high temperature, salinity, cold, heavy metals, mechanical wounding etc., under field conditions. In the global scenario, different cropping situations fall victim to a complex of stresses. Amongst many stresses, drought is being considered as a potent threat to crop establishment and growth (Boyer, 1982). Drought is a perennial and recurring feature in many parts of India. According to Government of India reports, about 68% of the country is prone to drought in varying degrees. The recent drought of 2002 in India ranks fifth in terms of magnitude (PACS Programme, 2001- 2008). Drought seems to cripple economic growth as it directly affects agriculture which contributes to GDP. Hence, water turns out to be the most critical factor for agricultural as well as economic development. Under this scenario, it is important to mitigate and reduce the harmful effect of drought on crop growth and productivity by using novel crop breeding approaches. To evolve novel crop types, it is necessary to understand the plant adaptive mechanisms associated with drought tolerance.

2.2 Abiotic stresses and their impact on crop yield

In a wide perspective, of late, the environment has shown to act rudely to all the different forms of life it supports, shedding apart its rosy and cosy portrait. The various environmental stresses that exist seem to choke both plant and animal life. The different stress realms prevail as both biotic as well as biotic stresses. Among these, the various abiotic stresses like temperature, salinity, drought and mechanical stresses on plants are found to create the maximum negative impact to crop productivity both in terms of quantity as well as quality (Bressan *et al*., 2009). Drought exhibits itself as the key critical factor that limits global productivity in the milieu of the recent drastic climatic change. A recent article reports on a major international joint effort in comparing upper-ocean heat content datasets and identifying the sources of uncertainties attached to them (Lyman *et al*., 2010). As per the study, it shows a robust evidence for a warming trend in the last decade which in turn would result in derogatory climatic patterns (Lyman *et al*., 2010).

There has been a mild and lukewarm response to the Copenhagen Accord which is the still born baby of the United Nations Climate Change Conference, 2009, Copenhagen (COP15) that went onto eventually fail in achieving a global climate change policy. The increasingly erratic weather is sure to affect plant survival adversely. Considering all these factors, it is highly significant to study into the depths of drought impact as well as normalising a particular performance level for each crop species to combat complex stresses.

2.3 Drought

Drought is broadly defined as a meteorological event which implies the absence of rainfall for a period of time long enough to cause moisture depletion in soil and water deficit with decrease of water potential in plant tissues (Kramer, 1980). Drought can be of three kinds: meteorological, hydrological and agricultural. Meteorological drought happens when the actual rainfall in an area is significantly less than the climatological mean of that area. Hydrological drought marks a depletion of surface water causing very low stream flow and drying of lakes, rivers and reservoirs. Agricultural drought occurs wherein an inadequate soil moisture results in acute crop stress and fall in agricultural productivity. In 2009, drought has resulted in the 29 per cent less rainfall, resulting in droughts in 246 districts in India (www.imd.gov.in). This will eventually lead to appalling climatic conditions mainly the rainfall pattern. Plant productivity relies upon striking a perfect balance between carbon assimilation and transpiration. For every gram of organic matter produced by plant, approximately 500g of water is lost by transpiration to atmosphere (Taiz and Zeiger, 2006). This loss and gain of water should in fact maintain an equilibrium failing which plant growth gets altered. Drought also has impeding impacts on soil fertility of which productivity has profound functionality (Moore, 2005).

Majority of soil water deficit situations directly lead to drought, which is closely linked with natural rainfall. Drought is a complex physico -chemical process, in which many biological macromolecules and small molecules are involved, such as nucleic acids (DNA, RNA, microRNA), proteins, carbohydrates, lipids, hormones, ions, free radicals, mineral elements (Liu *et al*., 2003; Chae *et al*., 2009; Ni *et al*., 2009). In addition, drought is also related to salt, cold, high temperature, acid and alkaline stresses, pathological reactions, senescence, growth, development, cell cycle, UV-B damage, wounding,

embryogenesis, flowering, signal transduction and so on, making the problem more complicated (Hsaio *et al*., 1973; Mittler *et al*., 2006; Ni *et al*., 2009).

2.4 Whole plant response under stress

In plant stress analysis, primary and secondary damage as well as primary and secondary stresses have been differentiated, the latter emerging as a consequence of the former. Oxidative stress damage caused by reactive oxygen species (ROS) is an example for secondary stress as well as damage (Randy, 1995).The secondary stress results from the impacts of cold or of excess of light energy, which are the primary stressors (Huner *et al.,* 1998). Stressors like cold or drought induce various primary effects on the cellular level, termed strains (Levitt, 1980) (Figure A), which in turn lead to uncontrolled (damage) or controlled effects (adaptation) on that level. Perception of the abiotic stressor via the strains and the subsequent signal transduction cascade are responsible for the specific responses of a plant; simultaneously the secondary strains often induce less specific responses (Figure B). Both types of responses mediate adaptive responses and tolerance to environmental stresses. Subsequently the molecular effects of that stressor are suppressed and signal transduction pathways are modified (Beck *et al*., 2007). Adaptations include short-term (biochemical), but also long term processes, like morphological changes with respect to the habit of a plant. In a varied angle, plant responses are exhibited both as rapid and emerging as well as in a slow and adaptive manner (Fujita *et al*., 2006; Yamaguchi- Shinozaki and Shinozaki, 2006).

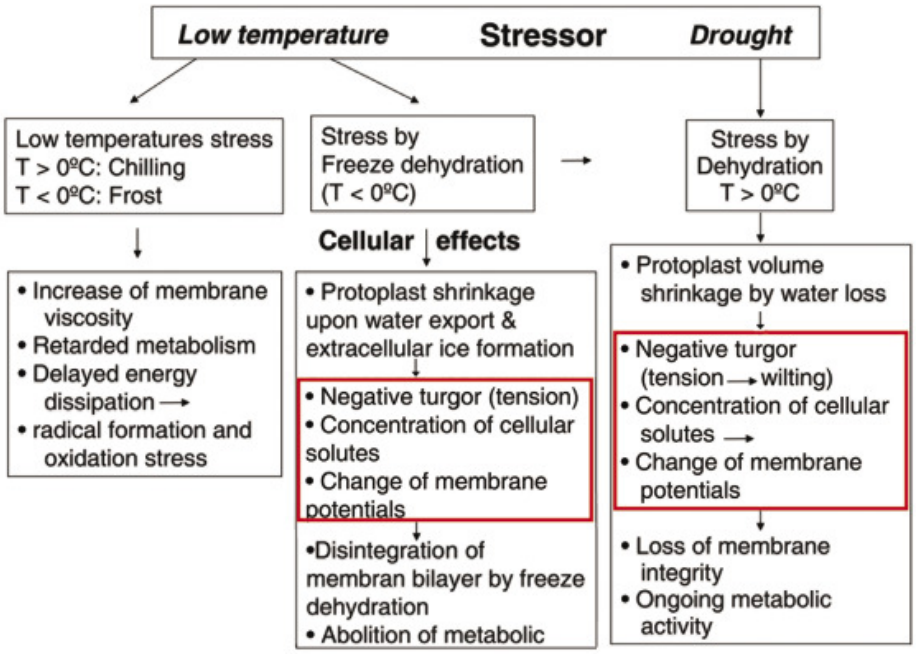


Figure A. Stresses and stress related strains in plant tissues (Beck *et al*; 2007)**.**

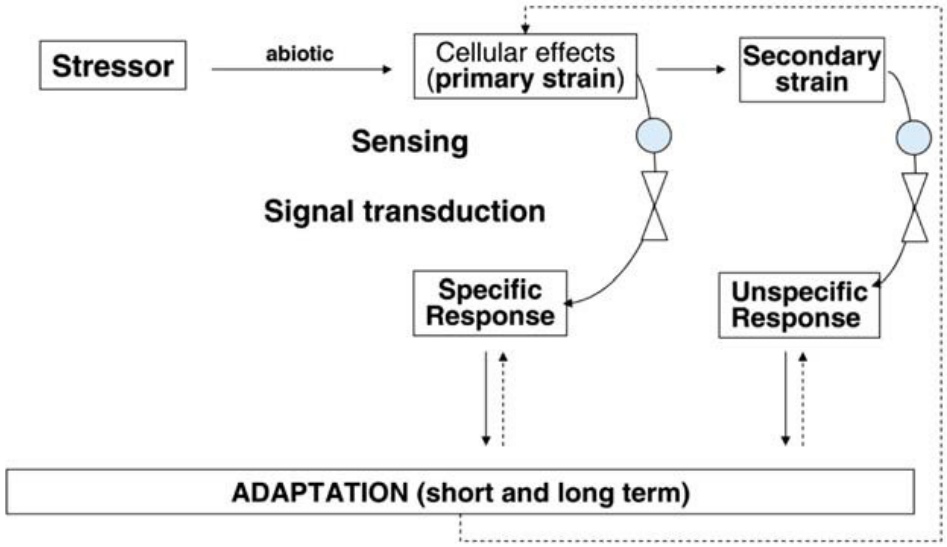


Figure B. General stress concept: emphasis on the incidence of a specific primary and a less specific secondary strain by a specific stressor
(Beck *et al*; 2007)**.**

2.4.1 Drought adaptive mechanisms

In response to the different stresses confronted by plants, several physiological, biochemical and molecular changes are initiated within them both at whole plant level and at cellular level. In due course of evolution, plants have also acquired certain adaptive mechanisms to combat these major threats. Of these, the major adaptations displayed to survive drought situations include dehydration postponement, desiccation tolerance and drought escape. Leaf area reduction can be considered as the first line of defence against drought. Wax deposition on the leaves, both on the surface and in the cuticle inner layer, in response to water deficit, controls the rate of water loss by complex mechanisms (Jenks *et al.,* 2002, Mamrutha *et al*., 2010). Evidences support the fact that plants alter their growth rates in response to any stress, by co-ordinately controlling many other important processes such as cell wall and membrane biosynthesis, cell division and protein synthesis (Burssens *et al*., 2000). Plants coordinate these stress responses by different signalling mechanisms. Abscisic acid (ABA), a stress hormone, along with other signalling components like pH and inorganic ion redistribution seems to play a considerable role in long distance signalling between roots and the shoots under drought leading to stomatal closure (Sauter *et al*., 2001; Davies *et al*., 2002). These adaptive mechanisms play a critical role in imparting stress tolerance.

Plant metabolites such as salicylic acid (SA), jasmonic acid (JA) and phytohormones such as ethylene (ET), and ABA are endogenous, low-molecular-weight molecules that primarily regulate the protective responses of plants against both biotic and abiotic stresses via synergistic and antagonistic actions, which are referred to as signaling crosstalk (Bostock, 2005; Lorenzo and Solano, 2005; Mauch- Mani and Mauch, 2005). Moreover, the generation of reactive oxygen species (ROS) has been proposed as a key process that is shared between biotic and abiotic stress responses (Apel and Hirt, 2004; Torres and Dangl, 2005). Rapidly accumulating data, resulting from large-scale transcriptome analyses with DNA microarray technology, strongly support the existence of such crosstalk between signalling network (Seki *et al*., 2002; Cheong *et al*., 2002; Davletova *et al*., 2005).

2.4.2 Genes regulating drought adaptive mechanisms

In living organisms, cellular response to environmental stimuli is mediated by three components of cellular organisation – the genome, transcriptome and proteome which coordinate together to program gene expression. Living organisms have naturally devised mechanisms to reprogram gene expression in order to cope up with the high degree of environmental plasticity. The transcriptome comprising of all the different classes of RNAs seems to be the significant factor which bridges between the gene and the protein. Among the different RNAs, messenger RNA (mRNA) is thought to be the key player whereby its abundance is regulated by its production as well as degradation kinetics (Shalem *et al.,* 2009).

Genes induced during stress conditions produce various metabolic proteins (functional proteins) and other proteins involved in the gene regulation (regulatory proteins) (Ramanjulu and Bartels, 2002; Sreenivasulu *et al.,* 2007; Govind *et al.,* 2008). The functional proteins include water channel proteins, chaperons, proteases, Late Embryogenesis Abundant (LEA) proteins as well as the enzymes for the synthesis of osmoprotectants (compatible solutes: sugars, proline, etc.), where as the regulatory proteins could be transcription factors, protein kinases, and enzymes for phosphoinositide (PI) turn- over, etc., (Bartels *et al.,* 1990; Bartels and Salamini, 2001; Waditee *et al.,* 2002; Bernacchia and Furini, 2004; Majee *et al*., 2004; Nakashima and Shinozaki, 2005; Bartels, 2005; Dastidar *et al.,* 2006).

Venn diagram analysis indicated the existence of significant cross talk between drought and high-salinity stress signalling processes (Seki *et al.*, 2002). There was a strong overlap of gene expression in response to drought, high-salinity and ABA and partial overlap of gene expression in response to cold and osmotic stress (Shinozaki and Yamaguchi-Shinozaki, 2000). Elucidation of the basic biochemical and molecular mechanisms for drought stress perception, transduction and tolerance is still a major challenge in biology. Plant modification for enhanced drought tolerance is mostly based on the manipulation of either transcription and/or signalling factors or genes that directly protect plant cells against water deficit (Valliyodan and Nguyen, 2006).

Evidences have been reviewed on the molecular mechanisms of abiotic stress responses and the genetic regulatory networks of drought stress tolerance in the immediate past (Figure C) (Wang *et al*., 2003; Vinocur and Altman, 2005; Chaves and Oliveira, 2004; Shinozaki *et al*., 2003; Yamaguchi- Shinozaki and Shinozaki, 2005). In all approaches, prime focus has been laid on the recent progress in understanding the mechanisms of gene regulation and the roles of protective metabolites in drought stress tolerance, and the progress in genetic or metabolic engineering for enhanced drought tolerance in crop plants. The genetic manipulation approaches have been categorised into two major groups: first, engineering cell signalling and gene regulation, and second, engineering for osmoprotectant accumulation (Valliyodan and Nguyen, 2006).

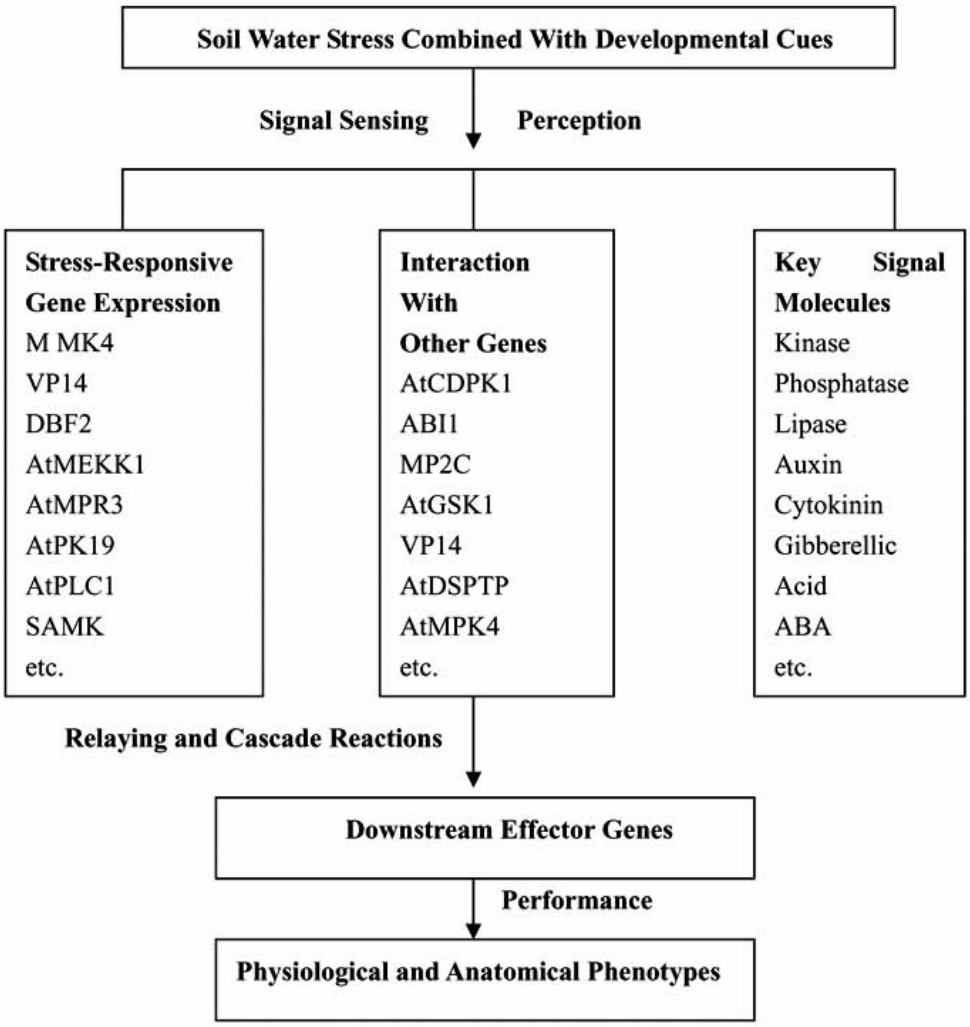


Figure C: Basic plant gene regulatory network system (Beck *et al*; 2007)**.**

In short, all plant genetic manipulative mechanisms utilises the functional relevance of

either of the above mentioned genes which will be targeted for transgenic approach by up regulation or down regulation events. The findings published as a research outcome establish that a cytosolic ascorbate peroxidase APX1 as a key component of the ROS gene network in Arabidopsis. Using a dominant-negative approach, it has been demonstrated that HSF21, a redox-sensitive transcription factor, regulates the expression of ZAT12, a multiple-stress-responsive transcription factor. The data also suggest that HSF21 is probably an initial sensor for H_2O_2 that is accumulated in response to various stresses (Davletova *et al*., 2005). Studies on stress regulated *HARDY* gene, showed an improvement of water use efficiency in rice (Karaba *et al*., 2007). It has been elaborately proven that Mitogen-Activated Protein Kinases 3 and 6 are required for full priming of stress responses in *Arabidopsis thaliana* (Beckers *et al*., 2009).

2.5 Strategies for gene discovery

In order to develop crop plants with enhanced tolerance to the much devastating drought stress, it is very vital to inculcate a basic understanding of physiological, biochemical and gene regulatory networks. A wide array of functional genomics tools have helped to advance our understanding of stress signal perception and transduction, and of the associated molecular regulatory network (Bray, 2004; Shinozaki and Yamaguchi-Shinozaki, 2007; Shao *et al*., 2008). These tools have helped in the successful revelation of several stress-inducible genes and various transcription factors that regulate the drought-stress-inducible systems. Translational genomics of these candidate genes using model plants provided encouraging results, but there is nominal progress in the direction of field testing of transgenic crop plants for better performance and yield. Better understanding of the specific roles of various metabolites in crop stress tolerance will give rise to a strategy for the metabolic engineering of crop tolerance of drought (Valliyodan and Nguyen, 2006).

Large-scale genome sequencing projects have been partially employed for identification of important genes in certain plant types like *Arabidopsis*, rice, maize, *Populus* sp., etc (Temsch and Greilhuber, 2000). When the genome sequence is not available and genome size is significantly large, alternative approaches have been followed. For example, groundnut genome size is significantly larger (2,800 Mb/1C) that some of the sequenced plants (Temsch and Greilhuber, 2000), such as *Arabidopsis* (128 Mb), rice (420 Mb), and

Medicago (500 Mb) (Sasaki, 2003; Bennett and Leitch, 2005). Under such conditions, comparative genomics is being attempted, which enables the application of data from one species to investigations of taxonomically disparate species, as has been achieved for example in the annotation of wheat sequences using information from *E. coli*, human, *A. thaliana* and rice (Heslop- Harrison, 2000). The BLAST result of groundnut Expresssed Sequence Tags (ESTs) revealed that significant number of unique groundnut seed ESTs match soybean (396), *Arabidopsis* (2952), rice (682), and other plant species (Guo *et al*., 2008).

Isolation of gene transcripts from drought tolerant crop species can be a good approach for gene discovery, which can be achieved by different techniques. Identification of the differentially expressed genes can be successfully done by different approaches like subtractive hybridization, suppressive subtractive hybridization (Diatchenko *et al*., 1996), differential display (Liang and Pardee, 1992; Cho *et al*., 2001), cDNA-AFLP (Kivioja *et al.,* 2005), microarray technology and other means (Seki *et al*., 2002; Shinozaki and Yamaguchi-Shinozaki, 2007). Similarly, Targeting-Induced Local Lesions IN Genomes (TILLING), a reverse-genetic strategy for the discovery and mapping of induced mutations can also be attempted. Once genomic DNA samples have been obtained, pooled and arrayed, thousands of samples can be screened in a short period (Till *et al.,* 2006).

Generation and characterization of stress specific ESTs can be a viable approach when genome information is limited. Such genomic resources are considered as important tools for plant functional genomics, which forms the basis to combine relevant traits of interest into a desired genetic background. Expresssed sequence tags and gain-in-function mutants are powerful tools in this regard. The ESTs represent a huge resource for the discovery of previously unknown genetic information and functional genome assignment. In plants, cDNA libraries are widely being used as the choice material for transcriptome profiling (Seki *et al*., 2002). Expression profiling of stress specific ESTs and full-length cDNA clones has lead to the identification of key regulators in plant signal transduction cascades and there has been a spurt in EST databases. A number of ESTs have been deposited, although the number of drought or abiotic stress specific ESTs are few.

EST projects in commercially important plant species have been initiated in several crops (Michalek *et al*., 2001; Miller *et al*., 2001; Fedorova *et al*., 2002; Fernandes *et al*., 2002; Shoemaker *et al*., 2002; Hoeven *et al*., 2002; Kikuchi *et al*., 2003; Ogihara *et al*., 2003; Ronning *et al*., 2003; Vettore *et al*., 2003; Ramı´rez *et al*., 2005). Currently, EST information is available for crops belonging to family Poaceae (Adams *et al*., 1991) including rice (Reddy *et al*., 2002; Rabbani *et al*., 2003) in addition to *Arabidopsis* (Seki *et al*., 1998, 2002), maize (Brendal *et al*., 2002), sugarcane (Abe *et al*.,1997; Narasuka *et al*., 2004; Tran *et al*.,2004), sorghum (Pratt *et al*., 2005), cotton (Joshua *et al*., 2006) and pearl millet (*Pennisetum glaucum*) (Mishra *et al*., 2007). On a broader angle, recently cDNA library construction and EST generation has been carried out across different genera for varied interests. A cDNA library of *Lathyrus sativus* inoculated with *Mycosphaerella pinodes* helped in the expression analysis of potential defence-related expressed sequence tags (ESTs) (Skiba *et al.,* 2005). Generation of a full-length cDNA Library from Chinese oak silkworm pupa aided in the identification of a KK-42-binding protein gene in relation to pupa-diapause termination (Li *et al.,* 2009). cDNA library of a green microalga *Ankistrodesmus convolutus* Corda was done for preliminary analysis of expressed sequence tags generated from it (Thanh *et al*., 2010). The expression sequence tag (EST) resource was useful in studying newt dedifferentiation and finally concluding that the expression of cancer- and apoptosis-related genes could be hallmarks during dedifferentiation (Maki *et al*., 2010). Recently, an EST database has been created for a commercially important bio diesel crop, *Jatropha curcas* (Easwaran *et* al., 2010).

Recent progress in molecular biology and bioinformatics (especially, DNA microarray technology), genomics, proteomics, metabolomics and transcriptomics) has provided insight into plant gene regulatory network system, which is mainly composed of inducible-genes (environmental factors and developmental cues), their expression programming and regulatory elements (cis-element and trans-element), corresponding biochemical pathways and diverse signal factors (Shinozaki *et al*; 2003; Shinozaki and Yamaguchi- Shinozaki, 2007). The genetic information for drought tolerance is expressed in many prokaryotes and lower eukaryotes, but only in very few higher plants. In higher plants, only seeds can survive for extended periods without water. Exceptional among higher plants is the small group of angiosperm plants termed 'resurrection plants' which can recover from complete dryness within one day of contact with water (Bartels *et al*; 1990; Bartels and Salamini, 2001; Bartels, 2005). Under the condition of soil water

deficits, related stress factors always result in overlapping responses, including anatomical, physiological, biochemical, molecular biological changes, which make plant gene regulatory network system more complicated and difficult to explore. A collection of genes involved in abiotic stress response belonging to different categories – transcriptional elements, potential genes mediating resistance soil water stress and osmoregulatory genes has been tabulated in tables 1, 2 and 3 respectively (Ni *et al.,* 2009).

2.6 Candidate gene analysis

Candidate gene analysis initiates with selection of some target genes based on biological pathway or genome location relative to a known QTL identified for the target trait (Byrne and McMullen 1996; Rothschild and Soller, 1997). A drought-related EST database, accurate microarray analysis and mutagenesis approaches for gene evaluation will yield a large number of valuable candidates for verification of their association with the drought tolerance traits. As an alternative approach, searches can be conducted for orthologs in existing literature and databases for information on drought and related abiotic stress genes. With the integration of QTL mapping, comparative mapping information, growing EST databases, expression (including microarray) results and the identification of numerous genes in the future, the candidate gene approach will emerge as an important and powerful tool to uncover the ambiguity behind the expression of quantitative traits (Nguyen) (Figure D).

Table 1: Typical transcriptional elements related to abiotic stresses in plants and crops (Ni *et al*., 2009)

Plant materials	Factors	Binding sites/ Factor types
Arabidopsis thaliana	ABI5/ AtDPBF	ABA response elements (ABREs)/ bZIP
A. thaliana	AtDPBF2	ABA response elements (ABREs)/ bZIP
A. thaliana	AtDPBF3/AREB3	ABA response elements (ABREs)/ bZIP
A. thaliana	AtDPBF4	ABA response elements (ABREs)/ bZIP
A. thaliana	AtDPBF5/ABF3	ABA response elements (ABREs)/ bZIP
A. thaliana	ABF1	ABA response elements (ABREs)/ bZIP
A. thaliana	ABF2/ AREB5	ABA response elements (ABREs)/ bZIP
A. thaliana	ABF4/ AREB2	ABA response elements (ABREs)/ bZIP
A. thaliana	GBF3	ABA response elements (ABREs)/ bZIP
A. thaliana	AB53	RY/sph elements/ B3 domain proteins
A. thaliana	ATMTB2	MTC
A. thaliana	ATHB6	HD- Zip
A. thaliana	ATHB7	HD- Zip
A. thaliana	ATHB12	HD- Zip
A. thaliana	ABI4	AP2
Oryza	TRAB1	ABA response elements (ABREs)/ bZIP
Oryza	OsVPI	RY/sph elements/ B3 domain proteins
Zea mays	VPI	MYB
Triticum	EmBP-1	ABA response elements (ABREs)/ bZIP
Avena	AtVPI	RY/sph elements/ B3 domain proteins
Helianthus	DPBF5,-2,-3	ABA response elements (ABREs)/ bZIP
Phaseolus	ROM2 (repressor)	ABA response elements (ABREs)/ bZIP
Phaseolus	PIARF	RY/sph elements/ B3 domain proteins
Craterestimma	Cpvp1	RY/sph elements/ B3 domain proteins
Daucus	C-ABI3	RY/sph elements/ B3 domain proteins
Populus	PtABI3	RY/sph elements/ B3 domain proteins

Table 2: Potential genes mediating resistance to soil water stress and related abiotic stresses in plants
(Ni *et al*., 2009; www.plantstress.com)

Gene	Gene Action	Species	References
adc	Polyamine synthesis	Rice	Ni *et al*., 2009; www.plantstress.com
Apo- lnv	Apoplastic invertase	Tobacco	
AtGoIS2	Galactinol and raffinose accumulation	Arabidopsis	
AtHAL3	Phosphoprotein phosphatase	Tobacco	
AtHAL3a	Phosphoprotein phosphatase	Arabidopsis	
ATP-PRT	ATP- Phosphoribosyl transferase	Alyssum	
AtTPS1	Trehalose- 6- phosphate synthase	Tobacco	
BADH1	Betaine aldehyde dehydrogenase	Carrot	
BADH1	Betaine aldehyde dehydrogenase	Rice	
BADH1	Betaine aldehyde dehydrogenase	Maize	
BADH1	Betaine aldehyde dehydrogenase	Tobacco	
BADH1	Betaine aldehyde dehydrogenase	Tobacco	
BADH1	Betaine aldehyde dehydrogenase	Tomato	
betA	Choline dehydrogenase (Glycine betaine synthesis)	Maize	
betA	Choline dehydrogenase (Glycine betaine synthesis)	Tobacco	
CHIT33, CHIT42	Endochitnase synthesis	Tobacco	
CMO	Choline monooxygenase (Glycine betaine synthesis)	Tobacco	
codA	Choline oxidase (Glycine betaine synthesis)	Arabidopsis	
codA	Choline oxidase (Glycine betaine synthesis)	Arabidopsis	
codA	Choline oxidase (Glycine betaine synthesis)	Arabidopsis	
codA	Choline oxidase (Glycine betaine synthesis)	*Brassica juncea*	
codA	Choline oxidase (Glycine betaine synthesis)	Rice	
codA	Choline oxidase (Glycine betaine	Rice	

	synthesis)		
codA	Choline oxidase (Glycine betaine synthesis)	Tobacco	
codA	Choline oxidase (Glycine betaine synthesis)	Tomato	
codA	Choline oxidase (Glycine betaine synthesis)	Tomato	
COR15a	Cold induced gene	Arabidopsis	
COX	Choline oxidase (Glycine betaine synthesis)	Rice	
EctA.ectC	Edtoin accumulation in chloroplasts	Tobacco	
GS2	Chloroplastic glutamine synthetase	Rice	
IMT1	Myo- inositol o- methyl transferase (D-ononitol synthesis)	Tobacco	
LWR1, LWR2	Solute accumulation (proline)	Arabidopsis	
M6PR	Mannose 6- phosphate reductase	Arabidopsis	
M6PR	Mannose 6- phosphate reductase	Arabidopsis	
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	Arabidopsis	
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	Petunia	Ni *et al*., 2009; www.plantstress.com
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	Tobacco	
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	Tobacco	
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	Wheat	
mt1D & GutD	Mannitol 1- phosphate dehydrogenase & glucitol 6- phosphate dehydrogenase	Lobolly pine	
mt1D	Mannitol 1- phosphate dehydrogenase (mannitol synthesis)	*Populus tomentosa*	
Osm1 -Osm4	Osmotin protein accumulation	Tobacco	
Osmyb4	Cold induced transcription factor	Arabidopsis	
Osmyb4	Specifically cold inducible	Tobacco	
Osmyb4	Cold induced transcription factor	Tomato	
OsP5CS2	Highly homologous to P5CS	Rice	

otsA	Trehalose 6 –phosphate synthase	Tobacco	Ni *et al*., 2009; www.plantstress.com
otsB	Trehalose 6 –phosphate synthase	Tobacco	
P5CR	Pyrroline carboxylate reductase (proline accumulation)	Soybean	
P5CR	Pyrroline carboxylate reductase (proline accumulation)	Soybean	
P5CS	Pyrroline carboxylate reductase (proline accumulation) (tomato)	Citrus	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Petunia	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Potato	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Rice	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Rice	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Rice	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Soybean	
P5CS	Pyrroline carboxylate reductase (proline accumulation) (tomato)	Soybean	
P5CS	Pyrroline carboxylate reductase (proline accumulation) (tomato)	Sugarcane	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Tobacco	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Tobacco	
P5CS	Pyrroline carboxylate reductase (proline accumulation)	Wheat	
P5CS	Pyrroline carboxylate reductase (proline accumulation) (tomato)	Yeast	
pdc1	Pyruvate decarboxylase over expression	Rice	
pdc1, pdc2	Pyruvate decarboxylase over expression	Arabidopsis	
PPO	Polyphenyl oxidases suppression	Tomato	
SAMDC	S- adenosyl methionine decarboxylase (polyamine synthesis)	Rice	

SAMDC	S- adenosyl methionine decarboxylase (polyamine synthesis)	Tobacco	Ni *et al*., 2009; www.plantstress.com
SMT	Selenocysteine methyl transferase	Arabidopsis, Indian mustard	
SPE	Spermidine synthase	Arabidopsis	
spe1-1, spe2-1	Spermidine non accumulating	Arabidopsis	
SST/FFT	Fructan accumulation	Potato	
TaCRT	Ca binding protein	Tobacco	
TPP1	Trehalose synthesis	Rice	
TPS, TPP	Trehalose synthesis	Arabidopsis	
TPS1	Trehalose synthesis	Tomato	
TPS1	Trehalose synthesis	Potato	
TPS1 & TPS2	Trehalose synthesis	Tobacco	
TPSP	Trehalose synthesis	Rice	
WCOR15	Cold induced gene	Tobacco	
WCOR15	Dehydrin	Strawberry	

Table 3: Some examples of osmoregulatory genes associated with abiotic stress tolerance

(Ni *et al*., 2009; www.plantstress.com)

Components	Metabolic functions	Genes / Proteins	References
ROS scavenging	Increases ROS scavenging enzymes	GP, PGPX	Ni *et al*., 2009; www.plantstress.com
Chaperones	Heat/ cold/ salt shock proteins; protein folding	Hsp, Csp, Ssp, *Dna J*	
Fructan	Osmoprotection	*SacB*	
Trehalose	Osmoprotection	*tps,tpp,*trehalase	
Glycine betaine	Protein protection and carbon sink	*codA*	
Proline	Substrate for mitochondrial respiration, redox control	P5CS/ P5CR	
Ectoine	Osmoprotection	*EctA,B,C*	
K transporters	High affinity K uptake	Hkt1, Hak1	
K channels	Low affinity or dual affinity k uptake	Akt1, Akt	
Water channel proteins	Membrane cycling control	TIP	

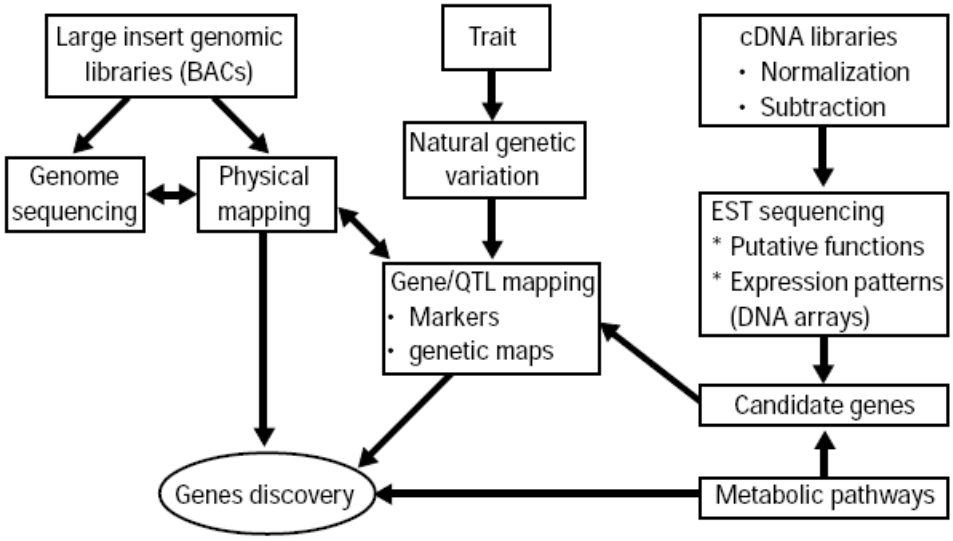


Figure D. Model showing events associated with gene discovery using specific traits (Nguyen)**.**

2.7 Benefits of gene discovery – up regulation and down regulation studies

The development and application of modern molecular biology have led to a better understanding of plant adaptations and responses to abiotic stress conditions. Genes responsible for adaptation processes to every example of given stress have been identified in *Arabidopsis thaliana* and other important model plants (Bressan *et al.,* 2009). The use of transgenic plants to over express or silence these genes is a powerful tool in determining if they are necessary or sufficient to induce stress tolerance. However, with stress tolerance being a quantitative character, it should not be expected that single genes would confer a high level of corresponding tolerance. Hence recent approaches aim at conferring drought stress tolerance through gene stacking strategies.

A recent review has laid adequate focus on the recent progress in genetic manipulation of important crops for the enhancement of drought tolerance (Yang *et al*., 2010). Tolerance to a given stress such as drought cannot be evaluated in isolation lest it be considered and reviewed in association with other major stresses such as high temperature, salt, and cold as a lot of cross talk exists among multiple stress situations (Chinnusamy *et al.,* 2004; Shinozaki and Yamaguchi- Shinozaki, 2007). A good number of genes that play a role in

drought tolerance also confer some degree of tolerance to salt and cold stresses. However, drought cum heat tolerant crops will be the major targets of attention in the event of the drastic global climate change (Battisti and Naylor, 2009). The challenge in the near future will be to identify the signalling elements missing in the current models of pathways and increasing our understanding of the cross-talk between pathways as our knowledge about stress signalling networks is still remains grey (Yang *et al*., 2010).

Hence, it can be concluded that under varied levels of stresses experienced by plants, there are different cellular as well as molecular mechanisms operating in order to enable the system to cope up with the adversities. These mechanisms enact their respective roles by conferring intrinsic cellular tolerance. The adapted crops have been successful enough in incorporating certain novel biosynthetic pathways by way of post translational modifications and so on. In this viewpoint, prospecting for drought tolerant species presumes significance.

MATERIALS AND METHODS

III. MATERIALS AND METHODS

In the global scenario, different cropping situations fall victim to a complex of stresses. Among the various abiotic stresses, drought cripples plant productivity the most. Genes induced during stress conditions produce various metabolic proteins (functional proteins) and other proteins involved in the gene regulation (regulatory proteins). The generation and characterization of stress specific ESTs can be a viable approach when genome information is limited as in the case of crops like finger millet (*Eleusine coracana* (L.) Gaertn.). With this outlook, an existing drought stress specific cDNA library was used as a rich resource to discover drought stress responsive genes which could be further used as candidate genes for improving drought susceptible species. The following objectives were framed in order to meet the broad target of discovering drought responsive genes from finger millet:

4. Screening and evaluation of finger millet stress cDNA library
5. Identification of novel genes by sequencing and annotation using bioinformatics tools
6. Expression analysis of selected genes by
    (i) e-northern using database information
    (ii) examining transcript levels in the tissues exposed to drought stress

Various experiments were designed to lead to the fulfilment of the major objectives. The broad details and protocols followed to carry out these experiments are explained elaborately in this chapter.

Location and plant species used in the study

All experiments were conducted in green house facility of the Department of Crop Physiology, University of Agricultural Sciences, GKVK campus, Bangalore, Karnataka. The site is situated at $12^0 58'$ North latitude, $77^0 35'$ East longitudes and an altitude of 930m above Mean Sea Level (MSL). The plant species used for the study was finger millet (*Eleusine coracana* (L.) Gaertn.).

3.1 Screening of drought stress specific cDNA library developed from finger millet leaf tissue

Construction of drought stress specific cDNA library

The stress cDNA library used in the present study was developed in the Department of Crop Physiology, UAS, GKVK, Bangalore. For the construction of cDNA library, finger millet (variety GPU-28) was used. Twenty day old seedlings were subjected to drought stress by maintaining different levels of field capacities viz., 100%, 80%, 60% and 30%, through gravimetric approach. Total RNA was isolated from the leaf tissues collected from plants maintained at different stress levels. The different RNA isolates were then pooled to have representative sample of transcripts of all the stress levels including moderate and severe stress. Messenger RNA (mRNA) was isolated from total RNA. Unidirectional double stranded cDNA was synthesised from mRNA by using ZAP-cDNA synthesis kit (StrataScript, Stratagene, CA, USA) as per protocols of the supplier. The resulting double stranded cDNA was size fractionated and the cDNA inserts were ligated into Uni-ZAP XR vector. The recombinant lambda phages were packaged with Gigapack III Gold packaging extract (ZAP-cDNA Gigapack III Gold Cloning Kit, StrataScript, Stratagene, CA, USA).

Screening of drought stress specific cDNA library

The screening of the library was done using different evaluation protocols. The overall screening procedure included five main steps:

1. Packaging and lifting of phage plaques each representing a recombinant clone
2. Analysis of plaques by polymerase chain reaction (PCR) with M13 forward and reverse primers for identification of recombinant clones
3. Selection of clones with insert DNA of size 500 bp or more
4. Purification of the PCR products from the selected clones and sequencing
5. Processing of DNA sequences and annotation

3.1.1 Packaging and plaque lifting

3.1.1.1 Testing packaging efficiency

Before proceeding with screening protocols, the packaging efficiency of the kit as well as the titer of the reaction was assessed. The packaging efficiency of the existing library was tested using XL1-Blue MRF′ cells. XL1-Blue MRF′ cells were streaked onto LB agar plates containing tetracycline (30 µg/ml) and incubated overnight at 37°C. Fifty millilitre

cultures of XL1-Blue MRF′ cells were prepared in LB broth with supplements (10 mM $MgSO_4$ and 0.2% (w/v) maltose). The culture was incubated at 37°C overnight while shaking at 160 rpm. The cells were pelleted at 5000 rpm for 5 min, resuspended in 25 ml sterile 10 mM $MgSO_4$, and diluted to an OD_{600} of 0.5 with sterile 10 mM $MgSO_4$. To determine the titer of the ligation product, 1 µl of the final packaged reaction and 200 µl of XL1-Blue MRF′ cells were mixed and the phage - bacteria mixture was incubated at 37°C for 15 min to allow the phage to attach to the cells. Three millilitre of NZY top agar (NZY broth: 5 g NaCl; 2 g $MgSO_4.7H_2O$; 5 g yeast extract; 10 g NZ amine (casein hydrolysate) + 0.4% (w/v) agarose), melted and cooled to ~48°C, was added and plated immediately onto dry, prewarmed NZY agar plates [5 g NaCl; 2 g $MgSO_4.7H_2O$; 5 g yeast extract; 10 g NZ amine (casein hydrolysate); 15 g of agar; pH 7.5]. Plates were allowed to set for 10 min, inverted and incubated at 37°C to allow phage to grow for 6-8 h or overnight.

3.1.1.2 Screening for recombinant phages

Massive screening procedures were adopted to elute out as many recombinant clones as possible. Two ml cultures of XL1-Blue MRF′ cells was prepared in NZY broth (5 g NaCl; 2 g $MgSO_4.7H_2O$; 5 g yeast extract; 10 g NZ amine (casein hydrolysate)) with supplements (10 mM $MgSO_4$ and 0.2% (w/v) maltose). The culture was incubated at 37°C for 6 -8 h while shaking at 160 - 200 rpm. The cells were pelleted at 5000 rpm for 5 min, resuspended in 500µl sterile 10 mM $MgSO_4$, and diluted to an OD_{600} of 0.5 with sterile 10 mM $MgSO_4$. About 1 µl of the final packaged reaction and 200 µl of XL1-Blue MRF′ cells were mixed and the phage - bacteria mixture was incubated at 37°C for 15 min to allow the phage to attach to the cells. The pre incubated phage – bacteria mixture was added to eight millilitre of NZY top agar (as mentioned earlier in 3.1.1.1) melted and cooled to ~48°C and plated immediately onto dry, pre warmed NZY agar plates [as mentioned earlier in 3.1.1.1]. Plates were allowed to set for 10 min, inverted and incubated at 37°C to allow phage to grow overnight. Phage growth was observed as transparent plaques the next day on the NZY agar plates. The plaques were then lifted using sterile Whatman filter paper strips and transferred to a sterile microcentrifuge tube containing 500 µl of SM buffer [5.8 g NaCl; 2 g $MgSO_4.7H_2O$; 50 ml of 1 M Tris-HCl (pH 7.5)]. The microcentrifuge tubes were incubated at 37°C for 15 – 20 min to release the phage particles into the SM buffer. An aliquot of 3 - 5 µl was used for PCR amplification to confirm the presence of insert.

3.1.2 Polymerase chain reaction (PCR) to identify recombinant plaques

Identification of the recombinant clones was done by PCR using universal M13 forward and reverse primers. The PCR components and reaction conditions are given in Appendix IG. The PCR amplified products were electrophoresed on agarose (1%) gel (Appendix IE) to visualise the band sizes. The gel images were documented using gel documentation unit (Herolab 440K, GmbH Laborgerate, Germany). All amplicons of length 500 bp or more were selected for further purification for sequencing accuracy. Purification of the PCR products of selected recombinant clones was done and electrophoresed (Appendix IE) to confirm purity prior to sequencing. The purified eluates were sequenced using M13 primers (Courtesy: Macrogen Inc., Korea).

Preservation of plaques

The plaque eluates in SM buffer were stored at 4°C. For long term storage at -70°C, Dimethyl Sulfoxide (DMSO 7%, v/v) was added and stored in aliquots.

3.1.3 DNA sequence analysis and annotation

Based on plaque PCR data, clones with insert sizes equal to or above 500 bp were selected, purified and subjected for sequencing to identify the genes getting expressed under stress. The sequence information obtained was first screened for vector contamination using VecScreen tool at the NCBI database (www.ncbi.nlm.nih.gov/blastx). The vector backbone clipped sequences were annotated using translated query-protein database (blastx) at the NCBI database (www.ncbi.nlm.nih.gov/blastx). The cut off value and the minimum sequence length considered for comparison were 1.5- 3.0 and 300 bp respectively. The resulting Expressed Sequence Tags (ESTs) were further analysed for their functional and regulatory significance and classified based on their operative relevance. On the basis of the stimulating functional significance of certain genes, a few novel genes were selected for further studies.

3.1.4 Excision of the pBluescript phagemid from the Uni-ZAP XR vector

The phage stocks were subjected to excision protocols in order to excise out the phagemids from Uni-ZAP XR vector and transfer to bacterial system for ease in handling. The Uni-ZAP XR vector system efficiently mediates the execution of *in vivo* excision and recircularization of any cloned insert contained within the lambda vector to produce a

phagemid containing the cloned insert. The placement of the DNA sequences within the lambda phage genome and the presence of a variety of proteins, including f1 bacteriophage-derived proteins ultimately determines excision efficiency. The f1 phage proteins recognize the f1 bacteriophage origin of replication on the DNA. This origin of replication containing two overlying parts – the sites of initiation and termination were subcloned separately into the Uni-ZAP XR vector system. Simultaneous infection of a strain of *E.coli* with both the lambda vector and the f1 bacteriophage facilitated the accessibility of the lambda phage (target) to the f1-derived proteins. The helper proteins (i.e., proteins from f1 or M13 phage) of *Escherichia coli* recognize the initiator DNA within the lambda vector and then nick one of the two DNA strands. At the nick, new DNA synthesis begins and duplicates whatever DNA exists in the lambda vector downstream (3′) of the nicking site. Synthesis of new single-strand of DNA continues through the cloned insert until a termination signal, positioned 3′ of the initiator signal, is encountered within the constructed lambda vector. The single-stranded DNA molecule is circularized by the gene II product from the f1 phage, forming a circular DNA molecule containing the DNA between the initiator and terminator. In case of the Uni-ZAP XR vector, this includes all sequences of the pBluescript SK(–) phagemid and the insert. This conversion is the sub-cloning step, since all sequences associated with normal lambda vectors are positioned outside of the initiator and terminator signals and are not contained within the circularized DNA. In addition, the circularization of the DNA automatically recreates a functional f1 origin as found in f1 bacteriophage or phagemids. The ExAssist helper phage with SOLR strain is designed to allow efficient excision of the pBluescript phagemid from the Uni-ZAP XR vector, while eliminating problems associated with helper phage co-infection. The ExAssist helper phage contains an amber mutation that prevents replication of the phage genome in a non-suppressing *E. coli* strain such as SOLR cells. This allows only the excised phagemid to replicate in the host, removing the possibility of co-infection from the ExAssist helper phage. SOLR cells are also resistant to lambda phage infection, thus preventing lambda phage contamination after excision.

Separate 3 ml cultures of XL1-Blue MRF′ and SOLR cells were grown in LB broth with supplements (10 mM $MgSO_4$ and 0.2% (w/v) maltose) at 37°C overnight at 165 rpm. Cells were pelleted at 5000 rpm for 5 min and the pellets were resuspended in 500µl of sterile 10 mM $MgSO_4$. The OD_{600} was measured and then the concentration of the cells was adjusted to an OD_{600} of 1 in 10 mM $MgSO_4$. About 200 µl of XL1-Blue MRF´ cells,

100 µl of phage stock (containing insert) and 1 µl of the ExAssist helper phage were combined and incubated at 37°C for 15 min to allow the phage to attach to the cells. Three millilitre of LB broth with supplements was added and incubated for overnight at 37°C with shaking. The tubes were heated at 65-70°C for 20 min to lyse the lambda phage particles and the cells. Tubes were centrifuged at 5000 rpm for 15 min to pellet the cell debris. The supernatant contains the excised pBluescript phagemid packaged as filamentous phage particles. 100 µl of the supernatent containing phagemid particles was mixed with 200 µl of freshly grown SOLR cells and incubated at 37°C for 15 min. About 50 µl of the cell mixture was standardised to be plated on LB-ampicillin agar plates (100 µg/ml) with X- gal (20 mg/ ml) mediated blue – white selection and incubated overnight at 37°C. Colonies appearing on plate contained pBluescript double-stranded phagemid (Appendix III) with cloned DNA insert. Another level of selection pressure was imposed by patching the colonies into LB-ampicillin agar plates (100µg/ml) with X- gal (20 mg/ ml) mediated blue – white selection and incubated overnight at 37°C. The recombinant clones alone would show growth as white colonies.

3.1.5 Confirmation of excised phagemids by colony PCR

The recombinant colonies were analyzed by PCR with universal M13 forward and reverse primers to confirm the excision of the pBluescript phagemid from the Uni-ZAP XR vector (Appendix III A). The PCR components and reaction conditions are given in Appendix IG. The PCR amplified products were electrophoresed on agarose (1%) gel (Appendix IE) to visualise the positive colonies. The gel images were documented using gel documentation system (Herolab 440K, GmbH Laborgerate, Germany).

3.1.6 Isolation of plasmids from recombinant colonies and confirmation

Plasmid isolation was done by following the Alkaline Lysis protocol which is explained in detail in Appendix IC (Sambrook *et al.,* 1989). Plasmid DNA was isolated from the recombinant positive colonies and confirmed by PCR using M13 primers. The PCR components and reaction conditions are given in Appendix IG. The amplified plasmid PCR products were electrophoresed on agarose (1%) gel (Appendix IE) and the gel images were documented using gel documentation system (Herolab 440K, GmbH Laborgerate, Germany). However for sequencing purpose, to ensure purity of isolated plasmids, plasmid DNA was isolated from recombinant colonies by using plasmid DNA isolation kit (Enzene Biosciences, Bangalore, India). The isolated plasmids were

sequenced to reconfirm the identity of the gene (Courtesy: Enzene Biosciences, Bangalore, India).

The sequence information was used to align with the existing sequences of the respective clones using ClustalW2 software (http://www.ebi.ac.uk/Tools/ clustalw2/index.html) to confirm the clone identity.

3.2 Gene expression studies

3.2.1 Expression analysis of selected genes by eNorthern

Electronic Northern (eNorthern) is a useful way to analyse how genes of interest are being expressed, using the gene expression data sets accumulated from publicly available micro array data. The ESTs developed from finger millet were analysed by translated BLAST to identify their Arabidopsis homologues. The Arabidopsis Genome Initiative (AGI) codes of the *Arabidopsis* homologues were used for eNorthern analysis. eNorthern analysis were conducted at the eNorthern/ Expression Browser tool of the Botany Array Resource (BAR) at http://bbc.botany.utoronto.ca/. The expression levels of genes in shoot and root of the *Arabidopsis* homologues under different abiotic stresses like osmotic, drought, cold, salinity, radiation, genotoxic, oxidative, wounding and heat stress at 0.5, 1, 3, 6, 12, 24 hours of stress duration were analysed. The log2 values obtained were transformed to actual fluorescence values by the formula 2^log value. The expression levels were indicated as whole number scores after eNorthern analysis. A score of 3.0 and above indicates that the genes are highly expressive. The fluorescence value obtained for each stress was plotted separately on the graph as fold increase over control against the different time periods. The graphs were studied for the expression pattern. Genes with consistently high expression levels were selected for further studies.

3.2.2 Expression studies using finger millet tissues exposed to abiotic stresses

Based on putative function and eNorthern results, a few clones coding for upstream/downstream genes were analyzed for their stress inducibility by semi quantitative reverse transcription PCR (RT-PCR) and quantitative real time RT-PCR (qPCR). The expression analyses were carried out on plants experiencing different levels of drought stress.

3.2.2.1 Imposition of drought stress at whole plant level

Healthy plants of finger millet (variety GPU-28) were raised, in battery pots containing 18- 20 kg of 2:1 garden soil and sand mixture, by adopting all timely plant production and protection measures. The water status of the pots was maintained at 100% field capacity (FC) by controlled irrigation. Water stress was imposed to 20 days old plants by withholding irrigation. Soil FCs of 80, 60, 50, 40, 30 and 20% were maintained by gravimetric approaches for stress imposition. During stress imposition, soil field capacity was monitored gravimetrically by checking the amount of water present in soil medium compared to maximum water it can hold (Sheshshayee *et al*., 2005).

To assess the extent of drought, soil water potential was monitored using dewpoint potentiometer (Govind *et al*., 2008). The measurement of soil water potential gives an idea of the water status of the soil. The soil sampling was done at a depth of 8 cm from the rhizospheres of plants subjected to different stress levels. The soil water potentials of the different samples were recorded using the WP4 Dewpoint Potentiameter (Decagon Devices Inc., USA). The soil samples were loaded onto the sample cup of the equipment and the sample drawer knob was turned to the read position to seal the sample cup with the chamber. The soil water potential measurement was given at the particular temperature after the reading cycle was over and the equipment gave a beep sound followed by flash of a green light. The experiment was performed in replicates of three. For comparison, control plants were maintained at 100% FC by regular watering.

After 10 days of stress imposition, the stressed as well as control leaf and root tissues were harvested, frozen in liquid nitrogen and stored at -70°C. The plant tissues collected were used to extract total RNA.

3.2.2.2 Characterisation of drought stressed plants by physiological analysis

Stressed and control plants were analysed by various physiological assays to assess the extent of stress induced effects. The following assays were done to assess the stress induced changes in leaf tissues.

The various assays undertaken were

(a) Relative water content (RWC)
(b) Membrane integrity

(c) Leaf solute potential and net osmotic adjustment

(d) Reduction in total chlorophyll content in stressed plants

(e) Cell viability

(a) Assessment of relative water content

Relative water content (RWC) is defined as the water content of leaf in relation to the water that the leaf can hold at saturation.

$RWC = \{(W_0)/(W_1)\} \times 100$

where, W_0 = Water content of the leaf , W_1 = Water the leaf can hold at saturation

The RWC of the drought stressed leaf samples was assessed using the protocol as described by Barrs and Weatherly, 1986; Mamrutha *et al*., 2010). The leaf samples collected were made into leaf discs and fresh tissue weight was recorded. The leaf discs were floated in water (10 ml) for 8- 10 hours to gain full turgidity of cells. Fully turgid condition was assessed based on the fact that two consecutive measurements recorded the same weight. The turgid weights of the leaf discs were recorded and the samples were then oven dried at 65- 70 °C overnight in the hot air oven. The dry weight of the sample was recorded and used for the estimation of relative water content. The experiment was performed in replicates of three. The RWC was calculated using the formula given belo

$RWC = \{(FW - DW) / (TW - DW)\} \times 100$

where, FW = Fresh weight , TW = Turgid weight , DW = Dry weight

(b) Assessment of loss of membrane integrity

Assessment of membrane integrity gives an indication of effect of stress experienced by the plant or tissue. A common method to assess the extent of loss of membrane integrity is to quantify the electrical conductivity which gives a quantification of total electrolyte leakage under stress (Govind *et al*., 2008). The electrical conductivity can be measured by using conductivity meters. For measuring electrical conductivity, leaf discs (10) were collected from plants and incubated in water (25 ml) overnight. The initial electrical conductivity was recorded using Conductivity – TDS Meter 307 (Systronix, India) at 25 °C. The leaf discs were then boiled for 30 min at 100 °C and the final leakage was recorded. The experiment was performed in replicates of three. The total electrolyte leakage was calculated and expressed in percent using the following formula.

Electrolyte leakage (%) = $\{(L_0) / (L_1)\}$ X 100
where, L_0 and L_1 are initial and final leakage readings respectively

(c) Measurement of leaf solute potential and net osmotic adjustment

Leaf samples from plants maintained at different field capacities were collected, wrapped in aluminium foil and frozen in liquid nitrogen. The samples were then thawed and centrifuged for 5 min at 12,000 rpm. The leaf solute potential (ψs) of the extracted sap was measured by VAPRO vapor pressure osmometer (Wescor Inc, Logan, USA). The experiment was performed in replicates of three. From the values of RWC and solute potential of control and stressed leaf tissues, osmotic adjustment was calculated using the formula (Senthil-Kumar and Udayakumar, 2006):

Osmotic adjustment (OA) = Stressed leaf Ψ_{s100} - Control leaf Ψ_{s100}.

where, $\Psi_{s100} = (\Psi s \text{ (MPa)} \text{ X RWC})/100$.

(d) Quantification of percent reduction in total chlorophyll content

Fresh leaf tissue (100 mg) was incubated in a mixture of 80 % acetone: DMSO (1:1) (10 ml) overnight in dark conditions to extract chlorophyll pigments. The next day absorbance was recorded at 663 nm, 645 nm and 652 nm using UV-visible spectrophotometer (UV- VIS 2450, Shimadzu Corporation, Kyoto, Japan). The experiment was performed in replicates of three. Total chlorophyll was estimated by following the methods described by Hiscox and Israelstam (1979). The calculations are given below.

$$\text{Chlorophyll a (mg/g FW)} = \frac{12.7\,(A_{663}) - 2.69\,(A_{645}) \text{ X Volume}}{\text{FW X } 1000}$$

$$\text{Chlorophyll b (mg/g FW)} = \frac{22.9\,(A_{645}) - 4.68\,(A_{663}) \text{ X Volume}}{\text{FW X } 1000}$$

$$\text{Total chlorophyll (mg/g FW)} = \frac{8.02\,(A_{663}) + 20.20\,(A_{645}) \text{ X Volume}}{\text{FW X } 1000}$$

Total chlorophyll (mg/g FW) = Chlorophyll a content + Chlorophyll b content

$$\text{Total chlorophyll (mg/g FW)} = \frac{A_{652} \text{ X Volume}}{34.5 \text{ X FW}} \quad \text{(formula used in the study)}$$

Reduction in chlorophyll content over control (R%) is expressed as given below:

$$R = \frac{\text{Total chlorophyll (control)} - \text{Total chlorophyll (stressed)}}{\text{Total chlorophyll (control)}} \times 100$$

(e) Estimation of cell viability

The estimation of cell viability is a good measure of stress induced damage to any tissue. The 2, 3, 5-triphenyl tetrazolium chloride (TTC) assay was performed to measure the extent of cell viability (Senthil-Kumar and Udayakumar, 2006). The viability of the tissue will be reflected as colour formation when TTC is oxidised to red formazone in live respiring tissues. TTC solution was prepared by dissolving 0.4% of TTC (Sigma Aldrich, Bangalore, India) in 50 mM sodium phosphate buffer (pH 7.4). Leaf discs from fully expanded leaves of control as well as stressed plants were incubated in TTC solution at room temperature for 5 h while shaking. The leaf discs were then washed to remove unbound formazan and boiled with 5 ml of 2-methoxy ethanol to dryness. The bound TTC was extracted with 5 ml of 2-methoxy ethanol and absorbance was measured at 485 nm using UV-visible spectrophotometer (UV - VIS 2450, Shimadzu Corporation, Kyoto, Japan). The experiment was performed in replicates of three. The reduction in cell viability over non-stressed plants was calculated using the following equation and expressed in per cent.

$$\text{Reduction in cell viability (\%)} = \{[(A_{485}) \text{ control} - (A_{485}) \text{ stressed}] / [(A_{485}) \text{ control}] \times 100$$

3.2.2.3 Imposition of abiotic stresses in seedling system

Finger millet seedlings were used for studying the expression of specific stress responsive genes. Stress response assays were performed where various degrees of abiotic stresses were imposed to seedlings on petriplates by using sodium chloride (NaCl : 300mM), polyethylene glycol (PEG : - 8 bars) and methyl viologen (MV : 10 µM). Seedlings assays were performed separately for finger millet variety (GPU-28). Seeds were germinated at 37^0C in an incubator for two days. Uniform seedlings of approximately 1.5cm root length were transferred to petriplates at different stress treatments. Seedlings were also maintained at control conditions by transferring the pre germinated seedlings to petriplates containing water. The shoot and root tissues from the different stress treatments as well as control plates were collected, frozen in liquid nitrogen and stored at – 70°C. These stressed and control plant tissues were used to extract total RNA.

3.2.3 Assessing drought effects by molecular assays

A. Isolation of total RNA

Total RNA was extracted according to the Phenol - Chloroform protocol described by Datta *et al* (1989) which is given in detail in Appendix IA and was used as the starter material for expression studies. Total RNA from the seedling tissues were isolated by Trizol method (Chomczynski *et al*., 1995; Haimov- Kochman *et al*., 2006) which is given in detail in Appendix IB.

B. Expression analysis of stress responsive gene - ELIP

Drought stressed leaf tissues and stressed seedling tissues were used as starter material to isolate RNA and reverse transcribed into cDNA to prepare the template for expression studies. The expression pattern of stress responsive gene ELIP was studied by semi quantitative RT-PCR was performed with varying PCR cycles (20, 25, 30, 35 cycles). The general protocol for semi quantitative RT-PCR is given in 3.2.4.1.

3.2.4 Expression pattern of selected genes under drought stresses

3.2.4.1 Semi-quantitative RT-PCR analysis

On the basis of available sequence information, gene specific primers were designed for gene expression analysis. The selected clones were subjected to homology analysis using BLAST (Basic Local Alignment Search Tool) bioinformatics tool (www.ncbi.nlm.nih.gov/). The homologous sequences were aligned using ClustalW2 software (http://www.ebi.ac.uk/Tools/ clustalw2/index.html). The well aligned regions were selected for primer designing. Primer designing was done with the aid of primer designing option equipped softwares like FastPCR and VectorNTI (www.invitrogen.com/site/Vector-NTI.html). Proper care was taken to see that the GC content of the primers was above 50% and there were no primer dimers formed. The primers were designed such that the annealing temperatures of the forward and reverse primers did not differ by more than five degrees.

Total RNA (5 µg) isolated from the leaf tissues was reverse transcribed into cDNA in an oligo-dT primed reaction using Molony-Murine Leukaemia Virus Reverse Transcriptase (MMLV-RT; MBI Fermentas, Hanover, MD, USA) (DeStefano *et al.,*

1991; Bustin., 2000). First-strand cDNA was used as template for the expression analysis. The reactions were performed under standard PCR conditions using gene specific primers. The DNA fragments were amplified from cDNA in a 20 µl reaction volume containing 100 ng of template DNA (1-3 µl), 2 µl PCR buffer (10X), 2 µl dNTPs (2 mM), 1 µl forward primer (3 pmol/µl), 1 µl reverse primer (3 pmol/µl), 1 U *Taq* DNA polymerase (0.3 µl) and volume made up to 20 µl with sterile water. PCR was performed in PCR machine (Master cycler Gradient, Eppendorf AG, Germany). The optimal number of PCR cycles and the annealing temperature (T_A) was determined empirically for each PCR. A standard PCR program followed is as follows:

Step 1: 94°C for 2 min (Initial denaturation)

Step 2: 94°C for 1 minute (Denaturation)

Step 3: T_A for 45 seconds (Primer annealing) where, T_A = Annealing temperature

Step 4: 72°C for 1 minute /kb fragment (DNA synthesis)

Step 5: Goto step 2; Repeat 29 cycles

Step 6: 72°C for 10 min (Final extension)

Step 7: Hold at 4°C; End

The PCR amplified products were electrophoresed on agarose (1.2 %) gel (Appendix IE) to visualise the band sizes (Sambrook *et al*., 1989). The gel images were documented using gel documentation unit (Herolab 440K, GmbH Laborgerate, Germany). The presence of the expected PCR products was determined with the help of DNA ladders (GeneRuler DNA ladders, MBI-Fermentas).

3.2.4.2 Quantitative RT-PCR analysis

The quantitative real-time RT-PCR was performed with the fluorescent dye SYBR-green (SYBR Premix Ex Taq, Perfect Real Time, TaKaRa, Japan) following the manufacturer's protocol and the reaction conditions using Opticon2, MJ research, USA. SYBR Green chemistry is an alternate method used to perform real-time PCR analysis (Morrison *et al.* 1998). SYBR Green dye binds to the minor groove of double stranded DNA and increases the intensity of the fluorescent emissions. As more double stranded amplicons are produced, dye signal will increase. SYBR Green dye will bind to any double stranded DNA molecule. The cDNA pool was used as template to perform quantitative real-time PCR in presence of SYBR-green dye. The optimization of the real-time PCR conditions was performed according to the manufacturer's instructions (MJ Research, USA & MJ

Bioworks, Inc.). The PCR was performed using the SYBR Premix Ex Taq. The conditions for the PCR were as follows: 94ºC for 2 min, 30 cycles of denaturation (at 94ºC for 1 min), annealing for 30 seconds (57ºC to 67ºC for gene as well as actin) extension for 30 seconds (72ºC) followed by plate read and final extension at 72ºC for 8 min, estimation of melting curve from 50ºC to 90ºC and incubation at 72ºC for 10 min. The amplification plots of the PCR reaction were used to determine the threshold cycle (C_T). The relative expression levels of the selected genes under a given stress condition compared to its corresponding control was calculated using the comparative ΔCt method. The data was normalized to actin expression levels. The estimation of expression levels and fold increase in expression over control was calculated using the following formulae.

Fold increase over control = $2^{(-\Delta\Delta Ct)}$

where, $\Delta\Delta Ct = \{\Delta Ct \text{ (treatment)} - \Delta Ct \text{ (control)}\}$

$\Delta Ct = \{Ct \text{ (target gene)} - Ct \text{ (control gene)}\}$

The PCR amplified products were electrophoresed on agarose (1.2 %) gel (Appendix IE) to visualise the band sizes. The gel images were documented using gel documentation unit (Herolab 440K, GmbH Laborgerate, Germany). The presence of the expected PCR products was determined with the help of DNA ladders (GeneRuler DNA ladders, MBI-Fermentas).

3.2.5 Confirmation of gene specific amplification by restriction digestion

The amplified PCR products were reconfirmed by restriction digestion using restriction endonucleases (MBI, Fermentas). The digestion components and conditions are given in Appendix ID. The digestion products were visualised by electrophoresing on agarose gel (2.5 %) (Appendix IE). The gel images were documented using gel documentation system (Herolab 440K, GmbH Laborgerate, Germany). The different restriction endonucleases used in the study is given in Appendix IV.

3.3 Construction of full-length genes

The available sequence information of the genes selected for full length cloning was analysed in RNA Analyzer (http://rnaanalyzer.bioapps.biozentrum.uni-uerzburg.de/server.html) to retrieve general information of the sequence (Bengert and Dandekar., 2003). This web interface helps to identify sequence characteristics like exon, UTRs, polyA signal, snRNP binding motifs etc.

3.3.1 Rapid Amplification of cDNA Ends (RACE)

RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) kit (Invitrogen Corporation, Carlsbad, CA, USA) was used to perform the RACE reaction according to manufacturer's instructions. The GeneRacer™ Kit ensures that only transcripts containing full-length cDNA ends are amplified. The advanced protocol starts at the RNA level by specifically targeting only the 5'capped mRNA.

RNA was isolated from tissues grown at a stress level of 50% FC by following the kit protocol (Nucleospin RNA plant, Machery – Nagel Ltd., Germany). As per the manufacturer's protocol, total RNA was treated Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure from full-length mRNA, leaving a 5′-monophosphate. RNA adapter oligonucleotide was ligated to the 5′ end of mRNA using T4 RNA ligase. The adapter cannot ligate to dephosphorylated RNA because these molecules lack the 5′-phosphate necessary for ligation. During the ligation reaction, the majority of the full-length decapped mRNA acquires the adapter sequence at its 5′ end. A random-primed reverse transcription reaction was performed using SuperScript III reverse transcriptase. The 3′ end of gene of interest was amplified using the cDNA template with GeneRacer™ primers and gene specific primers (Appendix II). The amplification was carried out by using a standardised touch up PCR profile as given below:

Step 1: 94°C for 2 min (Initial denaturation)
Step 2: 94°C for 1 minute (Denaturation)
Step 3: 60°C for 1minute with +1°C every cycle for 10 cycles
Step 4: 72°C for 1 minute /kb fragment (DNA synthesis)
Step 5: Go to step 2; Repeat 10 cycles for template enrichment
Step 6: 94°C for 1 minute (Denaturation)
Step 7: 65°C for 1minute for 30 cycles
Step 8: 72°C for 1 minute /kb fragment (DNA synthesis)
Step 9: Go to step 6; Repeat 30 cycles for
Step 10: 72°C for 10 min (Final extension)
Step 11: Hold at 4°C; End

The amplified products were then cloned into pTZ57R/T vector [T/A-vector; Fermentas, Hanover, MD, USA (Appendix III B)] for sequencing. The general protocol recommended by the manufacturer is given in Flowchart 1.

RESULTS

IV. RESULTS

Plants are sessile and hence are stricken the most by environmental coercion, because of the fact that they have mobility constraints which restricts them to evade unfavourable conditions like biotic and abiotic stresses. Certain crop species have placed them in the forefront in the path of evolution making them stress tolerant. This has been achieved mainly by equipping themselves with a wide array of physiological, molecular and ecological adaptations. Drought is a major challenge for crop productivity. The per capita food requirement still rises at alarming rates even when crop productivity is in the clutches of abiotic stress limitations. It is extremely significant to go deep into the tolerant species to study the adaptive mechanisms that have been accomplished at molecular level. This would in fact give an insight into the nature of stress responsiveness of different genes in tolerant species which can serve as an efficient means to combat stresses in susceptible species by possible genetic manipulation strategies. However, drought being a multigenic trait, has to be evaluated with extreme care so as to pitch upon the correct set of genes that in combination would bring out the most desirable result.

With this outlook, the present investigation was planned and designed in such way so as to suffice the broad objective "Prospecting candidate genes for drought tolerance from finger millet (*Eleusine coracana* (L.) Gaertn.) ". To realise this major objective the experiments were accurately designed by approaching the situation as follows:

7. Screening and evaluation of finger millet stress cDNA library
8. Identification of novel genes by sequencing and annotation using bioinformatics tools
9. Expression analysis of selected genes by
    (i) e-northern using database information
    (ii) examining transcript levels in the tissues exposed to drought stress

4.1 Screening of drought stress specific cDNA library developed from finger millet leaf tissue

The stress cDNA library developed in the Department of Crop Physiology, UAS, GKVK, Bangalore, was used for the screening and discovery of novel drought responsive genes. The screening of the library was done following the five main steps:

1. Packaging and lifting of phage plaques each representing a recombinant clone

2. Analysis of plaques by polymerase chain reaction (PCR) with M13 forward and reverse primers for identification of recombinant clones

3. Selection of clones with insert DNA of size 500 bp or more

4. Purification of the PCR products from the selected clones and sequencing

5. Processing of DNA sequences and annotation

4.1.1 Packaging and plaque lifting

The packaged phages were tested for their efficiency with XL1-Blue MRF′ cells on LB agar plates containing tetracycline (30 µg/ml). The screening was carried out by allowing the recombinant clones to grow as transparent plaques. Each transparent plaque was lifted and eluted into SM buffer. Aliquots from this eluate were used as template for PCR confirmation.

4.1.2 Polymerase Chain Reaction (PCR) to identify recombinant plaques

The recombinant plaques from among those lifted were identified by PCR confirmation with M13 forward and reverse primers (Plate 1). From many selected plaques, only the ones which gave single band amplification were purified and sequenced. Amplicons of length more than 500 bp were further subjected to PCR product column purification protocol. The purified products were electrophoresed (Plate 2) to confirm for purity before sending for sequencing. Around 1100 clones have been screened and nearly 600 subjected for sequence analysis. Out of the total screened clones, only around 70% had insert sizes of more than 500 bp and those were subjected for further purification and sequencing.

4.1.3 DNA sequence analysis and annotation

The blast analysis of the sequencing results was done using NCBI database (www.ncbi.nlm.nih.gov/blastx). Annotated results of the sequence information are given in Table 4. The resulting Expressed Sequence Tags (ESTs) were further analysed for their functional and regulatory significance and classified based on their operative relevance. The classification of the different genes into proteases, transcription factors, receptor proteins, kinases and phosphatases, catalytic genes, membrane proteins etc., is represented in Fig. 1. On the basis of the functional significance of certain genes and a preliminary analysis of expression patterns using e-northern, 22 genes associated with diverse drought resistance pathways were selected for further studies. The selected genes

are tabulated in Table 5. The sequence information of the clones/ genes selected for expression analysis is given in Appendix VA.

4.1.4 Excision of the pBluescript phagemid from the Uni-ZAP XR vector

The cDNA library was constructed such that most of the inserts were cloned into Uni-ZAP XR phage vector for enhanced stability for long term preservation. However, some of the inserts were cloned into bacterial vectors and were preserved as glycerol stocks of the respective clones when the library was constructed. The phage vectors are larger in size and they need bacterial host for multiplication. So, for convenience in handling, the Uni-ZAP XR phage vector was converted into phagemid vector which behaves as plasmid. The Uni-ZAP XR phage vector includes all sequences of the pBluescript SK (–) phagemid and the cloned inserts will fall with in the pBluescript SK (–) phagemid sequence with internal M13 and T_7 sequences. The phage vector was converted into phagemid vector containing the cloned insert, with the help of ExAssist helper phage and two *E. coli* strains (SOLR strain and XL1-Blue MRF′). The excised phagemid in the SOLR strain was plated on LB agar medium with ampicillin (100µg/ml). Colonies appearing on the plate contained the pBluescript double-stranded phagemid with the cloned cDNA insert.

Table 4: List of annotated genes cloned from drought stress specific cDNA library developed from finger millet leaf tissue

Sl no	Length	Accession no.	Gene	Bit	Raw	EValue
1	950	ACF06451.1	QM-like protein [Elaeis guineensis]	78.2	191	3e-20
2	840	CAQ43070.2	putative puroindoline b protein [Triticum aestivum subsp. macha]	75.9	185	4e-12
3	660	NP_001063134.1	Os09g0407200 [Oryza sativa (japonica cultivar-group)]	99.8	247	2e-19
4	950	ACL54645.1	unknown [Zea mays]	94.7	234	4e-18
5	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	330	845	2e-88
6	950	NP_001146666.1	hypothetical protein LOC100280266 [Zea mays]	44.3	103	2e-08
7	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	256	654	3e-79
8	904	NP_001146950.1	40S ribosomal protein S11 [Zea mays]	296	758	2e-78
9	950	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	66.6	161	3e-09
10	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	239	610	1e-66
11	655	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	60.8	146	9e-08
12	750	XP_678847.1	hypothetical protein [Plasmodium berghei strain ANKA]	61.2	147	9e-08
13	950	NP_001147937.1	AMP-binding protein [Zea mays]	537	1384	1e-151
14	950	ACG34185.1	60S ribosomal protein L10-3 [Zea mays]	426	1095	1e-117
15	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	174	442	2e-48
16	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	368	944	1e-100
17	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	368	944	1e-100
18	950	ACL54645.1	unknown [Zea mays]	333	855	1e-89
19	950	NP_001143056.1	hypothetical protein LOC100275525 [Zea mays]	132	333	4e-29
20	950	ACL54645.1	unknown [Zea mays]	134	336	6e-32

21	950	CAK26794.1	hypothetical protein [Sporobolus stapfianus]	210	534	2e-52
22	910	AAT11797.1	MT-like protein [Cynodon dactylon]	109	272	4e-22
23	950	ACG31749.1	serine-threonine kinase receptor-associated protein [Zea mays]	127	319	2e-27
24	386	XP_668857.1	hypothetical protein [Plasmodium berghei strain ANKA]	45.4	106	0.002
25	950	NP_001141193.1	hypothetical protein LOC100273280 [Zea mays]	120	301	2e-25
26	950	ACG30855.1	hypothetical protein [Zea mays]	185	469	6e-45
27	757	-	-	-	-	-
28	950	ACN29242.1	unknown [Zea mays]	81.6	200	1e-13
29	950	NP_001105217.1	glyoxalase I [Zea mays]	157	397	1e-36
30	950	CAA33364.1	dehydrin [Zea mays]	69.7	169	4e-10
31	921	XP_668857.1	hypothetical protein [Plasmodium berghei strain ANKA]	58.5	140	9e-07
32	950	ACG31749.1	serine-threonine kinase receptor-associated protein [Zea mays]	82.8	203	4e-14
33	950	ACG31749.1	serine-threonine kinase receptor-associated protein [Zea mays]	246	627	3e-63
34	950	NP_001052557.1	Os04g0367800 [Oryza sativa (japonica cultivar-group)]	40.4	93	2e-07
35	950	ABA98760.1	HAT family dimerisation domain containing protein [Oryza sativa (japonica cultivar-group)]	48.1	113	2e-11
36	950	EAY82913.1	hypothetical protein OsI_38128 [Oryza sativa Indica Group]	105	263	5e-21
37	950	YP_002387017.1	Minor capsid protein C from bacteriophage origin [Escherichia coli IAI1]	338	866	6e-91
38	950	NP_001141193.1	hypothetical protein LOC100273280 [Zea mays]	234	598	7e-60
39	885	XP_674265.1	hypothetical protein [Plasmodium berghei strain ANKA]	65.5	158	6e-09
40	950	NP_001147639.1	chlorophyll a-b binding protein 2 [Zea mays]	209	533	3e-56
41	950	ACG30677.1	histone H4 [Zea mays]	168	426	6e-40
42	950	ACN30669.1	unknown [Zea mays]	176	446	3e-42

43	950	XP_001942907.1	PREDICTED: similar to elongase, putative [Acyrthosiphon pisum]	135	340	6e-30
44	950	AAC49138.1	aldose reductase-related protein [Avena fatua]	98.6	244	8e-19
45	950	NP_001146950.1	40S ribosomal protein S11 [Zea mays]	325	832	5e-87
46	950	NP_001068065.1	Os11g0546300 [Oryza sativa (japonica cultivar-group)]	490	1262	1e-137
47	766	ABR26167.1	cytosol aminopeptidase [Oryza sativa (indica cultivar-group)]	139	349	4e-31
48	700	XP_668857.1	hypothetical protein [Plasmodium berghei strain ANKA]	56.6	135	2e-06
49	827	-	-	-	-	-
50	950	NP_001141193.1	hypothetical protein LOC100273280 [Zea mays]	207	527	1e-51
51	950	AAS48643.1	glutathione s-transferase II [Cynodon dactylon]	372	954	1e-101
52	793	-	-	-	-	-
53	811	ACG30677.1	histone H4 [Zea mays]	75.1	183	5e-28
54	950	ABR17728.1	unknown [Picea sitchensis]	145	365	4e-69
55	950	NP_001067844.1	Os11g0454300 [Oryza sativa (japonica cultivar-group)]	71.6	174	1e-10
56	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	125	315	6e-50
57	950	ACN26456.1	unknown [Zea mays]	367	943	1e-100
58	950	ACN26456.1	unknown [Zea mays]	67.0	162	9e-14
59	713	-	-	-	-	-
60	510	-	-	-	-	-
61	950	ACF80712.1	unknown [Zea mays]	57.8	138	2e-06
62	796	EEC79001.1	hypothetical protein OsI_19503 [Oryza sativa Indica Group]	95.1	235	6e-18
63	950	AAT11797.1	MT-like protein [Cynodon dactylon]	103	257	2e-20
64	950	-	-	-	-	-
65	862	-	-	-	-	-

66	950	NP_001105442.1	heat shock protein 17.2 [Zea mays]	140	352	2e-31
67	950	CAA70550.1	lacZ-PhoC [synthetic construct]	63.2	152	4e-08
68	950	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	62.4	150	6e-08
69	950	AAG15510.1|U56995_1	mutant green fluorescent protein [Cloning vector pGreenscript A]	63.5	153	3e-08
70	943	NP_001050795.1	Os03g0652100 [Oryza sativa (japonica cultivar-group)]	47.4	111	7e-07
71	880	ACG39447.1	hypoxia induced protein conserved region containing protein [Zea mays]	147	370	5e-40
72	950	AAX26921.2	SJCHGC09205 protein [Schistosoma japonicum]	54.7	130	1e-05
73	898	EEE61242.1	hypothetical protein OsJ_15294 [Oryza sativa Japonica Group]	204	518	1e-50
74	950	NP_001105346.1	ferredoxin3 [Zea mays]	241	616	6e-62
75	683	XP_001862383.1	hypothetical protein CpipJ_CPIJ012185 [Culex quinquefasciatus]	58.5	140	5e-07
76	676	ACG25593.1	thioredoxin H-type 5 [Zea mays]	140	352	1e-31
77	805	-	-	-	-	-
78	801	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	69.3	168	4e-10
79	950	ACG44127.1	translocon-associated protein alpha subunit precursor [Zea mays]	308	789	5e-82
80	950	NP_001145348.1	hypothetical protein LOC100278677 [Zea mays]	107	268	1e-21
81	870	NP_001150571.1	EF hand family protein [Zea mays]	267	683	8e-70
82	766	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	61.2	147	9e-08
83	746	NP_001141626.1	hypothetical protein LOC100273745 [Zea mays]	169	428	2e-40
84	872	XP_668857.1	hypothetical protein [Plasmodium berghei strain ANKA]	59.7	143	3e-07
85	388	-	-	-	-	-
86	950	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	59.7	143	4e-07
87	562	ABU39931.1	beta-galactosidase [Cloning vector pGreenII 0179]	58.2	139	4e-07

88	849	ABU39931.1	beta-galactosidase [Cloning vector pGreenII 0179]	89.4	220	4e-16
89	950	EEE68144.1	hypothetical protein OsJ_26249 [Oryza sativa Japonica Group]	105	263	8e-29
90	950	EAY96485.1	hypothetical protein OsI_18385 [Oryza sativa Indica Group]	81.3	199	1e-13
91	885	ACR37913.1	unknown [Zea mays]	171	432	1e-40
92	910	XP_668857.1	hypothetical protein [Plasmodium berghei strain ANKA]	59.7	143	4e-07
93	950	ABW84175.1	lacZ [Binary vector pCLEAN-G147]	64.7	156	1e-08
94	918	ACN26456.1	unknown [Zea mays]	64.7	156	1e-08
95	950	XP_678847.1	hypothetical protein [Plasmodium berghei strain ANKA]	61.6	148	1e-07
96	878	NP_001137133.1	hypothetical protein LOC100217315 [Zea mays]	230	587	1e-58

Table 4: List of annotated genes cloned from drought stress specific cDNA library developed from finger millet leaf tissue (continued)

Sl no	Clone no	Size (bp)	Accession no.	Gene	Organism	Score	E Value
1	1	950		Putative mitochondrial precursor, Nudix hydrolase	Ricinus communis	39.7	1.86
2	2	950		Iron-sulphur-binding reductase	Corynebacterium matruch	32.7	0.76
3	3	851		vacuolar protein sorting protein 25	Zea mays	118	8E - 25
4	5	950		GH11618	Drosophila grimshawi	41.4	0.13
5	7	950		unknown protein	Arabidopsis thaliana	38.5	0.023
6	9	887	SJCHGC09370	SJCHGC09370 protein	Schistosoma japonicum		
7	10	674		gamma gliadin	Triticum monococcum	35.8	3.1
8	12	890		no significant similarity found			
9	14	950		unknown protein, Zn finger protein	Zea mays		
10	16	628	Os06g0320500	Os06g0320500 , chlorophyll a/b binding protein	Oryza sativa japonica cult	129	1.00E-28
11	20	950		dehydrin1 , RAB 17	[Zea mays]	52.4	6.00E-05
12	23	705		PREDICTED: hypothetical protein	Gallus gallus	35.8	3.4
13	24	950	ref|NP_001154493.1|	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	Arabidopsis thaliana	38.1	1.3
14	26	950	gb|EEQ33321.1|	SMB [CBS 113480]	Microsporum canis	37.7	1.7
15	28	950	ref|ZP_01969332.1|	hypothetical protein RUMTOR_02920	Ruminococcus torques	39.3	0.57
16	30	950		No significant similarity found			

17	36	950		No significant similarity found			
18	35	877	gb|EEK31724.1|	transcriptional regulator, COG0583	Nocardiopsis dassonvillei	42	0.075
19	38	923	gb|ABS28523.1|	hypothetical protein	Anaeromyxobacter sp	44.7	0.013
20	39	748	gb|AAS88721.1|	metallothionein-like protein	Cynodon dactylon	48.1	8.00E-04
21	41	950	gb|ACG30117.1|	30S ribosomal protein S9	Zea mays	65.9	6.00E-09
22	43	825	ref|XP_002285104.1|	PREDICTED: hypothetical protein	Vitis vinifera	37.7	1.3
23	47	950	gb|EEF34255.1|	hypothetical protein RCOM_0146500	Ricinus communis	37.7	1.6
24	49	950	gb|EEQ87194.1|	actin associated protein Wsp1	Ajellomyces dermatitidis	36.6	3.7
25	50	767	dbj|BAF44481.1|	pyruvate orthophosphate dikinase	[Echinochloa frumentacea]	127	1.00E-27
26	51	950		No significant similarity found			
27	52	950		No significant similarity found			
28	33	513	gb|EAZ02114.1|	PREDICTED: similar to PCTAIRE protein kinase 1	Ornithorhynchus anatinus	38.1	0.35
29	54	911	sp|Q38793.1|RBS_AEGTA	RecName: Full=Ribulose bisphosphate carboxylase small chain,chloroplastic; Short=RuBisCO small subunit	Aegilops tauschii	50.1	3.00E-04
30	55	950		No significant similarity found			
31	57	950	gb|ACG33961.1|	calreticulin-3 precursor	Zea mays	42.7	0.051
32	66	616	gb|EDR07977.1|	predicted protein	Laccaria bicolor	42	0.032
33	70	950	gb|EDS40792.1|	conserved hypothetical protein	Culex quinquefasciatus	35.8	6.3
34	72	950	gb|AAI59253.1|	Unknown (protein for IMAGE:8994562)	Danio rerio	42	0.087
35	74	770	gb|EER95172.1|	hypothetical protein SORBIDRAFT_01g041040	Sorghum bicolor	51.6	7.00E-05

36	75	947	gb|AAV31119.1|	ribosomal protein S10p/S20e	Zea mays	85.9	5.00E-15
37	256	950	gb|ACX85641.1|	AT12646p	Drosophila melanogaster	39.7	0.43
38	80	729	ref|NP_001149295.1|	40S ribosomal protein SA	Zea mays	54.3	1.00E-05
39	257	950	gb|ACU65226.1|	delta 1-pyrroline-5-carboxylate synthetase 1	Sorghum bicolor	451	6.00E-125
40	258	950	gb|ABU39931.1|	beta-galactosidase	Cloning vector pGreenII 0179	40.4	0.25
41	86	516	ref|XP_001916107.1|	PREDICTED: similar to KIAA0174	Equus caballus	39.7	0.12
42	260	933	ref|XP_001214976.1|	conserved hypothetical protein	[Aspergillus terreus NIH2624]	36.2	5.1
43	88	885	ref|XP_001512208.1|	PREDICTED: similar to ROBO2	Ornithorhynchus anatinus	37.4	2.1
44	90	950	ref|XP_002582451.1|	predicted protein	[Uncinocarpus reesii 1704]	43.5	0.033
45	92	950		No significant similarity found			
46	94	950	gb|ACN35317.1|	unknown	[Zea mays]	36.2	5.3
47	95	950	ref|NP_104673.1|	ribose ABC transporter, ATP-binding protein	[Mesorhizobium loti	39.3	0.62
48	96	950	ref|XP_002302530.1|	predicted protein	[Populus trichocarpa	42	0.096
49	98	453		No significant similarity found			
50	110	950	ref|XP_001395851.1|	hypothetical protein An12g08000	Aspergillus niger	36.2	5.4
51	115	950		No significant similarity found			
52	117	928	ref|XP_002445129.1|	hypothetical protein SORBIDRAFT_07g004560	Sorghum bicolor	59.3	6.00E-07
53	118	950	ref|XP_002650503.1|	hypothetical protein EBI_26338		37.4	2.4
54	120	950	gb|EEE22467.1|	ankyrin repeat-containing protein, putative		32.3	0.018

55	123	609	ref|XP_002189949.1|	PREDICTED: similar to KIAA0268	Taeniopygia guttata	38.9	0.35
56	126	575	gb|ACC95865.1|	epsin 1	Lampetra fluviatilis	35.4	3.3
57	127	950		No significant similarity found			
58	129	950		No significant similarity found			
59	130	783	ref|YP_001505155.1|	O-antigen and teichoic acid-like export protein	Frankia sp.	38.9	0.6
60	132	950	sp|P02812.3|PRB2_HUMAN	Full=Basic salivary proline-rich protein 2,| proline-rich protein BstNI subfamily 2	Homo sapiens	30.8	0.33
61	133	897	ref|XP_002466696.1|	hypothetical protein SORBIDRAFT_01g012480	Sorghum bicolor	211	1.00E-52
62	136	950		No significant similarity found.			
63	139	950	ref|XP_539595.2|	PREDICTED: similar to procollagen, type VIII, alpha 2	Canis familiaris	40	0.38
			sp|Q04666.1|HES1_RAT	Full=Transcription factor HES-1, hairy and enhancer of split 1	Rattus norvegicus	38.1	1.4
64	144	950		No significant similarity found.			
65	147	633	gb|ACG46700.1|	chloride intracellular channel 6	Zea mays	96.3	2.00E-18
66	154	950		No significant similarity found.			
67	158	950		No significant similarity found.			
68	161	950		No significant similarity found.			
69	162	950		No significant similarity found.			
70	173	897	ref|XP_002087639.1|	GE15184 gene product from transcript GE15184-RA	Drosophila yakuba	29.3	0.002

71	181	920	gb|ACF81772.1|	unknown	Zea mays	119	4.00E-25
			b|AAA85273.1|	Ras-related protein Rab7	Cenchrus ciliaris	119	4.00E-25
72	182	463	gb|ACT53876.1|	S-adenosylmethionine decarboxylase	Saccharum officinarum	50.4	6.00E-05
73	183	900		No significant similarity found			
74	184	694		No significant similarity found			
75	193	950	gb|EEE30167.1|	subtilisin, putative	Toxoplasma gondii VEG	35.4	9.2
76	195	893		No significant similarity found.			
77	296	950		No significant similarity found.			
78	208	950		No significant similarity found.			
79	209	950	gb|ACF86670.1|	unknown	Zea mays	40.4	0.29
80	211	723	ref|NP_001131013.1|	asparagine synthetase 2	Zea mays	208	4.00E-52
81	212	889	ref|ZP_01969332.1|	hypothetical protein RUMTOR_02920	Ruminococcus torques	39.7	0.44
82	214	905	gb|AAX26662.2|	SJCHGC03128 protein	Schistosoma japonicum	40.8	0.003
83	215	950		No significant similarity found.			
84	217	950	gb|ACF84252.1|	unknown, LOC100282845	Zea mays	161	9.00E-38
85	220	724	sp|P08001.5|ACRO_PIG	Full=Acrosin	Sus scrofa	35	7.3
86	222	846	gb|EEU46987.1|	hypothetical protein NECHADRAFT_92045	Nectria haematococca	46.2	0.004
			sp|P02812.3|PRB2_HUMAN	Full=Basic salivary proline-rich protein 2,| proline-rich protein BstNI subfamily 2	Homo sapiens	42.1	0.021
87	224	950	gb|EFA05547.1|	hypothetical protein TcasGA2_TC015735	Tribolium castaneum	37	3.2
88	225	871	ref|NP_194848.1|	SWAP (Suppressor-of-White-APricot)/surp	Arabidopsis thaliana	37	2.7

				domain-containing protein			
89	227	950	gb|EFA02414.1|	hypothetical protein TcasGA2_TC008099	Tribolium castaneum	38.1	1.4
90	228	894		No significant similarity found.			
91	229	899	ref|XP_002468627.1|	hypothetical protein SORBIDRAFT_01g049305	Sorghum bicolor	39.3	0.001
92	230	950	ref|NP_959959.1|	Pra	Mycobacterium avium	36.6	4.1
93	239	950		. No significant similarity found			
94	246	598	gb|ABC02754.1|	PS II 10 kDa protein	Phyllostachys edulis	119	4.00E-49
95	250	554	ref|XP_002468442.1|	hypothetical protein SORBIDRAFT_01g045990	Sorghum bicolor	134	6.00E-30
			ref|NP_001049088.1|	hypothetical protein Os03g0168200	Oryza sativa (japonica	120	1.00E-25
			ref|XP_002315031.1|	predicted protein	Populus trichocarpa	93.2	1.00E-17
			ref|XP_002522483.1|	conserved hypothetical protein	Ricinus communis	90.1	1.00E-16
			ref|NP_001154334.1|	unknown protein	Arabidopsis thaliana	85.5	3.00E-15
96	254	950	gb|EAZ09217.1|	hypothetical protein OsI_31492	Oryza sativa Indica	36.2	5.40E+00
			gb|AAV98700.1|	BTH-induced ERF transcriptional factor 1	Oryza sativa Indica	36.2	5.40E+00

Table 5: Details of the clone identities of selected clones

Sl.no.	Clone no.	Seq. no.	Putative function	Organism	Acc. no	Score	E Value	Identities
1	(colony) 159	UAS_T3_E06_A_050.ab1	putative cyanase	Oryza sativa	AAG21913	131	2.00E-29	83%
2	(colony) 161	UAS_T3_G06_A_051.ab1	Saposin-like type B, region 1 family protein, expressed	Oryza sativa	ABA91162	169	1.00E-40	65%
3	(colony) 163	UAS_T3_A07_A_052.ab1	MT-like protein	Cynodon dactylon	AAT11797	109	1.00E-22	88%
4	(colony) 172	UAS_T3_A08_A_056.ab1	RISBZ4	Oryza sativa	BAB39174	495	3.00E-140	100%
5	(colony) 187	UAS_T3_A09_A_068.ab1	DIS1; protein binding / structural constituent of	Arabidopsis thaliana	NP_172777	435	1.00E-120	81%
6	(colony) 217	uas_55_155T3_H01_A_013.ab1	actin-related protein ARP3	Physcomitrella patens	CAL22910	134	2.00E-30	83%
7	(colony) 223	uas_55_155T3_F02_A_026.ab1	14-3-3 protein	Fagus sylvatica	CAE54082	33.1	0.29	93%
8	(colony) 237	uas_55_155T3_B04_A_032.ab1	sinapyl alcohol dehydrogenase	Plantago major	CAJ43715	45.8	0.002	100%
9	(colony) 246	uas_55_155T3_F04_A_042.ab1	putative TAF6	Oryza sativa	BAB92191	1081	0	100%
10	(colony) 295	uas_55_155T3_G08_A_067.ab1	farnesylated protein 1	Hordeum vulgare subsp.	CAD48128	184	5.00E-45	68%
11	(colony) 303	uas_55_155T3_D09_A_077.ab1	putative aldehyde dehydrogenase WIS1	Triticum aestivum	AAL70109	283	2.00E-86	73%
12	(colony) 311	uas_55_155T3_D10_A_089.ab1	trehalose-6-phosphate synthase /phosphatase	Cypripedium parviflorum var. pul	AAN86570	382	8.00E-105	71%
13	(colony) 321	uas_55_155T3_F11_A_094.ab1	putative signal recognition particle receptor	Oryza sativa	BAD09759	212	2.00E-53	60%
14	(colony) 351	UAS_A97_150R_A03_a_017.ab1	calumenin homologue	Ciona intestinalis	8556878|dbj|BAC5751	50.1	8.00E-05	88%
15	(old plq) 168	UAS193REST_B03_N_025.ab1	putative transcriptional regulator	Shewanella amazonensis SB2B	19768418|gb|ABM0098	37	3	27%
16	(old plq) 175	UAS193REST_A04_N_021.ab1	protein phosphatase	Arabidopsis thaliana	5222511|ref|NP_17715	160	1.00E-37	82%
17	(old plq) 254	UAS_A97_150R_D05_a_042.ab1	seven transmembrane helix receptor	Homo sapiens	1928439|dbj|BAC0581	45.1	0.013	30%
18	(old plq) 53	UAS1-96R_C02_A_006	Low molecular mass early light-inducible protein HV90,	Hordeum vulgare	p|P14897|ELI9_HORV	85.9	2.00E-15	78%
19	(old plq) 61	UAS1-96R_C06_A_038	Farnesyl pyrophosphate synthetase	Zea mays	p|P49353|FPPS_MAIZ	156	2.00E-36	70%
20	(old plq) 96	UAS1-96R_D03_A_026	DRP4 protein	Triticum turgidum	emb|CAF22029.1|	58.2	5.00E-07	67%
21	(old plq) 126	UAS1-96R_E01_A_003	monodehydroascorbate reductase	Plantago major	emb|CAJ43614.1|	44.7	0.006	64%
22	(new plq)175	UAS97-192R_D03_A_026	RAB17 protein	Zea mays	emb|CAM56275.1|	38.9	5.00E-05	94%

4.1.5 Confirmation of excised phagemids by colony PCR

All the 22 selected genes, were successfully converted into phagemids. The presence of the insert in the phagemid clones was confirmed through a colony PCR amplification using M13 forward and reverse primers (Plates 3 and 4). The amplification of expected size indicates that the inserts cloned into phage vector were retrieved into phagemid vector. This was further confirmed by a patch plate colony PCR amplification with M13 forward and reverse primers (Plate 5).

4.1.6 Isolation of plasmids from recombinant colonies and confirmation

Plasmid DNA with high level of purity was isolated from all the excised phagemid clones using plasmid DNA isolation kit for sequencing purpose for clone identity confirmation. The isolated plasmids were confirmed by PCR using M13 forward and reverse primers (Plate 6). The sequencing results were used to align with the existing sequences of the respective clones using ClustalW2 software (http://www.ebi.ac.uk/Tools/clustalw2/index.html) to confirm the clone identity. The sequence data and the alignment files are given in Appendix VB and VC respectively.

4.2 Gene Expression Studies

4.2.1 Expression analysis of selected genes by eNorthern

The *in silico* expression analysis of the genes was further performed using the freely available expression browser (www.bbc.utoronto.botany.ca). The Arabidopsis homologs of the fingermillet clones were obtained and their AGI codes were identified. This code was used to get the expression pattern of the genes under osmotic, drought, cold, salinity, radiation, genotoxic, oxidative, wounding and heat stresses in both shoot and root. The AGI codes of the Arabidopsis homologs corresponding to the finger millet clones are represented in Table 6.

Table 6: AGI codes of the homologous genes of *Arabidopsis thaliana* shortlisted for eNorthern analysis

AGI Code	Putative gene name	Putative function
At5g02380	Metallotheionin	Heavy metal homoeostasis by sequestration
At4g02640	RISBZ4	Transcription activating abilities
At1g13180	Actin Related Protein 3	Cytoskeletal structoraal component for regulating cell expansion
At3g21250	14.3.3 protein	Binding factors for TFs and other signalling proteins
At4g33930	Sinapyl Alcohol Dehydrogenase	Enhances drought and osmotic stress, Salinity tolerance
At1g04950	TAF6	TATA box binding protein associated factor interacting with TFII D
At5g66110	Farnesylated Protein 1	Response to ABA and reduces transcriptional loss
At5g62530	Alcohol Dehydrogenase WIS1	Upregulated under short term exposure to ABA
At4g30600	Signal Recognition Particle Receptor	Receptors involved in SRP binding in protein targeting
At1g66020	Calumenin homologue	Calcium signalling under drought
At1g64720	Transcriptionl Regulator	Signal transduction pathway under drought
At1g10430	Protein Phosphatase	Drought responding signal transduction pathway
At1g03160	Transmembrane Helix Receptor	Abscisic acid induced stomatal closure under drought
At4g14690	Light Inducible Protein HV 90	Adaptive response to stress induced photodanmage
At4g17190	Farnesyl Pyrophosphate Synthetase	Increased chlorophyll contents and delayed senescence
At5g22030	Mono Dehydro Ascorbate Reductase	ROS scavenging pathway
At1g59910	Superoxide Dismutase	ROS scavenging pathway

Note: The shortlisted genes with AGI codes available were used for e-Northern analysis.

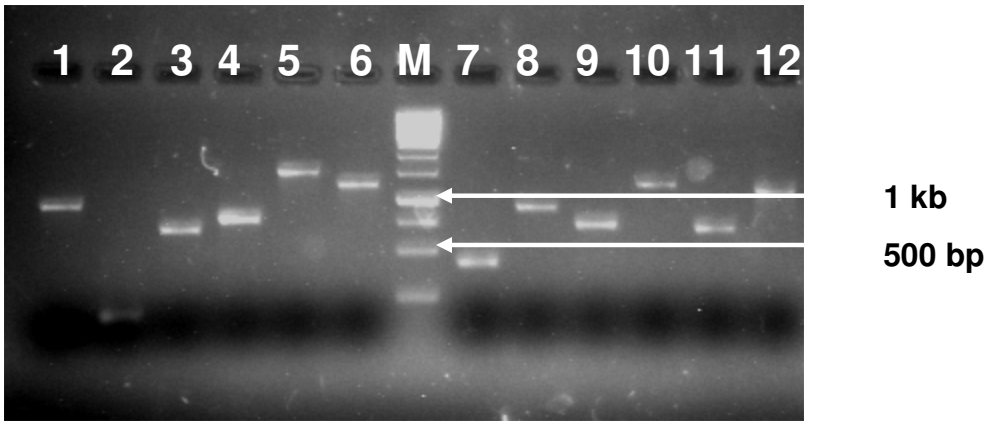


Plate 2: **Ethidium bromide stained agarose gel showing purified PCR amplified products.** Plaques grown on NZY agar plates were lifted and used as template for PCR amplification with M13 forward and reverse primers and the amplified products were separated on agarose gel (1%). . The PCR amplified products were then purified. Lanes 1-12 indicate purified PCR products of various clones; M – 1 kb DNA ladder.

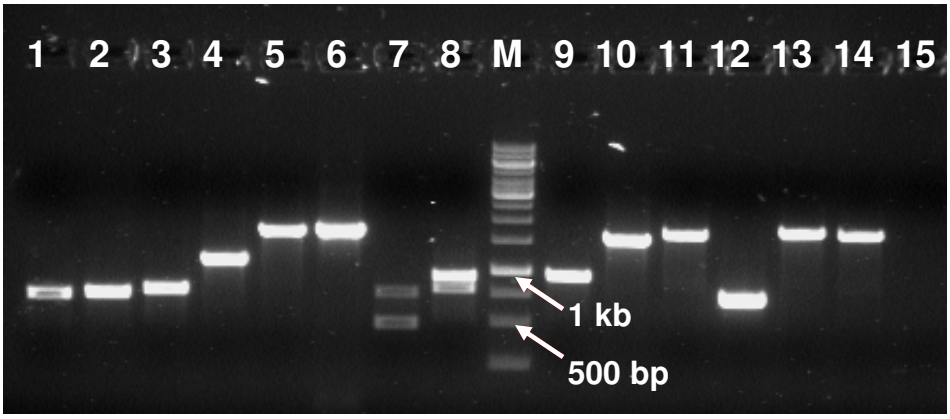


Plate 4: Ethidium bromide stained agarose gel showing PCR amplified products of converted phagemids after *in vivo* excision. Uni- ZAP XR vector containing finger millet drought stress specific genes were converted into phagemids, transformed into bacterial cells and selected on LB agar plates with ampicillin. Colony PCR was performed using M13 forward and reverse primers primers and the amplified products were separated on agarose gel (1%). Lanes 1- 14 indicate excised phagemids ; lane 15- negative control; lane M- 1 kb DNA ladder.

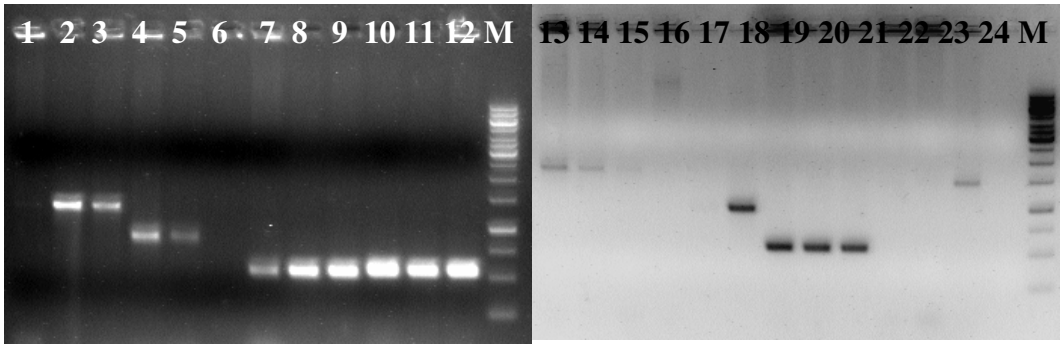


Plate 5 : Ethidium bromide stained agarose gel showing PCR amplified products of the selected genes in converted phagemids after *in vivo* excision. Uni- ZAP XR vector containing finger millet drought stress specific genes were converted into phagemids, transformed into bacterial cells and selected on LB agar plates with ampicillin. Colony PCR was performed using M13 forward and reverse primers primers and the amplified products were separated on agarose gel (1%). Lanes 1- 24 represent patched colonies; lane M- 1kd DNA ladder.

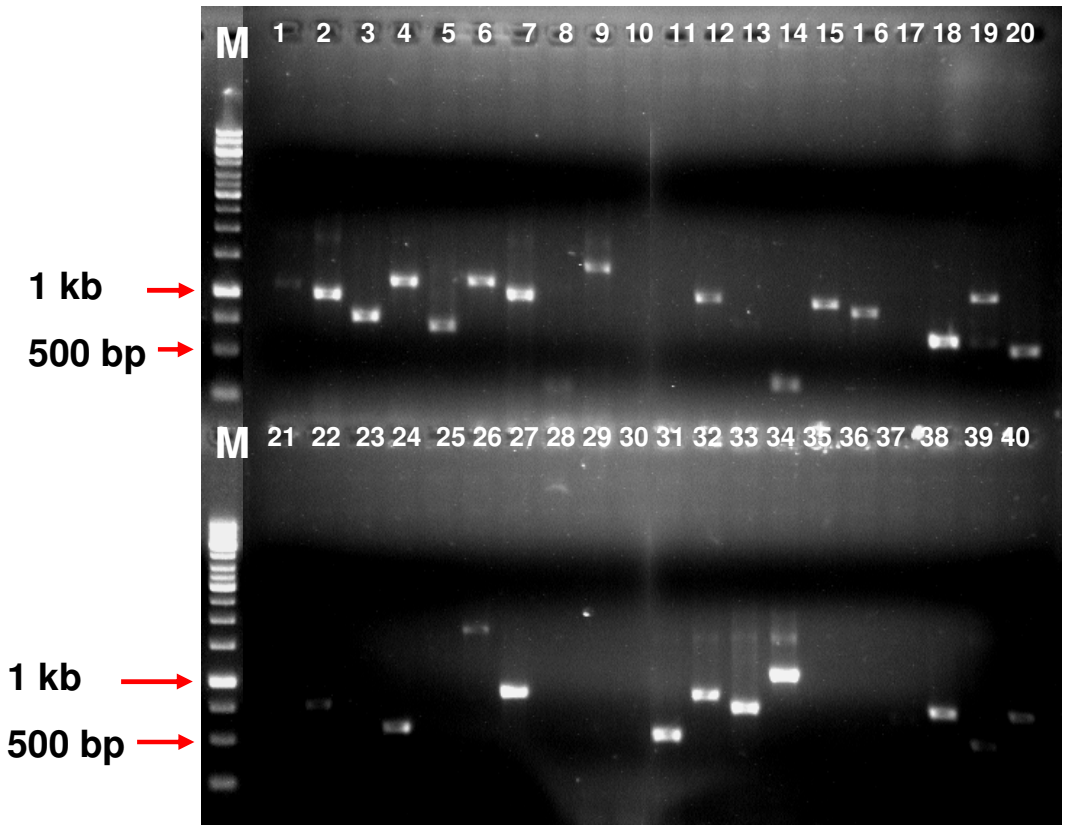


Plate 1: Ethidium bromide stained agarose gel showing PCR amplified products of lifted plaques . Plaques grown on NZY agar plates were lifted and used as template for PCR amplification with M13 forward and reverse primers and the amplified products were separated on agarose gel (1%). . Lanes 1- 40 indicate different clones; M – 1 kb DNA ladder

Under osmotic stress, both in root and shoot tissues, almost all the genes were found to be expressed, however, At5g66110 (homolog of farnesylated protein) showed high level of expression with 75 and 20 fold increase in shoot and root tissues respectively (Fig 2). However, under drought stress in shoot tissues, the fold increase in gene expression was seen at earlier stages which further reduced. In root tissues, At4g30600 (homolog of signal recognition particle receptor) was found to show increased expression towards the end of stress treatment (Fig 3). At4g14690 (homolog of Light inducible protein) was found to be highly expressive in both root and shoot tissues along with At4g33930 (homolog of sinapyl alcohol dehydrogenase) showing relatively high expression levels in root tissues (Fig 4). In case of salt stress, both in shoot and root tissues, At5g66110 (homolog of farnesylated protein) was found to express considerably along with homologs of sinapyl alcohol dehydrogenase and signal recognition particle receptor showing 3 and 20 folds increase in shoot and root tissues respectively (Fig 5). When exposed to UV-B radiation due to radiation stress, both in shoot and root tissues, At4g14690 (homolog of Light inducible protein) was found to express 80 and 15 fold respectively (Fig 6). Genotoxic stresses induced the relative increase in expression of At1g64720 (homolog of transcriptional regulator) both in shoot and root tissues but homolog of sinapyl alcohol dehydrogenase was found to show highest expression levels in root tissues (Fig 7). Oxidative stress led to a concomitant increase initially followed by a decrease in the expression levels of homologs of farnesylated protein, transcriptional regulator and sinapyl alcohol dehydrogenase (Fig 8). Homologs of signal recognition particle receptor and farnesylated protein were found to express highly in shoot tissues of wounded plants while in root tissues it was transcriptional regulator homolog that got expressed highly (Fig 9). However, when the plants were heat stressed, homolog of farnesylated protein was found to express highly in both shoot and root tissues (Fig 10). An overall evaluation helps to conclude that certain genes like farnesylated protein, transcriptional regulator, signal recognition particle receptor and sinapyl alcohol dehydrogenase are multiple stress responsive at varied levels. Hence, it was highly significant to elucidate their relevance under drought.

4.2.2 Expression studies using finger millet tissues exposed to abiotic stresses

The expression analyses were carried out on plants experiencing different levels of drought stress. Gravimetric approach was followed to impose moisture stress to finger millet plants grown in pots. Plants were maintained at different field capacities (FCs) and leaf tissue was used for expression analysis.

4.2.2.1 Imposition of drought stress at whole plant level

Drought stress was imposed to 20 days old plants by withholding irrigation. Soil field capacities of 80, 60, 50, 40, 30 and 20% were maintained by gravimetric approaches for stress imposition (Fig 11 & 12). During stress imposition, soil field capacity was monitored gravimetrically by checking the amount of water present in soil medium compared to maximum water it can hold. To assess the extent of drought, soil water potential was monitored using dewpoint potentiometer (Table 7). The soil water potential values varied from -0.13 Mpa at 80% FC to -1.68 MPa at 20% FC where as under well irrigated conditions, the value recorded was – 0.03 MPa. The plants grown at 20% FC were found to show soil water potential of -1.68 MPa.

4.2.2.2 Characterisation of drought stressed plants by physiological analysis

Soil water status has a direct effect on plant water relations. Hence, the extent of stress experienced was assessed by using indices like leaf osmotic potential and relative water content (RWC). The other biochemical changes like reduction in chlorophyll content, reduction in cell viability and loss of membrane integrity were also quantified to assess the stress effect.

(a) Assessment of relative water content

Relative water content (RWC) was found to decrease with the increase in severity of drought. The RWC was found to reduce from 95% under well irrigated conditions to 51% at 20% FC. But a drastic reduction in RWC commenced from 50% FC, continuing the decrease at an increasing rate, reaching critical RWC (51%) at 20% FC (Fig 13).

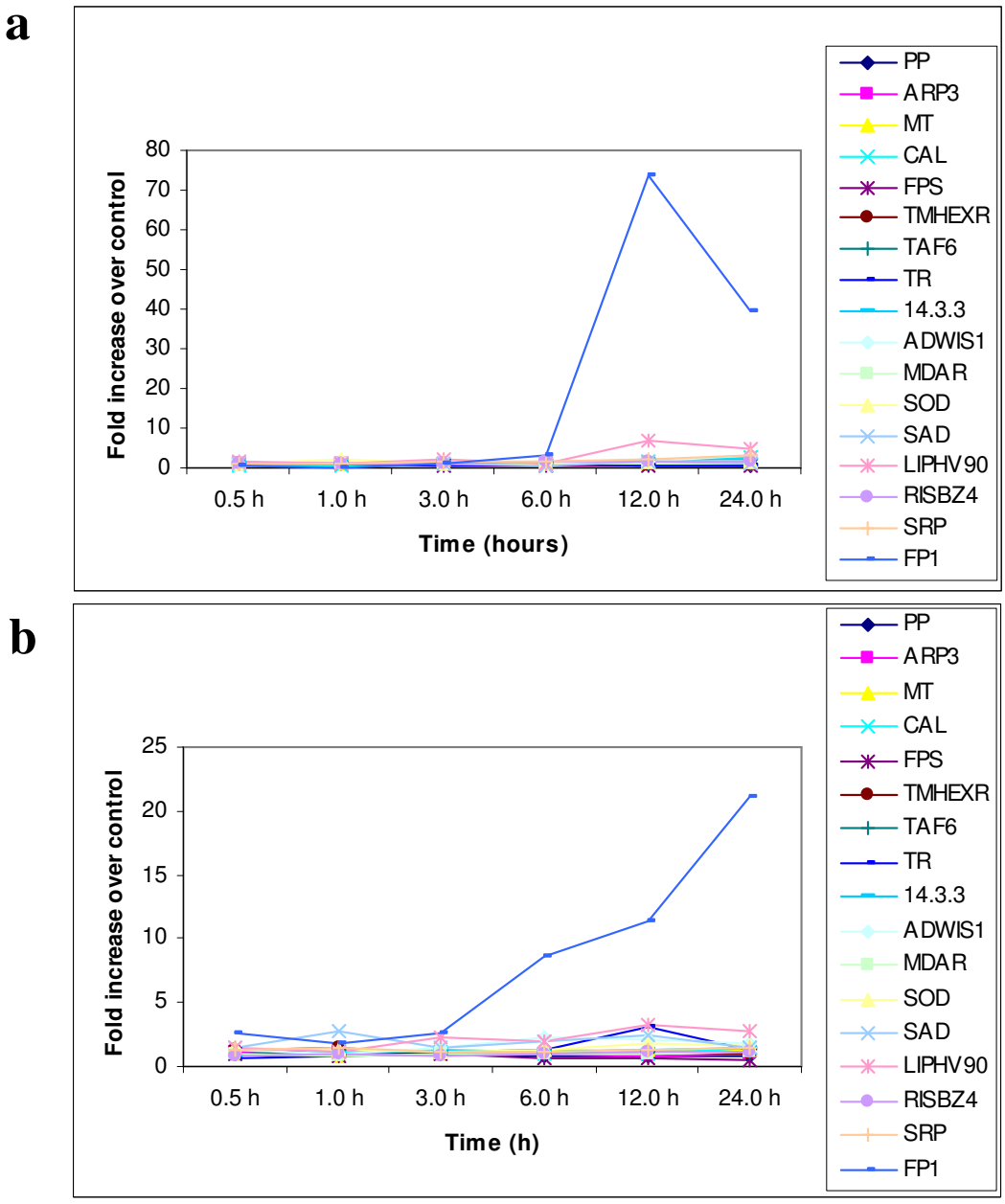


Fig 2: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under osmotic stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under osmotic stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

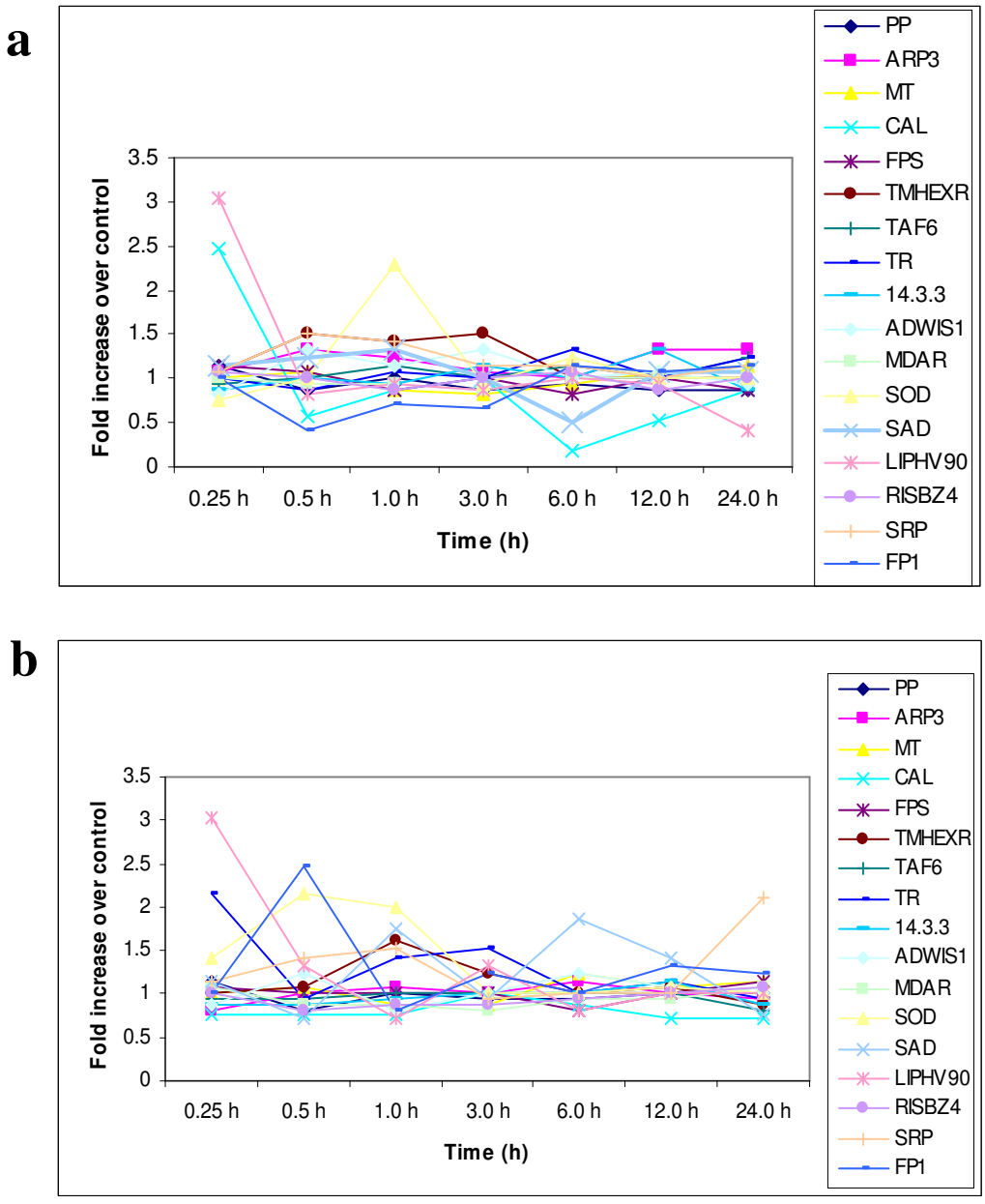


Fig 3: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under drought stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under drought stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

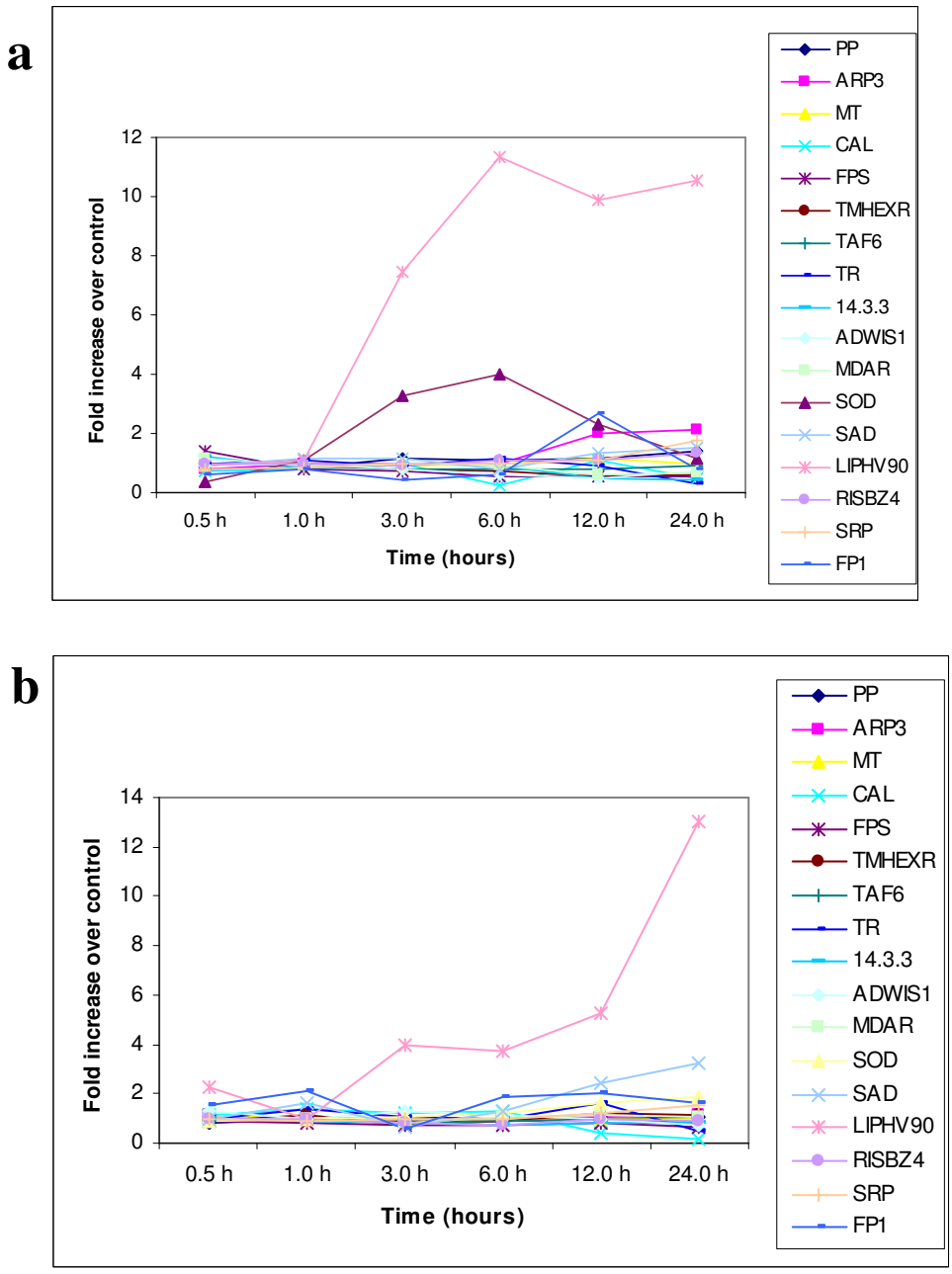


Fig 4: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under cold stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under cold stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

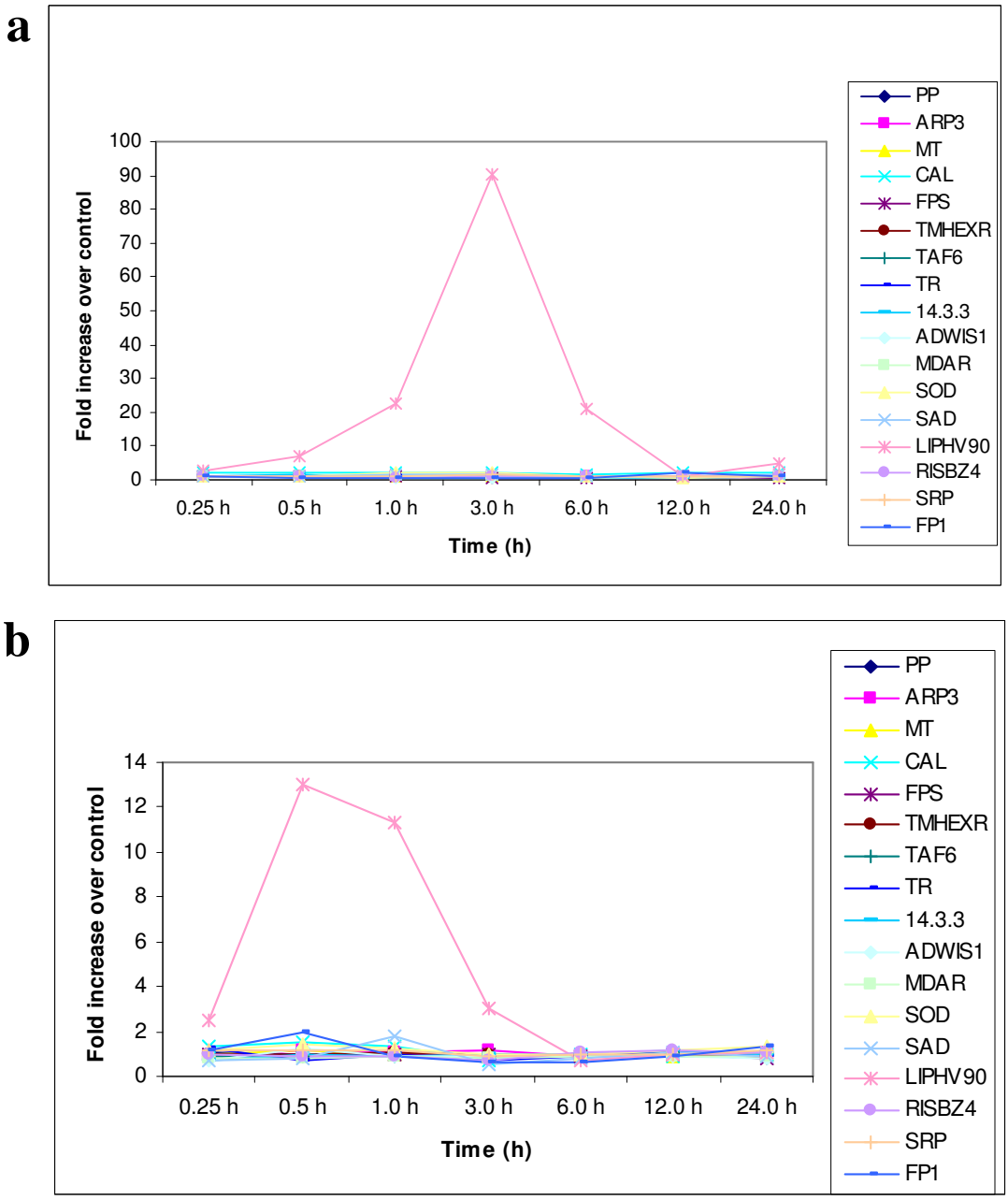


Fig 6: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under radiation stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under radiation stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

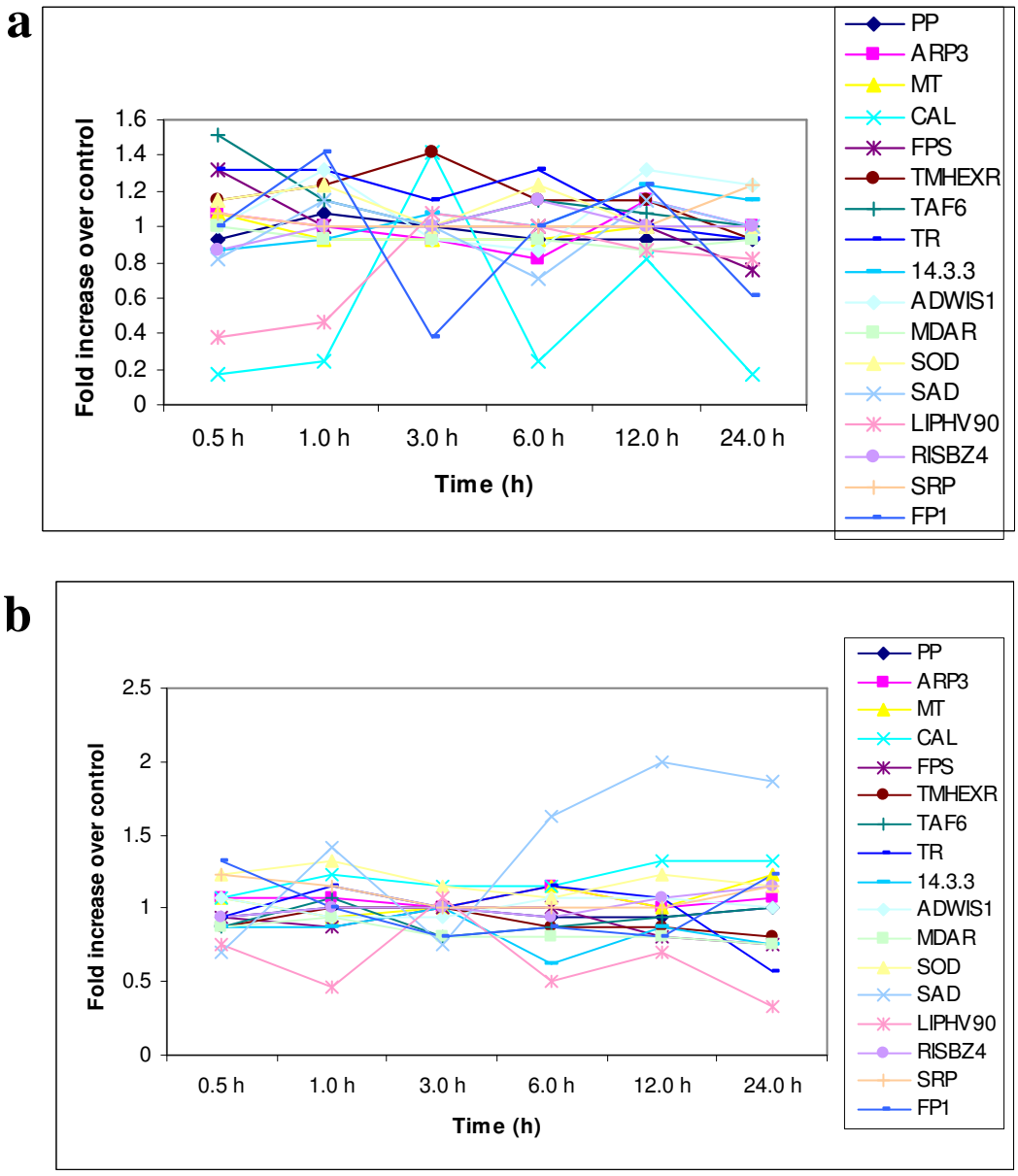


Fig 7: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under genotoxic stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under genotoxic stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

Table 7: Soil water potential of pots maintained at different field capacities to impose drought on finger millet.

p< 0.01

Sl. No.	Treatment (FC in %)	Soil Water Potential (ψ_w (Mpa))
1	20	-1.68
2	30	-0.80
3	40	-0.68
4	50	-0.42
5	60	-0.26
6	80	-0.13
7	100 (Control)	-0.03

Table 8: Leaf solute potential and net osmotic adjustment of finger millet plants maintained at various field capacities

p< 0.01

Sl. No.	Treatment (FC in %)	Leaf solute potential ψ_l (MPa)	Ψ_S 100	NET OSMOTIC ADJUSTMENT (MPa)
1	100 (Control)	-1.07	-1.01	
2	80	-1.25	-1.08	-0.06
3	60	-1.43	-1.14	-0.12
4	50	-1.55	-1.14	-0.13
5	40	-1.78	-1.19	-0.18
6	30	-2.04	-1.26	-0.25
7	20	-2.87	-1.47	-0.46

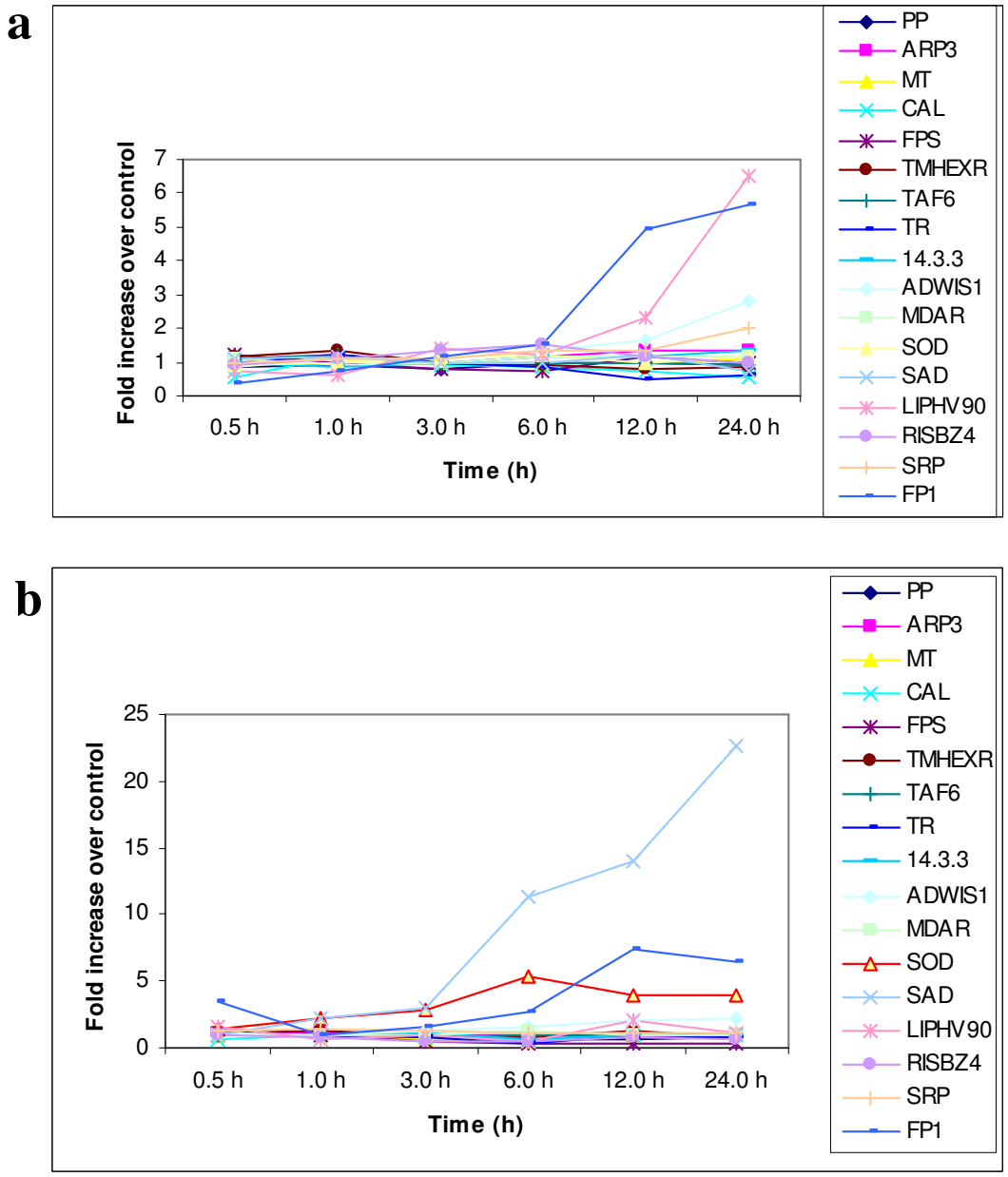


Fig 5: E-northern analysis of Arabidopsis homologs of the selected finger millet stress genes in (a) shoot and (b) root tissues under salt stress

The sequences of the selected genes were blast analyzed against the Arabidopsis genome to identify their homologs in Arabidopsis (AGI codes). The AGI codes obtained were analyzed for their expression pattern under salt stress in shoot and root tissues using Expression Browser tool (www.bbc.utoronto.botany.ca). Using the results obtained fold change over control was calculated using the formula 2^log value.

MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase, SOD- Super Oxide Dismutase, CAL- Calumenin, TMHEXR- Transmembrane Helix Receptor

(b) Measurement of leaf solute potential and net osmotic adjustment

As field capacity of the soil system supporting the crop reduced, due to alterations in the plant water status, the leaf osmotic potential as well as net osmotic adjustment increased (Table 8). The leaf solute potential decreased from -1.01 MPa under well watered situation to -2.87 MPa at 20% FC. However, the osmotic adjustment made by the crop was sufficient enough to cope up for the stress effect. This is a common adaptation seen in all crops in conjunction with any abiotic stress leading to a change in plant water status.

(c) Assessment of loss of membrane integrity

To analyze the extent of stress damage caused to the plant system, membrane integrity was estimated at the end of the stress treatment period. Damage to membranes increased with increase in stress intensity, with an extent of leakage varying from 11% under control conditions to about 25.60% at 20% FC (Fig 14). But there was a relatively lower range in the extent of leakage from approximately 11 to 18% across different stress levels.

(d) Quantification of percent reduction in total chlorophyll content

The total chlorophyll content also reduced on stress imposition. The reduction in total chlorophyll content over control increased from 18% at 80% FC to around 60% at 20% FC, giving a clear indication of stress effect (Fig 15).

(e) Estimation of cell viability

Along with membrane damage and reduction in total chlorophyll, there will generally be a simultaneous decrease in cell viability. As a result of stress imposition in this study, the reduction in cell viability was shown to be 16 % at 80% FC stably reducing but only up to 56% at 20% FC when compared to well irrigated control (Fig 16).

4.2.2.3 Imposition of abiotic stresses in seedling system

In another stress experiment, salt stress (NaCl: 300 mM), osmotic stress (PEG: -8 bars) and oxidative stress (methyl viologen: 10μM) treatments were imposed to finger millet seedlings grown in petri dishes. The shoot and root tissues of the treated seedlings were collected and used for expression studies. The root and shoot lengths were recorded to assess the extent of stress effect (Figure 17 and Plate 7).

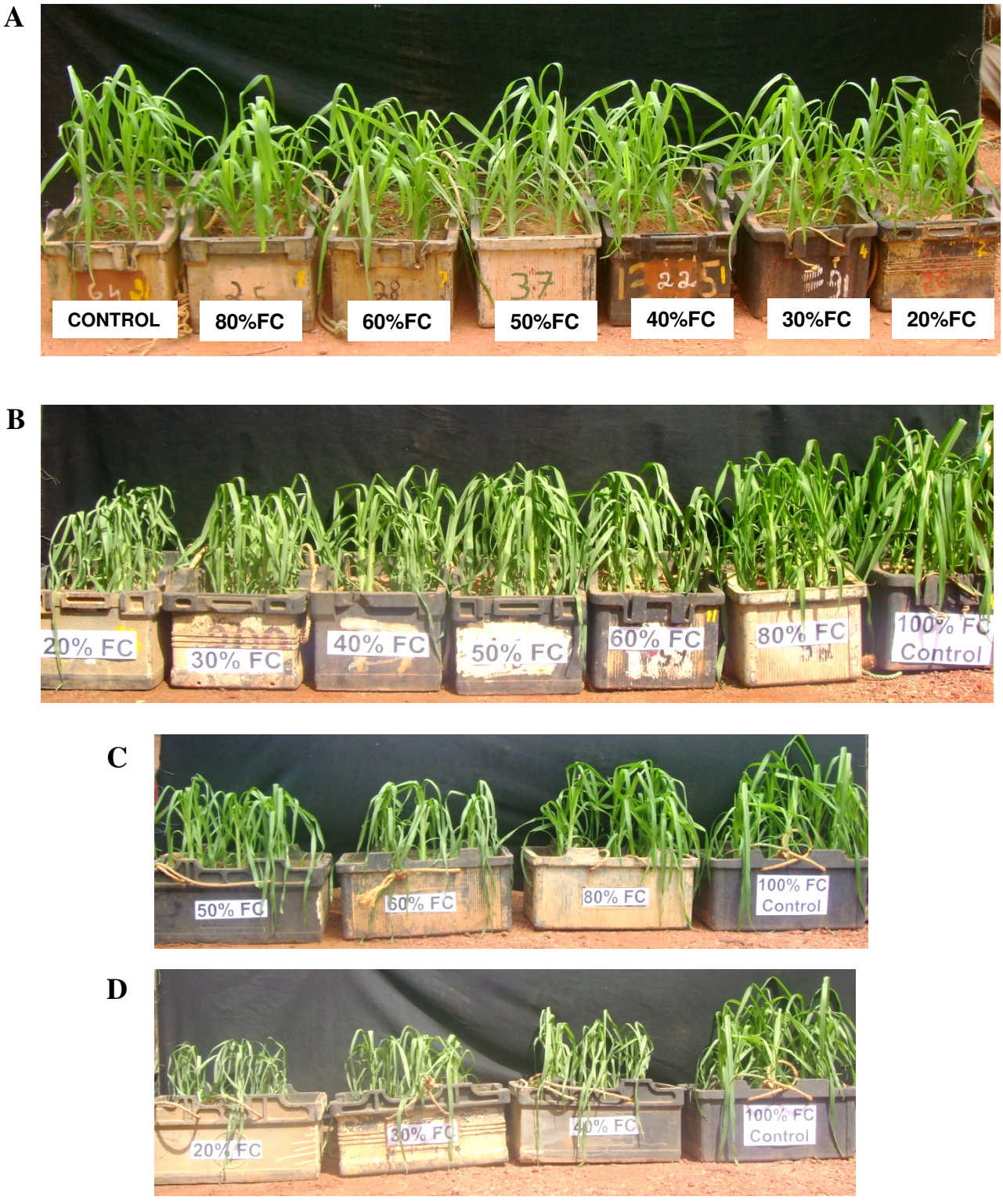


Fig 11: Imposition of drought stress in finger millet by gravimetric approach. A – Healthy plants before stress imposition; B- Stressed plants after stress imposition; C, D- Different stress treatments in comparison to control

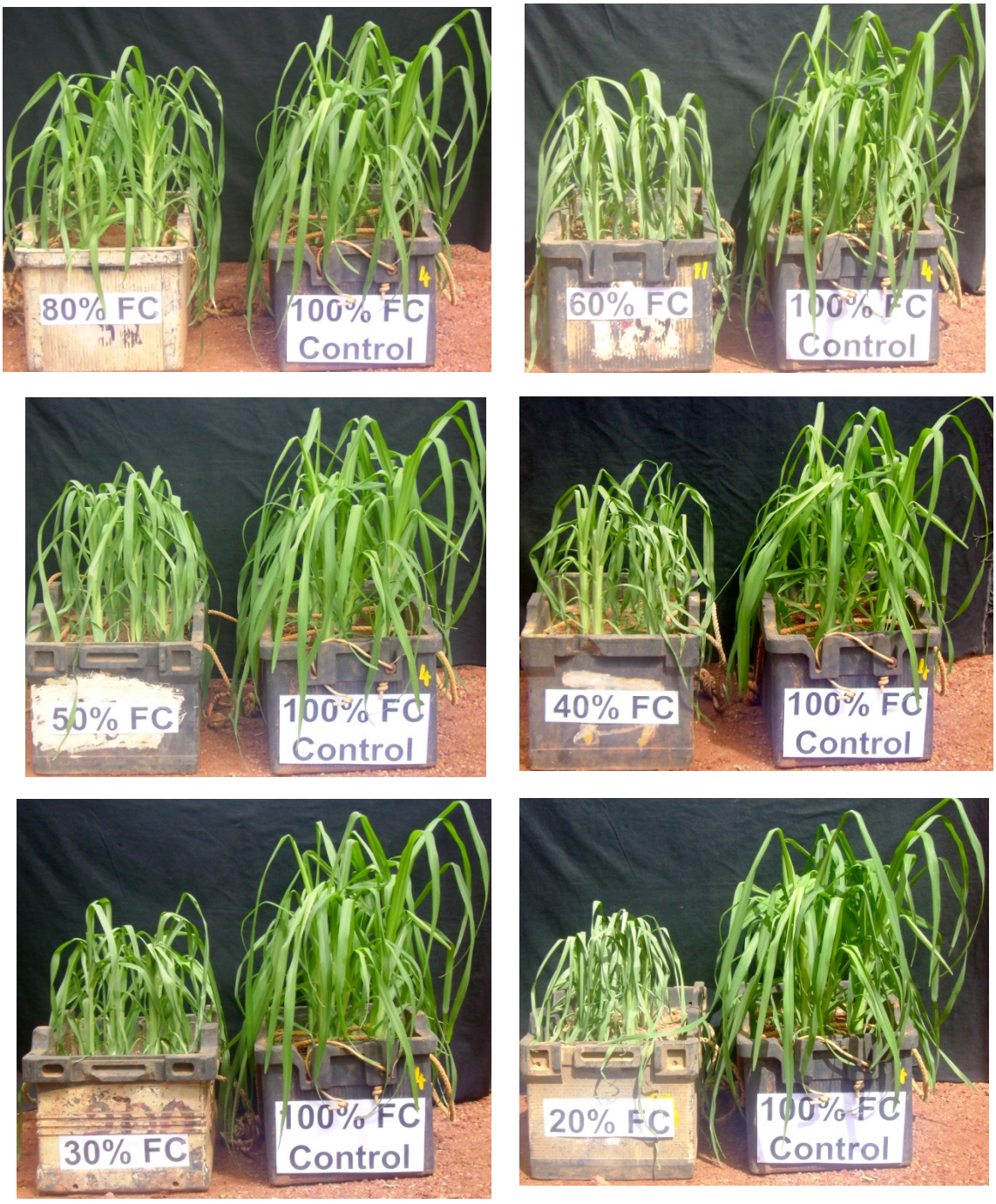


Fig 12: Phenotype of finger millet plants experiencing drought stress at whole plant level

4.2.3 Assessing drought effects by molecular assays

Expression analysis of stress responsive gene – ELIP

Total RNA was isolated from finger millet stressed as well as non stressed leaf tissues and reverse transcribed into cDNA. Semi quantitative RT-PCR was performed using the cDNA as template with varying PCR cycles (20, 25, 30, 35 cycles) using control and 20 % FC samples to look into the expression pattern of a stress responsive gene, early light inducible protein (ELIP) . ELIP was up regulated under stress when compared to well irrigated conditions (Plate 8). In all the RT PCR experiments, house keeping gene, actin was used as internal control.

4.2.4 Expression pattern of selected genes under drought stresses

4.2.4.1 Semi-quantitative RT-PCR analysis

On the basis of available sequence information, gene specific primers were designed. The selected clones were subjected to homology analysis using BLAST (Basic Local Alignment Search Tool) bioinformatics tool (www.ncbi.nlm.nih.gov/). The homologous sequences were aligned using ClustalW2 software (http://www.ebi.ac.uk/Tools/clustalw2/index.html). The primers used in the study are given in Table 9. Out of the 22 selected genes only 14 were selected for expression studies. However, due to the noise in PCR amplification, only ten genes were used for further expression studies. Annealing temperature for each primer pair was standardised using control cDNA as template. The standardised annealing conditions are presented in Table 9.

To standardise the number of cycles required for amplification, initially, signal recognition particle receptor was subjected to semi quantitative RT-PCR using cDNA as template with varying PCR cycles (15,20,23,25,30 cycles) using control and stressed (50 % FC) samples (Plate 9). It was found that clear and specific amplification happened at 30 cycles and was standardised for carrying out RT PCR for all genes. The reaction was further carried out for all the selected genes using first-strand cDNA, synthesised from well irrigated and stressed RNA samples, as template for 30 PCR cycles. Most of the genes were found to show enhanced expression at 60% FC indicating early stress responsiveness. The transcript levels of early light inducible protein, RISBZ and

Table 9: List of stress responsive genes and primers used for expression analysis

Primers for the expression analysis of selected clones					
Sl. No	**Name**	**Primer sequence 5' - 3'**	**Mer**	**Tm (^{0}C)**	**Product size (bp)**
1	MT F	GTTGCAAGTGCGGCTATATG	20	55	211
2	MT R	GAAGCAGCAAGCACCCATG	19		
3	RISBZ4 F	CCCCGACGGAGCATATGCAGAACTC	25	65	201
4	RISBZ4 R	GCTTCATCTCTTCACTTGGACAAGCC	26		
5	DIS1 F	CCCCTTACTGCTCCTGAGACCCGG	24	NS	209
6	DIS1 R	GTGGGTAGCTCCATCACCCACATC	24		
7	14-3-3 F	GATGATCCCTCAAGTTCCCCTGTAC	25	50	226
8	14-3-3 R	GGGTTTGACGGGTTCACCG	19		
9	SIN AD F	GCGTCGAAGACTGTGCTACTG	21	65	202
10	SIN AD R	CACACATGAAGCAACGCCCAAC	22		
11	TAF6 F	CCAACCCTGGAATGATGCAAGC	22	65	208
12	TAF6 R	GGTGGTTGAACAAAGGAGAAGATGGC	26		
13	F PROT F	GTGCCGCACCCTTACGTGC	19	56	201
14	F PROT R	CGACAGGCAACGAGACAGTGC	21		
15	AD WIS1 F	GGCGGTGAACCTTTGGAG	18	NS	178
16	AD WIS1 R	GCGCATTCATAATTTCACAGGC	22		
17	SIG RP F	GATAGGCTGATGACCAAGAATGTGGC	26	66	222
18	SIG RP R	GACATAAGGCCTCCCACGCTCC	22		
19	T REG F	GAACGTCAGACTCGTCGCC	19	66	258
20	T REG R	CGAAGTCTGTGGTGCGATCCAG	22		
21	PROT PHOS F	GGCGGCTACGGCTTCGGCG	19	58	210
22	PROT PHOS R	GGCAGCGCACGGGCTGCACG	20		
23	LIP HV 90 F	GGTGCCACTGCTCCAGGG	18	66	235
24	LIP HV 90 R	CCGCCTGATATACGTCATACACG	23		
25	F PYR SYN F	GACGAGCTCCGAGGAGTGC	19	62	203
26	F PYR SYN R	CAATAACAGAGAGCCCGCGATTGC	24		
27	MDAR F	GAGAAGGTGCTCGAGACTG	19	NS	231
28	MDAR R	GTACATAAGCAATCGTACATGGTC	24		

Note: F- Forward primer; R- Reverse primer; MT- Metallotheionin; RISBZ4- bZIP 9; DIS1- Distorted trichomes; 14-3-3- 14-3-3 protein; SIN AD- Sinapyl alcohol dehydrogenase; TAF6- TBP Associated Factor 6; FPROT- Farnesylated protein; ADWIS1- Alcohol dehydrogenase; SIG RP- Signal recognition particle receptor; T REG- Transcriptional regulator; PROT PHOS- Protein phosphatase; LIP HV90- Light inducible protein; F PYR SYN- Farnesyl pyrophosphate synyhetase; MDAR- Mono dehydro ascorbate reductase; Tm- Annealing temperature (standardised).

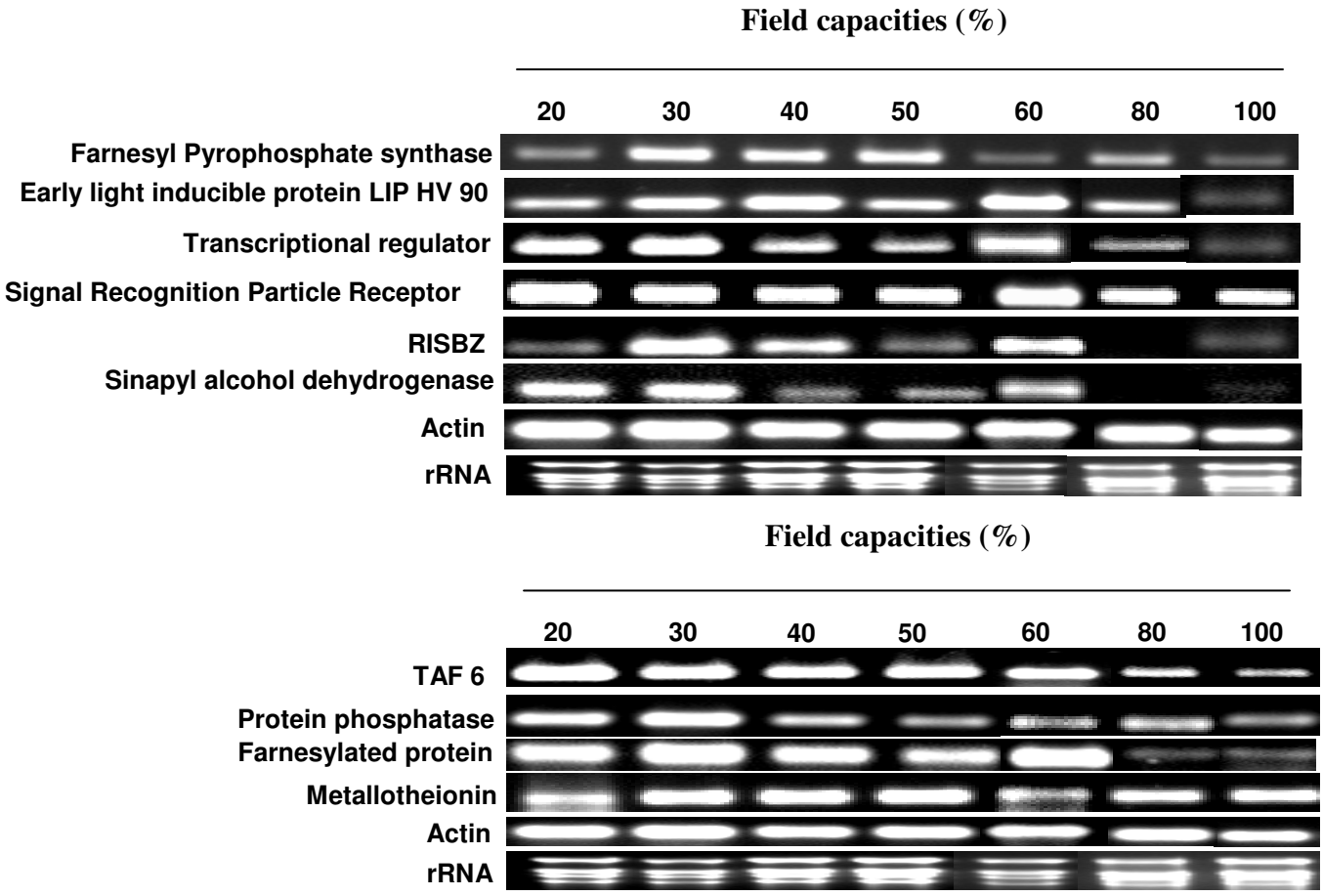


Plate 10 : Semi quantitative RT PCR to study the pattern of expression of stress responsive genes cloned from finger millet under 30 cycles. RNA was extracted from stressed (20, 30, 40, 50, 60 and 80% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.

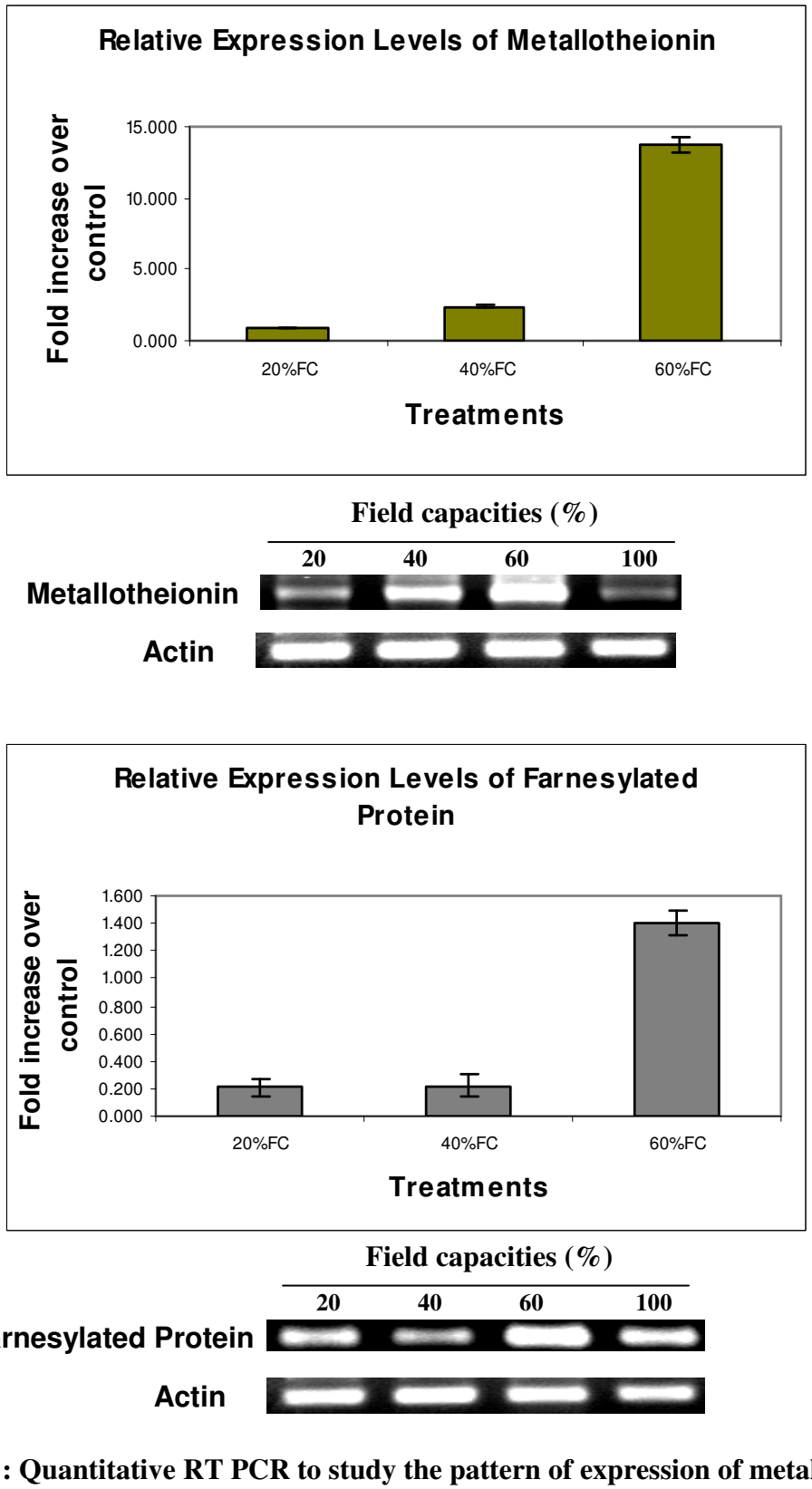
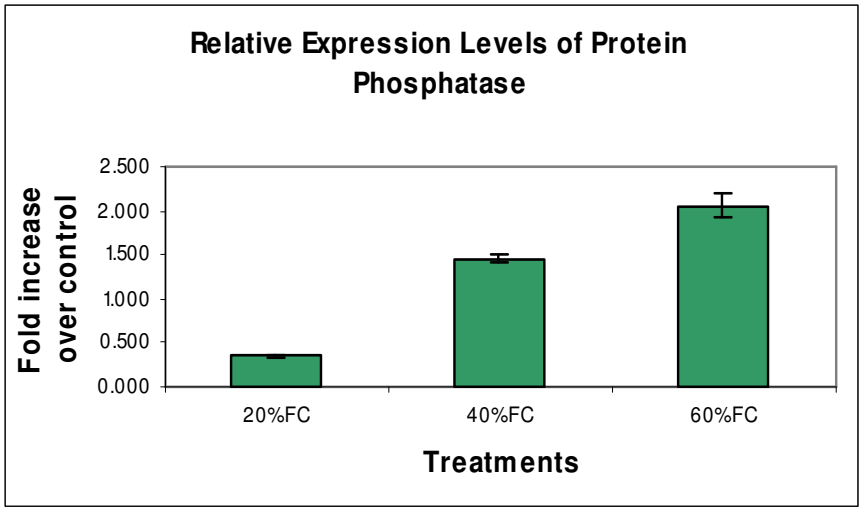
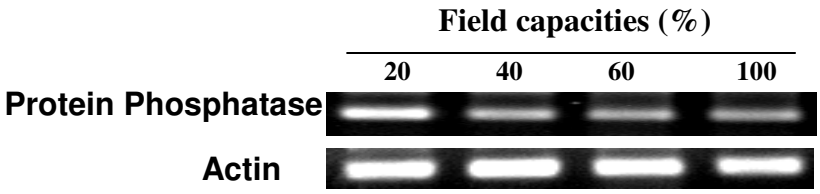
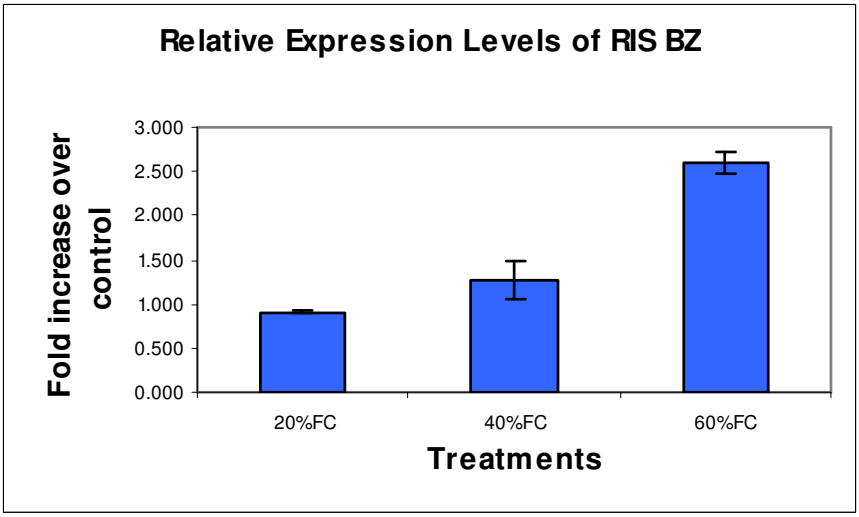


Plate 11 : Quantitative RT PCR to study the pattern of expression of metallotheionin and farnesylated protein cloned from finger millet. RNA was extracted from stressed (20%, 40%, and 60% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.







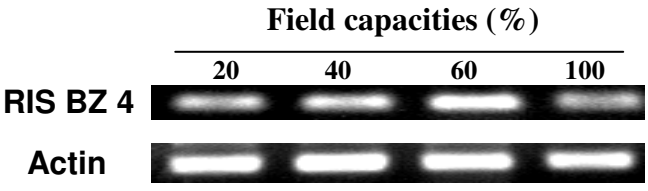
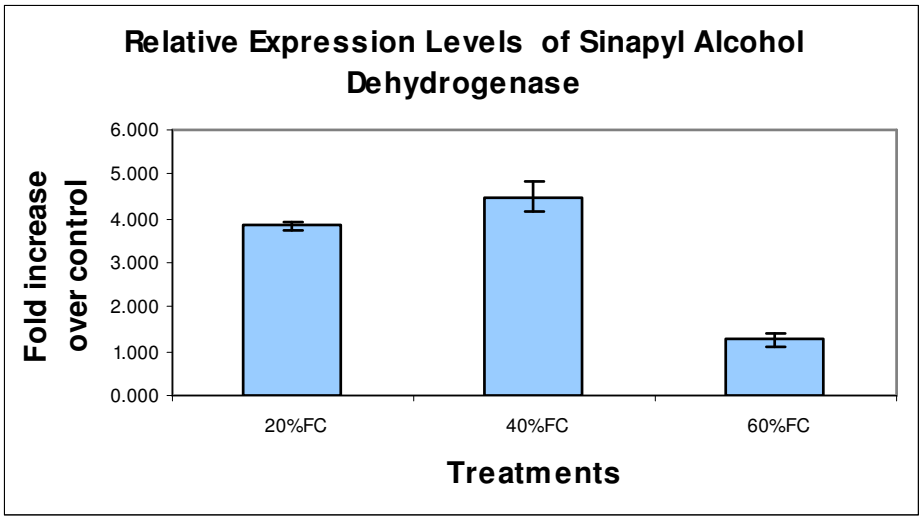
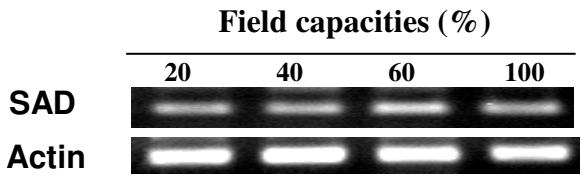
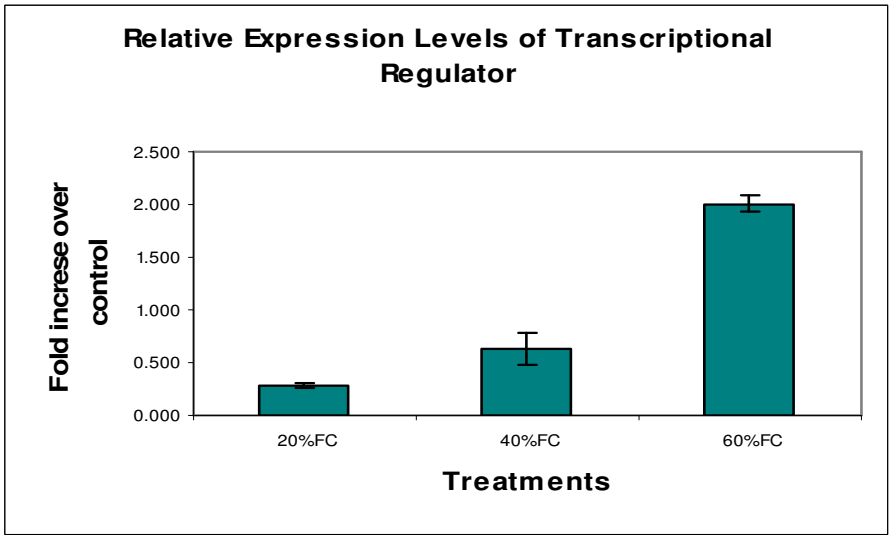


Plate 12 : Quantitative RT PCR to study the pattern of expression of protein phosphatase and RIS BZ4 cloned from finger millet. RNA was extracted from stressed (20%, 40%, and 60% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.







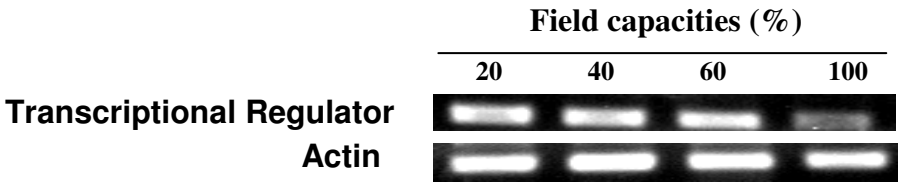
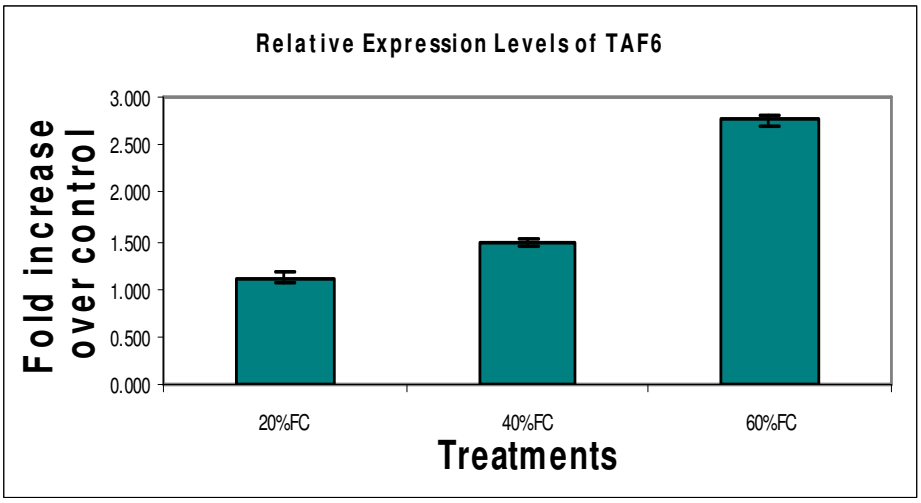
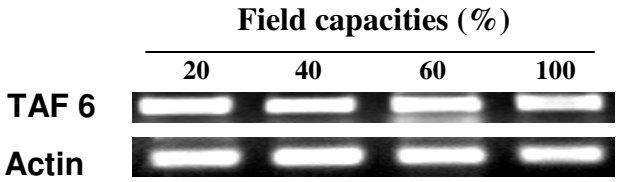
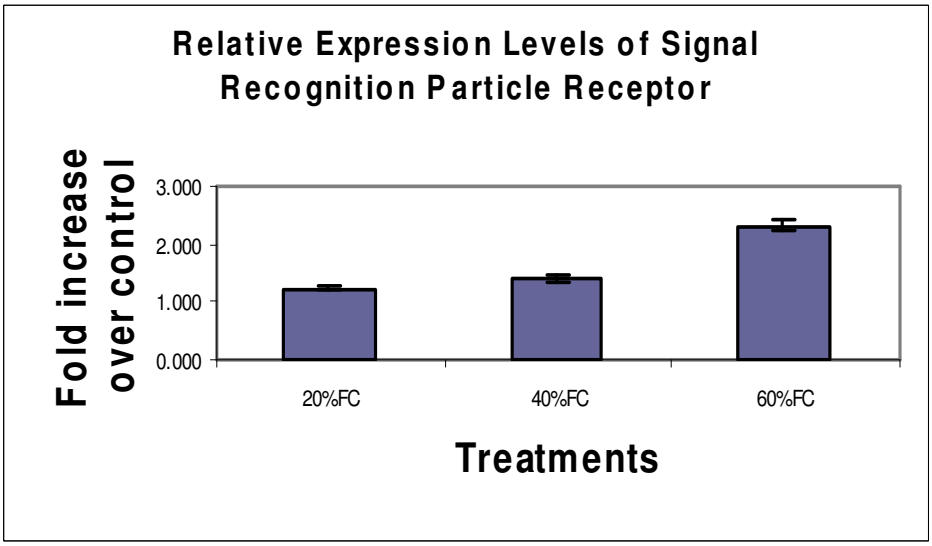


Plate 13 : Quantitative RT PCR to study the pattern of expression of sinapyl alcohol dehydrogenase and transcriptional regulator cloned from finger millet. RNA was extracted from stressed (20%, 40%, and 60%FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.







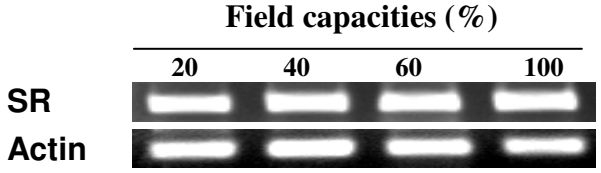


Plate 14 : Quantitative RT PCR to study the pattern of expression of TAF6 and Signal recognition particle receptor cloned from finger millet. RNA was extracted from stressed (20%, 40%, and 60% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.

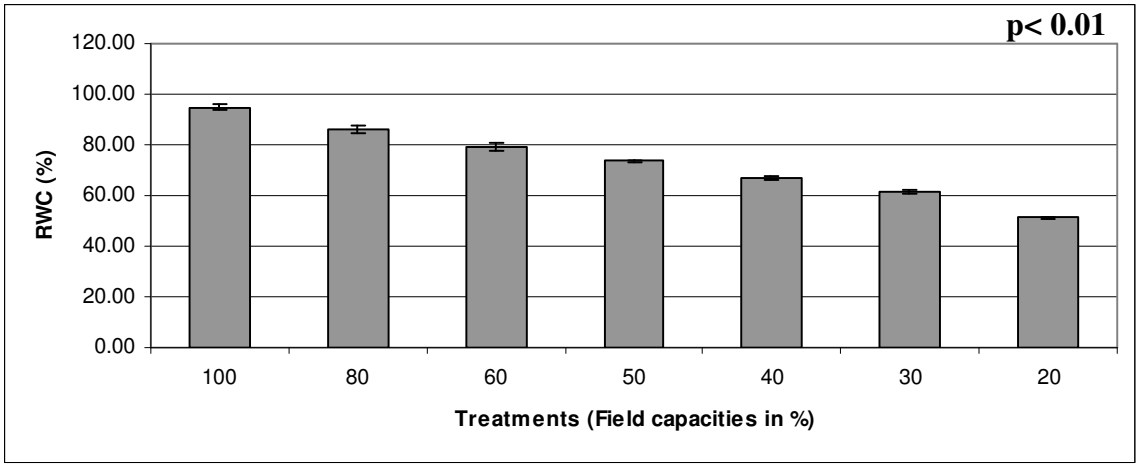


Figure 13 : Relative water content (RWC) of finger millet leaf tissue experiencing different levels of drought stress

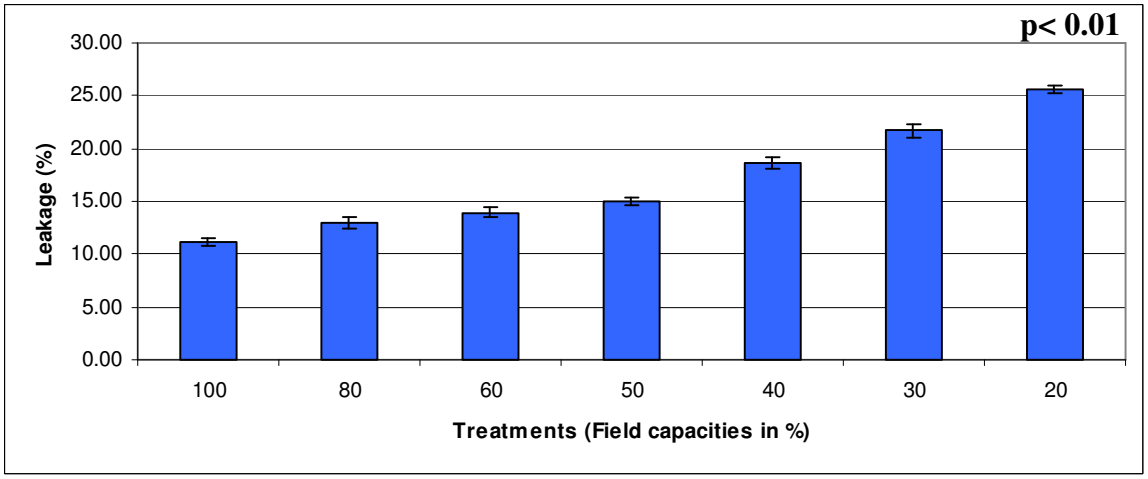


Figure 14 : Loss of membrane integrity of finger millet leaf tissue experiencing different levels of drought stress

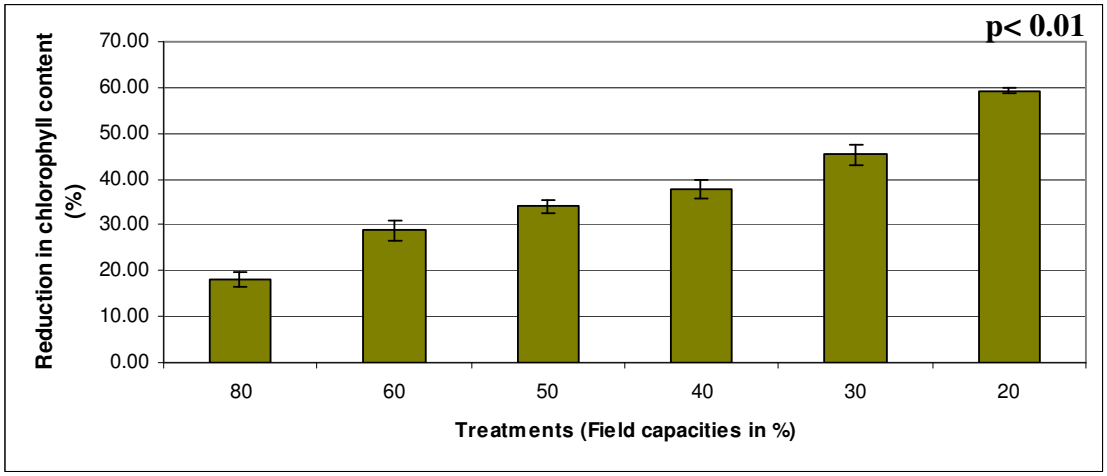


Figure 15 : Reduction in total chlorophyll content (R %) in finger millet leaf tissue experiencing different levels of drought stress

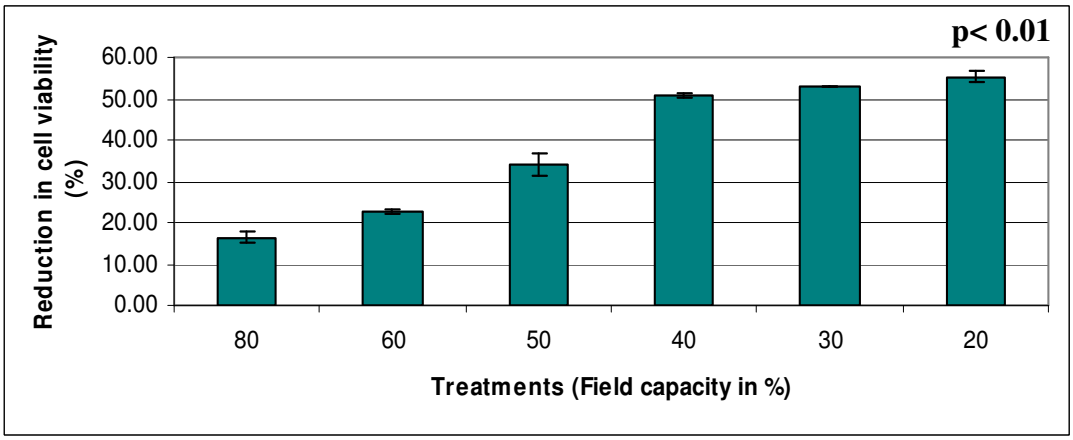


Figure 16 : Reduction in cell viability in finger millet leaf tissue experiencing different levels of drought stress

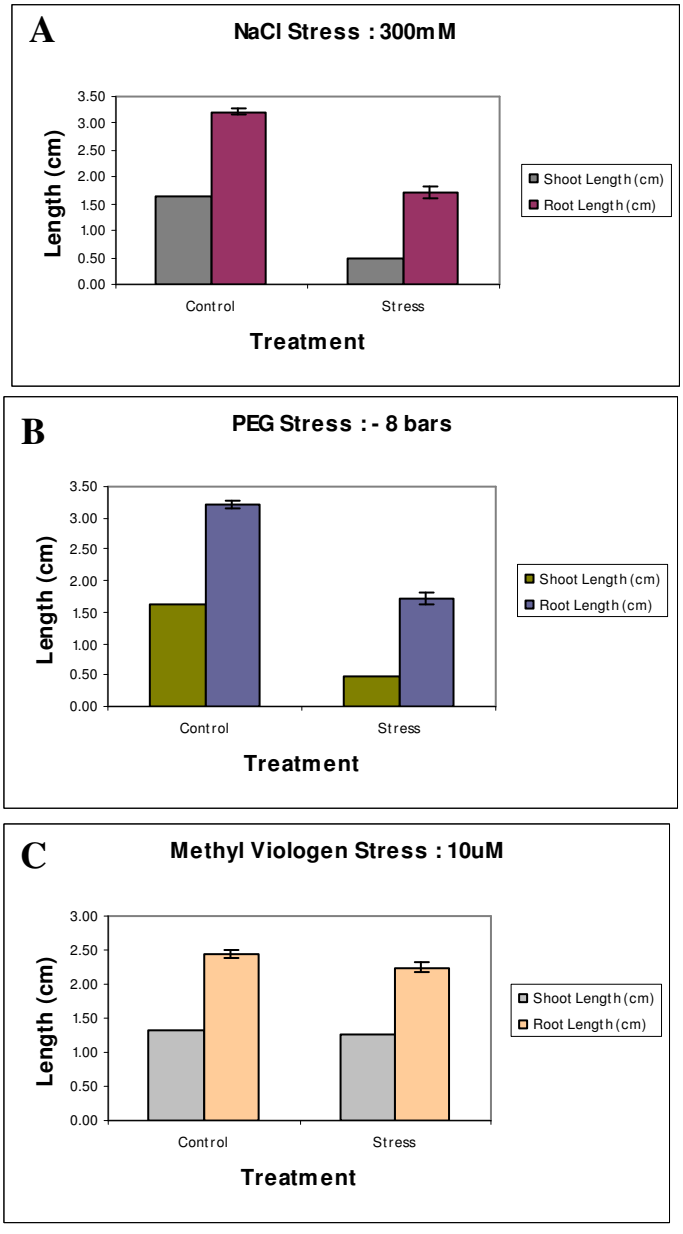


Figure 17 : Seedling assays: Reduction in growth on stress imposition on finger millet seedlings using NaCl (300 mM), PEG (- 8 bars) and Methyl Viologen (10 µM): Two day old finger millet seedlings were subjected to 24 hours stress treatment with 300 mM NaCl, - 8 bars PEG and 10 µM Methyl Viologen and the shoot and root tissues were harvested for expression studies.

A, B, C- Shoot and root lengths of control and stressed tissues under treatments : 300 mM NaCl, -8 bars PEG and 10 µM methyl viologen respectively.

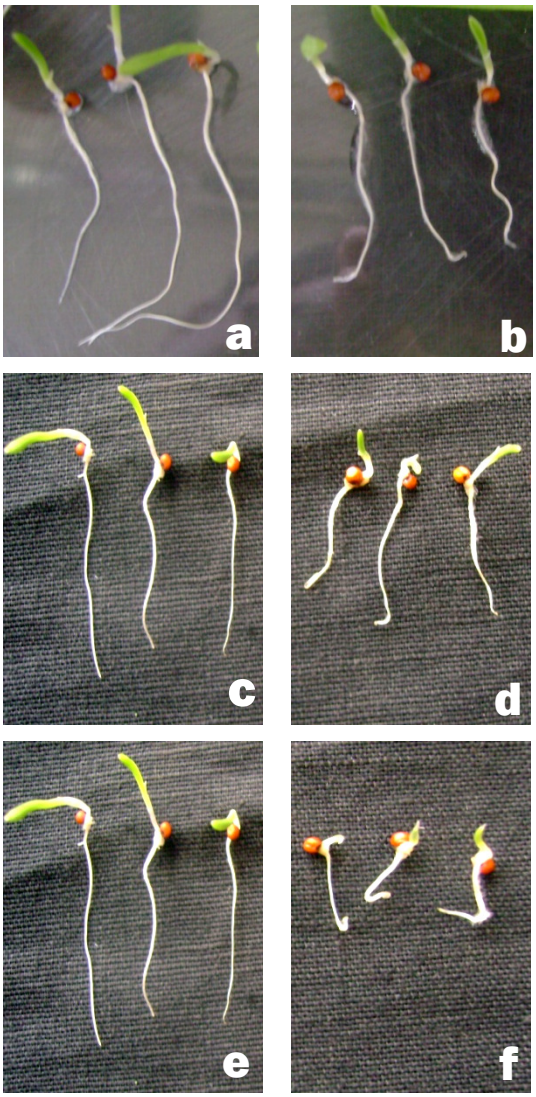


Plate 7 : Effect of stress imposition on finger millet seedlings using using NaCl (300 mM), PEG (- 8 bars) and Methyl Viologen (10 µM): Two day old finger millet seedlings were subjected to 24 hours stress treatment with 300 mM NaCl, - 8 bars PEG and 10 µM Methyl Viologen and the shoot and root tissues were harvested for expression studies.

a, c, e – seedlings grown in water (control); b, d, e- seedlings grown in 300 mM NaCl, -8 bars PEG and 10 µM Methyl Viologen

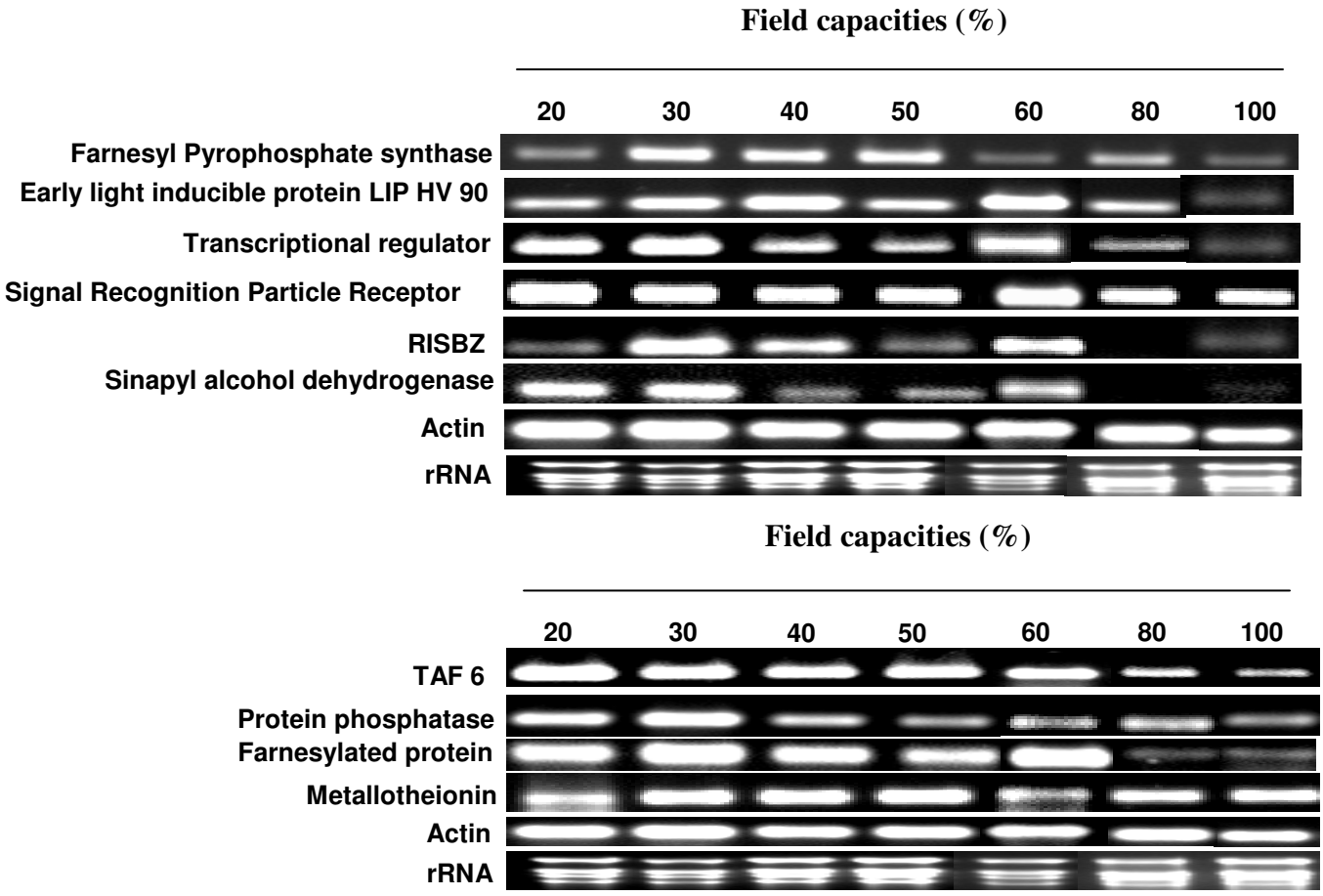


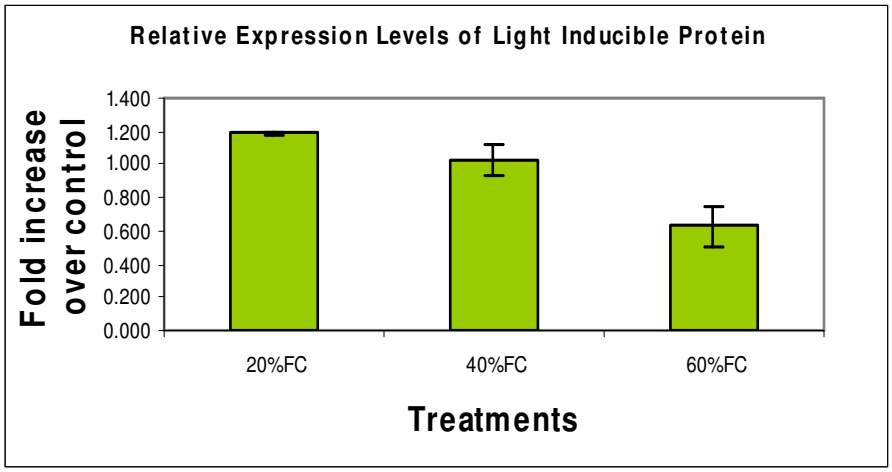
Plate 10 : Semi quantitative RT PCR to study the pattern of expression of stress responsive genes cloned from finger millet under 30 cycles. RNA was extracted from stressed (20, 30, 40, 50, 60 and 80% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.

transcriptional regulator were found to be high, relatively earlier, when plants reached 60% FC status. However, there were a few genes up regulated at 30% FC also. Genes like metallotheionin, farnesylated protein, protein phosphatase and farnesyl pyrophosphate synthase were found to show high expression levels initially at 60% FC and later at 30% FC further dropping down at 20% FC. The other genes like sinapyl alcohol dehydrogenase, signal recognition particle receptor and TAF6 were found to have a gradual increase as field capacity approached to 20% FC (Plate 10). To confirm the gene identity, restriction digestion analysis was performed for all the genes using PCR amplified products. The details regarding the various restriction enzymes used, their restriction sites as well as the restriction fragment sizes are given in Appendix IV. The gene identity was confirmed by the fragment sizes after restriction digestion (Appendix VI).

4.2.5 Quantitative RT-PCR analysis

For quantification of the expression levels of selected genes, quantitative real time PCR analysis was carried out for cDNA extracted form RNA of 20%, 40%, 60% and control samples. Real time PCR was performed using SYBR green dye binding assay in real time PCR machine (Opticon 2, M J Research, USA). Equal concentration of cDNA template was used and house keeping gene, actin was used as internal control for normalisation. The critical threshold cycle (C_t) value of the individual genes were normalised with internal control actin and fold increase or decrease in expression levels over control was computed. Melting curve analysis revealed and helped to confirm specific amplification ruling out background amplification (Plates 11- 15).

In conformity with semi quantitative PCR results, almost all genes like metallotheionin, protein phosphatase, farnesylated protein, transcriptional regulator etc. showed higher fold increase over control in samples collected from plants grown at 60% FC with their levels coming down in those at 20% FC. However, sinapyl alcohol dehydrogenase, light inducible protein, signal recognition particle receptor and TAF6 showed slightly altered expression patterns. The major advantage of executing quantitative RT PCR was that the relative expression levels of the various genes could be calculated as fold increase over control thereby throwing light over its stress responsiveness. Metallotheionin was shown to be highly up regulated at 60% FC with about 14 fold increase in transcript levels over



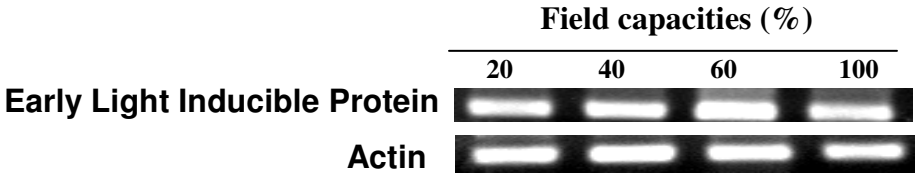


Plate 15 : Quantitative RT PCR to study the pattern of expression of Light inducible protein cloned from finger millet. RNA was extracted from stressed (20%, 40%, and 60% FC) and non stressed (control) leaf tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.

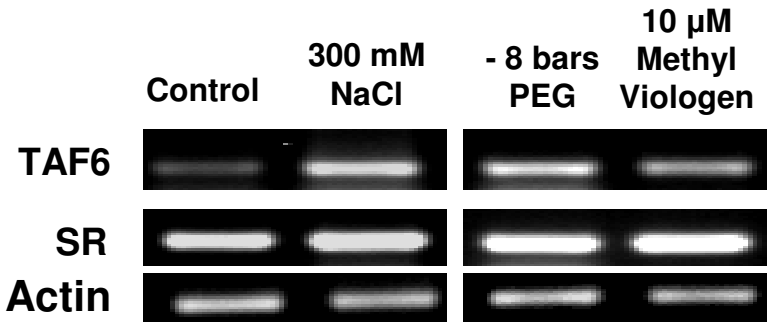


Plate 16: Semi quantitative RT PCR of TAF6 and SR in seedling (shoot) system. RNA was extracted from stressed (300 mM NaCl, -8 bars PEG and 10µM methyl viologen) and non stressed (control) shoot tissue, reverse transcribed into cDNA and used as template. Actin was used as internal control.

CTGCGTCTTCTATGGCAGGAAACATGCAAGGAACCATGGATGGGTTCTCCACCCAGTTAGCCAACCCTGGAATGATGCAAG
CTTCATCCACTTCACAAATGGTGGAAAACATTGCAGTTGCGGTTCGGAGAGAACAGGGCAGTGATCTTGCGCAGAGGGTTT
CCACGGTGCTGAGGCAAGCATGGAAGGAGGATCAAGATGCTGGCCATCTCCTCGGGTCCTTGTACGAGGTCTTTGGTGAAG
CCATCTTCTCCTTTGTTCAACCACCAGAGATATCCATCTTTGTGTAGAAAGGAAATGCGTAGGCTTTAGAGGATTTTAATC
TGCACTGTTGTATAATTAAGGGTATTCTTGAAATGGTAGACTGACATCAGAAAACAAGTAATGTAATGCTAGGTTTATATT
AATGTGAGGAATGAACATGTGTGTTGTACACTTGTACTATTATTACTTTACGTGACTGATTATCTTAGCCATTTGGCTACT
AATGTATACGTACATCGTTTAAA

Coding Sequence:
Pred. Protein1: none
Colored Sequence:
cugcgucuucuauggcaggaaacaugcaaggaaccauggauggguucucc 50
acccaguuagccaacccuggaaugaugcaagcuucauccacuucacaaau 100
gguggaaaacauugcaguugcgguucggagagaacagggcagugaucuug 150
cgcagaggguuuccacggugcugaggcaagcauggaaggaggaucaagau 200
gcuggccaucuccucggguccuuguacgaggucuuuggugaagccaucuu 250
cuccuuuguucaaccaccagagauauccaucuuuguguagaaaggaaaug 300
cguaggcuuuagaggauuuuaaucugcacuguuguauaauuaaggguauu 350
cuugaaaugguagacugacaucagaaaacaaguaauguaaugcuagguuu 400
auauuaaugugaggaaugaacaugugugguuguacacuuguacuauuauua 450
cuuuacgugacugauuaucuuagccauuuggcuacuaauguauacguaca 500
ucguuuaaa
Colored Structure:
(((((((.(((((((((....)))))....))))))))))).

Legend:
BOLD marks EXONS
UNDERLINED marks IRE or TRANSSPLICNG hits
ITALIC marks putative UTRs
RED marks SMSITES or snRNP-binding motifs
BLUE marks AU-rich regions
FUCHSIA with GREEN marks Stemm-GG-Pairs
LIME marks PolyA-signal
PURPLE marks a Promotor

Pr.A1.bin.site = Protein A1 binding site

Plate 17 : Partial sequence information and predicted structural features of TAF6. The partial length cDNA sequence of TAF6 was fed to RNA analyzer program which gives an idea about the full ORF with information on the start site, poly A tail and the UTR regions.

GGACAAGGGACAGGAGAACTATCAGCCCCGTTCTGCTCTTCTTGTCAAAGGAAAGGAGAATGGAGGTCCCAAGGACAACAACAAC
GGAGCCTTTGATGTAAATAAGTTACAGAAGCTGCGGAACAACAAGAAGAAAACTGTTCCCGCTGAAATCGGAACCAAGAAGAACA
CTACTAAACCTGACACTAAGAAGAAGGTCAACAGGGTCTGGGATGACAAACCTACCGACAACAAGAAGCTGGATTTTACTGACCC
AGCAGACAAGAGAGGGGATGATGAGGTGATCGACCATGCGCTCGTCAACCAGGGTGAGAGCATGATGGACAAGGATGAGAACGTC
AGCAGCGACAGCGAGGATGAAGAAGAAGAACAGGTGCAAGGTGTCCCCAAAAAGAAAGGATGGTTCTCCTCAATGTTTCAGAGTA
TTTCAGGTAACAATGTGCTGGAGAAGTCTGATCTGCAACCTGCGCTTAAAGCTCTTAAAGATAGGCTGATGACCAAGAATGTGGC
AGAGGAAATTGCAGAAAAGCTTTGTGAATCTGTTGCAGCTAGCCTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATCATCT
ACTGTACAGACAGCAATGGAGGAGGCTCTTGTTCGCATTTTGACCCCAAGGCGATCAATTGACATACTGAGAGATGTGCAGTCTG
CCAAGGAGCGTGGGAGGCCTTATGTCATTGTTTTCGTTGGGGTCAATGGAGTTGGCAAGTCCACCAATCTTGCAAAGGTCGCTTA
TTGGCTTCTTCAGCATAACCTCAGTGTAACACTTGCAGCGTGTGACACCTTCAGTCTGGAGCTGTTGAGCAGCTACGTACTCATG
CACGCAGCTTAAGATACCATATTTGAGAAGGTTATGAAAAGA

Coding Sequence:
Pred. Protein1: none
Colored Sequence:
ggacaagggacaggagaacuaucagccccguucugcucuucuugucaaag 50
gaaaggagaauggaggucccaaggacaacaacaacggagccuuugaugua 100
aauaaguuacagaagcugcggaacaacaagaagaaaacuguucccgcuga 150
aaucggaaccaagaagaacacuacuaaaccugacacuaagaagaagguca 200
acagggucugggaugacaaaccuaccgacaacaagaagcuggauuuuacu 250
gacccagcagacaagagaggggaugaugaggugaucgaccaugcgcucgu 300
caaccagggugagagcaugauggacaaggaugagaacgucagcagcgaca 350
gcgaggaugaagaagaagaacaggugcaaggugucccccaaaaagaaagga 400
ugguucuccucaauguuucagaguauuucagguaacaaugugcuggagaa 450
gucugaucugcaaccugcgcuuaaagcucuuaaagauaggcugaugacca 500
agaauguggcagaggaaauugcagaaaagcuuugugaaucuguugcagcu 550
agccuugagggaaaaaagcuagggucuuucacaagaauaucaucuacugu 600
acagacagcaauggauggaggcucuuguucgcauuuugaccccaaggcga 650
ucaauugacauacugagagaugugcagucugccagggagcgugggaggcc 700
uuaugucauuguuuucguugggucaauggaguuggcacguccacaau
Colored Structure:
(((((((.(((((((....)))))....))))))))).

Legend:
BOLD marks EXONS
UNDERLINED marks IRE or TRANSSPLICNG hits
ITALIC marks putative UTRs
RED marks SMSITES or snRNP-binding motifs
BLUE marks AU-rich regions
FUCHSIA with GREEN marks Stemm-GG-Pairs
LIME marks PolyA-signal
PURPLE marks a Promotor

Pr.A1.bin.site = Protein A1 binding site

Plate 18 : Partial sequence information and predicted structural features of Signal recognition particle receptor. The partial length cDNA sequence of signal recognition particle receptor was fed to RNA analyzer program which gives us an idea about the full ORF with information on the start site, poly A tail and the UTR regions.

.

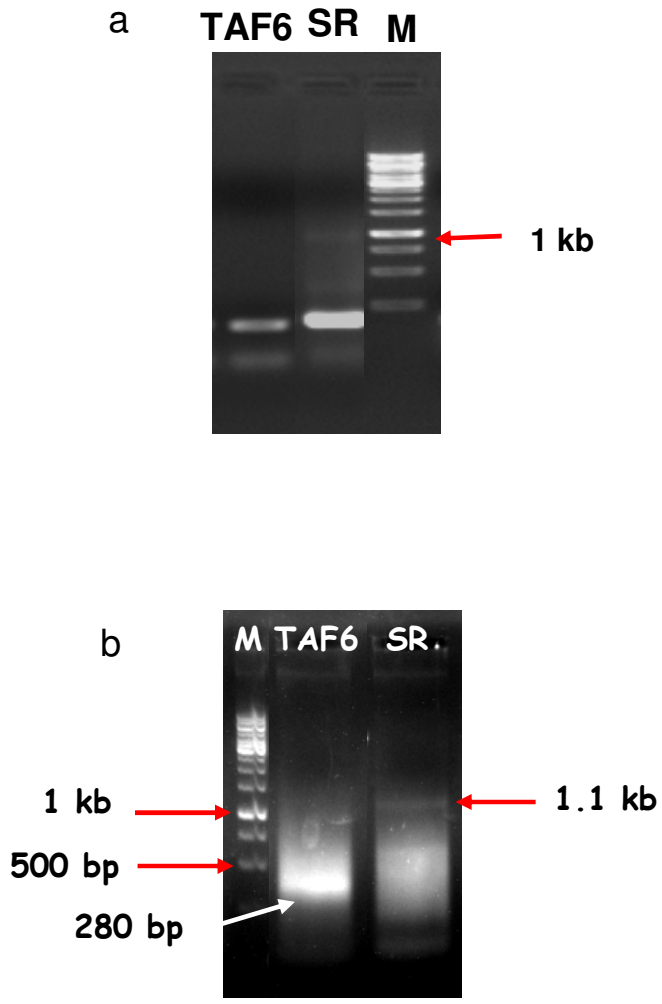


Plate 19 : Cloning of full length genes of TAF6 and signal recognition particle receptor through RACE approach a – Confirmation of gene specific fragments of TAF6 and Signal recognition particle receptor, b - 3′ end amplification of TAF6 and Signal recognition particle receptor. Total RNA was extracted from stressed finger millet leaf tissue , cDNA template synthesized as per Gene Racer protocol and random amplification of cDNA ends performed using Gene Racer primers and gene specific primers.

control. Farnesylated protein and protein phosphatase were also shown to be up regulated at 60% FC with 1.5 and 2.0 fold increase in transcripts in comparison to control. RIS BZ4 and transcriptional regulator also showed similar expression pattern with 2 to 2.5 fold increase in transcript levels. In the case of TAF6 and signal recognition particle receptor, the fold increase in transcripts doubled at 60% FC than it was at 40% FC. TAF6 transcript levels shot up to 3 fold at 60% FC from 1.5 fold at 40% FC whereas in case of signal recognition particle receptor, the fold increase in transcript levels were 1.4 and 2.3 at 40 and 60% FC respectively. In contrast to these expression patterns, sinapyl alcohol dehydrogenase and light inducible protein showed higher transcript levels at 40 than at 60% FC. In case of sinapyl alcohol dehydrogenase, the reduction in transcript levels was around 3 fold (4.4 fold at 40 to 1.2 fold at 60% FC). The quantitative data generated showed that all the genes were stress responsive and all were up regulated.

Out of the 10 selected genes, signal recognition particle receptor and TAF6 showed good expression profiles. Expression pattern of these two genes were also studied in seedling system under various abiotic stresses using semi quantitative and quantitative RT PCR. Actin was used as internal control. Both the genes were found to be significantly expressing under varying cycles under stress when compared to control as well as in stress imposed seedling system. Expression in root tissues collected from seedling tissues could not be detected significantly. In case of both the genes, there was around 2.5 fold increase in transcripts over control under NaCl stress (300 mM). However, TAF6 was found to be more up regulated under PEG stress (-8 bars) with about 2.5 fold increase over control than signal recognition particle receptor. However, the expression levels were comparatively lesser under methyl viologen stress (10 µM) although they were up regulated when compared to control (Plate 16, Figure 18).

4.3 Construction of full-length clones using RNA

The available sequence information of genes selected for full length cloning was analysed in RNA Analyzer (http://rnaanalyzer.bioapps.biozentrum.uni-uerzburg.de/server.html). The partial sequence information and the analysed data are given in Plates 17 and 18. The clustalW alignment files of the partial length sequence of finger millet aligned with full length complements of both the genes of close relative maize (*Zea mays*) is given in Appendix V D. The RACE cDNA template was synthesised as per the GeneRacer™ kit protocol and presence of the genes in the cDNA pool was confirmed by PCR

amplification with gene specific primers (Plate 19 a). The 3' ends of genes of interest – TAF6 and signal recognition particle receptor were amplified using the cDNA template with GeneRacer primers and gene specific primers (Plate 19 b). The details of the primers used for RACE are given in Appendix II. The expected 3′ end amplification of TAF6 is around 280 bp and that of signal recognition particle receptor is about 1.1 kb. The amplified products were re amplified and confirmed for their specificity. The products is to be cloned into T/A vector pTZ57R/T (Appendix IIIB) and sequenced to confirm the identity of genes cloned.

DISCUSSION

V. DISCUSSION

Plants have shown to exhibit more complexity than animals with respect to their explicit ability to withstand and forego adverse environmental circumstances. This unique phenomenon is thought to be achieved by the various adaptations that they accomplish. Among the various adaptations integrated by plants into their system, those that are realised at the cellular and molecular level are thought to be over the more intangible. However, by looking into tolerant or adapted species, one can probably probe into the basic molecular events that frame certain tolerance mechanisms under complex environmental plasticity. Adding to the overall intricacy of the issue, investigating drought tolerant mechanisms becomes tricky in the sense that drought is a polygenic trait.

Under an array of stressors, be it drought, cold, high temperature and so on, plant cell has equipped itself to display varied levels of resistance mechanisms thereby granting intrinsic cellular tolerance (ICT). Tolerance is essentially gained through regulations laid at all critical cellular checkpoints especially at transcriptional level. Hence transcript profiling under specific environmental as well as plant conditions would be the key for resourceful gene prospecting. Prospecting genes from stress adapted species than from model systems is advantageous owing to the fact that the adaptations would have occurred as a result of a novel biosynthetic pathway or a novel modified protein (Gopalakrishna *et al*., 2001).

Molecular defence mechanisms can be assessed by carefully looking into the expression patterns of certain genes whose identity is known to us. The significance of possessing the sequence information of genes, be it partial length or full length before glancing over their expression profiles is immense (Kreps *et al*., 2002; Gorantala *et al*., 2007). In the light of this fact, one can authentically proclaim the importance of identifying Expressed Sequence Tags (ESTs). Expresssed sequence tags and gain-in-function mutants are powerful tools in this regard. ESTs from drought adapted crops like finger millet is thought to be a valuable asset, where complete genome information is unavailable. In plants, cDNA libraries are widely being used as the choice material for transcriptome profiling (Seki *et al*., 2002). EST generation is made possible through cDNA library construction which has to be thoroughly screened and evaluated (Mundy and Chua, 1988; Bartels *et al.,*1990; Godoy *et al*., 1990; Fernandez *et al*., 2003; Michelmore *et al*., 2003).

From this context, a drought stress specific cDNA library was developed from finger millet, a drought adapted crop plant, maintained at different levels of drought stress (Ramegowda, 2008). Since the cDNA library was developed from plants subjected to drought stress simulating field situations, it forms a more viable and realistic approach to induce stress response than stress imposition at laboratory conditions. In this scenario, screening, evaluation and validation of stress responsive genes from the available cDNA library assumes significance.

The identification of stress specific genes alone does not answer the issue of stress tolerance but understanding their functionality and importance is more essential. There are innumerable functional and regulatory genes that act as key players in specific signal transduction cascades that would advance our knowledge leading to a better understanding of the regulation of the plant stress transcriptome. In house experiments in the department of Crop Physiology have proved groundnut, finger millet and horse gram to be better adapted to drought (Gopalakrishna, 2001; Geetha-Govind, 2006). Hence, highly stress responsive novel genes identified from finger millet, a drought adapted crop can be used as candidate genes for improving drought susceptible species. Thus, this study has aimed at discovering stress responsive genes from stress adapted species finger millet.

Discovering drought stress responsive genes from finger millet

A very important approach in functional genomics is identification of stress responsive genes of functional significance based on ESTs generated out of cDNA libraries developed from plants subjected to specific stresses. The factor which stands out with extreme priority is the precision with which stress imposition is excercised followed by accurate tissue harvest stage and efficient RNA isolation from stressed tissues for transcriptome profiling. For clear cut stress imposition, a feasible protocol which mimics field situation was adopted which involves exposing the plants to stress by gradually withholding irrigation using gravimetric approach (Sheshshayee *et al*., 2005). Hence in the present study, cDNA library constructed out of an mRNA pool from all stress levels imposed - 80, 60 and 30% field capacities were used for screening and evaluation.

Identification of moisture stress responsive genes from drought stress specific cDNA library:

The cDNA library was screened to identify diverse clones which would have a putative function under drought stress. Identification of stress responsive genes by plaque lifting followed by PCR confirmation and sequencing was resourceful enough. The functionality of the different clones was to be carefully studied. The annotation of the sequencing results revealed that the genes cloned out of drought stressed tissues belonged to varied categories like protective proteins (10%) like chaperones, heat shock proteins, dehydrins etc.; cytoskeletal structural genes (10%) like actin related proteins; catalytic genes (9%) like kinases, phosphatases etc.; transcription factors (4%) and other upstream regulatory genes (Fig 1). All these would be having roles in favouring and enhancing ICT.

Validation of stress responsiveness of the cDNA clones by expression studies

The expression profile of more than 50 finger millet cDNA clones identified was examined by *in silico* analysis, using the web based Expression Browser (www.bbc.utoronto.botany.ca) which utilises the micro array data of *Arabidopsis* cDNAs (Toufighi *et al*., 2005). Expression profile of *Arabidopsis* homologs of finger millet clones expressed under drought, radiation, genotoxic, oxidative, cold, salinity, osmotic, and heat stresses was studied. The results of e-northern analysis suggest that At5g66110 (homolog of farnesylated protein), At4g30600 (homolog of signal recognition particle receptor, At4g14690 (homolog of Light inducible protein), At4g33930 (homolog of sinapyl alcohol dehydrogenase), At5g66110 (homolog of farnesylated protein), At4g14690 (homolog of Light inducible protein), At1g64720 (homolog of transcriptional regulator) were found to be induced and highly expressible under various stresses across different time points considered in the study (Fig 2 -10). The expression analysis of various genes across time periods helps us to arrive at an accord that the library generated diverse genes responsive to a range of abiotic stresses and hence may have specific role in various physiological processes that understate stress acclimation. Microarray data coupled with semi quantitative RT-PCR or quantitative real time (qRT-PCR) or RNA gel blot analysis would reconfirm the stress responsiveness of the genes (Rabbani *et al*., 2003). In this regard, further analysis of stress responsive nature of the genes was undertaken after confirming the clone identities by re-sequencing. The sequence information of the clones selected was used to design primers for semi-quantitative and quantitative RT PCR (Table 9).

Evaluation of expression pattern of selected genes in stressed tissues

To examine the expression pattern of specific stress responsive genes, drought stress was imposed at whole plant level in battery pots by gravimetric approach to simulate field conditions (Fig 11 &12). Evaluation of stress effect on overall plant water status and other biochemical and physiological processes was done to ascertain the stress effects. The most important parameter for assessing the imposed drought stress was to measure soil water potential. The soil water potential was shown to decrease as the stress intensity increased and was very low at 20% FC (Table 7). Plants water relations were assessed by estimating RWC, leaf solute potential and net osmotic adjustment. The RWC was found to reach critical levels of 51% under severe drought stress (Fig 13). The leaf solute potential and osmotic adjustment was found to increase as stress levels increased (Table 8). The membrane damage due to stress effect was was observed phenotypically (Fig 11 & 12). There was also significant reduction in total chlorophyll content and cell viability under drought stress (Fig 15 & 16). However, the leakage due to loss of membrane integrity and reduction in chlorophyll content were not in congruence. This suggests that finger millet plants are able to maintain turgor driven processes due to enhanced osmotic adjustment thereby maintaining membrane integrity under severe stress conditions. These observations clearly indicated the accuracy of drought stress imposition by gravimetric approach at whole plant level in finger millet, which led to the activation of a set of genes associated with drought response. From earlier studies it has been shown that Early Light Inducible Proteins (ELIP), a thylakoid protein gets up regulated under stress (Meyer and Kloppstech, 1984; Adamska, 2001; Hutin *et al*., 2003). In the present study, ELIP expression pattern was analysed in stressed leaf tissue in comparison to control. The gene expression was found to be highly up regulated in stressed tissues as evidenced by RT PCR assay (Plate 8). This gives an indication that the cells are experiencing stress and the perception of stress signals activated stress responsive genes.

Semi quantitative expression analysis of selected drought stress responsive genes

Gene expression was carried in tissues exposed to varying degrees of drought stress by semi quantitative RT PCR and quantitative Real Time PCR (qRT-PCR). The PCR conditions were standardized by using cDNA as template (Plate 9) to avoid non-specific amplification.

The expression patterns of genes across the different stress treatments were compared with the control after 30 cycles of PCR amplification (Plate 10). There was up regulation of eight out of the ten selected genes in tissues exposed to 60% FC showing their early responsive nature. The genes like metallothionein, farnesylated protein, protein phosphatase and farnesyl pyrophosphate synthase seems to be highly stress responsive as there was significant increase in the transcript level under moderate as well as severe stress. Metallothioneins (MTs) are one among the different metal binding factors that play a major part in the processes of plant metal homeostasis and detoxification (Cobbett and Goldsbrough, 2000) and are also reported to be involved in ROS scavenging (Wong *et al*., 2004). Multiple metallothioneins have been cloned from hybrid poplar (Kohler *et al*., 2004). MTs are found to be expressed in barley during senescence and heavy metal treatment (Heise *et al*., 2007). They were also found to be differentially expressed in mangrove species under heavy metal stress (Usha *et al*., 2007). *Arabidopsis thaliana* metallothionein, AtMT2a, is reported to mediate ROS balance during oxidative stress (Zhu *et al*., 2009). Farnesyl pyrophosphate synthase has been reported to be an essential cytosolic enzyme in eukaryotes that confers resistance to unfavourable chemicals (Ortiz-Gomez *et al*., 2006). Farnesylated proteins have been reported earlier to be involved in attachment of proteins to membranes thereby maintaining memebrane integrity and having regulatory roles in plant growth and development (Taylor, 1996). Genetic studies in *Arabidopsis* have shown that farnesylation is important in ABA signalling (Nambara *et al*., 1999). Protein farnesylation also plays a role in flower development by regulating the activity of a protein that determines flower meristem identity in *Arabidopsis* (Yalovsky *et al*., 2000). At the same time, sinapyl alcohol dehydrogenase, signal recognition particle receptor and TAF6 were found to have a gradual increase as field capacity approached to 20% FC. However sinapyl alcohol dehydrogenase is found to be involved in lignol biosynthesis which occurs during plant cell wall formation (Li *et al*., 2001). Recent reports from sorghum suggest that these genes are part of the cinnamyl alcohol dehydrogenase family involved in lignin biosynthesis (Saballos *et al*., 2009). Hence, most of the genes are found to be significant under adverse conditions by their productive or protective functions thereby granting intrinsic cellular tolerance.

q PCR analysis of the selected genes

The quantification of expression levels of any gene can be accurately and efficiently evaluated by qRT- PCR. In the present study, to confirm the expression pattern of certain genes related to drought stress, qPCR was performed using the DNA binding dye SyBR Green under standardized PCR conditions (Sathik *et al*., 2005). The dsDNA binding dye, SYBR-Green used for real-time RT-PCR, exhibits enhanced fluorescence upon binding to a dsDNA. The SYBR Green dye gets excited at a wavelength of 485 nm and the emission is generally measured at a wavelength of 520 nm. The fluorescence signal was monitored at the end of each PCR cycle, which was required to detect the product of interest during the linear range of amplification for efficient detection of the threshold cycle (Ct). The increase in fluorescence emission above background is dependent on the initial template concentration.

The primers for qPCR were specifically designed to avoid non specific target amplification and the PCR conditions were standardised. The melting curve analysis revealed single peak plots thereby confirming that gene specific primers amplified a single product. A single dissociation peak with a Tm with in variation of 2^{o}C of the expected temperature indicated the amplification of gene of interest

Almost all genes like metallotheionin, protein phosphatase, farnesylated protein, farnesyl pyrophosphate synthase, transcriptional regulator etc., showed higher fold increase over control in samples collected from plants grown at 60% FC with their levels coming down in those at 20% FC indicating that these genes are responsive to drought stress (Plates 11-15). However certain genes like signal recognition particle receptor and TAF6 were shown to have good expression profiles. The signal recognition particle (SRP) and the SRP receptor (SR) were reported to cotranslationally deliver newly synthesized proteins from the cytosol to target membranes (Walter and Johnson, 1994; Keenan *et al*., 2001). A novel SRP-dependent protein targeting pathway has already been reported in chloroplast (Schuenemann *et al*., 1998). Chloroplast specific cpSRP pathway was characterised to be involved in posttranslational mode of targeting thereby transporting fully translated LHCPs from stroma to the thylakoid (Delille *et al*., 2000; Tu *et al*., 2000 ; Chevolleau *et al*.,2007). Another clinching evidence reports that early light inducible protein synthesized in the cytoplasm is imported into the chloroplast and inserted in thylakoids

via a pathway involving cpSRP43 (Hutin *et al*., 2002). However, an interaction between the SRP receptor and translocon was evidenced to be critical during co translational protein translocation (Jiang *et al*.,2008). The major core complex, TFIID which forms the important factor in the formation of basal transcription apparatus is recognized by TAF6 and TAF9 subunits of human TFIID (Lagoa *et al.,* 2004; Shao *et al.,* 2005). The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth (Lagoa *et al.,2005*). TAF6δ was shown to control apoptosis and gene expression. The core TFIID subunit TAF6 is essential for viability in yeast , plants, insects and fish which can induce apoptosis in response to genotoxic stress (Wilhem *et al*., 2008).

Although all the genes studied were stress responsive and roles of a few have been reported, the evidences for a certain set of genes like TAF6 and signal recognition particle receptor were lacking under stress. Hence it was interesting to know whether the expression patterns of these genes were developmentally regulated or were found to express under other abiotic stresses as well. In this direction, expression patterns of these genes were monitored in tissues collected after conducting seedling assays by imposing salinity stress (300 mM), PEG stress (-8 bars) and methyl viologen stress (10 µM) to two day old seedlings. The stress was imposed to the seedlings and the expression patterns of these two genes were monitored. Semi quantitative and quantitative RT PCR was performed using cDNA reverse transcribed from RNA extracted from stressed and control shoot tissues. The results indicated that both TAF6 and signal recognition particle receptor were up regulated under all the three stresses in comparison to control (Plates 16, 17 & Fig 18). This suggests that the two genes are shown to be responsive to multiple abiotic stresses and can be considered as potential candidate genes. Since the genes are novel and stress responsive, functional validation of the genes by over expression studies might yield fruitful information. From this context, attempts have been made to clone full length cDNA of these genes from finger millet.

Full length amplification was attempted after scrutinising the partial length sequence information available (Plates 17 and 18). Using stress cDNA pool, full length amplification was done by rapid amplification of cDNA ends (RACE) as per a kit protocol by Invitrogen (www,invitrogen.com), the 3′ ends of both the genes were successfully amplified whose identity is yet to be confirmed (Plate 19). The full length genes to be cloned would be useful for functional validation by over expression studies.

Salient findings:

Drought resistance is governed by multiple genes and hence it is significant to prospect candidate genes regulating these traits. With this objective, a drought stress specific cDNA library was screened and evaluated. The salient findings of the study are :

- Over 1100 clones were screened and the cDNA library was found to be diverse which would serve as a potent source for prospecting candidate genes
- Functional validation of certain genes by different approaches like e-northern, semi quantitative and q RT PCR revealed the stress responsive expression patterns of certain set of genes associated to intrinsic cellular tolerance namely, metallothionein, sinapyl alcohol dehdrogenase etc.
- Two novel upstream regulatory genes TAF6 and signal recognition particle receptor were found to be responsive to drought as well as other abiotic stresses like salt, osmotic and oxidative stresses. Hence, TAF6 and signal recognition particle receptor are the prospected candidate genes up regulated under stress which could be used for manipulating intrinsic cellular tolerance.

Future line of work

The two upstream regulatory genes TAF6 and signal recognition particle receptor seems to be novel candidate genes associated with drought tolerance. Since little information is available on the relevance of TAF6 and signal recognition particle receptor in imparting stress tolerance, it will be highly rewarding to clone full length genes and functionally validate through transgenic approach.

SUMMARY

VI. SUMMARY

Mother Nature has been strangled by vagaries that exist within her own niche. Human life has been supported by the luxuries prevailing in nature. Ever since the evolution of human settlement, the prominent aspect that has abetted the survivability of mankind is the varied crop produce harvested out of domesticated as well as wild species. This life sustaining factor has been observed to decrease at alarming rates over the past few decades. This decline in crop productivity has been mainly due to limitations that are instigated by adverse environmental conditions. One of the major threats to crop productivity is drought which spans around the entire globe. However, global climatic scenario predicts an existence of increased abiotic stress conditions along with elevated levels of evaporation of moisture from plant rhizospheres. This would eventually lead to diminution in crop yield and productivity.

Most of the plants fall victim to these unforeseen vagaries. Admist these limitations, certain plant species have shown to acquire an array of strategies at whole plant level to cope up with the adverse circumstances due to various molecular mechanisms. The different molecular biology tools like gene discovery, transcriptome profiling, promoter analysis etc., have proved to be worthy enough to give a better understanding of the molecular basic of stress tolerance. The various virtues at cellular level increase the intrinsic cellular tolerance levels. All these adaptations will be based on cellular regulation of transcription which ultimately determines gene expression. Hence, transcript profiling assumes acute significance. The stress transcriptome forms the strong foundation for deciphering stress responsive molecular mechanisms.

With this outlook, screening and evaluation of a drought stress specific cDNA library developed from drought adapted crop, finger millet (*Eleusine coracana* (L.) Gaertn.) was undertaken. Over 1100 clones were screened by plaque lifting and nearly 600 clones sequenced. Annotation of the sequenced clones gave a revelation that the cDNA library was a rich source of stress responsive clones which would have functional implications under drought stress. The ESTs generated belonged to diverse classes of genes like cytoskeletal structures (10%), protective proteins (10%), catalytic genes (9%) etc., which are thought to be involved in maintaining and enhancing intrinsic cellular tolerance.

Based on an *in silico* expression analysis tool, e-northern (www.bbc.utoronto.botany.ca), using the AGI codes of *Arabidopsis* homologs of finger millet clones, the expression pattern of around 50 genes was overviewed under different stress situations like osmotic, drought, cold, salinity, radiation, genotoxic, oxidative, wounding and heat stresses. This analysis led to pouncing upon 22 relevant genes whose functional novelty was also taken into account. The expression levels were recorded as fold increase over control and it was found that certain genes like farnesylated protein, transcriptional regulator, signal recognition particle receptor and sinapyl alcohol dehydrogenase are multiple stress responsive at varied levels. Hence, it was highly significant to elucidate their relevance under drought. The 22 selected genes were further subjected to single clone *in vivo* excision into phagemids and sequenced to reconfirm the clone identities. Based on sequencing results, 14 clones were targeted for further studies. Accurate primer designing was exercised for the short listed genes.

The pattern of gene expression under stress was studied through semi quantitative and quantitative RT PCR, using leaf tissue harvested from finger millet plants experiencing drought stress imposed by gravimetric approach. After drought stress was imposed, the stress effect was assessed by various assays. Soil water potential, leaf solute potential and osmotic adjustment gave indication of the alteration in water relations of the system due to drought. Stress damage was also assessed by analysing certain key biochemical aspects like reduction in total chlorophyll content, reduction in cell viability and loss of membrane integrity. Ten days post stress imposition, the tissue was harvested for gene expression analysis. The effect of stress at molecular level was assessed by looking into the expression pattern of a known drought stress responsive gene, early light inducible protein (ELIP) by semi quantitative RT PCR. ELIP was found to be up regulated under stress (20% FC) when compared to control. The expression analysis of the 10 genes out of those 14 selected, for which the reaction conditions were standardised, was done initially by semi quantitative RT PCR. Most of the genes were found to show enhanced expression at 60% FC, indicating their early stress responsive nature. The transcript levels of early light inducible protein, RISBZ and transcriptional regulator were found to be high, relatively earlier, when plants reached 60% FC status. However, there were a few genes getting up regulated at 30% FC also. Genes like metallotheionin, farnesylated protein, protein phosphatase and farnesyl pyrophosphate synthase were found to show

high expression levels initially at 60% FC and later at 30% FC further dropping down at 20% FC. To quantify the response of different genes under drought stress, quantitative RT PCR analysis was taken up. Most of the genes like metallotheionin, protein phosphatase, farnesylated protein, farnesyl pyrophosphate synthase, transcriptional regulator, sinapyl alcohol dehydrogenase etc. showed higher fold increase over control in samples collected from plants grown at 60% FC with their levels coming down in those at 20% FC indicating that these genes are early responsive to drought stress. However, expression profiles of TAF6 and signal recognition particle receptor in both semi quantitative and quantitative RT PCR analysis were impressive enough. Due to their unknown role under stress, expression studies for these two genes were done using shoot tissues from finger millet seedling system experiencing varied abiotic stresses like salt (sodium chloride), osmotic (poly ethylene glycol) as well as oxidative stress (methyl viologen). Although expressive under all the three stresses, both were highly up regulated under salt and osmotic stresses over and above oxidative stress. This gave crucial lead into probing more towards the functional relevance of the two genes. Since only partial length sequence information was available for both the genes, full length cloning of these two genes was attempted. Efforts were made to clone 3′ ends of the two genes using RACE approach. The 3′ ends of both the genes were amplified by adding on around 100 bp in case of TAF6 and nearly 850 bp in case of signal recognition particle receptor on the 3′ end whose identity has to be confirmed by sequencing.

The cDNA library developed is novel and can serve as a potent source for prospecting drought responsive genes. Amongst the selected genes, upstream regulatory genes like signal recognition particle receptor and TAF6 could be the best candidate genes for further study.

REFERENCES

VII. REFERENCES

ABE, H., YAMAGUCHI-SHINOZAKI, K., URAO, T., IWASAKI, T., HOSOKAWA, D., AND SHINOZAKI, K., 1997, Role of *Arabidopsis* MYC and MYB homologs in drought- and abscisic acid-regulated gene expression. *The Plant Cell,* **9(10):**1859–1868.

ADAMS, M. D., KELLEY, J. M., GOCAYNE, J. D., DUBNICK, M., POLYMEROPOULOS, M. H., XIAO, H., MERRIL, C. R., WU, A., OLDE, B., MORENO, R. F., 1991, Complementary DNA sequencing: expressed sequence tags and human genome project. *Science,* **252:** 1651–1656.

ADAMSKA, I., 2001, The Elip family of stress proteins in the thylakoid membranes of Pro- and Eukaryota. In: Aro E-M, Andersson B, eds. *Regulation of photosynthesis*. Dordrecht: Kluwer Academic Publishers, 487–505.

APEL, K. AND HIRT, H., 2004, Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol*, **55**: 373-399.

BARRS H. D. AND WEATHERLY, P. E., 1986, A re-examination of the relative turgidity for estimating water deficits in leaves. *Aust J Biol Sci.,* **15**: 413–428.

BARTELS, D. AND SALAMINI, F., 2001, Desiccation tolerance in the resurrection plant *Craterostigma plantagineum*: a contribution to the study of drought tolerance at the molecular level. *Plant Physiol*., **127**: 1346–1353.

BARTELS, D., 2005, Desiccation tolerance studied in the resurrection plant *Craterostigma plantagineum*. *Integr. Comp. Biol.,* **45**: 696–701.

BARTELS, D., SCHNEIDER, K., TERSTAPPEN, G., PIATKOWSKI, D. AND SALAMINI, F., 1990, Molecular cloning of ABA-modulated genes from the resurrection plant *Craterostigma plantagineum* which are induced during desiccation. *Planta,* **181**: 27–34.

BATTISTI, D. S. AND NAYLOR, R. L., 2009, Historical warnings of future food insecurity with unprecedented seasonal heat. *Science*, **323**: 240–244.

BECK, E. H., FETTIG, S., KNAKE, C., HARTIG, K. and BHATTARAI, T., 2007, Specific and unspecific responses of plants to cold and drought. *J. Biosci.* **32(3)**: 501–510.

BECKERS, G. J. M., JASKIEWICZ, M., LIU, Y., UNDERWOOD, W. R., HE, S. Y., ZHANG, S., AND CONRATHA, U., 2009, Mitogen-Activated Protein Kinases 3 and 6 Are Required for Full Priming of Stress Responses in *Arabidopsis thaliana. The Plant Cell*, **21**: 944–953.

BENNET, M.D. AND LEITCH, I.J., 2005, Plant Genome Size Research: A field in focus. *Annals of Botany,* **95(1):** 1-6.

BERNACCHIA, G. AND FURINI, A., 2004, Biochemical and molecular responses to water stress in resurrection plants. *Physiol. Plant.,* **121**: 175–181.

BOSTOCK, R. M., 2005, Signal crosstalk and induced resistance: straddling the line between cost and benefit. *Annu Rev Phytopathol*, **43**: 545-580.

BOYER, J.S., 1982, Plant productivity and environment. *Science,* **218(4571):** 443-448.

BRAY, E.A., 2004, Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. *J. Exp. Bot.*, **55**: 2331-2341.

BRENDEL, V., KURTZ, S. AND WALBOT, V., 2002, Comparative genomics of *Arabidopsis* and maize: prospects and limitations *Genome Biology,* **3(3)**: 1005.1–1005.6.

BRESSAN, R., BOHNERT, H. AND ZHU, J. K., 2009, Abiotic stress tolerance: from gene discovery in model organisms to crop improvement. *Mol. Plant.,* **2**: 1–2.

BRESSAN, R., BOHNERT, H. AND ZHU, J.K., 2009, Abiotic stress tolerance: from gene discovery in model organisms to crop improvement. *Mol.Plant*, ***2*:** 1-2.

BURSSENS, S., HIMANEN, K., VAN DE COTTE, B., BEECKMAN, T., VAN MONTAGU, M., INZE, D. AND VERBRUGGEN, N., 2000, Expression of cell cycle regulatory genes and morphological alterations in response to salt stress in *Arabidopsis thaliana. Planta,* **211**: 632- 640.

BUSTIN, S. A., 2000, Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology,* 25: 169–193.

BYRNE, P. F. AND MCMULLEN, M. D., 1996, Defining genes for agricultural traits: QTL analysis and the candidate gene approach. *Probe* **7**: 24–27.

CAMPALANS, A., PAGES, M. AND MESSEGUER, R., 2001, Identification of differentially expressed genes by the cDNA-AFLP technique during dehydration of almond (*Prunus amygdalus*). *Tree Physiol,* **21**: 633–643.

CHAE, L., SUDAT, S., DUDOIT, S., ZHU, T. AND LUAN, S., 2009, Diverse transcriptional programs associated with environmental stress and hormones in the Arabidopsis receptor-like kinase gene family. *Mol. Plant*, **2**: 84-107.

CHAVES, M. M. AND OLIVEIRA, M. M., 2004, Mechanisms underlying plant resilience to water deficits: prospects for water-saving agriculture. *J Exp Bot,* **5**: 2365-2384.

CHAVES, M.M., MAROCO, J.P. AND PEREIRA, J.S., 2003, Understanding plant responses to drought – from genes to whole plant. *Functional Plant Biology,* **30**: 239- 260.

CHEN T. H. AND MURATA, N., 2002, Enhancement of tolerance of abiotic stress by metabolic engineering of betaines and other compatible solutes. *Curr Opin Plant Biol.* **5:** 250–257.

CHEONG, Y. H., CHANG, H. S., GUPTA, R., WANG, X., ZHU, T. AND LUAN, S., 2002, Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. *Plant Physiol*, **129**: 661-677.

CHEVOLLEAU, T.T., HUTIN, C., NOE¨ , L. D., GOFORTH, R., CARDE, J. P., CAFFARRI, S., SINNING, I., GROVES, M., TEULON, J. M., HOFFMAN, N. E., HENRY, R., HAVAUX, M. AND NUSSAUMEA, L., 2007, Canonical Signal Recognition Particle Components Can Be Bypassed for Post translational Protein Targeting in Chloroplasts. *The Plant Cell*, **19:** 1635–1648.

CHINNUSAMY, V., SCHUMAKER, K. AND ZHU, J. K., 2004, Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *J. Exp. Bot*., **55**: 225–236.

CHO, Y.J., MEADE, J.D., WALDEN, J.C., CHEN, X., GUO, Z. AND LIANG, P., 2001, Multicolor fluorescent differential display. *Biotechniques* , **30**: 562–572.

CHOMCZYNSKI, P. AND MACKEY, K., 1995, Short technical report: Modification of the TRIZOL reagent procedure for isolation of RNA from Polysaccharide-and proteoglycan-rich sources. *Biotechniques*, **19(6)**: 942-945.

COBBETT, C. S. AND GOLDSBROUGH, P. B., 2000, Mechanisms of metal resistance: Phytochelatins and metallothioneins. p. 247–269 *In* I. Raskin and B.D. Ensley (ed.) Phytoremediation of toxic metals: Using plants to clean up the environment. John Wiley & Sons, New York.

CUSHMAN, J.C. AND BOHNERT, H.J., 2000, Genomic approaches to plant stress tolerance. *Curr. Opin. Plant Biol.,* **3:** 117–124.

DASTIDAR, K.G., MAITRA, S., GOSWAMI, L., ROY, D., DAS, K.P. AND MAJUMDER, A.L., 2006, An insight into the molecular basis of salt tolerance of L-*myo*inositol 1-P synthase (PcINO1) from *Poteresia coarctata* (Roxb.) Tateoka, a halophytic wild rice. *Plant Physiol*., **140**: 1279–1296.

DATTA, K., SCHIMIDT, A. AND MARCUS, A., 1989, Characterization of two soyabean repetitive proline rich proteins and a cognate cDNA from germinated axes. *Plant Cell.,* **1:** 945-952.

DAVIES, W.J., WILKINSON, S. AND LOVEYS, B., 2002, Stomatal control by chemical signalling and the exploitation of this mechanism to increase water use efficiency in agriculture. *New Phytol.,* **153**: 449- 460.

DAVLETOVA, S., RIZHSKY, L., LIANG, H., SHENGQIANG, Z., OLIVER, D. J., COUTU, J., SHULAEV, V., SCHLAUCH, K. AND MITTLER, R., 2005, Cytosolic Ascorbate peroxidase 1 is a central component of the reactive oxygen gene network of Arabidopsis. *Plant Cell,* **17**: 268-281.

DELILLE, J., PETERSON, E. C., JOHNSON, T., MORRE, M., KIGHT, A. AND HENRY, R., 2000, A novel precursor recognition element facilitates posttranslational binding to the signal recognition particle in chloroplasts. *Proc. Natl. Acad. Sci.*, **97:** 1926–1931.

DESTEFANO, J. J., BUISER, R. G., MALLABER, L. M., MYERS, T. W., BAMBARA, R. A. AND FAY, P. J., 1991, Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *Journal of Biological Chemistry,* **266:** 7423–7431.

DIATCHENKO, L., LAU, Y.F., CAMPBELL, A.P., CHENCHIK, A., MOQADAM, F., HUANG, B., LUKYANOV, S., LUKYANOV, K., GURSKAYA, N., SVERDLOV, E.D. AND SIEBERT, P.D., 1996, Suppression subtractive hybridization: a method for generating differentially regulated or tissue specific cDNA probes and libraries. *Proc. Natl. Acad. Sci., USA* , **93**: 6025–6030.

EASWARAN, N., PARAMESWARAN, S., SATHRAM, B., ANANTHARAMAN, B., KRISHNAKUMAR, G. R. AND TANGIRALA, S. J., 2010, Yeast functional screen to identify genetic determinants capable of conferring abiotic stress tolerancein *Jatropha curcas*. *BMC Biotechnology,* **10**: 23.

FEDOROVA, M., MORTEL, J. V. D., MATSUMOTO, P. A., CHO, J., TOWN, C. D., VANDENBOSCH, K. A., GANTT, J. S. AND VANCE, C. P., 2002, Genome-wide identification of nodule-specific transcripts in the model legume Medicago truncatula. *Plant Physiol.*, **130:** 519–537.

FERNANDES, J., BRENDEL, V. AND GAI, X., 2002, Comparison of RNA expression profiles based on maize expressed sequence tag frequency analysis and microarray hybridization. *Plant Physiology,* **128:** 896–910.

FERNÁNDEZ, P., PANIEGO, N., LEW, S., HOPP, H. E. AND HEINZ, R. A., 2003, Differential representation of sunflower ESTs in enriched organ-specific cDNA libraries in a small scale sequencing project. *BMC Genomics*, **4**: 1-9.

FUJITA, M., FUJITA, Y., NOUTOSHI, Y., TAKAHASHI, F., NARUSAKA, Y., YAMAGUCHI-SHINOZAKI, K. AND SHINOZAKI, K., 2006, Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Current Opinion in Plant Biology*, **9**: 436–442.

GAO, W. R., WANG, X. S., LIU, Q. Y., PENG, H., CHEN, C., LI, J. G., ZHANG, J. S., HU, S. N. AND MA, H., 2008, Comparative analysis of ESTs in response to drought stress in chickpea (*C. arietinum* L.). *Biochem. Biophy. Res. Comm.*, **376:** 578-583.

GEETHA GOVIND. 2006, Profiling Trancriptome of *Arachis Hypogaea* L under drought and monitoring the expression pattern. Ph.D thesis.

GODOY, G., STEADMAN, J. R., DICKMAN, M.B. AND DAM, R., 1990, Use of mutants to demonstrate the role of oxalic acid in pathogenicity of *Sclerotinia sclerotiorum* on *Phaseolus vulgaris*. *Physiol Mol Plant Pathol*., **37**: 179-191.

GOPALAKRISHANA, R. 2001, Characterization of stress responsive genes in groundnut. Ph.D thesis.

GORANTLA, M., BABU, P. R., REDDY LACHAGARI, V. B., REDDY, M., RAMAKRISHNA, WUSIRIKA JEFFREY. BENNETZEN, L., AND REDDY, A.R. 2007, Identification of stress-responsive genes in an indica rice (Oryza sativa L.) using ESTs generated from drought-stressed seedlings. *Journal of Experimental Botany,* **58(2):** 253–265.

GOVIND, G., THAMMEGOWDA, H.V., KALAIARASI, P.J., IYER, D.R., MUTHAPPA, S.K., NESE, S. AND MAKARLA, U.K., 2008, Identification and functional validation of a unique set of drought induced genes preferentially expressed in a response to gradual water stress. *Mol. Genet. Genomics.,* doi 10.1007/ s 00438- 009- 0432- z.

GUO, B.Z., CHEN, X., DANG, P., SCULLY, B.T., LIANG, X., HOLBROOK, C.C., YU, J. AND CULBREATH, A.K., 2008, Peanut gene expression profiling in developing seeds at different reproduction stages during *Aspergillus parasiticus* infection. *BMC Dev. Biol.,* **8**: 12.

HAIMOV-KOCHMAN, R., FISHER, S. J. AND WINN, V. D., 2006, Modification of the Standard Trizol-Based Technique Improves the Integrity of RNA Isolated from RNase-Rich Placental Tissue. *Clinical Chemistry*, **52**: 159-160.

HEISE, J., KREJCI, S., MIERSCH, J., KRAUSS, G. J. AND HUMBECK, K., 2007, Gene Expression of Metallothioneins in Barley during Senescence and Heavy Metal Treatment. *Crop Science*, **47**: 1111-1118.

HESLOP- HARRISON, Y., 2000, Control gates and micro ecology : The pollen stigma interaction in perspective. *Annals of Botany,* **85 (Supp A):** 5-13.

HOEVEN, R. V. D., RONNING, C., GIOVANNONI, J., MARTIN, G. AND TANKSLEY, S., 2002, Deductions about the number, organization, and evolution of genes in the tomato genome based on analysis of a large expressed sequence tag collection and selective genomic sequencing. *Plant Cell,* **14:** 1441–1456.

HSIAO, T. C., 1973, Plant response to water stress. *Ann. Rev. Plant Physiol.*, **24**: 519-534.

HUNER, N. P. A., ÖQUIST, G. AND SARHAN, F., 1998, Energy balance and acclimation to light and cold, *TIPS,* **3:** 224–230.

HUTIN, C., NUSSAUME, L., MOISE, N., MOYA, I., KLOPPSTECH, K. AND HAVAUX, M., 2003, Early light-induced proteins protect *Arabidopsis* from photooxidative stress. *Proceedings of the National Academy of Sciences,* **100**: 4921–4926.

ITURRIAGA, G., GAFF, D.F. AND ZENTELLA, R., 2000, New desiccation-tolerant plants, including a grass, in the central highlands of Mexico, accumulate trehalose. *Aus. J. Bot.,* **48(2):** 153–158.

JENKS, M. A., EIGENBRODE, S. AND LEMEIUX, B., 2002, Cuticular waxes of *Arabidopsis.* In the *Arabidopsis Book,* C. Somerville and E. Meyerowitz, eds., American Society of Plant Physiologists, Rockville MD, doi /10.1199 / tab.0009, http:// www.aspb.org/ publications /arabidopsis /.

JIANG, Y., CHENG, Z., MANDON, E. C. AND GILMORE, R., 2008, An interaction between the SRP receptor and the translocon is critical during cotranslational protein translocation. *The Journal of Cell Biology*, **180 (6):**1149–1161.

JOSHUA, A. U., JORDAN, M., SWANSON, HALLER, k., RYAN A., RAPP, MICHAEL, E., HATFIELD, S. J., YU, Y., WU, Y., DOWD, C., ARPAT, A.B., BRAD, A., SICKLER, THEA, A., WILKINS, GUO, J. Y., CHEN, X. A., SCHEFFLER, A., TALIERCIO, E., TURLEY, R., MCFADDEN, H., PAYTON, P., KLUEVA, N., ALLEN, R., ZHANG, D., HAIGLER, C., WILKERSON, C., SUO, J., STEFAN, R., SCHULZE, MARGARET, L., PIERCE, ESSENBERG, M., KIM, H., DANNY, J., LLEWELLYN, DENNIS, E. S., KUDRNA, D., WING, R., PATERSON, A. H., SODERLUND, C. AND WENDEL, J. F., 2006, A global assembly of cotton ESTs. *Genome Res.,* **16**: 441-450.

KARABA, A., DIXIT, S., GRECO, R., AHARONI, A., TRIJATMIKO, K. R., MARSCH-MARTINEZ, N., KRISHNAN, A., NATARAJA, K., UDAYAKUMAR, M. AND PEREIRA, A., 2007, Improvement of water use efficiency in rice by expression of HARDY, an *Arabidopsis* drought and salt tolerance gene. *PNAS, USA*, **104**: 15270-15275.

KEENAN, R. J., FREYMANN, D. M., STROUD, R. M. AND WALTER, P., 2001, The signal recognition particle. *Annu. Rev. Biochem.,* **70**: 755–775.

KIKUCHI, S., SATOH, K., NAGATA, T., KAWAGASHIRA, N., DOI, K., KISHIMOTO, N., YAZAKI, J., ISHIKAWA, M., YAMADA, H. AND OOKA, H., 2003, Collection, mapping, and annotation of 28,000 full-length cDNA clones from Japonica rice. *Science,* **301:** 376–379.

KIVIOJA, T., ARVAS, M., SALOHEIMO, M., PENTTILÄ, M. AND UKKONEN, E., 2005, Optimization of cDNA-AFLP experiments using genomic sequence data. *Bioinformatics,* **21(11**): 2573–2579.

KOHLER, A., BLAUDEZ, D., CHALOT, M. AND MARTIN, F., 2004, Cloning and expression of multiple metallothioneins from hybrid poplar. *New Phytologist,* **164** : 83–93.

KRAMER, P. J., 1980, Drought, stress, and the origin of adaptations. *In* Adaptation of Plants to Water and High Temperature Stress. Eds. N.C. Turner and P.J. Kramer. John Wiley and Sons, New York, pp 7- 20.

KREPS, J.A., WU, Y., CHANG, H.S., ZHU, T., WANG, X., HARPER, J.F., 2002, Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant Physiology,* **230:** 2129–2141.

LAGO, C., CLERICI, E., DRENI, L., HORLOW, C., CAPORALI, E., COLOMBO, L. AND KATER, M. M., 2005, The Arabidopsis TFIID factor AtTAF6 controls pollen tube growth. *Developmental Biology,* **285**: 91 – 100.

LAGO, C., CLERICI, E., MIZZI, L., COLOMBO, L. AND KATER, M. M., 2004, TBP-associated factors in Arabidopsis. *Gene*, **342**: 231–241.

LEVITT, J., 1980, Responses of plants to environmental stresses. *Chilling, freezing, and high temperature stress*, **Vol. I**, London, New York, Toronto: Academic Press.

LI, L., CHENG, X. F., LESHKEVICH, J., UMEZAWA, T. AND HARDING, S. A., 2001, The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. *Plant Cell,* **13**: 1567–1585.

LI, Y. P., XIA, R. X., WANG, H., LI, X. S., LIU, Y. Q., WEI, Z. J., LU, C. AND XIANG, Z. H., 2009, Construction of a full-length cDNA Library from Chinese oak silkworm pupaand identification of a KK-42-binding protein gene in relation topupa-diapause termination. *Int. J. Biol. Sci.,* **5(5)**: 451-457.

LIANG, P. AND PARDEE, A. B., 1992, Differential display of eukaryotic messenger RNA by means of polymerase chain reaction. *Science,* **257**: 967– 971.

LIU, X. A., BAIRD, V. AND WM.V., 2003, Differential expression of genes regulated in response to drought or salinity stress in sunflower. *Crop Sci.*, **43**: 678-687.

LORENZO, O. AND SOLANO, R., 2005, Molecular players regulating the jasmonate signaling network. *Curr Opin Plant Biol.*, **8**: 532-540.

LYMAN, J.M., GOOD, S. A., GOURETSKI, V. V., ISHII,M., JOHNSON, G. C., PALMER, M. D., SMITH, D. M. AND WILLIS, J. K., 2010, Robust warming of the global upper ocean**.** *Nature,* **465**: 334–337.

MAJEE, M., MAITRA, S., DASTIDAR, K. G., PATTNAIK, S., CHATTERJEE, A., HAIT, N. C., DAS, K. P. AND MAJUMDER, A.L., 2004, A novel salt-tolerant L-*myo*-inositol-1-P synthase from *Poteresia coarctata* (Roxb.) Tateoka, a halophytic wild rice, molecular cloning, bacterial overexpression, characterization, and functional introgression into tobacco conferring salt tolerance phenotype. *J. Biol. Chem.,* **279**: 28539–28552.

MAKI, N., MARTINSON, J., NISHIMURA, O., TARUI, H., MELLER, J., TSONIS, P. A. AND AGATA1, K., 2010, Expression profiles during dedifferentiation in newt lens regeneration revealed by expressed sequence tags. *Molecular Vision,* **16:** 72-78.

MAMRUTHA, H. M., MOGILI, T., LAKSHMI, K. J., RAMA, N., KOSMA, D., UDAYAKUMAR, M., JENKS, M. A. AND NATARAJA, K. N., 2010, Leaf cuticular wax amount and crystal morphology regulate post-harvest water loss in mulberry (*Morus* species). *Plant Physiology and Biochemistry*, **48 (8)**: 690-696.

MAUCH-MANI, B. AND MAUCH, F., 2005, The role of abscisic acid in plant–pathogen interactions. *Curr Opin Plant Biol.*, **8**: 409-414.

MCCREE, K. J. AND RICHARDSON, S. G., 1987, Stomatal closure vs. osmotic adjustment: a comparison of stress responses. *Crop Sci.,* **27**: 539- 543.

MEYER, G. AND KLOPPSTECH, K., 1984, A rapidly light-induced chloroplast protein with a high turnover coded for by pea nuclear DNA. *Eur. J. Biochem.*, **138**: 201–207.

MICHALEK, W., SMILDE, D., PEROVIC, D., PLEISSNER, K. P., WILLSCHER, U., POTOKINA, J. AND GRANER, A., 2001, ESTs as a resource for barley genomics (abstract no. w30). In Plant and Animal Genome Conference IX, January 13–17, 2001, San Diego.

MICHELMORE, R., KOZIK, A., TRUCO, M. J., MATVIENCHO, M., OCHOA, O., DAMME, M. V., LAVELLE, D., LIN, H., PANDE, B., MCHALE, L., SUDARSHANA, P., ARGYRIS, J., ELLISON, P., BRADFORD, K., JACKSON, L. AND KESSELI, R., 2003, ESTs and candidate gene approaches in the Compositae Genome Project. *Eucarpia Leafy Vegetable,* **13:** 131-136.

MILLER, R., CHAO, S., BUTLER, E., KANG, Y., RAUSCH, C., SEATON, C., WILSON, C., HSIA, C., TONG, J. AND HUMMEL, D., 2001, Progress of the Triticum genome project in the U.S.: EST generation and evaluation (abstract no. P30). In Plant and Animal Genome Conference IX, January 13–17, 2001, San Diego.

MISHRA, R. N., REDDY, P. S., NAIR, S., MARKANDEYA, G., REDDY, A. R., SOPORY, S. K. AND REDDY, M. K., 2007, Isolation and characterization of expressed sequence tags (ESTs) from subtracted cDNA libraries of *Pennisetum glaucum* seedlings.*Plant Mol Biol.,* **64:** 713-732.

MITTLER, R., 2006, Abiotic stress,the field environment and stress combination. *Trends Plant Sc.*, **11**: 15-19.

MITTLER, R., MERQUIOL, E., HALLAK-HERR, E., KAPLAN, A. AND COHEN, M., 2001, Living under a "dormant" canopy: a molecular acclimation mechanism of the desert plant *Retama raetam*. *Plant J.,* **25**: 407-416.

MOORE, E. L., 2005, Drought Management Factsheet - No. 665.000-7.*Canada-British Columbia Water Supply Expansion Program,* Canada local Ministry of Agriculture and Land. www.infobasket.gov.bc.ca.

MORRISON, T. B., WEIS, J. J. AND WITTWER, C. T., 1998, Quantification of low-copy transcripts by continuous SYBR Green I monitoring during amplification. *Biotechniques,* **24:** 954–962.

MUNDY, J. AND CHUA, N. H., 1988, Abscisic acid and water-stress induce the expression of a novel rice gene. *EMBO J.*, **7**: 2279-2286.

NAKASHIMA, K. AND SHINOZAKI, K. Y., 2005, Regulons involved in osmotic stress responsive and cold stress responsive gene expression in plants. *Physiologia Plantarum,* **126 (1):** 62-71.

NAMBARA, E. AND MCCOURT, P., 1999, Protein farnesylation in plants: a greasy tale. *Current Opinion in Plant Biology*, **2**: 388–392.

NARUSAKA, Y., NARUSAKA, M., SEKI, M. 2004, Crosstalk in the responses to abiotic and biotic stresses in *Arabidopsis*: analysis of gene expression in *cytochrome P450* gene superfamily by cDNA microarray. *Plant Molecular Biology,* **55(3):** 327–342.

NI, F. T., CHU, L. Y., HONG-BO SHAO, H. B. AND LIU, Z. H., 2009, Gene Expression and Regulation of Higher Plants Under Soil Water Stress. *Current Genomics*, **10**: 269-280.

NYUGEN, H. T., Molecular Dissection of Drought Resistance in Crop Plants: from Traits to Genes. pp 36 -40.

OGIHARA, Y., MOCHIDA, K., NEMOTO, Y., MURAI, K., YAMAZAKI, Y., SHIN, I. T. AND KOHARA, Y., 2003, Correlated clustering and virtual display of gene expression patterns in the wheat life cycle by large-scale statistical analyses of expressed sequence tags. *Plant J,* **33:** 1001–1011.

ORTIZ-GO´MEZ, A., JIME´NEZ, C., M. ESTE´VEZ, A. M., CARRERO-LE´RIDA, J., RUIZ-PE´REZ, L. M. AND GONZA´LEZ-PACANOWSKA, D., 2006, Farnesyl Diphosphate Synthase Is a Cytosolic Enzyme in *Leishmania major* Promastigotes and Its Overexpression Confers Resistance to Risedronate. *Eukaryotic Cell*, **5 (7)**: 1057–1064.

PETER BENGERT AND THOMAS DANDEKAR, 2003, A software tool-box for analysis of regulatory RNA elements. *Nucl. Acids. Res.,* **3**: 3441-3445.

POOREST AREAS CIVIL SOCIETY PROGRAMME Report, 2001-2008, Drought in India : Challenges and initiatives.

PRATT, L. H., LIANG, C., SHAH, M., SUN, F., WANG, H., REID, P., ALAN, R., GINGLE, ANDREW, H., PATERSON, WING, R., DEAN, R., KLEIN, R., NGUYEN, H. T., MA, H. M., ZHAO, X., MORISHIGE, D. T., MULLET, J. E. AND PRATT, M. M. C., 2005, Sorghum Expressed Sequence Tags Identify Signature Genes for Drought, Pathogenesis, and Skotomorphogenesis from a Milestone Set of 16,801 Unique Transcripts. *Plant Physiology,* **139:** 869–884.

RABBANI, M.A., MARUYAMA, K., ABE, H., KHAN, M.A., KATSURA, K., ITO, Y., YOSHIWARA, K., SEKI, M., SHINOZAKI, K., YAMAGUCHI-SHINOZAKI, K. 2003, Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA get-blot analyses. *Plant Physiology,* **133**: 1755–1767.

RAMANJULU, S. AND BARTELS, D., 2002, Drought- and desiccation-induced modulation of gene expression in plants. *Plant Cell Environ.,* **25**: 141–151.

RAMEGOWDA, H. V., 2008, Cloning and Functional Characterization of Moisture Stress Responsive Transcription Factors from Finger Millet [*Eleusine coracana* (L.) Gaertn.], Ph.D Thesis**.**

RAMI´REZ, M., GRAHAM, M. A., BLANCO-LO´PEZ, L., SILVENTE, S., MEDRANO-SOTO, A., BLAIR, M. W., HERNA´NDEZ, G., VANCE, C. P. AND LARA, M., 2005, Sequencing and analysis of common bean ESTs : building a foundation for functional genomics. *Plant Physiol.,* **137:** 1211–1227.

RANDY, R. D., 1995, Dissection of oxidative stress tolerance using transgenic plants. *Plant Physiol.,* **107:** 1049–1054.

REDDY, A. R, 2002b, Novel EST enrichment with normalized cDNA libraries from drought-stressed rice (Oryza sativa L. cv. Nagina 22). www.ncbi.nlm.nih.gov/dbest.

REDDY, A. R., RAMAKRISHNA, W., SEKHAR, A. C., ITHAL, N., BABU P. R., BONALDO, M. F., SOARES, M. B. AND BENNETZEN, J. L., 2002a., Novel genes are enriched in normalized cDNA libraries from drought-stressed seedlings of rice (Oryza sativa L. subsp. indica cv. Nagina 22). www.ncbi.nlm.nih.gov/dbest.

RONNING, C. M., STEGALKINA, S. S., ASCENZI, R. A., BOUGRI, O., HART, A. L., UTTERBACH, T. R., VANAKEN, S. E., RIEDMULLER, S. B., WHITE, J. A. AND CHO, J., 2003, Comparative analyses of potato expressed sequence tag libraries. *Plant Physiol.,* **131:** 419–429.

RONNING, C. M., STEGALKINA, S. S., ASCENZI, R. A., BOUGRI, O., HART, A. L., UTTERBACH, T. R., VANAKEN, S. E., RIEDMULLER, S. B., WHITE, J. A. AND CHO, J., 2003, Comparative analyses of potato expressed sequence tag libraries. *Plant Physiol.,* **131:** 419–429.

ROTHSCHILD, M. F. AND SOLLER, M., 1997, Candidate gene analysis to detect genes controlling traits of economic importance in domestic livestock. *Probe* **8**: 13–20.

SABALLOS, A., EJETA, G., SANCHEZ, E., KANG, C. AND VERMERRIS, W., 2009, A Genomewide Analysis of the Cinnamyl Alcohol Dehydrogenase Family in Sorghum [*Sorghum bicolor* (L.) Moench] Identifies SbCAD2 as the Brown midrib6 Gene. *Genetics*, **181**: 783–795.

SAMBROOK, J., FRITSCH, E. F., AND MANIATIS, T., 1989, Molecular Cloning: A Laboratory Manual, Second Edition (Plainview, New York: Cold Spring Harbor Laboratory Press).

SASAKI, T., 2003, Rice genome analysis: understanding the genetic secrets of the rice plant. *Breed. Sci.,* **53**: 281-289.

SATHIK, M., NATARAJA, K. N., MOLLY- THOMAS AND JAMES, J., 2005, An efficient method to isolate good quality RNA from bark tissues of mature *Hevea brasiliensis* trees. *J. Rubber Research*, **8(3)**: 182-189.

SAUTER, A., DAVIES, W.J. AND HARTUNG, W., 2001, The long distance abscisic acid signal in the droughted plant : The fate of the hormone on its way from the root to the shoot. *J. Exp.Bot.,* **52**: 1-7.

SCHUENEMANN, D., GUPTA, S., PERSELLO-CARTIEAUX, F., KLIMYUK, V. I., JONES, J. D. G., NUSSAUME, L. AND HOFFMAN, N. E., 1998, A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proc. Natl. Acad. Sci.,* **95:** 10312–10316.

SEKI, M., CARNINCI, P., NISHIYAMA, Y., HAYASHIZAKI, Y., AND SHINOZAKI, K., 1998, High-efficiency cloning of *Arabidopsis* full-length cDNA by biotinylated CAP trapper. *Plant J*., **15:** 707–720.

SEKI, M., NARUSAKA, M. AND KAMIYA, A., 2002, Functional annotation of a full length *Arabidopsis* cDNA collection. *Science,* **296**: 141–145.

SEKI, M., NARUSAKA, M., ISHIDA, J., NANJO, T., FUJITA, M., OONO, Y., KAMIYA, A., NAKAJIMA, M., ENJU, A. AND SAKURAI, T., 2002, Monitoring the expression profiles of 7000 Arabidopsis genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. *Plant J.*, **31**: 279-292.

SENTHIL-KUMAR, M. AND UDAYAKUMAR, M., 2006, High-throughput virus-induced gene-silencing approach to assess the functional relevance of a moisture stress-induced cDNA homologous to *lea4. Journal of Experimental Botany*, **57 (10)**: 2291–2302.

SHALEM, O., DAHAN, O., LEVO, M., MARTINEZ, M. R., FURMAN, I., SEGAL, E. AND PILPEL, Y., 2008, Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation. *Mol Syst Biol.,* **4**: **223**: 1-10.

SHAO, H. B., CHU, L. Y. AND SHAO, M. A., 2008, Calcium as a versatile Plant Signal Transducer under Soil Water Stress. *BioEssays*, **30:** 634-641.

SHAO, H. B., CHU, L. Y., JALEEL, C. A. AND ZHAO, C. X., 2008, Water-deficit stress-induced anatomical changes in higher plants. *C. R. Biol.*, **331**: 215-225.

SHAO, H. B., CHU, L. Y., LU, Z. H. AND KANG, C. M., 2008, Primary antioxidant free radical scavenging and redox signaling pathways in higher plant cells. *Int. J. Biol. Sci*., **4**: 8-14.

SHAO, H. B., CHU, L. Y., SHAO, M. A. AND ZHAO, C. X., 2008, Advances in functional regulation mechanisms of plant aquaporins: Their diversity, gene expression, localization, structure and roles in plant soil-water relations. *Mol. Membr. Biol.*, **25:** 179-191.

SHAO, H. B., CHU, L. Y., SHAO, M. A., LI, S. Q. AND YAO, J. C., 2008, Improving Plant Resistance to abiotic Stresses by the global Calcium Signal System. *Biotech. Adv.*, **26:** 503-510.

SHAO, H. B., LI, B., WANG, B. C., TANG, K. AND LIANG, Y.L., 2008, A study on differentially expressed gene screening of *Chrysanthemum* plants under sound stress. *C.R. Biol*., **331**: 329-333.

SHAO, H., REVACH, M., MOSHONOV, S., TZUMAN, Y., GAZIT, K., ALBECK, S., UNGER, T. AND DIKSTEIN, R., 2005, Core promoter binding by histone-like TAF complexes. Mol. *Cell. Biol.,* **25:** 206– 219.

SHESHSHAYEE, M. S., BINDUMADHAVA, H., RAMESH, R., PRASAD, T. G., LAKSHMINARAYANA, M. R., AND UDAYAKUMAR, M., 2005, Oxygen isotope enrichment ($_{18}$O) as a measure of time-averaged transpiration rate. *Journal of Experimental Botany,* **56(422)**: 3033-3039.

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(2):** 221–227.

SHINOZAKI, K., YAMAGUCHI-SHINOZAKI, K. AND SEKI, M., 2003, Regulatory network of gene expression in the drought and cold stress responses. *Curr Opin Plant Biol.*, **6**: 410-417.

SHOEMAKER, R., KEIM, P., VODKIN, L., RETZEL, E., CLIFTON, S. W., WATERSTON, R., SMOLLER, D., CORYELL, V., KHANNA, A. AND ERPELDING, J., 2002, A compilation of soybean ESTs: generation and analysis. *Genome*, **45:** 329–338.

SKIBA, B., FORD, R. AND PANG, E., 2005, Construction of a cDNA library of *Lathyrus sativus* inoculated with *Mycosphaerella pinodes* and the expression of potential defence-related expressed sequence tags (ESTs). *Physiological and Molecular Plant Pathology*, **66:** 55-67.

SREENIVASULU, N., SOPORY, S.K. AND KISHOR, P.B.K., 2007, Deciphering the regulatory mechanisms of abiotic stress tolerance in plants by genomic approaches. *Gene,* **388(1–2):** 1-13.

TAIZ, L. AND ZEIGER, E., 2006, *Plant Physiology,* FOURTH EDITION, Sinauer Associates Inc. ISBN 0-87893-856-7 2.

TAYLOR, C., 1996, Farnesylation -A Sticky Business. *The Plant Cell*, **8**: 2151-2153.

TEMSCH, E.M. AND GREILHUBER, J., 2000, Genome size variation in *Arachis hypogaea* and *A.monticola* re-evaluated. *Genome,* **43**: 449- 451.

THANH, T., CHI, V. T. Q., ABDULLAH, M. P., OMAR, H., NOROOZI, M., KY, H. AND NAPIS, S., 2010, Construction of cDNA library and preliminary analysis of expressed sequence tags from green microalga *Ankistrodesmus convolutus* Corda. *Mol Biol Rep,* DOI 10.1007/s11033-010-0092-4.

TILL, B.J., ZERR, T., COMAI, L. AND HENICOFF, S., 2006, A protocol for TILLING and Ecotilling in plants and animals. *Nature Protocols,* **1**: 2465- 2477.

TORRES, M. A. AND DANGL, J. L., 2005, Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. *Curr Opin Plant Biol.*, **8**: 397-403.

TOUFIGHI, K., BRADY S. M., AUSTIN, R., LY, E. AND PROVART N. J., 2005, The Botany Array Resource: e-Northerns, Expression Angling, and promoter analyses. *Plant J.,* **43(1)**: 153-63.

TRAN, L. S. P., NAKASHIMA, K. AND SAKUMA, Y., 2004, Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive *cis* element in the *early responsive to dehydration stress 1* promoter. *The Plant Cell*, **16(9):** 2481–2498.

TU, C. J., PETERSON, E. C., HENRY, R. AND HOFFMAN, N. E., 2000, The L18 domain of light-harvesting chlorophyll proteins binds to chloroplast signal recognition particle 43. *J. Biol. Chem*., **275:** 13187–13190.

USHA, B., PRASHANTH, S. R. AND PARIDA, A., 2007, Differential expression of two metallothionein encoding genes during heavy metal stress in the mangrove species, *Avicennia marina* (Forsk.) Vierh. *Current Science*, **93 (9)**: 1215- 1219.

VALLIYODAN, B. AND NGUYEN, H. T., 2006, Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Current Opinion in Plant Biology*, **9**: 1–7.

VETTORE, A. L., DA SILVA, F. R., KEMPER, E. L., SOUZA ,G. M, DA SILVA, A. M., FERRO, M. I., HENRIQUE-SILVA, F., GIGLIOTI, E. A., LEMOS, M. V. AND COUTINHO, L. L., 2003, Analysis and functional annotation of an expressed sequence tag collection for tropical crop sugarcane. *Genome Res.,* **13:** 2725–2735.

VINOCUR, B. AND ALTMAN, A., 2005, Recent advances in engineering plant tolerance to abiotic stress: achievements and limitations. *Curr Opin Biotechnol*, **16**: 123-132.

WADITEE, R., HIBINO, T., NAKAMURA, T., INCHAROENSAKDI, A. AND TAKABE, T., 2002, Over expression of Na+/H+ antiporter confers salt tolerance on fresh water *Cyanobacterium*, making it capable of growth in sea water. *Proc. Natl. Acad. Sci., USA,* **99(6):** 4109–4114.

WALTER, P. AND JOHNSON, A. E., 1994, Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane. *Annu. Rev. Cell Biol*., **10:** 87–119.

WANG, W., VINOCUR, B. AND ALTMAN, A., 2003, Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta*, **218**: 1-14.

WILHELM, E., PELLAY, F. X., BENECKE, A. AND BELL, B., 2008, TAF6δ Controls Apoptosis and Gene Expression in the Absence of p53. *PLoS ONE*, **3(7): e2721:** 1-13.

WONG, H. L., SAKAMOTO, T., KAWASAKI, T., UMEMURA, K. AND SHIMAMOTO, K., 2004, Down-Regulation of Metallothionein, a Reactive Oxygen Scavenger, by the Small GTPase OsRac1 in Rice. *Plant Physiology*, **135**: 1447–1456.

YALOVSKY, S., KULUKIAN, A., RODRÍGUEZ-CONCEPCIÓN, M., YOUNG, C. A. AND GRUISSEM, W., 2000, Functional Requirement of Plant Farnesyl transferase during Development in Arabidopsis. *The Plant Cell*, **12**: 1267–1278.

YAMAGUCHI-SHINOZAKI, K. AND SHINOZAKI, K., 2005, Organization of cis-acting regulatory elements in osmotic and cold stress responsive promoters. *Trends Plant Sci.,* **10**: 88-94.

YAMAGUCHI-SHINOZAKI, K. AND SHINOZAKI, K., 2006, Transcriptional Regulatory Networks in Cellular Responses and Tolerance to Dehydration and Cold Stresses. *Annu. Rev. Plant Biol*., **57**: 781–803.

YANG, S., VANDERBELD, B., WAN, J. AND HUANG, Y., 2010, Narrowing down the Targets: Towards Successful Genetic Engineering of Drought-Tolerant Crops. *Molecular Plant* , **3(3)**: 469–490.

ZHU, W., ZHAO, D. X., MIAO, Q., XUE, T. T., LI, X. Z. AND ZHENG, C. C., 2009, *Arabidopsis thaliana* Metallothionein, AtMT2a, Mediates ROS Balance during Oxidative Stress. *J. Plant Biol.,* **52**: 585–592.

APPENDICES

APPENDIX I

DETAILED PROTOCOLS USED IN THE STUDY

A. Isolation of total RNA by phenol - chloroform method

A modified phenol – chloroform protocol was followed to extract total RNA with the basic protocol being described by Datta *et al.,* (1989). All precautions were taken to use RNase free material during RNA isolation (Sambrook *et al.,* 1989). Glass wares and plastic wares to be used for extraction were treated with 0.1% DEPC (diethyl pyrocarbonate). The solutions were prepared using DEPC treated water and they were autoclaved. About 1 g of frozen tissue was taken in mortar and pestle and ground to a fine powder in liquid nitrogen. The tissue was homogenized with 10 ml of RNA extraction buffer (0.1 M Tris-HCl, pH 9.0; 0.25 M sucrose; 0.2 M NaCl and 0.01 M $MgCl_2$) to which 10 ml of water saturated phenol:chloroform (1:1) was added followed by 1 ml of 0.5 M EDTA and 1 ml of 20% SDS. β-mercaptoethanol was added to the final concentration of 0.2% *i.e.,* 144 µl . The homogenized mixture was transferred to oakridge tubes and incubated at 4°C for 20 min while shaking. The tubes were centrifuged at 16000 rpm for 30 min at 4°C. The supernatant was transferred into a fresh tube and equal volume phenol:chloroform was added, vortexed and centrifuged at 15000 rpm for 15 min at 4°C. The aqueous supernatant was transferred to a new tube and an equal volume of chloroform: isoamyl alcohol (49:1) was added, mixed and centrifuged at 15000 rpm for 15 min at 4°C. The aqueous phase was further transferred to another mew tube and 8 M $LiCl_2$ was added to get a final concentration of 3 M and incubated at -70°C overnight. The thawed tubes were centrifuged for 30 min at 16000 rpm at 4°C. The pellet was then washed with 2 M $LiCl_2$ (5 ml) by centrifuging at 15000 rpm for 20 min at 4°C. This step was followed by one wash with 75% ethyl alcohol for a 20 minute spin at 15000 rpm at 4°C. The resulting pellet was air dried at room temperature and dissolved in 100 µl RNase free DEPC water. The concentration and purity of the total RNA was ascertained spectrophotometrically (Nano drop, USA) at 260/280nm. The quality was also checked on formaldehyde denaturing agarose (1%) gel.

A.1 Preparation of DEPC treated water (0.1%)

Diethyl pyrocarbonate (DEPC) [O $(COOC_2H_5)_2$] inactivates RNAses in buffers and glassware. One millilitre of DEPC (Sigma Aldrich, USA) was taken into a washed conical flask or Schott duran bottle containing 999 ml of water (Double distilled). The bottle was kept on magnetic stirrer overnight for proper mixing. The DEPC treated water

was further used for treating glassware and other plasticwares. The remaining DEPC treated was autoclaved and further used for preparing solutions for RNA extraction. The autoclaved DEPC treated water was stored at room temperature.

A.2 Preparation of water saturated phenol

Water saturated phenol was prepared as per the protocol cited in EMBO journal (1987). The phenol was melted at 65 °C and 200 ml of melted phenol was thoroughly mixed with an equal volume of DEPC water (200 ml) by keeping it over magnetic stirrer overnight for equilibration. The equilibrated mixture was then transferred to a separating funnel and left for 2 – 2.5 h to separate out the lower water saturated phenol phase. The lower phenol phase was collected and stored at room temperature in dark coloured bottles.

A.3 RNA gel casting and loading

Composition of running gel buffer for RNA [MOPS-EDTA 10X] (pH 7.0)

Molecular Weight (g)	Required conc.	Chemical
209.3	200mM	MOPS
82.03	50mM	Sodium acetate
372.24	10mM	EDTA

A.3.1 Preparation of 10X MOPS electrophoresis buffer

20.93 g of MOPS, 0.205 g Na acetate and 1.861 g EDTA was dissolved in 400ml of DEPC water and mixed well. The pH was adjusted to 7.0 using 1N NaOH and the volume was made up to 500 ml. The prepared solution was stored in dark coloured bottles, autoclaved and kept at room temperature.

A.3.2 Composition of RNA Loading Dye

160 µl 10X MOPS Buffer

260 µl 37% Formaldehyde

160 µl DEPC treated water

40 µl Ethidium bromide (10 µg / µl)

80 µl Glycerol (100%)

80 µl Saturated Bromo Phenol Blue (BPB) (10 mg / ml sterile water)

720 µl Formamide

A.3.3 RNA Gel Preparation

Sl no.	Volume of gel (ml)	Volume of DEPC Water (ml)	Volume of 10X MOPS (ml)	Volume of formaldehyde (ml)	Amount of agarose (g)
1.	25	21.15	2.5	1.35	0.25
2.	40	33.84	4.0	2.6	0.40
3.	50	42.30	5.0	2.7	0.50
4.	60	50.75	6.0	3.24	0.60
5.	70	59.22	7.0	3.78	0.70
6.	80	68.70	8.0	4.32	0.80
7.	100	84.60	10.0	5.4	1.0

A.3.4 RNA loading

2-4 µl RNA sample was mixed with equal volume of RNA loading dye and kept at 65^0C for 10 min to remove RNA secondary structures. The samples were then loaded onto formaldehyde denaturing agarose (1%) gel.

B. Trizol RNA isolation protocol

50-100mg of tissue was ground in liquid nitrogen and extracted in 1ml of Trizol (Tri Reagent, Sigma Aldrich, USA) and incubated for 5 min on ice. To this extract, 200 µl of chloroform was added and incubated for 5 min on ice. The mixture was then centrifuged at 12,000 rpm for 15 min at 4 ºC. The aqueous phase was taken into new sterile eppendorf and was precipitated by adding 500µL of isopropyl alcohol. RNA was pelleted by centrifuging at 12,000rpm for 10 min at 4 ºC. The pellet was then washed with 75% alcohol (1ml) by centrifuging at 7500rpm for 5 min at 4 ºC. The pellet was air dried and dissolved in 30µL DEPC water.

C. Plasmid DNA isolation

A modified alkaline-lysis protocol (Sambrook *et al.,* 1989) was used for isolating plasmid DNA. The bacterial colony (*E. coli*) containing the appropriate plasmid was inoculated in

3 ml liquid LB medium (10 g tryptone, 5 g yeast extract, 10 g NaCl per litre pH 7.0) with appropriate antibiotics and allowed to grow for 16 h at 37°C and 160 rpm. About 1.5 ml the grown culture was taken into centrifuge tube and centrifuged twice at 4000 rpm for 7 min at 4°C. The pellet was resuspended in 100 µl of Solution I (25 mM Tris, pH 8.0; 10 mM EDTA and 50 mM glucose) by vortexing thoroughly and further incubated on ice for 5 min without shaking. To the reaction tube, 200 µl of Solution II (0.2 M NaOH and 1% SDS freshly prepared) was added to the suspension, carefully mixed by inverting the tubes and further incubated for 5 min at room temperature. In order to obtain high quality plasmid DNA, 200 µl of ice cold Solution III [5 M potassium acetate (60 ml), glacial acetic acid (11.5 ml) and water (28.5 ml)] was added to the suspension and carefully mixed by inverting the tubes to avoid breaking of DNA. The mixture was again incubated on ice for 10 min and centrifuged at 11000 rpm for 15 min at 4°C. The supernatant containing the plasmid DNA was carefully taken to a fresh tube and 0.75 volumes of chilled isopropanol was added and immediately centrifuged at 11000 rpm for 15 min at 4°C. The resulting supernatent was discarded, the pellet was completely dried and dissolved in 500 µl of TE buffer (10:1 of 1M Tris-HCl pH 8.0 and 0.5 M EDTA) by tapping gently. To this resuspended mixture, 2 µl of RNase A (2.5 mg/ml) was added and incubated at 37°C for 45 min. Further, equal volume (500 µl) of chloroform:isoamylalcohol (24:1) was added and centrifuged at 11000 rpm for 15 min at 4°C. The plasmid DNA in the supernatant was then precipitated in a fresh tube with 2 – 2.5 volumes of absolute alcohol and kept at -70°C overnight to ensure complete precipitation. The plasmid DNA was recovered in the pellet after centrifugation at 11000 rpm for 15 min at 4°C the next day. The plasmid DNA pellet was washed in 70% (v/v) ethanol by centrifuging at 11000 rpm for 15 min at 4°C. The pellet was then air dried at room temperature and dissolved in 20 µl of sterile water and stored at -20°C until further use. About 5 µl of plasmid prep was tested on an agarose gel (1%) to assess the purity of extraction. The absorbance of sample was measured at 260 nm to determine the DNA concentration.

D. Restriction endonuclease treatments

Restriction digestion of amplified PCR products was carried out by appropriate restriction endonucleases in presence of suitable 1X buffer using five units of restriction enzymes per 1 µg of DNA to be digested. The digested samples were resolved on agarose (2.5 %) gel and the gel images were documented using using gel documentation unit

(Herolab 440K, GmbH Laborgerate, Germany).

D.1 **Restriction Digest Components and conditions**

20 µl digests were set up using the following components :

PCR Product : 6 µl

10 X Buffer : 2 µl

Enzyme : 1 µl

Sterile Water : 11 µl

The digests were generally incubated at 37 °C for 45 min to 1 h with a exception for a few which requires 30 °C incubation (eg. SmaI)

E. Agarose gel separation of DNA

Horizontal gel electrophoresis was employed to separate out nucleic acid fragments. Depending on the requirement 0.8 to 1% agarose solutions were prepared in 1X TAE buffer (40 mM Trisacetate, 2 mM EDTA, 57.1 ml glacial acetic acid, pH 8.0) by heating in a microwave oven. 0.5 µg/ml of ethidium bromide was used for staining. DNA samples were mixed with 0.1 volume loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol and 15% glycerol) and separated electrophoretically. The detection of the DNA fragments was carried out under UVtransilluminator (302 nm) and photographed using a gel documentation system (Herolab, GmbH Laborgerate, Germany).

F. Screening of recombinant clones – library screening

The recombinant clones either in cloning vector or binary vector were initially selected on appropriate antibiotic medium. The survived colonies were further tested for their recombinant nature by colony PCR using appropriate primers. A typical colony PCR was performed similar to that of a standard PCR. The amplified products were separated on agarose (1 %) gel.

G. Polymerase chain reaction

The DNA fragments were amplified from various plasmid DNA/genomic DNA/cDNA in a 20 µl reaction volume containing 100 ng of template DNA (1-3 µl), 2 µl PCR buffer (10X), 2 µl dNTPs (2 mM), 1 µl forward primer (3 pmol/µl), 1 µl reverse primer (3 pmol/µl), 1 U *Taq* DNA polymerase (0.3 µl) and volume made up to 20 µl with sterile water. PCR was performed in Master cycler Gradient (Eppendorf AG, Germany). The optimal number of PCR cycles and the annealing temperature (T_A) was determined

empirically for each PCR. A standard PCR program followed is as follows :

Step 1: 94°C for 5 min [2 min for cDNA] (Initial denaturation)

Step 2: 94°C for 1 minute (Denaturation)

Step 3: T_A for 45 seconds (Primer annealing) where, T_A = Annealing temperature

Step 4: 72°C for 1 minute /kb fragment (DNA synthesis)

Step 5: Goto step 2; Repeat 29 cycles

Step 6: 72°C for 10 min (Final extension)

Step 7: Hold at 4°C; End

The PCR samples were analyzed by agarose electrophoresis. The presence of the expected PCR products was determined with the help of DNA ladders (GeneRuler DNA ladders, MBI-Fermentas).

H. Quantitative real-time RT-PCR conditions

The cDNA pool was used as template to perform quantitative real-time PCR in presence of SYBR-green dye. The optimization of the real-time PCR conditions was performed according to the manufacturer's instructions (MJ Research, USA & MJ Bioworks, Inc.). The PCR was performed using the SYBR Premix Ex Taq (Perfect Real Time) (SYBR Premix Ex Taq (Perfect Real Time), TaKaRa). The conditions for the PCR were as follows: 94ºC for 2 min, 30 cycles of denaturation (at 94ºC for 1 min), annealing for 30 seconds (57ºC to 67ºC for gene as well as actin) extension for 30 seconds (72ºC) followed by plate read and final extension at 72ºC for 8 min, estimation of melting curve from 50ºC to 90ºC and incubation at 72ºC for 10 min. The amplification plots of the PCR reaction were used to determine the threshold cycle (C_T).

APPENDIX II

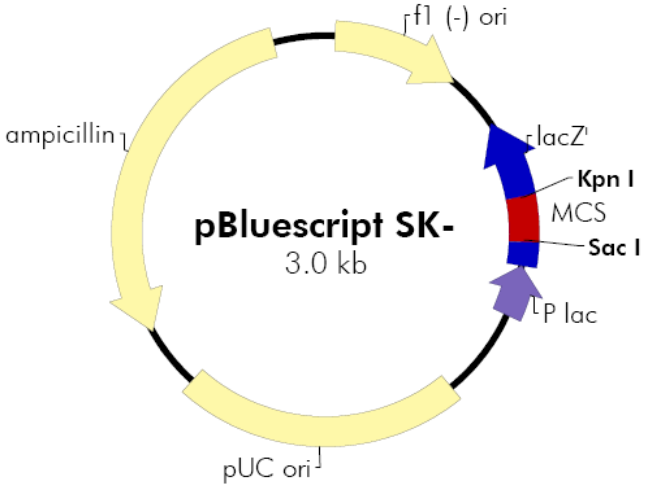
PRIMERS USED IN THE STUDY

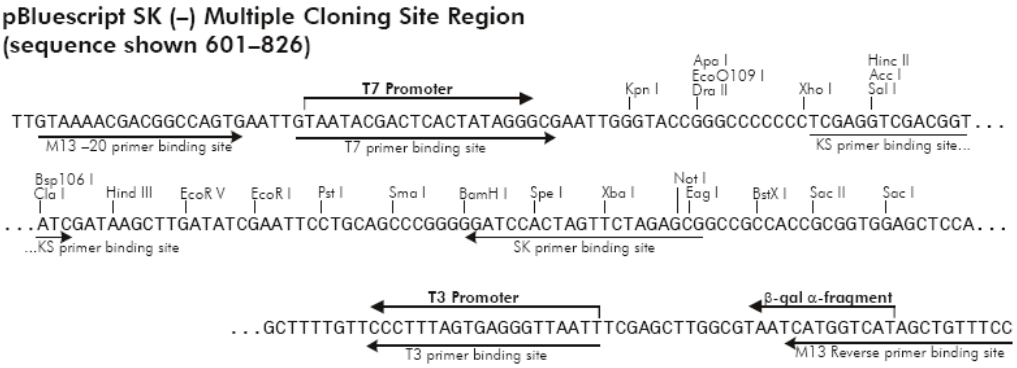
Primers for the expression analysis of selected clones					
Sl. No	**Name**	**Primer sequence 5′ - 3′**	**Mer**	**Tm (^{0} C)**	**Product size (bp)**
1	MT F	GTTGCAAGTGCGGCTATATG	20	55	211
2	MT R	GAAGCAGCAAGCACCCATG	19		
3	RISBZ4 F	CCCCGACGGAGCATATGCAGAACTC	25	65	201
4	RISBZ4 R	GCTTCATCTCTTCACTTGGACAAGCC	26		
5	DIS1 F	CCCCTTACTGCTCCTGAGACCCGG	24	NS	209
6	DIS1 R	GTGGGTAGCTCCATCACCCACATC	24		
7	14-3-3 F	GATGATCCCTCAAGTTCCCCTGTAC	25	50	226
8	14-3-3 R	GGGTTTGACGGGTTCACCG	19		
9	SIN AD F	GCGTCGAAGACTGTGCTACTG	21	65	202
10	SIN AD R	CACACATGAAGCAACGCCCAAC	22		
11	TAF6 F	CCAACCCTGGAATGATGCAAGC	22	65	208
12	TAF6 R	GGTGGTTGAACAAAGGAGAAGATGGC	26		
13	F PROT F	GTGCCGCACCCTTACGTGC	19	56	201
14	F PROT R	CGACAGGCAACGAGACAGTGC	21		
15	AD WIS1 F	GGCGGTGAACCTTTGGAG	18	NS	178
16	AD WIS1 R	GCGCATTCATAATTTCACAGGC	22		
17	SIG RP F	GATAGGCTGATGACCAAGAATGTGGC	26	66	222
18	SIG RP R	GACATAAGGCCTCCCACGCTCC	22		
19	T REG F	GAACGTCAGACTCGTCGCC	19	66	258
20	T REG R	CGAAGTCTGTGGTGCGATCCAG	22		
21	PROT PHOS F	GGCGGCTACGGCTTCGGCG	19	58	210
22	PROT PHOS R	GGCAGCGCACGGGCTGCACG	20		
23	LIP HV 90 F	GGTGCCACTGCTCCAGGG	18	66	235
24	LIP HV 90 R	CCGCCTGATATACGTCATACACG	23		
25	F PYR SYN F	GACGAGCTCCGAGGAGTGC	19	62	203
26	F PYR SYN R	CAATAACAGAGAGCCCGCGATTGC	24		
27	MDAR F	GAGAAGGTGCTCGAGACTG	19	NS	231
28	MDAR R	GTACATAAGCAATCGTACATGGTC	24		
Primers used for RACE					
1	TAF6 5' RACE	GGCCAGCATCTTGATCCTCCTTCCATGC	28	64	-
2	TAF6 3' RACE	CTGCACTGTTGTATAATTAAGGGTATTC	28	58	-

3	SRP 5' RACE	GCTTCTTGTTGTCGGTAGGTTTGTCATCC	29	62	-
4	SRP 3' RACE	GATAGGCTGATGACCAAGAATGTGGCAG	29	61	-
5	GeneRacer™ 5′ Primer	CGACTGGAGCACGAGGACACTGA	23	74	-
6	GeneRacer™ 5′ Nested Primer	GGACACTGACATGGACTGAAGGAGTA	26	78	-
7	GeneRacer™ 3′ Primer	GCTGTCAACGATACGCTACGTAACG	25	76	-
8	GeneRacer™ 3′ Nested Primer	5′-CGCTACGTAACGGCATGACAGTG	23	72	-
Other primers used in the study					
Sl No.	**Name**	**Sequence 5′ - 3′**	**Mer**	**Tm (^{0} C)**	**Product size (bp)**
1	M13 Forward	GTAAAACGACGGCCAGT	17	58	**-**
2	M13 Reverse	CAGGAAACAGCTATGGAC	18		
3	Finger millet Actin Forward	TCCATAATGAAGTGTGATGT	20	58	200
4	Finger millet Actin Reverse	GGACCTGACTCGTCATACTC	20		
5	ELIP-Forward	GGAGCTCGGAGTGCAATTTGATACACAC	28	55	750
6	ELIP-Reverse	CGGAGCTCTATACATACACTAAGATAAG ATAAAAGG	36		

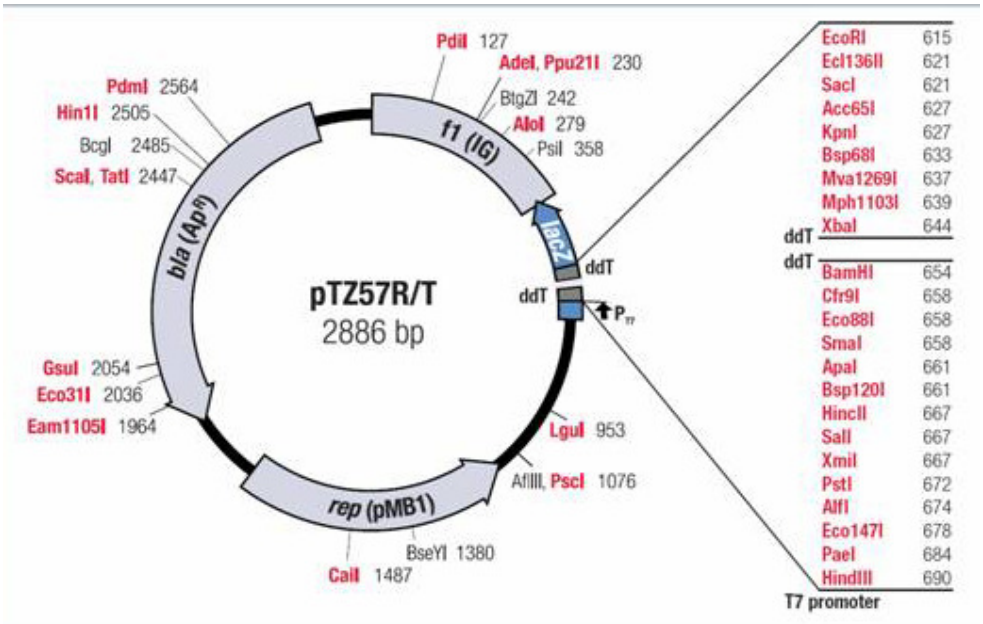
APPENDIX III

VECTORS USED IN THIS STUDY





A. Circular map and poly-linker sequence of the pBluescript SK(–) phagemid (Stratagene, CA, USA)



B. Circular map of vector pTZ57R/T (T-vector). The pTZ57R/T vector was obtained from MBI Fermentas. Initially the PCR products were prepared with 'A' overhang and cloned into the pTZ57R/T vectors adopting manufacturers protocols (MBI Fermentas, Hanover, MD, USA).

APPENDIX V

Sequence information and alignment files

A. Sequence information of selected genes

Clone no. (C)159 (putative cyanase)

GGCTCTGAACTAGTGGATCCCCCGGGCTGCAGGCTGTTATGCATTTCGGAGAGAGCATAAAGGAAATAATCAACGAAG
ACTTTGGTGATGGGATCATGTCGGCCATAGATTTCTATTGCTCAGTTGACAAGGTTAAAGGTGCCGATGGAAAAGATCG
TGTTGTGGTCACATTTGATGGGAAGTATCTGCCTTACACTGAACAGAAATCTGAACACATGATGTCAAGGATGAATCTC
AAGTCATCTTAAAACAAGTTACAGCTACTTAATGCATTTTTTTTACACGGAAACAGGAAGCGAAGCCCCTGCTCATAAT
GCGTTCGGTGAACGGAGCATATGTGTTGTCTTCTTAGACCTTTGGGTCTGGAAAAGAATTTAGCTATTTTCTCTCTCTTTT
CTTTCAAAGCGATCAACAAAATAAATGGGATTGTAATAATGACAATAAAGTCGATCACGCTCGTCCACGAGGTGAACA
ATAAACTCTTGGACTGCATCTGCTCGTGTTGAAATGATAAAGAATTTGGCCAAA

Clone no. (C) 161(Saposin-like type B, region 1 family protein, expressed)

GGCTCTAGACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGCAAAGACATAGCCCTTTTCTCTGGTGTGAGAA
GTGACAGCACATGTGTATTTTGCCGCCACCTGCTTGACGAAGTTATGTCCAAACTGAAAGATCCTGATGCAGAGTTTGA
GATAATTCAAATTCTTATCAAGGAGTGTAATAAGATTGAAGGCCATGTACAGCAGTGCAAGAGACTGGTCCTGCAGTAC
ATTCCTCTGATCTTGGTCAATGGTGAGAAGTTCCTTGAGAAGAATGATGTCTGTGCTCTCGTCCAAGCTTGTGACGCCAG
CCAAAAGAGGGCAGTCAGCTCATTCTTGGAGGCCGGGTTGCTGAGTGATGCATGAAGAAAACTGGAAGGCAAAAATAG
TGTCCAGTTATTTGTGTTGTACATAATAAGGCAATAAATGGCTGTTGTCTTATGCTTCTAGTCTTTGCACCCTTGTTGCAT
TATAAAGGGAAGTGAGCGTTATGAATACCGATCACTGTGTCAAGTTAGTGTGTA

Clone no. (C) 163 (MT-like protein)

TGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGCCAGAGTCTTCCTTCGTCTCACAG
ATCAACCAAGACAGGATGTCGTGCAGCTGCGGATCAAGCTGCAACTGCGGTTCCAGCTGCAAGTGCGGCAAGATGTAC
CCTGACATCGAGGAGAAGAGCACCGGCGCCCAGACCACCGTCATCCTCGGCGTTGCAACGGAGCAGAAGGTCCAGTTC
GAGGCGGCCACCGAGTCCGGCGAGACCGCCCACGCCTGCGGCTGCGGCGCCAACTGCAAGTGCAACCCCTGCACCTGC
TGATCTGATCAGTTTCAGGGGTGCTTGCTGCTTCCTGCTTGTGTCATGACATGCGTGTGAGCACAACAGTATATAATAAG
GAGGAGCTAGCCATGTACCGACGATCTGGCTCTGGAGCTTGCATCAGTTCTTTGTGTGGTTTGAAACCTTAATTTTACCT
CTTTTGACGACGTGTGTGTTTGAGGAGTTGTTTGGTGTGATGGCAGCATCACAAACACATCTCCCTGTCTGTATGTTTGT
GTGTCGCTCTGTAACGAC

Clone no. (C) 172 (RISBZ4)

TGGCGGCCGCTTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGCTCGAGGAAGAGGAAGCAAGCT
CACTTAACAGACCTTGAAACACAGGTTGACCAGCTCCGAGGTGAGAATGCTTCACTCTTCAAGCAGCTGACAGACGCCA
ACCAGCAATTCACAACCTCAGTCACGGACAACAGGATCCTGAAATCAGACGTCGAGGCCCTCAGAGTCAAGGTGAAGC
TGGCAGAGGACATGGTGGCGCGCGGCACAATGTCATGCGGGCTGGGCCACCTGGGCCTGTCGCCG
GTGCTGAACCCACGGCAGGCGTGCCGCGGCCCGGACGTACTCGTGGGCCTGGACTTCCCCGCCGACGACGCGTGCTTCA
CGGGCCTCTCCCCGACGGAGCATATGCAGAACTCGCCGTTGCAGAGCATCGCCACCAGCCTCGAGAGCCTC
GACAACCGCATGGCCAGCGAGGTGACGAGCTGTGGCCGCGGCGGCGCCGGCGTCGACGTCTGGCCCTGGGACTCTCAC
TCCAACGGTGGCTTGTCCAAGTGAAGAGATGAAGCTGAAGCTTGGAATTCGCTGGCGGCAGCTGCTGTTCTG
TACGCGTGTTATATGAGCTGCAAGAATTTGTGAGCTGCTTCTGTGTTAGCTGGAGTCATGAACTGAACGGACTTGAGTTT
TTCAGAGTCCTCGTGCTGGTATATTGATTGCCTAGTCCTGTTATAGACCTGATGACTAATAGACATATCTATGTGCTATG
ATACTGAACTCATAAGGATAAGCATTATTATTCGTATCCCTGGTGT

Clone no. (C) 187 (DIS1; protein binding / structural constituent of cytoskeleton)

GGGGCGCTCCTGAACTGTGGATCCCCCGGGCTGCAGGCGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACA
TTGGGGAGGAGGCGCTGGCGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGC
CAAGTGGAGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTGACCCGGAAGATC
ACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGGAGTACACCGGAGAGATCATGTTCGAGACGTTC
AATGTTCCTGGCCTCTATATCGGAGTGCAACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAAATGAC
AGGTGTTGTAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGTTATGTCATCGGGAGCAGCATC
AGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAACTTATGAAGGAAAGAGGTGAACACATTCCAG
CAGAAGAATCTTTGGATGTAGCAAGGAGGCTTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAATTCAA
TAAGCATGACAGGGAGCCCTCAAAGTACATTAAACACTGGAATGGCATCAAACCAAAAACTGGTACTAAGTATTCATG
TGACATGGGATATGAGCGCTTCTTGGGGCCTGAGATCTTCTTCCACCCAGAAATTTACAACAATGACTTCACCACTCCTT
TGCAAGTTGTTATTGACAAGTGCAT

Clone no. (C) 217 (actin-related protein ARP3)

TCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGCGCAGCACAGCGCCGGCGTCATGGCCGACCT
CGACTTCTACATTGGGGAGGAGGCGCTGGCGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCAT
CCGCAATGGCCAAGTGGAGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTGAC
CCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCCTTTTA

Clone no. (C) 223 (14-3-3 protein)

GCCGCTCTAGAACTAGTGGAATCCCCCGGGCTGCAGGAATTCGGCACGAGGATGATCCCTCAGGTTCCGCTGTACCAGA

CCTTCAACCATGGCTAGGAGGCTCCACTTGGCTGCGAAAGTTGGAAGCAACTGTTAAAACTTGCTATGTTGTGTTGTCAT
CTATTAAATGGAGTCTATTTCAGAGTCGATTTCTATTACAAAAATTCGTGTAGCGTCAAACGTAGTTGTATAAGGAGTCC
TAATTCGTTGAGACTATTATGGTGAACCGGACTAAAACACTGGGGCATGACTAATGTCTAGCTTTCTGCTCATAAAAAA
AAAAAAAAAAAATCTAACGATTTTTGAGGAGGGCAAACCCACCAAAAAAAAAACCAGGTCTCTTTTTGCTGAACCGAT
TGTTGCTTTTGGGGACGGGCCTATCAATGCGGTGGTGGTCATATCCCAAAGGACACAAGGAACGCTATAAGATATTCCA
CGTTTAGCGACTATGCTATAGAGGCTACGAGATTTTACAGCCCCCAAAATTTATTTCACTCCTGACACGTGATTAGAGTC
AATTCTTATTACTCTCTAGAGAGAGAAATTAATTTATCTTATAGTCCTGAGAGTCTTCTCTGATAGTGATTCTCATTCATA
TCCATCTAAGTTGACTATCCTAGATTCACACTTTATATTAAAGATAAGCCACATAGCTGAGACCCTCAGTACAATGACA
ATCTGTACACAGTACATATCATCAGCACGTGATGCACTACGTAACTATATAACATTAGCTCGTCTACACTATCTATACTA
TCTTCGATCGCACTGAGCATCCACTCTCCGC

Clone no. (C) 237 (sinapyl alcohol dehydrogenase)

GTCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGAAAGTTTCCTAAGT
ACACACACCACAATTCCTTCATAAGGGACCAGCGGGCTCACCTGTGTTACAAACAATGTGCTTCTTCCTCCTGCTGCTGC
GCCGAAAATGGCTGCCATCTGCACGATAGTACTGTGCGTTGCCTTGGGACCTCATCCGGCGGCCGTTGTTCATGGTCGG
GCGACGCCGGAAAACGAGCCACCAGAGTGCTACTGAGACATGCCCCTGCCTGCGGATGCTCGCCCCATGCTGCTCTCCC
AACCCCATCCCCCCTAAACGATCGCATCAAATACTGGGCTACTGACGCCTGCAAGATTTGCTAATCAAATTGCCGGATT
TGCTGCCGCTTTGTACCATCATGATATCCCCACTGCCCCAAGAAATTTGAATGCGAATTCGAATTGAGCATGGATGTTGC
CAAAAACTGACCCGCTGCTGTTATGATAACTCGCCCTGCAACACGAGTTGGGCGTTGCTTCATGTGTGTCCAGACTCGT
GACTCTGGAGGCCTCTCGGCTTTGCCGCCTTCCTTTCAATAAACAAGTATTGTGTTTTATTGAATCAACAGAACATGTGT
TCTTACTAGTTATCGTACTATGGTAAACTCAGAGCATTGTGTTTATGCTATTTTGCTATGATGTATCTATTTAACATTGCG
GG

Clone no. (C) 246 (putative TAF6)

TGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGCTGCGTCTTCTATGGCAGGAAACATGCAAGGAACCATG
GATGGGTTCTCCACCCAGTTAGCCAACCCTGGAATGATGCAAGCTTCATCCACTTCACAAATGGTGGAAAA
CATTGCAGTTGCGGTTCGGAGAGAACAGGGCAGTGATCTTGCGCAGAGGGTTTCCACGGTGCTGAGGCAAGCATGGAA
GGAGGATCAAGATGCTGGCCATCTCCTCGGGTCCTTGTACGAGGTCTTTGGTGAAGCCATCTTCTCCTTTGTTAGAGGAT
TTTAATCTGCACTGTTGTATAATTAAGGGTATTCTTGAAATGGTAGACTGACATCAGAAAACAAGTAATGTAATGCTAG
GTTTATATTAATGTGAGGAATGAACATGTGTGTTGTACACTTGTACTATTATTACTTTACGTGACTGATTATCTTAGCCAT
TTGGCTACTAATGTATACGTACATCGTTTAAA

Clone no. (C) 254 (putative arp3 (protein complex)

CGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGCGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTC
TACATTGGGGAGGAGGCGCTGGCGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAA
TGGCCAAGTGGAGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTGACCCGGAA
GATCACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGGAGTACACCGGAGAGATCATGTTCGAGA
CGTTCAATGTTCCTGGCCTCTATATCGGAGTGCAACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAA
ATGACAGGTGTTGTAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGTTATGTCATCGGGAGCA
GCATCAGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAACTTATGAAGGAAAGAGGTGAACACAT
TCCAGCAGAAGAATCTTTGGATGTAGCAAGGAGGCTTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAA
TTCAATAAGCATGACAGGGAGCCCTCAAAGTACATTAAACACTGGAATGGCATCAAACCAAAAACTGGTACTAAGTAT
TCATGTGACATGGGATATGAGCGCTTCTTGGGGCCTGAGATCTTCT

Clone no. (C) 295 (farnesylated protein 1)

GTCCCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGCAGTAAAACGCTAG
TACTAAAAACAGAGCGGCTTTCCTTCTCGATCACCCAGCAAGCACCTCTTTGGCCAAACGCCATCAGCCACA
CCAGCACAGCCAAAATGCTTGGCGCGGACCTCATCACTGAGCTATGCACGGCGCCGACCAAGATCCTCGGCAGGAAGA
AGCGCAAGCAGTTCCAGAAGGTGGAACTGAAGGTGCGGATGGACTGCGAAGGTTGCGAGCGCAAGGTGAGGAAGGCG
CTGGAAGACATGAAAGGCGTGAGCAGCGTGGAGGTGGACACGAAGCAGAACAAGGTGACGGTGTCGGGGTACGTGGA
GACGAAGGAGGTGATAGAGCGGCTGCGCCACCGGGCGGGGAAGAAGGCGGAGCCGTGGCCGTACGTGCCCTACGACG
TGGTGCCGCACCCTTACGTGCCGGGTGCCTACGACAAGAAGGCGCCGCCCGGGTACGTGCGCAACGTGCTCGACGACC
CTGACGTCGGCCCGCTCGCGCGGGCAAACTCCATGGAGGAGAAGTACACGTCCGCCTTCAGTGACGAGAATCCCAACG
CGTGCGTCGTCATGTGATCGGCTCGTGGCACTGTCTCGTTGCCTGTCGAATACGTACCATATAGGTTCCTATGATACTCC
GCGTTTGTAATACAGAGTACTGTACAGTTTTTTGCTCCATCGATAGCTGTAAAAT
TTTGTAATAGTGTTTCCAG

Clone no. (C) 303 (putative aldehyde dehydrogenase WIS1)

GGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGGCTTATGCTTGCAGTGGTCAGAAATGCT
CTGCCCAGTCTGTTCTATTCATGCACAAGAATTGGTCATCTAGTGGGTTGCTTCGAGAATATGAAGAAACTATCTGAAA
GAAGGAAGCTCGAAGATTTGACAATTGGCCCGGTCCTTACTGTTACTACAGAAGCTATGATGGAGCACATGAACAACCT
TCTTAAAATACCAGGATCCAAGGTTCTTTTCGGCGGTGAACCTTTGGAGAATCACTCTATTCCGAAAGTATATGGTGCTT
TGAAGCCAACTGCTGTATTTGTTCCTATAGAGGAAATCCTTAAAAGCGGGAACTTTGAGCTTGTGACAAAGGAGATATT
TGGTCCATTCCAGAATACTCCGAGGATCAGCTTGACTTGGTATTAGAAGCCTGTGAAATTATGAATGCGCATCTGACAG
CTGCTGTAGTTTCAAACAACCCGTTATTCCTGCAGGAAGTACTTGGGATGATCTGTCAACGGTACAACATATGCTGGGA
TTCGAGCAAGGACCACGGGCGCTCCACAAAACCACTGGTTTGGGCCAGCCGGTGACCCGAGAGGTGCAGGCATCGGAA
CTCCGGAAGCCATCAAAATTGTCTGGTCTTGCCACAGAGAGATCATATATGACATCGGTCCCGTGCCCAAGAACTGGGG
CGCTTCCTTCCGCCACTTAATTTTGTGGAGGTTCCCAGTTTAGTAAGGAAGTATAGTTTGCGCCAAGATA

Clone no. (C) 311 (trehalose-6-phosphate synthase/phosphatase)

TCCCGCGGTGGCGGCCGCTCTAGCAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGAAGACTTGGAAAGG

GCTTGCAAGGACCACTTTCGGAGGACCTGTTGGGGTATAGGTTTGGGTTTTGGTTTTAGGGTTGTGGCCTTGGACCCTCA
TTTCACAAAGCTTAACATGGATTCAATTGTTAATGCTTACGAGATGTCAGAGAGCAGAGCTATATTTCTTGATTACGATG
GAACTCTGGTTCCTCAGACTTCCATCAACAGGACACCTAGCCCTGATGTTTTGAACATCATCAACACCCTCTGCTCAGAT
AGGAGGAACATTGTTTTTCTTGTCAGTGGACGCGAAAAGGATAAATTGGGAGAGTGGTTTTCCTCATGTCCTAAGCTGG
GCATTGCGGCAGAGCATGGTTACTTCTTAAGGTGGTCTAGAGATGAGGAGTGGCAAACCTGCGCCCAGGCTTTAGACTT
TGGATGGATGCAAATGGCAAAGCCAGTGATGAATTTGTATACGGAAGCAACTGATGGATCATACATTGAGACCAAGGA
AAGTGCTCTGGTTTGGCACCATCAGGATGCTGACCCAGGCTTTGGATCATCACAGGCGAAGGAGATGCTTGATCATCTG
GAAAGTGTGCTAGCAAATGAACCAGTCTCTGTAAAGAGTGGTCAGTTCATTGTTGAAGTCAAGCCTCAGGGAGTAAGC
AAGGGAGTAGTTGCTGAGAGGATACTTGCAACAATGAAGGAGAGGGGAAGGCCAGCTGATTTTGTGCTATGCATTGGT
GATGACAGGTCTGACGA

Clone no. (C) 321 (putative signal recognition particle receptor)

GCTCCCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGGACAAGGGACAGG
AGAACTATCAGCCCCGTTCTGCTCTTCTTGTCAAAGGAAAGGAGAATGGAGGTCCCAAGGACAACAACAACG
GAGCCTTTGATGTAAATAAGTTACAGAAGCTGCGGAACAACAAGAAGAAAACTGTTCCCGCTGAAATCGGAACCAAGA
AGAACACTACTAAACCTGACACTAAGAAGAAGGTCAACAGGGTCTGGGATGACAAACCTACCGACAACAAGAAGCTG
GATTTTACTGACCCAGCAGACAAGAGAGGGGATGATGAGGTGATCGACCATGCGCTCGTCAACCAGGGTGAGAGCATG
ATGGACAAGGATGAGAACGTCAGCAGCGACAGCGAGGATGAAGAAGAAGAACAGGTGCAAGGTGTCCCCAAAAAGAA
AGGATGGTTCTCCTCAATGTTTCAGAGTATTTCAGGTAACAATGTGCTGGAGAAGTCTGATCTGCAACCTGCGCTTAAA
GCTCTTAAAGATAGGCTGATGACCAAGAATGTGGCAGAGGAAATTGCAGAAAAGCTTTGTGAATCTGTTGCAGCTAGC
CTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATCATCTACTGTACAGACAGCAATGGATGGAGGCTCTTGTTCGC
ATTTTGACCCCAAGGCGATCAATTGACATACTGAGAGATGTGCAGTCTGCCAGGGAGCGTGGGAGGCCTTATGTCATTG
TTTTCGTTGGGTCAATGGAGTTGGCACGTCCACAAT

Clone no. (C) 351 (calumenin homologue)

CCACCGCGGTGGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGCAAGTGCGGCAAGTGAGCCCTGGCCGCC
GCCGGAGATGGAAGGACTTGATAGTCGGAACACCTAATGGCGTCTATCATGTTGATCATACTTGTTTAAGCGTGTGTAG
CAAATGCGTGGCATGCTTAATAACATGCATGCTGCCAGTGGTAGTAATTAAGTACTGAGAATAATAAGGCAAAGCAGA
GAGTATGGCGTGGCCGTGTCCTTCCTCTCCAGCTAGAGTGAGGCTCTAGCTGTTTGTGGTCTAATATCTGCAGTGTGCCC
TGTGCATGTTAGTATGGTGTCTCATGAGGGGGGCCCCGGTACCCAATTCCCCCTAAAGGGAGTCTTATTACAATTCTG

Clone no. (OP) 168 (putative transcriptional regulator)

AGTACATTTTTGAGACACGAATATATCCTTTCCTGACATTAACCTTTCTACACACTGTGATAGTTATAAGAGATTAAACT
AGCTACGCCGATCAAAGTACGTTAACATATCAGCGACAGCGCGCACGAGAAAAAGTAAATACCCAAAAGAGACACCGC
CCCGAGCTTGTATTCGCTGTGTTTATGCCCTCATGGGTCTTATTCTTGACCTGCTTCGGACTCCGCTTGTTCCGCCGGTTG
CGGAAAAAAAATTCGGGGCCGCTCCAACTTTTCCTTGCAAATGATTTGGGATACGCGCCGGTTTTCCTGGCGGGCGATGG
TAAAAAAACAATCTAGGGCCCATACGTATAACTCGGGATAACAAAAGGTATGCGACACGGGTACGTCACCCAAGGGGG
ATTGTATCCCCACTTTGGGTATCTATGCTTAACGGGAGTGAACTGCATTGGACCGGACTCGGTGCTGCTCATATAATCAA
CGTACCAAAAAAACACATTTTATTACTCCCTACTCATCTACAATTTACTCTTCATGGTTCAATAAATCCTACGTTTTCTTC
TTCCATAAACGACCTCTCTTCGTAAATATGGTCCTACCGTACGATCGAACCACGGTCGCGCTTCCACCCTAAATCTGAAA
TGTAAACTTTGCTTCGTATCTTTATAACTTTTTTCTCACTTCTAAATCATATATACACTCAGAAAGAACCCAACGTAAAG
GGCAGCAGCCCCACGCAGCCGTCAGCTATAGACGCAGTCGCTGCGTTATCTGCGATTATGCGACGCGCTTGCTCATACG
AGCGCACGCGCGATAGACTGTGTGACTGTTGTGCGTCCGCTGCTGAGCACGCGCGCGTAGCACAGCGCGAGTTGTGAGC
AACAGTGCGTGCCTTGTCACTCGCGCTAGTATACAGAGTGATGTGACATGCGCTGCAGCAGTCTGATTATACCATGCTA
GTGAGATGACATGCGCGTGTGTCGCAGAGTAGACTGACTGCGCGCATGACCGCGTGTCCGACCCTGTAGATACAGCGA
CGTGAGTGTGCTACACAGTGAGAGCGTCTAGTGACGCACGCAAGTATGCTGTGCACTGCGCGCGTCATCACGCGCGCGC
GCTCAGTGAGACGACTCTGACACATAGTGACGACGCCACAGATGTGTGTAGTGTGAGAGTACCAGTGAGTGAGTCTGC
GCGTGAGCACTGTGAGCTCGCCTGAGATCTGACGACGCGCGTCTGTGGAGTACAGACACCCGCATAGCGTAGCGCCAG
TGTGTACTGTCTGCGTCACTGAGATGTGAGTGCAGAGAAGTGTGTCAGTAGCCTAGAGCGCAAGTGCGTCGAGCTGATG
TGCAAGAGTCGATAGTAGAGTTCTGTGTGAGTGTGTAGAGTCAGCAGAGCATGTCTTGAGCGTGTCTGACTGCTCGCTG
TAGTGCATGTAGAGACAGTAGTGAGTGATCGACTGCGCGCGAGTGCTTCGTAGCGACCTGATGTGAGACGTTGTGACGC
GTGTCTA

Clone no. (OP) 175 (protein phosphatase)

CCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGCATGCACATCGCTGCCCCCGAGTGTCCTCT
TCCGCTGTCTCCTGGCCGGAGATTCTTCGATTCGCATCTCAGTTCTAGCCGTTACACAGTCCTGTGTCGAGCGGCTACGG
CTTCGGCGTTACAGTCTGACTAGGGCCAGGTGAAAAGCGGGTTACATGCGAAGACGGCGGACGAACGAATGCACGTCC
TTGCGCTGGATCTGTACCGGAGGATCTGGCCTTTGCGCAAGTGCAATCTCCTCCCCGAAGGCAGAGCTGCAAGTCCCAC
TCTAGAGGAAGGCCAAAGGAAGATCGCTAGATGAAGATAAACAAAGCGAGCAGTCCCGGTGCCGACTTCCCCATCACC
ATTGTGTGGCTGCTTTCAACGATTTCTTCTGCTCACCTCATTGAAGCTCTTCCGCAACTATGTGTGACGATGCCCTGACG
ACCAACTACCTCTTCATGGTCGACTACGTCGACCGTGGCTACTACTCAGTGGAGACTGTGTCGTTGTTAGTGGCTCTAAA
AGTGCGTTATAGAGACAGAATCACAATATTGAGAGGAAATCATGAGAGCAGGCAAATAACTCAAGTGTATGGCTTCTA
TGATGAATGCTTACGGAAGTATGGAAATGCAAATGTATGGAAACACTTCCTCGATTTGTTTGATTATTCGGCCTCTCAAC
CAGCTCTTATAGAAAAACCAGTATCTTCTGGCGCTACATGGGTGGGTCTCTCTCCATCACTGGGAACACATTTGAGGAT
AATGATTTCCGTGCCCTGTTGATCGCATACAAGAAGAGTTCCCTCA

Clone no. (OP) 254 (seven transmembrane helix receptor)

TACCTAGTTATCATTTTTTTCTTACGGGCTGGTCAGTTCCGGTAAGCGTTTTTTTTATAAGAAAATAACTAAGATCGTCTC
AACCAAAGCTAAGTCTAAAAACAAGTAGCGCAAGGATCAGATCTAAAGTAGTAGCATTAGCAGGCTTTAGGTATTAAA

TAGGTAGAACCCCTTTTGAATATGCATAAGTAGAAGAAGCAGTGTTGAGAAGCGTCTCTGGACTGTTCTACTCCCCAAG
TTCATGTTATTTGTGGATGGAAGAATGTATTGGGAGCATTGCAGGAGGTAGAATATTATATACAATATTTGGTGTCGTA
GAGAGGGTGATGGGCACGTCGTTGTATCTATGATCTTGTAGGGTAGTAGTATATGTTGTATCTCAAGTACATTTCCAGCC
CGTTGTTCGGCCGTTTTACTGCAGTTAGTTTGATTTCTCGGAGGTCTGCAGAACAGTAGTTGTTCCTGGGACGTAGGTAT
TGAATTTTTCAGTTTGTGCCATACTGTGCTACTAATAGTATGTCGCCACGAGTAAACTGATTGCGGGTAATGTTGTCCTT
CGCTATAACCCATGAAATGAATGTGCCCCTAAAGCAACATTTGTTGCTGCAATGTATCAGCTGCTGGGTGTGGGGCGTC
TGTATATTGCTCGCGCTCTCTGTTTTAGTGTCTATTCTGTGTGCACGTCTTTGCTGTTCGTGCGCACCGTGTAGCAGTGTA
GCTGATTACGATAGATTATATGATGATGATGAGTGAGGTATGTGTGAGTGCGACGGAGATGAGCTACTCTTGTATTATT
GTTTGTAGATTGAGTGCATACGTGTCGTTCGATGGGCCGCGCGCAGTGCGTAGCTGCGTGTCAGTCGTTGCGATCTTGTC
GTGTTATAGTGTGATTGTGACTACGTCCGTGTGTAGCGGTACGCTGTATACGCTAAGTGCTGTATTTGCTTGTGTATTAG
ATTGTTGTCAGCATAGTGTCTGTGTGCTAGTCGATTGCGTGAGTATAGTGTGTTGTTTCAGACTAGTGATGTACTATACG
CTATCTACTTACGCTTGTGATGTGCATGCGTAGCTGTCGCGGCCCGTGTCGTGTCTTTGTGTATGTGAGATGTCGTATAC
GGAGCGTGTTGCAGTCGCGCGTACGAGAGCGTGTGTGCTTTTGTTAGCAGCTGTGTTGAGATAGTAGTGTGGCGAGTAC
TGTGTGTGATTTTGATCCTACTCCTCTATCGTCTGACTACCTACTGATTGGTGTATGCTGAGTTGTTAGTTGCTAGCTAGT
GCGCTAGTGTAGCCGCTAGATATTGTGTTAGTGCGCGTGTACTGAGAGAGCGTGTGAGATACGAGCGAGTGCGTGAGTA
GTGAGAGCAGAGTGCGTGCGTGTGTGTGATAAGTACATATATCTGCATGTGAGTATTGTGAGTGAGTAGTTGTGCTTGT
GTGTATAGTGTGTGTGACGTGTTTGTATCGTGTGAGTGTGTGTATGCGCTGCGTGTGAGCGCGGAACGACGATGAGTAT
GTGCCGCGCGTAGCGTGTGTGTGTGTGTATGAGATGTTAGTGTGTGTGAGTGCGCTACGGATTTGTGATGTGCGTGTGTG
CGCTGTTGCACGTACGTGTACTGTGTGATTGCGCGCTTGTGTGCTCAGTTTGCGAGACTAAAGCGAGATGAGAGACGTC
ACCCCGAGCGACATTAGTTGTAGTGTAGGCTCAGTGAGTACGATCGTTATGAGTGTGTTGTTGAGTGTGAATGAG

Clone no. (OP) 53 (Low molecular mass early light-inducible protein HV90,)

GGCGGCCGCTCTGAACTAGTGGATCCCCCGGGCTGCAGGCGTCGCTGGTGCCACTGCTCCAGGGGGAGACCGCCGAGG
CCAAGACCAGCGGACTCATGAACGCCAATGCGGAGCTCTGGAATGGCCGCCTCGCCATGCTCGGGCTTGTTGCGCTCGC
AGTGACCGAATACCTCACCGGCGCGCCCTTCATCAACGCGTAATTAATCATGTCTTATGTGTACATAGCTTAGGCATTCT
CCGTATCAAATTTCCCGGGGCACGTGTATGACGTATATCAGGCGGTGTAATTGTACTAAGTGTACCATCCTGGAAACAT
TAAAAATAGACTGTTTTATGGCCTTCAATTTGCGCTGCAAAAAACATGTTACACTAGTTAAAAAAAAAAAAAAAAAAA
CTCGGGGGGGGGGCCCGGTCCCCATTTCCCCCTATGGGGGGGCCGAATTACAATTCCGGCCCCGGGGGGTTTTATTCAAAAA
GGGGGGGGGGGGGCGCCGAAGCGGGAGGGGAGGGGGGGGGGGGGGGGGGCCGCCGGCCGGCGGGCCCACGCCGCC
GCCAGTCCCCGGGAAGCAGGAGGGGGGGAGAGAAAGAACGACAAAAAAAAAAAAAACAAACAAAAAAACAAAAACAA
AAAAAACCAAACCCTCAAATAATTTATCTTTTATAGAAGTAGGGGGGGGGCGGGCGCGAGCGGAAAAAATCAAATATC
TAAATAAACAGGGTGAATTAAAATAGGAAAAAGATAGTGAGTGAGTATGCAGCTAGGTGATCATACTGAGTCAGCGTG
ATG

Clone no. (OP) 61 (Farnesyl pyrophosphate synthetase)

AAAGGACGGAAAGAGGCAAACACAGAACGAGAAACACAGAGAGAGAGGGAGAGAGAAGGCCGGCCGCCCGCTAC
ATCCCTTGAGTGATGTTTACGTTCTTGTTTCAAGATAAACGACCTTTCCCAAAAAAACCTACAGACGATAAGAAACGGG
CGCGCCTGCCCCCATACTAAAGTGGGATCCCGCGCGCGCGCGTGCTAGGCAATTCTTGCACGAAGGCACCCTGCTTCCG
CGCCTGCGCGGTACCTGTACGTGGGGAATGAGCTCCGAGGAGTGCGCAATGGCGACTGTGATGGCGGCGGCCAACGGC
TCGGGCGGCGACACCAAGGCCGCCTTCGCGGAGATATACGACAAGCTCAAGCAGGAGATGCTCGAGGACCCCTCTTTC
GAGTTCACCGACGAGTCGCTCCAGGGGATCGACCGCATGGTGGATTACAATGTGCTTGGAGGAAAGTGCAATCGCGGG
CTCTCTGTTATTGATAGCTACAAGATACTGAAGGGTGTTGATGTCCTGAGCAAGGAAGAGACATTTCTTGCTTGCACCCT
TGGATTGGTGTATTGAATGGATTCAAAGCTTATTTTCTTGGCGCTTGATGATATTATGGACAACTCTCAAACAAGGCGTG
GAACAACCTTGCTGGAACAAGGTGCCTCGGTGATTGGAAGCTCAATTGCGGTAAATGCATAGAGAATTATCCTCTCCCA
ACCTTTATATCTCGAAATCCTATCTAGACGACCAGCATATGCTGAGAAGGGAATAGCACCAGTAGTGAGATGGAGCATG
AATCGTGCCGATCTGAAGTACTAGTGGATAGCAAGAGGAGAGAGTGCCAAGAGAGTGCTGTGAGACCATCAGCTATCA
GTAGGTGAGTGAGCAGCGACTCGCGAGATCCTGTATCGACAGTAGACCAGCGAGTAGATGGATAGAAGACAGAGGACC
AGAACGCAACAGAAGAAGAGAGATATGTGAGAGAGAGATGAGAGACCGAGCGAGAGTGATGTGGTATAGCATAGCTA
CTGAGTGAGCAGTCCAGAAGAGAGTAGCAGAGTGACGTAAACCTACACACTGAGATAGAAGATAGAGTGTAGAG

Clone no. (OP) 96 (DRP4 protein)

CCTCTGAATCTTGGTGGATCCCCCGGGCTGCTGGAATTCGGCACGAGGTACAACTAGTCACCATGGAGCACCAGGGACA
GTACGGCCACGGTCTACCACCGGCCGCGTCGACAAGTACGGCAACCCGGTTGTCGTGACACGGAACCACCG
GAACCGGCGGTATGGGCATGGGGACACACGGGACCACCGGCACTGGTACCGGCTTGGGCATGGGAACTCACGGAACCG
GTACCGGCGCCGGCATGGGCGGGCAGTTCCAGCCCACCAGGGAGGAGCACAAGACCGGGGGCATCCTGCACCGCTCTG
GCAGCTCCAGCTCCAGCTCGTCTGAGGACGACGGCATGGGTGGCCGGAGGAAGAAGGGCATCAAAGACAAGATCAAG
GATAAGCTCCCCGGTGGCCACAAGGATGACCAGCAGCAGATGGGTACCGGCGGAACCTACGGCCAGCAAGGTACACAC
CGGCATGACCGGTACCACTGGCACTGGCACCGGCACCACTGGCACCTACGGCCACGAGACCGGCGATAAGAAGGTGCA
TCTTGGGACAAGTATCAAGGAAAAAGCTCCCCGGGCCAGCACTAAAGCGGGGTCACGCCGGCGTACTTCCAAAACGGC
TCTGTTGGCCACTAACCTGGTCACGGTACGACGTTTCCTTGGGAACATAATAAGATGGAAGCAGCCCTGGTAATTAAAA
ATGTCCCTTGAAGCACGGCTCCACACGTGCGTGTGGTTATATGTGAGAACAGCGTGTTGTGTCGCTCTGGTCATAGGTTG
GAGCTAATGATGTGTGTGAAAA

Clone no. (OP) 126 (monodehydroascorbate reductase)

GGCGGCCGCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAATTCGGCACGAGGGGATCCCAAGCTGGCTCTTACGCTG
GCAGTGGGCATGGCAAGCCGTGAGCGAGCTTGCCAAGAGAAGGTGCTCGAGACTTTACCAAGATTATTATCTCGCATTA
TTGTGTACTGATTGTATTGCTTCGTCCATGTGTTTCAGCCACCCAGCATCATTTGTGTTTATAACCGCAAATAAGGCTAC
AGCCGGCAGATTAAGAATAAGGTGCTAGCATATGTCTGTCCATGTAAACATTTTATATGGGTGTGTGTCTCACATTTATA
TGTGACCATGTACGATTGCTTATGTACATTTTGCCTAAATAGTAGTATTCAGTTAAGAGAGAAAAAAAAAAAAAAAAAA
ACCCCGGGGGGGGGGCCCGGACCCCATTTCCCCCTAAGGGGGGCCGAATTAAAATTCCGGCCGGCCCGGTTTTTAAAAA
AAGGGGGGGGGGGGCGCGAGCGACACGGGGGGGGCTGCCACTTGCTTCTCTCGTCTCATTCCCCTACGTATATAATTGG

GCGGGCAGGCGCGTATACAAATAAAGACAAAAATAAATATACAAAAAAAATCAATTCCGGGGCGCGAAATTTGGGGG
AGAGTGAGGGGGAAAGAGAAAAAAAAAAAAAATAAAGGGGGGGCGACCAATGTACGGAGAGGGGGGGGGGTAACCA
TAATAGAAGCAGAGAGCGAGAGAAGAGTAATACCAGCGATGATAGAAGATGACAACAAGAGCGCGAAGTAGCTGAGT
GAGAG

Clone no. (NP)175 (RAB17 protein)

GGCGCGTCTTTTTTAATGGATCCCCCGGGCTGCAGGAACCTAGTGATTTGAGTGTTTGTGAGAGCAGCTTACAACTAGTC
ACCATGGAGCACCAGTGGACAGTACGGCCACGTGCACCACTCGGCCGCGTCGACGAGTACGTGCAACCCG
GTTGCCGGACACGGCACCACCGGAACCGGCGGCATGGGCATGGGGACACACGGGACCACCGGCACCGGTACCGGCATG
GGCATGGGAACTCACGGAACCGGTACCGGCGCCGGTATGGGTGGGCAGTTCCAGCCCACCAGGGAGGAGCACAAGACT
GGGGGCATACTGCACCGCTCTGCCAGTTCCAGCTCCAGCTCGTCTGAGGACGACGGCTTGGGTGGCCGAAGGGAAGAA
GGGCATCAAAGACAATATCAAGGATTATCTCCCCGGTGGCCACAAGGTATGGACCAGCTGCATATGGGTTCTGGGCGG
AACCTTCTGCCTTTTAAGGACACACTCTGCTTGACCGGTACCTCTGGCACCTGTTCCGGTTTCCACTGGATACCTACGGG
CACCTTTACTTTGTGTAGTAATTAACGTCTATCATTGTCCCAATATCAAGTGTAAATC
TTCCCCCGTGCCTAATACTTTATCTGTTCTCTCCCCCTTTCTATTCCTAGCTATTCTTTGGCCCTTACCTGGTTTAACTTAT
CTTGGTTATCTACTGTTATAGTATCGGAGCCTGCACCGTATCTGTGTATAATTATTTATTCCCTTTGATATCCACTGGTCC
ACCATAGTCGTTGTCATTCGTATATCGTATGAGTATGTACATGCCTGGTCTTGTTCGGGCTATGTTGTGTGAATACTGTCT
CTGTTGCTTGTAT

B. Sequence information of excised phagemid clones

Clone no. (C)159 (putative cyanase)

TTTACTGGACATAAACACAAACATGGTAAAAACAAACTCAGTCCTCCACTCAAACAAGTGAAGAAACAAGCAGGAAAT
ACCGTACCGATGGATACAATACAACAACACAACACGAGTAAATGGATGGATCCTCGCTCTCTCGCACGAACACACACA
CTGAGAACACTCTTCTACGGAGCGTATAGGAGTGGCAGTGTACACGAGGGAGGGATTTTATTCATCATCAGTGGAGCTG
GTGGTCGGGGAGTGGACTCATCATCATCATCATCATCGCACAGACGATGACACTACGTATATATAGTGGAATAATGGCG
ATCATCGATGGAAGTGTGGTTTAACTTAATTAATTTAGCTGACCCTGGAGCAGTCGATGGAGGCGCTGATGGCGTAGGG
GAGGCTGACGCCGCACTTTCCGGGGATGCTGGAGGCCTTGCCGGCGTTGAGGCCGCTGACCCTGGACGCGGCGGACTTG
AGGCAGTTGCAGGCGGCGCGGCGGTCGGCGGTGGTCCTGGCGGC

Clone no. (C) 161(Saposin-like type B, region 1 family protein, expressed)

Reaction failed

Clone no. (C) 163 (MT-like protein)

TCTAGAAAAATGAAATGAGGATAGGAAAACGAATGATATATCGTTTTGCTTCAATCCATGACAAGACAAAAGGGTGAA
GCAGCAAGCACCCATGGTTCTGAGCTAATAGGCAGGTGCAGGATAGCAAAACATAAGGAGCCGCAGCGGCAGTAGAA
GACGATCTAGCATGTTTNGGTAGAAGCAAAGAGCGGGACGTTGAGATCCGTTGCAACNCCGAAGATGACGGTGGACGG
GGCGCCGGTGGTTTTCTCCTGGATATCAGGATACATATAGCCGCACTTGCAACTGGAACCACAGTTGCAGCTTGATCCG
CAGCTGCAAGACATCCTGTCTTGGNTGATNTGTGAGACGAAGGAAGACTCTGGC

Clone no. (C) 172 (RISBZ4)

AAGCTAAAATTACTTGAAGCCACAAGAAATGCTGCTTCAGAATGCCAGGGTTTCACTTAGCTGCTGGCCACGATTTTCA
CCACCGAGTATTCAGAACTACTAACTGCATATATCTGCAGTCAAACTAGAAGAGCGCATATAACACCAGGGATACGAA
TAATAATGCTTATCCTTATGAGTTCAGTATCATAGCACATAGATATGTCTATTAGTCATCAGGTCTATAACAGGACTAGG
CAATCAATATACCAGCACGAGGACTCTGAAAAACTCAAGTCCGTTCAGTTCATGACTCCAGCTAACACAGAAGCAGCTC
ACAAATTCTTGCAGCTCATATAACACGCGTACAGAACAGCAGCTGCCGCCAGCGAATTCCAAGCTTCAGCTTCATCTCT
TCACTTGGACAAGCCACCGTTGGAGTGAGAGTCCCAGGGCCAGACGTCGACGCCGGCGCCGCCGCGGCCACAGCTCGT
CACCTCGCTGGCCATGCGGTTGTCGAGGCTCTCGAGGCTGGTGGCGATGCTCTGCAACGGCGAGTTCTGCATATGCTCC
GTCGGGGAGAGGCCCGTGAAGCACGCGTCGTCGGCGGGGAAGTCCAGGCCCACGAGTACGTCCGGGCCGCGGCACGCC
TGCCGTGGGTTCAGCACCGGCGACAGGCCCAGGTGGCCCAGCCCGCATGACATTGTGCCGCGCGCCACCATGTCCTCTG
CCAGCTTCACCTTGACTCTGAGGGCCTCGACGTCTGATTTCAGGATCCTGTTGTCCGTGACTGAGGTTGTGAATTGCTGG
TTGGCGTCTGTCAGCTGCTTGAAGAGTGAAGCATTCTCACCTCGGAGCTGGTCAACCTGTGTTTCAAGGTCTGTTAAGTG
AGCTTGCTTCCTCTTCCTCGAGC

Clone no. (C) 187 (DIS1; protein binding / structural constituent of cytoskeleton)

CNAGTATCAGTACATAGNATATCGGANCTNAGATTCGCTATCTTTTNGATTGTGATGTAACTGATATGGAAATCTTTCAC
ATTGTCTACCCCAAACCAGGCAAAATATGATCAGAGACGTCATCGTTTTGATTGGGGAGATTGGTGCACTTTTCATTAG
AACTTGCAATGCATAGTTATTCATGTTGTAATTCTGTGTGGAGAAGATTCAGGCTCCAGAAGCGTTCATATCCCATGTCA
CATGAATACTTGTACCAGTTTTAGTTTGATGCCATTCCAGTGTTTAATGTACTTTGAGGGCTCCCTGTCATGNTTATTGAA
TTCCTTCACAATGTCTGAACAAGTATAGCAATACATTTCTTTAAGCCTCCTTGCTACATCCAAAGATTCTTTTGCTGGAA
TGTGTTCACCTCTTTCCTTCATAAGTTGCTGAATAAACTGGGTAACATCCTTGCCTGTAATTGGAATAGATCTGATGCTG
CTCCCGATGACATAACCATCAGCAACAGGAACGATGTGGGTAGCTCCATCACCCACATCAACTACAACACCTGTCATTT
CACACTTGGTAGTAGTGTATCCGGCAGCAAGAGCAAGGACCGGTTGCACTCCGATATAGAGGCCAGGAACATTGAACG
TCTCGAACATGATCTCTCCGGTGTACTCCCGGGTCTCAGGAGCAGTAAGGGGACTCTCGGTTAGAAGAAAATAGTGATC
TTCCGGGTCACACCGGAGATAATTGAAGATGCACTGCTGCCAGAACCTCTCCATAGTGTCCCAATTCTCCACTTGGCCAT
TGCGGATGGGGTAGCTGAGGCTGTACGTGCTGCTGGCACGGGACCGCGCCAGCGCCTCCTCCCCAATGTAGAAGTCGA
GGTCGGCCATGACGCCGGCGCTGTGCTGCG

Clone no. (C) 217 (actin-related protein ARP3)

AGAACGGACCCGTGCATCCACTATCTTTTAAGGTCTCGTGCAATCTCTATGGAAATCCTTAAACATAGTCGAACCCCCA
GACAAGACAATATTCTTATAAAGAGACCTCCTTGTATCGATTGGGGATGATTGGATGCACTTGTCAATAACAACTTGCA
AAGGAGTGGTGAAGTCATTGTTGTAAATTTCTGGGTGGAAGAAGATCTCAGGCCCCAAGAAGCGCTCATATCCCATGTC
ACATGAATACTTAGTACCAGTTTTTGGTTTGATGCCATTCCAGTGTTTAATGTACTTTGAGGGCTCCCTGTCATGCTTATT
GAATTCCTTCACAATGTCTGAACAAGTATAGCAATACATTTCTTTAAGCCTCCTTGCTACATCCAAAGATTCTTCTGCTG
GAATGTGTTCACCTCTTTCCTTCATAAGTTGCTGAATAAACTGGGTAACATCCTTGCCTGTAATTGGAATAGATCTGATG
CTGCTCCCGATGACATAACCATCAGCAACAGGAACGATGTGGGTAGCTCCATCACCCACATCAACTACAACACCTGTCA
TTTCACACTTGGTAGTAGTGTATCCGGCAGCAAGAGCAAGGACCGGTTGCACTCCGATATAGAGGCCAGGAACATTGA
ACGTCTCGAACATGATCTCTCCGGTGTACTCCCGGGTCTCAGGAGCAGTAAGGGGACTCTCGGTTAGAAGAAAATAGTG
ATCTTCCGGGTCACACCGGAGATAATTGAAGATGCACTGCTGCCAGAACCTCTCCATAGTGTCCCAATTCTCCACTTGGC
CATTGCGGATGGGGTAGCTGAGGCTGTACGTGCTGCTGGCACGGGACCGCGCCAGCGCCTCCTCCCCAATGTAGAAGTC
GAGGTCGGCCATGACGCCGGCGCTGTGCTGCG

Clone no. (C) 223 (14-3-3 protein)

TGGAAGGGTAAAAGCACGGCCAGGAAAGTGTAATGTGACTCATTAAGGATCGCATCGGGTACCTGGCCCCCCCTCTTCC
CCATTAATTCTTTTTTTGAGAGCAGAAAGCCAGACGTTTGACATTCACCCGTGGGTTTGACGGGTTCACCGTCATAGTAT
AAGCGAACTAGAACTCTGTATAAAACAAGGGTCGACGGTACGCGAAGGGCGTAATAGAAATGTCGGCCTGAAATAGAC
TCCATTTAATAGATGACAACACAACATAGCAAGTTTTGCCCGCCACTCCCAACTTTCCACCAAGTGGATCACCCTAGCC
ATGGTTGAAGGACTGGTACAGGGGAACTTGAGGGATCATC

Clone no. (C) 237 (sinapyl alcohol dehydrogenase)

TAGTGATGCAACTTTAATCAATTCGTGTTATGCATCACAGTCATGCAGTCATCAGTAGCACCCGCAATGTTTAAATAGG
ATTACATCATAGCAAAATAGCATAAACACAATGCTCTGAGTTACATAGTACGATAACTAGTAAGAATACATGTTCTGTT
GATACAATAAAACACAATACTTGCTTATTGAAAGGAACGCGGCAAAGCCGAGCGGACTCACGAGTCACGAGTCACGAC
ACACATGAAGCAACGCCCAACTGTGTTGCAGGGCGAGGTATCCTAACAGCAGCGGGACACTCGCTGGCAACATGCATG
CTCACTTCGAATCGCATTCAAATTGCTCGGCGCAGTCGGGATAACATGACGGAACAAAGCGGCACCAAATCCGGCAATT
CGCGTAGCAAGTTTTGCACGCCTCAGTAGCACAGTCTTCGACGCCGCCGTTTTTGGGGATCAGCTGGGAGAGCACATCG
GCGAGCTCCTGCGCCAGGGGCATCTGCAGTAGCACCTGCTGCTCCTTCTCCGGCGTCGCCCGACCATGAACAACGGCCG
CCAGCTGAGGTCCCAAGACAACGCACAGGACCAGCGCGCAGATGGCAACCATCTTCGGCGCAGCAGCGGAGGAAGAA
GCCATTGTTTGTAACACAGGTGAGCTCGCTGGCTACTAACGAAGGACTGGGTGTGTGTACTTAGGAAACTTTC

Clone no. (C) 246 (putative TAF6)

TTTAAACGATGTACGTATACATTAGTAGCCAAATGGCTAAGATAATCAGTCACGTAAAGTAATAATAGTACAAGTGTAC
AACACACATGTTCATTCCTCACATTAATATAAACCTAGCATTACATTACTTGTTTTCTGATGTCAGTCTACCATTTCAAG
AATACCCTTAATTATACAACAGTGCAGATTAAAATCCTCTAAAGCCTACGCATTTCCTTTCTACACAAAGATGGATATCT
CTGGTGGTTGAACAAAGGAGAAGATGGCTTCACCAAAGACCTCGTACAAGGACCCGAGGAGATGGCCAGCATCTTGAT
CCTCCTTCCATGCTTGCCTCAGCACCGTGGAAACCCTCTGCGCAAGATCACTGCCCTGTTCTCTCCGAACCGCAACTGCA
ATGTTTTCCACCATTTGTGAAGTGGATGAAGCTTGCATCATTCCAGGGTTGGCTAACTGGGTGGAGAACCCATCCATGG
TTCCTTGCATGTTTCCTGCCATAGAAGACGCAG

Clone no. (C) 254 (putative arp3 (protein complex)

Reaction failed

Clone no. (C) 295 (farnesylated protein 1)

TAAGAATACTGGAAACACTATTACAAAATTTTACAGCTATCGATGGAGCAAAAAACTGTACAGTACTCTGTATTACAAA
CGCGGAGTATCATAGGAACCTATATGGTACGTATTCGACAGGCAACGAGACAGTGCCACGAGCCGATCACATGACGAC
GCACGCGTTGGGATTCTCGTCACTGAAGGCGGACGTGTACTTCTCCTCCATGGAGTTTGCCCGCGCGAGCGGGCCGACG
TCAGGGTCGTCGAGCACGTTGCGCACGTACCCGGGCGGCGCCTTCTTGTCGTAGGCACCCGGCACGTAAGGGTGCGGCA
CCACGTCGTAGGGCACGTACGGCCACGGCTCCGCCTTCTTCCCCGCCCGGTGGCGCAGCCGCTCTATCACCTCCTTCGTC
TCCACGTACCCCGACACCGTCACCTTGTTCTGCTTCGTGTCCACCTCCACGCTGCTCACGCCTTTCATGTCTTCCAGCGCC
TTCCTCACCTTGCGCTCGCAACCTTCGCAGTCCATCCGCACCTTCAGTTCCACCTTCTGGAACTGCTTGCGCTTCTTCCTG
CCGAGGATCTTGGTCGGCGCCGTGCATAGCTCAGTGATGAGGTCCGCGCCAAGCATTTTGGCTGTGCTGGTGTGGCTGA
TGGCGTTTGGCCAAAGAGGTGCTTGCTGGGTGATCGAGAAGGAAAGCCGCTCTGTTTTTAGTACTAGCGTTTTACTGC

Clone no. (C) 303 (putative aldehyde dehydrogenase WIS1)

TCGAGTGCCACAGAACAATACTACATCACAGTCACAGCCTACAAGATCCATATCTGGCGTTTATCAATGTTTCTGGAGG
GTATGGAGAATAGCTACACGGGCAAATTATTCCTATACACTCCGTTCCTGACTACTCCGGCCCTGTCAGCTTTTCTTCCC
CATTTATTCTTGGCGCAAACTATACTTCTTTACTAACTGGGAACCTCCACAAATTAAGTGGCGGAAGGAAGCGCCCAGT
TCTTGGGCACGGGACCGATGTCATATATGATCTCTCTGTGGCAAGACCAGACAATTTTGATGGCTTCCGGAGTTCCGAT
GCCTGCACCTCTCGGGTCACCGGCTGGCCCAAACCAGTGGTTTTGTGGAGCGCCCGTGGTCCTTGCTCGAATCCCAGCA
TATGTTGTACCGTTGACAGATCTCCCAAGTACTTCCTGCAGGAATAACGGGTTGTTTGAAACTACAGCAGCTGTCAGAT
GCGCATTCATAATTTCACAGGCTTCTAATACCAAGTCAAGCTGATCCTCGGAGTATTCTGGAATGGACCAAATATCTCCT
TTGTCACAAGCTCAAAGTTCCCGCTTTTAAGGATTTCCTCTATAGGAACAAATACAGCAGTTGGCTTCAAAGCACCATAT
ACTTTCGGAATAGAGTGATTCTCCAAAGGTTCACCGCCGAAAAGAACCTTGGATCCTGGTATTTTAAGAAGGTTGTTCA
TGTGCTCCATCATAGCTTCTGTAGTAACAGTAAGGACCGGGCCAATTGTCAAATCTTCGAGCTTCCTTCTTTCAGATAGT
TTCTTCATATTCTCAAGCAACCCACTAGATGACCAATTCTTGTGCATGAATAGAACAGACTGGGCAGAGCATTTCTGAC
CACTGCAAGCATAAGCC

Clone no. (C) 311 (trehalose-6-phosphate synthase/phosphatase)

Reaction failed

Clone no. (C) 321 (putative signal recognition particle receptor)

TCTTTTCATAACCTTCTCAAATATGGTATCTTAAGCTGCGTGCATGAGTACGTAGCTGCTCAACAGCTCCAGACTGAAGG
TGTCACACGCTGCAAGTGTTACACTGAGGTTATGCTGAAGAAGCCAATAAGCGACCTTTGCAAGATTGGTGGACTTGCC
AACTCCATTGACCCCAACGAAAACAATGACATAAGGCCTCCCACGCTCCTTGGCAGACTGCACATCTCTCAGTATGTCA
ATTGATCGCCTTGGGGTCAAAATGCGAACAAGAGCCTCCTCCATTGCTGTCTGTACAGTAGATGATATTCTTGTGAAAG
ACCCTAGCTTTTTTCCCTCAAGGCTAGCTGCAACAGATTCACAAAGCTTTTCTGCAATTTCCTCTGCCACATTCTTGGTCA
TCAGCCTATCTTTAAGAGCTTTAAGCGCAGGTTGCAGATCAGACTTCTCCAGCACATTGTTACCTGAAATACTCTGAAAC
ATTGAGGAGAACCATCCTTTCTTTTTGGGGACACCTTGCACCTGTTCTTCTTCTTCATCCTCGCTGTCGCTGCTGACGTTC
TCATCCTTGTCCATCATGCTCTCACCCTGGTTGACGAGCGCATGGTCGATCACCTCATCATCCCCTCTCTTGTCTGCTGGG
TCAGTAAAATCCAGCTTCTTGTTGTCGGTAGGTTTGTCATCCCAGACCCTGTTGACCTTCTTCTTAGTGTCAGGTTTAGTA
GTGTTCTTCTTGGTTCCGATTTCAGCGGGAACAGTTTTCTTCTTGTTGTTCCGCAGCTTCTGTAACTTATTTACATCAAAG
GCTCCGTTGTTGTTGTCCTTGGGACCTCCATTCTCCTTTCCTTTGACAAGAAGAGCAGAACGGGGCTGATAGTTCTCCTG
TCCCTTGTCC

Clone no. (C) 351 (calumenin homologue)

GAACGGACCCGTGCATCCACTATCTTTTAAGGTCTCGTTGCAATCTCCTATGGAAATCTTAAACATAGTCGAACCCCCAG
ACAAGACAATATTCTTATAAAGAGACCTCCTTGTATCGATTGGGGATGATTGGATGCACTTGTCAATAACAACTTGCAA
AGGAGTGGTGAAGTCATTGTTGTAAATTTCTGGGTGGAAGAAGATCTCAGGCCCCAAGAAGCGCTCATATCCCATGTCA
CATGAATACTTAGTACCAGTTTTTGGTTTGATGCCATTCCAGTGTTTAATGTACTTTGAGGGCTCCCTGTCATGCTTATTG
AATTCCTTCACAATGTCTGAACAAGTATAGCAATACATTTCTTTAAGCCTCCTTGCTACATCCAAAGATTCTTCTGCTGG
AATGTGTTCACCTCTTTCCTTCATAAGTTGCTGAATAAACTGGGTAACATCCTTGCCTGTAATTGGAATAGATCTGATGC
TGCTCCCGATGACATAACCATCAGCAACAGGAACGATGTGGGTAGCTCCATCACCCACATCAACTACAACACCTGTCAT
TTCACACTTGGTAGTAGTGTATCCGGCAGCAAGAGCAAGGACCGGTTGCACTCCGATATAGAGGCCAGGAACATTGAA
CGTCTCGAACATGATCTCTCCGGTGTACTCCCGGGTCTCAGGAGCAGTAAGGGGACTCTCGGTTAGAAGAAAATAGTGA
TCTTCCGGGTCACACCGGAGATAATTGAAGATGCACTGCTGCCAGAACCTCTCCATAGTGTCCCAATTCTCCACTTGGCC
ATTGCGGATGGGGTAGCTGAGGCTGTACGTGCTGCTGGCACGGGACCGCGCCAGCGCCTCCTCCCCAATGTAGAAGTCG
AGGTCGGCCATGACGCCGGCGCTGTGCTGCG

Clone no. (OP) 168 (putative transcriptional regulator)

TCTCCCGTCTGATGCTGGAAGAACACTGCACATAACTGGTTCAACAGACATGGCTAGCTTCTCCTGAAGGCCGTGTGAC
AGCCGTCTCCGTAGTGATTACCGTGGAATGAATTTTGTTAATGTTTTCAGCTCTGGACATTTAGAATAAGTCCCACTGCC
AGTATATCTTCTCGACTATGCCTCCGCCGCTCCGCACTGCAACATGGAGGGTCTCTGAGGCTGGAGGAACACCCCCACAG
CAATGGTACTGCAACAGACATGGTTAGCTTCTCCTGAAGGCCATCGTTGCTTGCCTTCTCTGCCGTGACAATCTCACCGC
CATGCTCCAATACTGTATCTTCCACTACATCGCTTAGCATTTCAACAGAACGACAAAAGCCAAATATACATAATCTTCTG
CCCATGTCTAGGCCATTTGTTGCTGCGATGTTAAACAGTGCTCCAAGGCAAAGGCTTTGATTTGGGTCCATGTAAGAGCT
CTCCTGAAAACGGCGAAGTTTATGTGCACCTTTCACTTTTCGTTGCCTCTTCTCAGAAGCACGAAGGCGACGAGGAACA
AGGGAACCCCCAATTTGGATCAATCCATCAGAAGTGAACAGTGTGACCCTGTCTTCTGGTACCCACTGAATGTCGAAGT
CTGTGGTGCGATCCAGGGGCTGTTCCCAGGGGTAGGGGCCGTAGAGGTCCTCCGGATCATCCTTCCGTGCCCGAGCCTG
CAACCGGTGGAGCGACCGCGGCCCGAACCGGCAGCAGCGCGTGTGCGGGGCCCCGGCGTAGAGAGGGGTGGGCCTCCT
GGACGCCAGCGAGCGGGGAAGGGTGGAGGAGGAGGCGGTGGAGAGCGCGGCGGAATCGGCCATGATGGGAAACGGG
CGACGAGTCTGACGTTCTGAGGC

Clone no. (OP) 175 (protein phosphatase)

CACCAATTGAACCTTCCATTACAAGCTGATGAGCCCTTGAAATAAGACTAAGCCCGTGGTATGGTTGAATTGTGTGCAA
TATCTTGCCCAAATGTGTACCTGCTCCTCTTGGTGAAATTCCCCATCCACATCGGTCATCCGGGTCAGACCACAAAAGAT
CACACATAGGTCTTCATGAGGGACCTCTTGTATGCGATCAAGGGCACGAATATTATCCAATGTGTCCAGTGATGGAGAG
AGACCACCATGTAGGCAGAAGATCTGGTTTTCTATAAGAGCTGTGAGAGGCAAATAATCAAACAAATCGGTGAAGTGC
TTCCATACATTTGCATTTCCATACTTCCGTAAGCATTCATCATAGAAGCCATACACTTGAGTTATTTGCCTGCTCTCATGA
TTTCCTCTCAATATTGTGATTCTGTCTCTATAACGCACTTTTAGAGCCACTAACAACGACACAGTCTCCACTGAGTAGTA
GCCACGGTCGACGTAGTCGCCCATGAAGAGGTAGTTGGTGTCGGGCGTGTCACCGCCGATGCGGAAGAGCTCGATGAG
GTCGTAGAACTGGCCGTGGATGTCACCGCAGACGGTGACGGGGCAGCGCACGGGCTGCACGTTCCATTCCTCCATGAG
GATCGCCTTCGCCTTCTCGCAGAGTGCTTTCACCTCTGCCTCGGGGAGGAACTTGCACTCCCGCAGCTGCGAGATCTGCC
GGTCCAGATCCGCGTGCGACGACATCGCCGCTGCCGCTCGACGTCACCCTCCCTTCCCTTACCTAGTCCCCTGTAACGCC
GAAGCCGTAGCCGCCGCCAGCGTTGCCGCAGAAGCCGAGGAGGATCAAGGGTCCCGGCGAGCCGGCCGGCGATGTGG
AAACCGGGGCCGCGAGGGATC

Clone no. (OP) 254 (seven transmembrane helix receptor)

TTTGAAGAATAAGATGTAATTATTGTAGGACCTATCATCCAAAGAATTAAGTGAACATAATCAAAAGTATGTACAAAAG
TGAAACCTAGTTATCCATGAATATTTTGGATTTCAACTTTCAAAGTTTGTAGCTTACGTTCAACACTGGACAACTCACCC
TGAATACCCTGTAACCAATCTACCTTTTGTTGACATGCATCCTGGTATGGTTTGCTGATAGTCTCAACAAAATGGACTAG
TCTATTGGCACATAGCGATATATCCTTCTCGATAGCTTCGTCAACTTTCCGAGATAGTGCATCAGCAGACTTGGTGC

Clone no. (OP) 53 (Low molecular mass early light-inducible protein HV90)

TAACTAGTGTAACATGTTTTTTGCAGCGCAAATTGAAGGCCATAAAACAGTCTATTTTTAATGTTTCCAGGATGGTACAC
TTAGTACAATTACACCGCCTGATATACGTCATACACGTGCCCCGGGAAATTTGATACGGAGAATGCCTAAGCTATGTAC
ACATAAGACATGATTAATTACGCGTTGATGAAGGGCGCGCCGGTGAGGTATTCGGTCACTGCGAGCGCAACAAGCCCG
AGCATGGCGAGGCGGCCATTCCAGAGCTCCGCATTGGCGTTCATGAGTCCGCTGGTCTTGGCCTCGGCGGTCTCCCCCT
GGAGCAGTGGCACCAGCGACG

Clone no. (OP) 61 (Farnesyl pyrophosphate synthetase)

CTCTCATCAGCATGCTCAAGAGCTTGCACGACTAACCAGGAACACTTATTATCTTCAATGTCAGTTCCGATCTGCCAATG
AATTCAGGATCACCAAAACAATCTAGATAATCATCTGAACTTGAAAGTATGTCCCCATGTCAACAAGAATGTTCTTGAC
ATCACCAAAGTTATCCAAGTTCTCGCCAGAAAGCAGCAGTGCACATGCAACAGGAAGATAGAATGAATAGTAGGCTGT
CTTGTATTGAACAATGCGGCGGTGAACTGTCAAGTTATATTTAGATAGGTCCTTTTCACCCTCATGAGTAGTGATAAGAT
CCAGCAACTGCCCTGACGCTGTCTTGAACTCAACCTCATTGAACAAATCAATGAGATCAACATAATATGGCTTTCCTTTG
AAGTGGCGTTGAAGGATCCGAGATATATGGTTGCGAAGGATAATCCCATCATTTACAGCAATGAGACCAACCTGAGGC
ACCCTGAACCAGCAAGGCTGACCACGCCTTGTTTGAGAGTTGTCCATAATATCATCAAGCACAAGAAAATAAGCTTGAA
GCCATTCAATACACCAACCAAGGGTGCAAGCAAGAAATGTCTCCTCCTTGCTCAGGACATCAACACCCTTCAGTATCTT
GTAGCTATCAATAACAGAGAGCCCGCGATTGCACTTTCCTCCAAGCACATTGTAATCCACCATGCGGTCGATCCACTGG
AGCGACTCGTCGGTGAACTCGAAAGAGGGGTCCTCGAGCATCTCCTGCTTGAGCTTGTCGTATATCTCCGCGAAGGCGG
CCTTGGTGTCGCCGCCCGAGCCGTTGGCCGCCGCCATCACAGTCGCCATTGCGCACTCCTCGGAGCTCGTCCCCACGTCA
GGTCCCGCAGGCGCGGAGAGGGTC

Clone no. (OP) 96 (DRP4 protein)

Reaction failed

Clone no. (OP) 126 (monodehydroascorbate reductase)

TTCTCTCTTAACTGAATACTACTATTTAGGCAAAATGTACATAAGCAATCGTACATGGTCACATATAAATGTGAGACAC
ACACCCATATAAAATGTTTACATGGACAGACATATGCTAGCACCTTATTCTTAATCTGCCGGCTGTAGCCTTATTTGCGG
TTATAAACACAAATGATGCTGGGTGGCTGAAACACATGGACGAAGCAATACAATCAGTACACAATAATGCGAGATAAT
AATCTTGGTAAAGTCTCGAGCACCTTCTCTTGGCAAGCTCGCTCACGGCTTGCCATGCCCACTGCCAGCGTAAGAGCCA
GCTTGGGATCCC

Clone no. (NP)175 (RAB17 protein)

GAACAACTTAAACACAGTGACCATACATGAGACACCATACTACATGCACAGGGCACACTGCAGATATTAGACCACAAA
CAGCTAGAGCCTCACTCTAGCTGGAGAGGAAGGACACGGCCACGCCATACTCTCTGCTTTGCCTTATTATTCTCAGTACT
TAATTACTACCACTGGCAGCATGCATGTTATTAAGCATGCCACGCATTTGCTACACACGCTTAAACAAGTATGATCAAC
ATGATAGACGCCATTAGGTGTTCCGACTATCAAGTCCTTCCATCTCCGGCGGCGGCCAGGGCTCACTTGCCGCACTTGC
AGGGGTTGCAGCTGCAGTTGGGGCCGCACTTGCAACCGTCGTTCTCAGCTCCGGCAGCCTCGAACCCGTCCTCAGCGTG
CCCCTTGGAAGGCGAAACACCCATGATGACGGTCTGGGTGGTGGTGCTCACCTCCTCGGCCATCTCCGGGTACATCTTG
CACCCTCCGCAGCCGCTGCCGCACTTGCAGCCAGAGCCGCACCCGCAGTTGCCTCCGCAGCACGACATCTTTCAGATCG
AGAGATCAACAACAGAACTCAATACAAG

C. Alignment files of sequenced plasmids with existing sequence information

1. **Clone no. (C)159 (putative cyanase)**

```
CLUSTAL 2.0.12 multiple sequence alignment

Old             ------------------AGACATAGCCCTTTTCTCTGGTGTGAGAAGTGACAGCACATG 42
new             GCCGCCAGGACCACCGCCGACCGCCGCGCCGCCTGCAACTGCCTCAAGTC-CGCCGCGTC 59
                                  *   ** *      *   **    ****  *  * * *

Old             TGTATTTTGCCGCCACCTGCTTGACGAAGTT----ATGTCC----AAACTGAAAGATC-- 92
new             CAGGGTCAGCGGCCTCAACGCCGGCAAGGCCTCCAGCATCCCCGGAAAGTGCGGCGTCAG119
                     *  ** *** *      * * * *         ***    *** **     **

Old             ---CTGATGCAGAGTTTGAGATAAT-TCAAATTCTT----ATCAAGGAGTGTAATAAGAT144
new             CCTCCCCTACGCCATCAGCGCCTCCATCGACTGCTCCAGGGTCAGCTAAATTAATTAAGT179
                   *   * *    *  * *      ** * * **      ***   *   **** *  *

Old             TGAAGGCCATGTACAGCAGTGCAAGAGACTGGTCCTGCAGTACATTCCTCTGATCTTGGT204
new             TAAACCACACTTCCATCGATGATCGCCATTATTCC-ACTATATATACGTAGTGTCATCGT238
                * **   **  * ** *  **   *  * *  ***  *  ** ** * *    ** * **

Old             CAATGGTGAGAAGTTCCTTGAGAAGAATGATGTCTGTGCTC-TCGTCCAAGCTTGTGACG263
new             CTGTGCGATGATGATGATGATGATGAGTCCACTCCCCGACCACCAGCTCCACTGATGATG298
                *  **    ** * *  *   ** ** *    **   *  *  *  *    **  *** *
```

```
Old             CCAG-----CCAAAAGAGGGCAGTCAGCTCATTCTTGGAGGCCGGGTTGCTGAGTGAT--316
new             AATAAAATCCCTCCCTCGTGTACACTGC-CACTCCTATACGCTCCGTAGAAGAGTGTTCT357
                         **      * * *  * ** ** ** *  * **   ** *  ***** *

Old             --GCATGAAGA---AAACTGGAAGGCAAAAATAGTGTCCAGTTATTTGTGTTGTAC-ATA370
new             CAGTGTGTGTGTTCGTGCGAGAGAGCGAGGATC-CATCCATTTACTCGTGTTGTGTTGTT416
                  *  **          *  **  ** *  **    **** *** * *******    *

Old             ATAAGGCAATAAATGGCTGTTGTCTTATGCTTCTAGTCTTTGCACCCTTGTTGCATTATA430
new             GTATTGTA-TCCATCGGTACGGTATT-TCCTGCTTGTTTCTTCAC--TTGTTTGAGTGGA472
                 **  * * *  ** * *   ** ** * ** ** ** * * ***  *****  * *  *

Old             AAGGGAAGTGAGCGTTATGAATACCGATCACTGTGTCAAGTTAGTGTGT 479
new             GGACTGAGTTTGTTTTTACCATGTTTGTGTTTATGTCCAGTAAA----- 516
                      ***  *  **    **     *   * **** *** *
```

2. Clone no. (C) 161(Saposin-like type B, region 1 family protein, expressed)

Reaction failed

3. Clone no. (C) 163 (MT-like protein)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GCCAGAGTCTTCCTTCGTCTCACAGATCAACCAAGACAGGATGTCGTGCAGCTGCGGATC 60
new             GCCAGAGTCTTCCTTCGTCTCACANATCANCCAAGACAGGATGTCTTGCAGCTGCGGATC 60
                ************************ **** ************** **************

Old             AAGCTGCAACTGCGGTTCCAGCTGCAAGTGCGGCAAGATGTACCCTGACATCGAGGAGAA120
new             AAGCTGCAACTGTGGTTCCAGTTGCAAGTGCGGCTATATGTATCCTGATATCCAGGAGAA120
                ************ ******** ************ * ***** ***** *** *******

Old             GAGCACCGGCGCCCAGACCACCGTCATCCTCGGCGTTGCAACGGAGCAGAAGGTCCAGTT180
new             AACCACCGGCGCCCCGTCCACCGTCATCTTCGGNGTTGCAACGGATCTCAACGTCCCGCT180
                 * *********** * *********** **** *********** *  ** **** * *

Old             CGAGGCGGCCACCGAGTCCGGCGAGACCGCCCACGCCTGCGGCTGCGGCGCCAACTGCAA240
new             CTTTGCTTCTACCNAAACATGCTAGATCGTCTTCTACTGCCGCTGCGGCTCCTTATGTTT240
                *   **  * *** *  *  ** *** ** *  *  **** ******** **   **

Old             GTGCAACCCCTGCACCTGCTGATCTGATCAGTTTCAGGGGTGCTTGCTGCTTCCTGCTTG300
new             -TGCTATCC-TGCACCTGCCTATTAGCTCAGAACCATGGGTGCTTGCTGCTTCACCCTTT298
                 *** * ** *********  **  * ****   ** ****************   ***

Old             TGTCATGACATGCGTGTGAGCACAACAGTATATAATAAGGAGGAGCTAGCCATGTACCGA360
new             TGTCTTGTCATGGATTGAAGCAAAACGATATATCATTCGTTTTC-CTATCCTCATTTCAT357
                **** ** ****  *   **** ***  ***** **  *      *** **   *  *


Old             CGATCTGGCTCTGGAGCTTGCATCAGTTCTTTGTGTGGTTTGAAACCTTAATTTTACCTC420
new             TTTTCTAGA----------------------------------------------------366
                   *** *

Old             TTTTGACGACGTGTGTGTTTGAGGAGTTGTTTGGTGTGATGGCAGCATCACAAACACATC480
new             ------------------------------------------------------------

Old             TCCCTGTCTGTATGTTTGTGTGTCGCTCTGTAACGACGGGATACGGAGTGATATATAAAT540
new             ------------------------------------------------------------


Old             TTATATACAAAAAAAAAATGGTTTGCTA 568
new             ----------------------------
```

4. Clone no. (C) 172 (RISBZ4)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GCTCGAGGAAGAGGAAGCAAGCTCACTTAACAGACCTTGAAACACAGGTTGACCAGCTCC 60
new             GCTCGAGGAAGAGGAAGCAAGCTCACTTAACAGACCTTGAAACACAGGTTGACCAGCTCC 60
                ************************************************************

Old             GAGGTGAGAATGCTTCACTCTTCAAGCAGCTGACAGACGCCAACCAGCAATTCACAACCT120
new             GAGGTGAGAATGCTTCACTCTTCAAGCAGCTGACAGACGCCAACCAGCAATTCACAACCT120
                ************************************************************

Old             CAGTCACGGACAACAGGATCCTGAAATCAGACGTCGAGGCCCTCAGAGTCAAGGTGAAGC180
new             CAGTCACGGACAACAGGATCCTGAAATCAGACGTCGAGGCCCTCAGAGTCAAGGTGAAGC180
                ************************************************************
```

```
Old             TGGCAGAGGACATGGTGGCGCGCGGCACAATGTCATGCGGGCTGGGCCACCTGGGCCTGT240
new             TGGCAGAGGACATGGTGGCGCGCGGCACAATGTCATGCGGGCTGGGCCACCTGGGCCTGT240
                ************************************************************

Old             CGCCGGTGCTGAACCCACGGCAGGCGTGCCGCGGCCCGGACGTACTCGTGGGCCTGGACT300
new             CGCCGGTGCTGAACCCACGGCAGGCGTGCCGCGGCCCGGACGTACTCGTGGGCCTGGACT300
                ************************************************************

Old             TCCCCGCCGACGACGCGTGCTTCACGGGCCTCTCCCCCGACGGAGCATATGCAGAACTCGC360
new             TCCCCGCCGACGACGCGTGCTTCACGGGCCTCTCCCCCGACGGAGCATATGCAGAACTCGC360
                ************************************************************

Old             CGTTGCAGAGCATCGCCACCAGCCTCGAGAGCCTCGACAACCGCATGGCCAGCGAGGTGA420
new             CGTTGCAGAGCATCGCCACCAGCCTCGAGAGCCTCGACAACCGCATGGCCAGCGAGGTGA420
                ************************************************************

Old             CGAGCTGTGGCCGCGGCGGCGCCGGCGTCGACGTCTGGCCCTGGGACTCTCACTCCAACG480
new             CGAGCTGTGGCCGCGGCGGCGCCGGCGTCGACGTCTGGCCCTGGGACTCTCACTCCAACG480
                ************************************************************

Old             GTGGCTTGTCCAAGTGAAGAGATGAAGCTGAAGCTTGGAATTCGCTGGCGGCAGCTGCTG540
new             GTGGCTTGTCCAAGTGAAGAGATGAAGCTGAAGCTTGGAATTCGCTGGCGGCAGCTGCTG540
                ************************************************************

Old             TTCTGTACGCGTGTTATATGAGCTGCAAGAATTTGTGAGCTGCTTCTGTGTTAGCTGGAG600
new             TTCTGTACGCGTGTTATATGAGCTGCAAGAATTTGTGAGCTGCTTCTGTGTTAGCTGGAG00
                ************************************************************

Old             TCATGAACTGAACGGACTTGAGTTTTTCAGAGTCCTCGTGCTGGTATATTGATTGCCTAG660
new             TCATGAACTGAACGGACTTGAGTTTTTCAGAGTCCTCGTGCTGGTATATTGATTGCCTAG660
                ************************************************************

Old             TCCTGTTATAGACCTGATGACTAATAGACATATCTATGTGCTATGATACTGAACTCATAA720
new             TCCTGTTATAGACCTGATGACTAATAGACATATCTATGTGCTATGATACTGAACTCATAA720
                ************************************************************

Old             GGATAAGCATTATTATTCGTATCCCTGGTGT------------------------------751
new             GGATAAGCATTATTATTCGTATCCCTGGTGTTATATGCGCTCTTCTAGTTTGACTGCAGA780
                ********************************

Old             ------------------------------------------------------------
new             TATATGCAGTTAGTAGTTCTGAATACTCGGTGGTGAAAATCGTGGCCAGCAGCTAAGTGA840

Old             -----------------------------------------------------
new             AACCCTGGCATTCTGAAGCAGCATTTCTTGTGGCTTCAAGTAATTTTAGCTT 892
```

5. **Clone no. (C) 187 (DIS1; protein binding / structural constituent of cytoskeleton)**

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             CGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACATTGGGGAGGAGGCGCTGG 60
new             CGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACATTGGGGAGGAGGCGCTGG 60
                ************************************************************


Old             CGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGCCAAGTGG120
new             CGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGCCAAGTGG120
                ************************************************************

Old             AGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTG180
new             AGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTG180
                ************************************************************

Old             ACCCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGG240
new             ACCCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGG240
                ************************************************************

Old             AGTACACCGGAGAGATCATGTTCGAGACGTTCAATGTTCCTGGCCTCTATATCGGAGTGC300
new             AGTACACCGGAGAGATCATGTTCGAGACGTTCAATGTTCCTGGCCTCTATATCGGAGTGC300
                ************************************************************

Old             AACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAAATGACAGGTGTTG360
new             AACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAAATGACAGGTGTTG360
                ************************************************************
```

```
Old             TAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGTTATGTCATCG420
new             TAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGTTATGTCATCG420
                ************************************************************

Old             GGAGCAGCATCAGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAAC480
new             GGAGCAGCATCAGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAAC480
                ************************************************************

Old             TTATGAAGGAAAGAGGTGAACACATTCCAGCAGAAGAATCTTTGGATGTAGCAAGGAGGC540
new             TTATGAAGGAAAGAGGTGAACACATTCCAGCAAAAGAATCTTTGGATGTAGCAAGGAGGC540
                ******************************** ***************************

Old             TTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAATTCAATAAGCATGACA600
new             TTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAATTCAATAANCATGACA600
                ************************************************** *******

Old             GGGAGCCCTCAAAGTACATTAAACACTGGAATGGCATCAAACCAAAAACTGGTACTAAGT660
new             GGGAGCCCTCAAAGTACATTAAACACTGGAATGGCATCAAACTAAAA-CTGGTAC-AAGT658
                ****************************************** **** ******* ****

Old             ATTCATGTGACATGGGATATGAGCGCTTCTTGGGGCCTGAGATCTTCTTCCACCCAGAAA720
new             ATTCATGTGACATGGGATATGAACGCTTCT-GGAGCCTGA-ATCTTCT-CCACACAGAA-714
                ********************** ******* ** ****** ******* **** *****

Old             TTTACAACAATGACTTCACCACTCCTTTGCAAGTTGTTATTGACAAGTGCAT--------772
new             -TTACAACA-TGAAT--AACTATGCATTGCAAGTT-CTAATGAAAAGTGCACCAATCTCC769
                 ******** *** *  * *  * * *********  ** *** *******

Old             ------------------------------------------------------------
new             CCAATCAAAACGATGACGTCTCTGATCATATTTTGCCTGGTTTGGGGTAGACAATGTGAA829


Old             ------------------------------------------------------------
new             AGATTTCCATATCAGTTACATCACAATCNAAAAGATAGCGAATCTNAGNTCCGATATNCT889


Old             ---------------
new             ATGTACTGATACTNG 904
```

6. **Clone no. (C) 217 (actin-related protein ARP3)**

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             CGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACATTGGGGAGGAGGCGCTGG 60
new             CGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACATTGGGGAGGAGGCGCTGG 60
                ************************************************************

Old             CGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGCCAAGTGG120
new             CGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGCCAAGTGG120
                ************************************************************

Old             AGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTG180
new             AGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTG180
                ************************************************************

Old             ACCCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCCTTTTAC--------CTTGGC232
new             ACCCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGG240
                ************************************ *  * *        *  **

Old             CTTCCCCTGGGAGGAACCCCGGGGGAAGGTT--ACCCCCCCCCCCCCGCCCGCCGG-----285
new             AGTACACCGGAGAGATCATGTTCGAGACGTTCAATGTTCCTGGCCTCTATATCGGAGTGC300
                  * * * **   ** *      *  * ***  *    **   **       ***

Old             --TCGGTCCGCCCTCTTTTTTTTTTTTTTTCTTTCTCTGGTTTCCTCCTCCCCCCCCCTC343
new             AACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAAATGACAGGTGTTG360
                   ******   *****  *      *   **    *    *      *  *      *

Old             TCCCCCCCTCTGCTCTCCGGTTTTCGGTTGGCGCGTGGGGGGGCG-TGTGCGGGGCGCGG402
new             TAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGTTATGTCATCG420
                *          * *    *   * *      **     *  *  * **     * *   *

Old             GGGGGCGCGGGGGGGGGGCCCGGGGCTCGGCGGGGGGCGT------GTGTGCGGTGTGGT456
new             GGAGCAGCATCAGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAAC480
                ** *  **    *       *      * *    **  **      ** *     *
```

```
Old             TTGTGTGCGTGTGTGGTGCGTGGGTTTGGTGCGTGTGTGTGGTGCGGGTGGCGGGGT---513
new             TTATGAAGGAAAGAGGTGAACACATTCCAGCAGAAGAATCTTTGGATGTAGCAAGGAGGC540
                ** **   *   * ****      **      *         **   ** **  **


Old             ------------------------------------------------------------
new             TTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAATTCAATAAGCATGACA600


Old             ------------------------------------------------------------
new             GGGAGCCCTCAAAGTACATTAAACACTGGAATGGCATCAAACCAAAAACTGGTACTAAGT660


Old             ------------------------------------------------------------
new             ATTCATGTGACATGGGATATGAGCGCTTCTTGGGGCCTGAGATCTTCTTCCACCCAGAAA720


Old             ------------------------------------------------------------
new             TTTACAACAATGACTTCACCACTCCTTTGCAAGTTGTTATTGACAAGTGCATCCAATCAT780


Old             ------------------------------------------------------------
new             CCCCAATCGATACAAGGAGGTCTCTTTATAAGAATATTGTCTTGTCTGGGGGTTCGACTA840


Old             ------------------------------------------------------------
new             TGTTTAAGGATTTCCATAGAGATTGCACGAGACCTTAAAAGATAGTGGATGCACGGGTCC900


Old             -----
new             GTTCT 905
```

7. Clone no. (C) 223 (14-3-3 protein)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             TCGGCACGAGGATGATCCCTCAGGTTCCGCTGTACCAGACCTTCAACCATGGCTAGGAGG 60
new             ----------GATGATCCCTCAAGTTCCCCTGTACCAGTCCTTCAACCATGGCTAGGGTG 50
                          ************ ***** ********* ******************  *

Old             CTCCACTTGGCTGCGAAAGTTGGAAGCAACTGTTAAAACTTGCTATGTTGTGTTGTCATC120
new             ATCCACTTGGTGG--AAAGTTGGGAGTGGCGGGCAAAACTTGCTATGTTGTGTTGTCATC108
                 *********  *  ******** **   * *  **************************

Old             TATTAAATGGAGTCTATTTCAGAGTCG-ATTTCTATTACAAAAATTCGTGTAGCGTCAAA179
new             TATTAAATGGAGTCTATTTCAGGCCGACATTTCTATTACGCCC-TTCGCGTACCGTCGAC167
                ***********************      ***********     **** *** **** *

Old             CGTAGTTGTATAAGGAGTCCTAATTCGTTGAGACTATTATGGTGAACCGGACTAAAACAC239
new             CCTTGTTTTATACAGAGTTCTAGTTCGCTTATACTATGACGGTGAACCCGTCAAACCCAC227
                * * *** ****  **** *** **** * * ***** * ******** * * **  ***

Old             TGGGGCATGACTAATGTCTAGCTTTCTGCTCATAAAAAAAAAAAAAAAAAAAATCTAACGAT299
new             GGGTGAATGTCAAACGTCTGGCTTTCTGCTCTCAAAAAAAGAATTAATGGGGAAGAGGGG287
                 ** * *** * ** **** ***********  ******* **  **        * *

Old             TTTTGAGGAGGGCAAACCCACCAAAAAAAAAAACCAGGTCTCTTTTTGCTGAACCGATTGT359
new             GGGCCAGGTACCCGATGCGATCCTTAATGAG-TCACATTACACTTTCCTGGCCGTGCTTT346
                     ***    * *  * * *   **  *   **  *  *  *** ***  *    * *

Old             TGCTTTTGGGGACGGGCCTATCAATGCGGTGGTGGTCATATCCCAAAGGACACAAGGAAC419
new             TACCCTTCCA----------------------------------------------------356
                * *  **

Old             GCTATAAGATATTCCACGTTTAGCGACTATGCTATAGAGGCTACGAGATTTTACAGCCCC479
new             ------------------------------------------------------------


Old             CAAAATTTATTTCACTCCTGACACGTGATTAGAGTCAATTCTTATTACTCTCTAGAGAGA539
new             ------------------------------------------------------------


Old             GAAATTAATTTATCTTATAGTCCTGAGAGTCTTCTCTGATAGTGATTCTCATTCATATCC599
new             ------------------------------------------------------------
```

```
Old             ATCTAAGTTGACTATCCTAGATTCACACTTTATATTAAAGATAAGCCACATAGCTGAGAC659
new             ------------------------------------------------------------


Old             CCTCAGTACAATGACAATCTGTACACAGTACATATCATCAGCACGTGATGCACTACGTAA719
new             ------------------------------------------------------------


Old             CTATATAACATTAGCTCGTCTACACTATCTATACTATCTTCGATCGCACTGAGCATCCAC779
new             ------------------------------------------------------------


Old             TCTCCGC 786
new             -------
```

8. Clone no. (C) 237 (sinapyl alcohol dehydrogenase)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GAAAGTTTCCTAAGTACACACACCACAATTCCTTCATAAGGGACCAGCGGGCTCACCTGT 60
new             GAAAGTTTCCTAAGTACACACACCC--AGTCCTTCGTTAGTAGCCAGCGAGCTCACCTGT 58
                ************************   * ****** * **   ****** **********

Old             GTTACAAACAATGTGCTTCTTCCTCCTGCTGCTGCGCCGAAAATGGCTGCCATCTGCACG120
new             GTTACAAACAATG-GCTTCTTCCTCC-GCTGCTGCGCCGAAGATGGTTGCCATCTGCGCG116
                ************* ************ ************** **** ********** **
Old             ATAGTACTGTGCGTTGCCTTGGGACCTCATCCGGCGGCCGTTGTTCATGGTCGGGCGACG180
new             CTGGTCCTGTGCGTTGTCTTGGGACCTCAGCTGGCGGCCGTTGTTCATGGTCGGGCGACG176
                 * ** ********** ************ * ****************************

Old             CCGGAAAACGAGCCACCAGAGTGCTACTGAGACATGCCCCTGCCTGCGGATGCTCGCCCC240
new             CCGGAGAAGGAGCAGC--AGGTGCTACTGCAG-ATGCCCCTGGC-GCAGGAGCTCGCCG-231
                ***** ** ****  *    *********    ********** * ** *  *******

Old             ATGCTGCTCTCCCAACCCCATCCCCCCTAAACGATCGCATCAAATACTGGGCTACTGACG300
new             ATG-TGCTCTCCCAGCT--GATCCCCAAAAACGGCGGCGTCGAAGACTGTGCTACTGAGG288
                *** ********** *      ****  *****   ** ** ** **** ******** *

Old             CCTGCAAGATTTGCTAATCAAATTGCCGGATTTGCTGCCGCTTTGTACCATCATGATATC360
new             CGTGCAAAACTTGCTACGCGAATTGCCGGATTTGGTGCCGCTTTGTTCCGTCATGTTATC348
                * ***** * ******  * ************** ************ ** ***** ****

Old             CCCACTGCCCCAAGAAATTTGAATGCGAATTCGAATTGAGCATGGATGTTGCCAAAAACT420
new             CCGACTGCGCCGAGCAATTTGAATGCGA-TTCGAAGTGAGCATGCATGTTGCCAGCGAGT407
                ** ***** ** ** ************* ****** ******** *********   * *

Old             GACCCGCTGCTGTTATGATAACTCGCCCTGCAACACGAGTTGGGCGTTGCTTCATGTGTG480
new             GTCCCGCTGCTGTTAGGATACCTCGCCCTGCAACAC-AGTTGGGCGTTGCTTCATGTGTG466
                * ************* **** *************** ***********************

Old             TCCAGACTCGTGACTCTGGAGGCCTCTCGGCTTTGCCGCCTTCCTTTCAATAAACAAGTA540
new             TCGTGACTCGTGACTCGTGAGTCCGCTCGGCTTTGCCGCGTTCCTTTCAATAAGCAAGTA526
                **  ************  *** ** ************** ************* ******

Old             TTGTGTTTTATTGAATCAACAGAACATGTGTTCTTACTAGTTATCGTACTATGGTAAACT600
new             TTGTGTTTTATTGTATCAACAGAACATGTATTCTTACTAGTTATCGTACTATG--TAACT584
                ************* *************** ***********************   ****

Old             CAGAGCATTGTGTTTATGCTATTTTGCTATGATGTA--TCTATTTAA-CATTGCGGGGGG657
new             CAGAGCATTGTGTTTATGCTATTTTGCTATGATGTAATCCTATTTAAACATTGCGGG---641
                ************************************   ******** *********

Old             GGTTTTTTTAAAAAACGGAAGGCTGGGAATGCCAAACCCGAATTTGGTTAAAGGTGTGCC717
new             --TGCTACTGATGACTGCATGACTGTGA-TGCATAACACGAATT-GATTAAAGTTGCATC697
                  *  *  * *  *  * * * *** ** ***  *** ****** * ****** **   *

Old             TCCCTAAAAAAAAAAGAAAAAAAACCCCCCG 748
new             ACTA--------------------------- 701
                 *
```

9. Clone no. (C) 246 (putative TAF6)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             CTGCGTCTTCTATGGCAGGAAACATGCAAGGAACCATGGATGGGTTCTCCACCCAGTTAG 60
new             CTGCGTCTTCTATGGCAGGAAACATGCAAGGAACCATGGATGGGTTCTCCACCCAGTTAG 60
                ************************************************************

Old             CCAACCCTGGAATGATGCAAGCTTCATCCACTTCACAAATGGTGGAAAACATTGCAGTTG120
new             CCAACCCTGGAATGATGCAAGCTTCATCCACTTCACAAATGGTGGAAAACATTGCAGTTG120
                ************************************************************

Old             CGGTTCGGAGAGAACAGGGCAGTGATCTTGCGCAGAGGGTTTCCACGGTGCTGAGGCAAG180
new             CGGTTCGGAGAGAACAGGGCAGTGATCTTGCGCAGAGGGTTTCCACGGTGCTGAGGCAAG180
                ************************************************************

Old             CATGGAAGGAGGATCAAGATGCTGGCCATCTCCTCGGGTCCTTGTACGAGGTCTTTGGTG240
new             CATGGAAGGAGGATCAAGATGCTGGCCATCTCCTCGGGTCCTTGTACGAGGTCTTTGGTG240
                ************************************************************

Old             AAGCCATCTTCTCCTTTGTTCAACCACCAGAGATATCCATCTTTGTGTAGAAAGGAAATG300
new             AAGCCATCTTCTCCTTTGTTCAACCACCAGAGATATCCATCTTTGTGTAGAAAGGAAATG300
                ************************************************************

Old             CGTAGGCTTTAGAGGATTTTAATCTGCACTGTTGTATAATTAAGGGTATTCTTGAAATGG360
new             CGTAGGCTTTAGAGGATTTTAATCTGCACTGTTGTATAATTAAGGGTATTCTTGAAATGG360
                ************************************************************

Old             TAGACTGACATCAGAAAACAAGTAATGTAATGCTAGGTTTATATTAATGTGAGGAATGAA420
new             TAGACTGACATCAGAAAACAAGTAATGTAATGCTAGGTTTATATTAATGTGAGGAATGAA420
                ************************************************************

Old             CATGTGTGTTGTACACTTGTACTATTATTACTTTACGTGACTGATTATCTTAGCCATTTG480
new             CATGTGTGTTGTACACTTGTACTATTATTACTTTACGTGACTGATTATCTTAGCCATTTG480
                ************************************************************

Old             GCTACTAATGTATACGTACATCGTTTAAAAAAAAAAAAAAAAAAAAAACTCGGGGGGGGGC540
new             GCTACTAATGTATACGTACATCGTTTAAA--------------------------------509
                ****************************

Old             CCGGACCCCAATTCCCCCTATAGGGAGCGTATTTCAATTCCGGGGGGGGGGGCCTTTTTTT600
new             ------------------------------------------------------------

Old             AAAAAAAAAAGGGGGGGGGGGGGCCCGGGTGTGTCCCGGGGGGGGGGGGGGGGGGGGGG660
new             ------------------------------------------------------------

Old             GGGGGGGGGGGGGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 717
new             ---------------------------------------------------------
```

10. Clone no. (C) 295 (farnesylated protein 1)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             GCAGTAAAACGCTAGTACTAAAAACAGAGCGGCTTTCCTTCTCGATCACCCAGCAAGCAC 60
new             GCAGTAAAACGCTAGTACTAAAAACAGAGCGGCTTTCCTTCTCGATCACCCAGCAAGCAC 60
                ************************************************************

Old             CTCTTTGGCCAAACGCCATCAGCCACACCAGCACAGCCAAAATGCTTGGCGCGGACCTCA120
new             CTCTTTGGCCAAACGCCATCAGCCACACCAGCACAGCCAAAATGCTTGGCGCGGACCTCA120
                ************************************************************

Old             TCACTGAGCTATGCACGGCGCCGACCAAGATCCTCGGCAGGAAGAAGCGCAAGCAGTTCC180
new             TCACTGAGCTATGCACGGCGCCGACCAAGATCCTCGGCAGGAAGAAGCGCAAGCAGTTCC180
                ************************************************************

Old             AGAAGGTGGAACTGAAGGTGCGGATGGACTGCGAAGGTTGCGAGCGCAAGGTGAGGAAGG240
new             AGAAGGTGGAACTGAAGGTGCGGATGGACTGCGAAGGTTGCGAGCGCAAGGTGAGGAAGG240
                ************************************************************

Old             CGCTGGAAGACATGAAAGGCGTGAGCAGCGTGGAGGTGGACACGAAGCAGAACAAGGTGA300
new             CGCTGGAAGACATGAAAGGCGTGAGCAGCGTGGAGGTGGACACGAAGCAGAACAAGGTGA300
                ************************************************************
```

```
Old             CGGTGTCGGGGTACGTGGAGACGAAGGAGGTGATAGAGCGGCTGCGCCACCGGGCGGGGA360
new             CGGTGTCGGGGTACGTGGAGACGAAGGAGGTGATAGAGCGGCTGCGCCACCGGGCGGGGA360
                ************************************************************

Old             AGAAGGCGGAGCCGTGGCCGTACGTGCCCTACGACGTGGTGCCGCACCCTTACGTGCCGG420
new             AGAAGGCGGAGCCGTGGCCGTACGTGCCCTACGACGTGGTGCCGCACCCTTACGTGCCGG420
                ************************************************************

Old             GTGCCTACGACAAGAAGGCGCCGCCCGGGTACGTGCGCAACGTGCTCGACGACCCTGACG480
new             GTGCCTACGACAAGAAGGCGCCGCCCGGGTACGTGCGCAACGTGCTCGACGACCCTGACG480
                ************************************************************

Old             TCGGCCCGCTCGCGCGGGCAAACTCCATGGAGGAGAAGTACACGTCCGCCTTCAGTGACG540
new             TCGGCCCGCTCGCGCGGGCAAACTCCATGGAGGAGAAGTACACGTCCGCCTTCAGTGACG540
                ************************************************************

Old             AGAATCCCAACGCGTGCGTCGTCATGTGATCGGCTCGTGGCACTGTCTCGTTGCCTGTCG600
new             AGAATCCCAACGCGTGCGTCGTCATGTGATCGGCTCGTGGCACTGTCTCGTTGCCTGTCG600
                ************************************************************

Old             AATACGTACCATATAGGTTCCTATGATACTCCGCGTTTGTAATACAGAGTACTGTACAGT660
new             AATACGTACCATATAGGTTCCTATGATACTCCGCGTTTGTAATACAGAGTACTGTACAGT660
                ************************************************************

Old             TTTTTGCTCCATCGATAGCTGTAAAATTTTGTAATAGTGTTTCCAGTATTCTTAAAAAAA720
new             TTTTTGCTCCATCGATAGCTGTAAAATTTTGTAATAGTGTTTCCAGTATTCTTA------714
                ********************************************************

Old             AAAAAAAAAAAACACCGCGGA 741
new             ---------------------
```

11. Clone no. (C) 303 (putative aldehyde dehydrogenase WIS1)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GGCTTATGCTTGCAGTGGTCAGAAATGCTCTGCCCAGTCTGTTCTATTCATGCACAAGAA 60
new             GGCTTATGCTTGCAGTGGTCAGAAATGCTCTGCCCAGTCTGTTCTATTCATGCACAAGAA 60
                ************************************************************

Old             TTGGTCATCTAGTGGGTTGCTTCGAGAATATGAAGAAACTATCTGAAAGAAGGAAGCTCG120
new             TTGGTCATCTAGTGGGTTGCTT-GAGAATATGAAGAAACTATCTGAAAGAAGGAAGCTCG119
                ********************** ***************************************

Old             AAGATTTGACAATTGGCCCGGTCCTTACTGTTACTACAGAAGCTATGATGGAGCACATGA180
new             AAGATTTGACAATTGGCCCGGTCCTTACTGTTACTACAGAAGCTATGATGGAGCACATGA179
                ************************************************************

Old             ACAACCTTCTTAAAATACCAGGATCCAAGGTTCTTTTCGGCGGTGAACCTTTGGAGAATC240
new             ACAACCTTCTTAAAATACCAGGATCCAAGGTTCTTTTCGGCGGTGAACCTTTGGAGAATC239
                ************************************************************
Old             ACTCTATTCCGAAAGTATATGGTGCTTTGAAGCCAACTGCTGTATTTGTTCCTATAGAGG300
new             ACTCTATTCCGAAAGTATATGGTGCTTTGAAGCCAACTGCTGTATTTGTTCCTATAGAGG299
                ************************************************************

Old             AAATCCTTAAAAGCGGGAACTTTGAGCTTGTGACAAAGGAGATATTTGGTCCATTCCAGA360
new             AAATCCTTAAAAGCGGGAACTTTGAGCTTGTGACAAAGGAGATATTTGGTCCATTCCAGA359
                ************************************************************

Old             ATACTCCGAGGATCAGCTTGACTTGGTATTAGAAGCCTGTGAAATTATGAATGCGCATCT420
new             ATACTCCGAGGATCAGCTTGACTTGGTATTAGAAGCCTGTGAAATTATGAATGCGCATCT419
                ************************************************************

Old             GACAGCTGCTGTAGTTTCAAACAACCCGTTATTCCTGCAGGAAGTACTTGGGATGATCTG480
new             GACAGCTGCTGTAGTTTCAAACAACCCGTTATTCCTGCAGGAAGTACTTGGGA-GATCTG478
                ************************************************** ******

Old             TCAACGGTACAACATATGCTGGGATTCGAGCAAGGACCACGGGCGCTCCACAAAACCACT540
new             TCAACGGTACAACATATGCTGGGATTCGAGCAAGGACCACGGGCGCTCCACAAAACCACT538
                ************************************************************

Old             GGTTTGGGCCAGCCGGTGACCCGAGAGGTGCAGGCATCGGAACTCCGGAAGCCATCAAAA600
new             GGTTTGGGCCAGCCGGTGACCCGAGAGGTGCAGGCATCGGAACTCCGGAAGCCATCAAAA598
                ************************************************************

Old             TTGTCTGGTCTTGCCACAGAGAGATCATATATGACATCGGTCCCGTGCCCAAGAACTGGG660
new             TTGTCTGGTCTTGCCACAGAGAGATCATATATGACATCGGTCCCGTGCCCAAGAACTGGG658
                ************************************************************
```

```
Old             GCGCTTCCTTCCGCCACTTAATTTTGTGGAGGTTCCCAGTTTAGTAAGGAAGTATAGTTT720
new             -CGCTTCCTTCCGCCACTTAATTT-GTGGAGGTTCCCAGTT-AGTAAAGAAGTATAGTTT715
                 ********************** ********** ***** ***********

Old             GCGCCAAGA-TA-------------------------------------------------731
new             GCGCCAAGAATAAATGGGGAAGAAAAGCTGACAGGGCCGGAGTAGTCAGGAACGGAGTGT775
                ********* **
Old             ------------------------------------------------------------
new             ATAGGAATAATTTGCCCGTGTAGCTATTCTCCATACCCTCCAGAAACATTGATAAACGCC 835


Old             ------------------------------------------------------
new             AGATATGGATCTTGTAGGCTGTGACTGTGATGTAGTATTGTTCTGTGGCACTCGA 890
```

12. Clone no. (C) 321 (putative signal recognition particle receptor)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             GGACAAGGGACAGGAGAACTATCAGCCCCGTTCTGCTCTTCTTGTCAAAGGAAAGGAGAA 60
new             GGACAAGGGACAGGAGAACTATCAGCCCCGTTCTGCTCTTCTTGTCAAAGGAAAGGAGAA 60
                ************************************************************

Old             TGGAGGTCCCAAGGACAACAACAACGGAGCCTTTGATGTAAATAAGTTACAGAAGCTGCG120
new             TGGAGGTCCCAAGGACAACAACAACGGAGCCTTTGATGTAAATAAGTTACAGAAGCTGCG120
                ************************************************************

Old             GAACAACAAGAAGAAAACTGTTCCCGCTGAAATCGGAACCAAGAAGAACACTACTAAACC180
new             GAACAACAAGAAGAAAACTGTTCCCGCTGAAATCGGAACCAAGAAGAACACTACTAAACC180
                ************************************************************

Old             TGACACTAAGAAGAAGGTCAACAGGGTCTGGGATGACAAACCTACCGACAACAAGAAGCT240
new             TGACACTAAGAAGAAGGTCAACAGGGTCTGGGATGACAAACCTACCGACAACAAGAAGCT240
                ************************************************************

Old             GGATTTTACTGACCCAGCAGACAAGAGAGGGGATGATGAGGTGATCGACCATGCGCTCGT300
new             GGATTTTACTGACCCAGCAGACAAGAGAGGGGATGATGAGGTGATCGACCATGCGCTCGT300
                ************************************************************

Old             CAACCAGGGTGAGAGCATGATGGACAAGGATGAGAACGTCAGCAGCGACAGCGAGGATGA360
new             CAACCAGGGTGAGAGCATGATGGACAAGGATGAGAACGTCAGCAGCGACAGCGAGGATGA360
                ************************************************************

Old             AGAAGAAGAACAGGTGCAAGGTGTCCCCAAAAAGAAAGGATGGTTCTCCTCAATGTTTCA420
new             AGAAGAAGAACAGGTGCAAGGTGTCCCCAAAAAGAAAGGATGGTTCTCCTCAATGTTTCA420
                ************************************************************

Old             GAGTATTTCAGGTAACAATGTGCTGGAGAAGTCTGATCTGCAACCTGCGCTTAAAGCTCT480
new             GAGTATTTCAGGTAACAATGTGCTGGAGAAGTCTGATCTGCAACCTGCGCTTAAAGCTCT480
                ************************************************************

Old             TAAAGATAGGCTGATGACCAAGAATGTGGCAGAGGAAATTGCAGAAAAGCTTTGTGAATC540
new             TAAAGATAGGCTGATGACCAAGAATGTGGCAGAGGAAATTGCAGAAAAGCTTTGTGAATC540
                ************************************************************

Old             TGTTGCAGCTAGCCTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATCATCTACTGT600
new             TGTTGCAGCTAGCCTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATCATCTACTGT600
                ************************************************************

Old             ACAGACAGCAATGGATGGAGGCTCTTGTTCGCATTTTGACCCCAAGGCGATCAATTGACA660
new             ACAGACAGCAATGGA-GGAGGCTCTTGTTCGCATTTTGACCCCAAGGCGATCAATTGACA659
                *************** ******************************************

Old             TACTGAGAGATGTGCAGTCTGCCAGGGAGCGTGGGAGGCCTTATGTCATTGTTTTCGTTG720
new             TACTGAGAGATGTGCAGTCTGCCAAGGAGCGTGGGAGGCCTTATGTCATTGTTTTCGTTG719
                ************************ **********************************

Old             GG-TCAATGGAGTTGGCACGTCCAC-AAT-------------------------------747
new             GGGTCAATGGAGTTGGCAAGTCCACCAATCTTGCAAAGGTCGCTTATTGGCTTCTTCAGC779
                ** *************** ****** ***

Old             ------------------------------------------------------------
new             ATAACCTCAGTGTAACACTTGCAGCGTGTGACACCTTCAGTCTGGAGCTGTTGAGCAGCT 839
```

```
Old             ----------------------------------------------------
new             ACGTACTCATGCACGCAGCTTAAGATACCATATTTGAGAAGGTTATGAAAAGA 892
```

13. Clone no. (C) 351 (calumenin homologue)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             ------------------------------------------------------------
new             CGCAGCACAGCGCCGGCGTCATGGCCGACCTCGACTTCTACATTGGGGAGGAGGCGCTGG 60

Old             ------------------------------------------------------------
new             CGCGGTCCCGTGCCAGCAGCACGTACAGCCTCAGCTACCCCATCCGCAATGGCCAAGTGG120

Old             ------------------------------------------------------------
new             AGAATTGGGACACTATGGAGAGGTTCTGGCAGCAGTGCATCTTCAATTATCTCCGGTGTG180

Old             ------------------------------------------------------------
new             ACCCGGAAGATCACTATTTTCTTCTAACCGAGAGTCCCCTTACTGCTCCTGAGACCCGGG240

Old             ------------------------------------------------------------
new             AGTACACCGGAGAGATCATGTTCGAGACGTTCAATGTTCCTGGCCTCTATATCGGAGTGC300

Old             ----------------------------------------CAAGTGCGGCAAGTGAG---CCC 20
new             AACCGGTCCTTGCTCTTGCTGCCGGATACACTACTACCAAGTGTGAAATGACAGGTGTTG360
                                                 ****** *  * *  **

Old             TGGCCGCCGCCGGAGATGGAAGGACTTG-ATAGTCGGAACACCTAATGGCGTCTATCATG 79
new             TAGTTGATGTGGGTGATGGAGCTACCCACATCGTTCCTGTTGCTGATGGT-TATGTCATC419
                * *  *  *  ** ******   **    ** **        ** ****  * * ****

Old             TTGA------TCATACTTGTTTAAG---CGTGTGTAGCAAATGCGTGGCATGCTTAATAA130
new             GGGAGCAGCATCAGATCTATTCCAATTACAGGCAAGGATGTTACCCAGTTTATTCAGCAA479
                  **      *** *  * **  *    *  *    *    * *   *  *  * *  **

Old             C--ATGCATGCTGCCAGTGGTAGTAATTAAGTACTGAGAATAATAAGGCAAAGCAGAGAG188
new             CTTATGAAGGAAAGAGGTGAACACATTCCAGCAG-AAGAATCTTTGGATGTAGCAAGGAG538
                *  *** * *      ***     * *  ** *   *****  *  *    ****  ***

Old             ---TATGGCGTGGCCGTGTCCTTCCTCTCCAGCTAGAGTGA-GGCTCTAGCTGTTTGTGG244
new             GCTTAAAGAAATGTATTGCTATACTTGTTCAGACATTGTGAAGGAATTCAATAAGCATGA598
                   **  *    *   **   * * * * ***  *  **** **   *   *     **

Old             TCTAATATCTGCAGTGTGCCCTGTGCATGTTAGTATGGTGTCTCATGTATGGTCCCTGTG304
new             CAGGGAGCCCTCAAAGTACATTAAACA--CTGGAATGGCATCAAACCAAAAAC-----TG651
                        *  **  ** *  *   **   * * ****  **  *   *         **

Old             GTTTAAGTTGTTCTAGTCTTCTGTGTTATAACTGTTA-----------------------341
new             GTACTAAGTATTCATGTGACATGGGATATGAGCGCTTCTTGGGGCCTGAGATCTTCTTCC711
                **   *  * ***  **    ** * *** *  * *

Old             ------------------------------------------------------------
new             ACCCAGAAATTTACAACAATGACTTCACCACTCCTTTGCAAGTTGTTATTGACAAGTGCA771

Old             ------------------------------------------------------------
new             TCCAATCATCCCCAATCGATACAAGGAGGTCTCTTTATAAGAATATTGTCTTGTCTGGGG831

Old             ------------------------------------------------------------
new             GTTCGACTATGTTTAAGATTTCCATAGGAGATTGCAACGAGACCTTAAAAGATAGTGGAT891

Old             --------------
new             GCACGGGTCCGTTC 905
```

14. Clone no. (OP) 168 (putative transcriptional regulator)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             AGTACATTTTTGAGACACGAATATATCCTTTCCTGACATTAACCTTTCTACACACTGTGA 60
new             ------------------------------------------------------------

Old             TAGTTATAAGAGATTAAACTAGCTACGCCGATCAAAGTACGTTAACATATCAGCGACAGC120
new             ------------------------------------------------------------

Old             GCGCACGAGAAAAAGTAAATACCCAAAAGAGACACCGCCCCGAGCTTGTATTCGCTGTGT180
new             ------------------------------------------------------------

Old             TTATGCCCTCATGGGTCTTATTCTTGACCTGCTTCGGACTCCGCTTGTTCCGCCGGTTGC240
new             ------------------------------------------------------------

Old             GGAAAAAAATTCGGGGCCGCTCCAACTTTTCCTTGCAAATGATTTGGGATACGCGCCGGT300
new             ------------------------------------------------------------

Old             TTTCCTGGCGGGCGATGGTAAAAAAACAATCTAGGGCCCATACGTATAACTCGGGATAAC360
new             ------------------------------------------------------------

Old             AAAAGGTATGCGACACGGGTACGTCACCCAAGGGGGATTGTATCCCCACTTTGGGTATCT420
new             ------------------------------------------------------------

Old             ATGCTTAACGGGAGTGAACTGCATTGGACCGGACTCGGTGCTGCTCATATAATCAACGTA480
new             ------------------GCCTCAGAACGTCAGACTCGTCGCC-CGTTTCCCATCA-TGGC 41
                                    *  **      * ******  **  *   *   ****  *

Old             CCAAAAAAACACATTTTATTACTCCCTACTCATCTACAATTTACTCTTCATGGTTCAATA540
new             CGATTCCGCCGCGCTCT-CCACCGCCTCCTCCTCCAC-------CCTTCCCCGCTCGCTG 93
                * *      * *  * *   **  *** *** ** **        ****   * **  *

Old             AATCCTACGTTTTCTTCTTCCATAAACGACCTCTCTTCGTAAATATGGTCCTACCGTACG600
new             GCGTCCAGGAGGCCCACCCCTCTCTACGCCGGGGCCCCGCACACGCGCTGCTGCCG---G150
                    * * *    *  *  *  *  *** *    *  ** * *   * * ** ***   *

Old             ATCGAACCACGGTCGCGCTTCCACCCTAAATCTGAAATGTA-AACTTTGCTTCGTA---T656
new             TTCGGGCCGCGGTCGCTCCACCGGTTGCAGGCTCGGGCACGGAAGGATGATCCGGAGGAC210
                 ***  ** ******* *  **      *  **          **   ** * ** *

Old             CTTTATAACTTTTTTCTC--------ACTTCTAAATCATAT-ATACACTCAGAAAGAACC707
new             CTCTACGGCCCCTACCCCTGGGAACAGCCCCTGGATCGCACCACAGACTTCGACATT-CA269
                ** **   *   *  * *         *  **  ***  *  * * ***  ** *   *
Old             CAACGTAAAGGGCAGCAGCCCCACGCAGCCGTCAGCT-ATAGACGCAGTCG--CTGCGTT764
new             GTGGGTACCAGAAGACAGGGTCACACTGTTCACTTCTGATGGATTGATCCAAATTGGGGG329
                    ***   *    ***   *** * *    *  ** ** **   *  *    ** *

Old             ATCTGCGATTATGCGACGCGCTTGCTCATACGAGCGCACGCGCGATAGACTGTGTGACTG824
new             TTCCCTTGTTCCTCGTCGCCTTCGTGCTTCTGAGAAGAGGCAACGAAAAGTGAAAGGTGC389
                 **     **   ** ***  * *  * *  ***   * **     * * **   *

Old             TTGTGCG-TCCGCTGCT-----GAGCACGCGCGCGTAG-CACAGCGCGAGTTGTGAGCAA877
new             ACATAAACTTCGCCGTTTTCAGGAGAGCTCTTACATGGACCCAAATCAAAGCCTTTGCCT449
                   *    * *** * *     ***  * *   * * * * **   * *    *  **

Old             CAGTGCGTGCCTTGTCACTCGCGCTAGTATACAG---AGTGATGTGACATGCGCTGCAGC934
new             TGGAGCACTGTTTAACA-TCGCAGCAACAAATGGCCTAGACATGGGCAGAAGATTATGTA508
                  * **     **  ** ****   *  * *  *   **  *** *        *

Old             AGTCTGA-TTATACCATGCTAGTGAGATGACATGCGCGTGTGTCGCAGAGTAGACTGACT993
new             TATTTGGCTTTTGTCGTTCTGTTGAAATGCTAAGCG---ATGTAGTGGAAGATACAGTAT565
                  * **  ** *  * * * **  *** ***  * ***    *** *  **  * ** *  *

Old             GCGCGCATGACCGCGTGTCCGACCCTGTAGATACAGCGACGTGAGTGTGCTACACAGTGA1053
new             TGGAGCATGGCGGTGAGATTGTCACGGCAGAGAAGGCAA-GCAACGATGGCCTTCAGGAG624
                  * ***** * * * *   * * * * *** *  ** * *  *   **     ***
```

```
Old             GAGCGTCTAGTGACGCACGCAAGTATGCTG-TGCACTGCGCGCGTCATCACGCGCGCGCG1112
new             AAGCTAACCATGTCTGTTGCAGTACCATTGCTGTGGGGTGTTCCTCCAGCCTCAGAGACC684
                 ***      ** *    ***       ** **    * *  * **    * *     *

Old             CTCAGTGAGACGACTCTGACACATAGTGACGACGCCACAGATGTGTGTAGTGTGAGAGTA1172
new             CTCCATG------TTGCAGTGCGGAGCGGCGGAGGCATAGTCGAGAAGA-TATACTGGCA737
                ***  **       *      *  ** * **  * ** **  * *   * * *    * *

Old             CCAGTGAGTGAGTCTGCGCGTG-AGCACTGTGAGC-TCGCCTGAGATCTGACGACGCGCG1230
new             GTGG-GACTTATTCTAAATGTCCAGAGCTGAAAACATTAACAAAATTCATTCCACGGTAA796
                   * ** * * ***    **  **  ***  * * *   *  *  **   * ***
Old             TCTGTGGAGTACAGACACCCGCATAGCGTAGCGCCAGTGTGTACTGTCTGCGTCACTGAG1290
new             TCACTACGGAGACGGCTGTCACACGGCCT----TCAGGAGAAGCTAGCCATGTC--TGTT850
                **  *   *    * *   * **  ** *     ***      **  *   ***  **

Old             ATGTGAGTGCAGAGAAGTGTGTCAGTAGCCTAGAGCGCAAGTGCGTCGAGCTGATGTGCA1350
new             GAACCAGTTATGTGCAGTGTTCTTCCAGCATCAGACGGGAGA------------------892
                     ***   * * *****      *** *    **  **

Old             AGAGTCGATAGTAGAGTTCTGTGTGAGTGTGTAGAGTCAGCAGAGCATGTCTTGAGCGTG1410
new             ------------------------------------------------------------

Old             TCTGACTGCTCGCTGTAGTGCATGTAGAGACAGTAGTGAGTGATCGACTGCGCGCGAGTG1470
new             ------------------------------------------------------------

Old             CTTCGTAGCGACCTG 1485
new             ---------------
```

15. Clone no. (OP) 175 (protein phosphatase)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GCATGCACATCGCTGCCCCCGAGTGTCCTCTTCCGCTGTCTCCTGGCCGGAGATTCTTCG 60
P1              ---GATCCCTCGCGGCCCC--GGTTTCCACATCGCCGGCCGGCTCGCCGGGA--CCCTTG 53
                       * **** *****   ** *** * **  * * *  ** *****     * * *

Old             ATTCGCATCTCAGTTCTAGCCGTTACACAGTCCTGTGTCGAGCGGCTACGGCTTCGGCGT120
P1              ATCCTC--CTCGGCTTCTGCGGCAACGCTGGC--------GGCGGCTACGGCTTCGGCGT103
                ** * *  *** * *   ** *  ** * * *         ******************

Old             TACAGTCTGACTAGGGCCAGGTGAAAAGCGGGTTACATGCGAAGACGGCGGACGAACGAA180
P1              TACAGGG-GACTAGG--TAAGGGAAGGGAGGGTGACGTCGAGCGGCAGCGG-CGA-----154
                *****   *******   * * ***  * **** ** *     * * **** ***

Old             TGCACGTCCTTGCGC-TGGATCTGTACCGGAGGATCTGGCCTTTGCGCAAGTGCAATCTC239
P1              ----TGTCGTCGCACGCGGATCTGGACCGGCAGATCTCGCAGCTGCGGGAGTGCAAGTTC210
                     *** * ** *  ******* *****  ***** **   ****  *******  **

Old             CTCCCCGAAGGCAGAGCTGCAAGTCCCACTCTAGAGGAAGGCCAAAGGAAGATCGCTAGA299
P1              CTCCCCGA-GGCAGAGGTGAAAG---CACTCTGCGAGAAGGCGAAGG--CGATCCTC--A262
                ******** ******* ** ***   ******    ****** ** *   ****     *

Old             TGAAGATAAACAAAGCGAGCAGTCCCGGTGCCGACTTCCCCATCACCATTGTGTGGCTGC359
P1              TGGAGG--AATGGAACGTGCAGCCCGTGCGCTG----CCCCGTCACCGTC-TGCGG-TGA314
                ** **   **   * ** **** **  * ** *    **** ***** *  ** ** **

Old             TTTCAACGATTTCTTCTGCTCACCTCATTGAAGCTCTTCCGCAACTATGTGTGACGATGC419
P1              CATCCACGGCCAGTTCTACG-ACCTCATCGA-GCTCTTCCGCATCGGCG-GTGAC-ACGC370
                  ** ***     **** *  ******* ** *********** *   * ***** * **

Old             CCTGACGACCAACTACCTCTTCATGGTCGACTACGTCGACCGTGGCTACTACTCAGTGGA479
P1              CC-GAC-ACCAACTACCTCTTCATGGGCGACTACGTCGACCGTGGCTACTACTCAGTGGA428
                ** *** ****************** ********************************

Old             GACTGTGTCGTTGTTAGTGGCTCTAAAAGTGCGTTATAGAGACAGAATCACAATATTGAG539
P1              GACTGTGTCGTTGTTAGTGGCTCTAAAAGTGCGTTATAGAGACAGAATCACAATATTGAG488
                ************************************************************

Old             AGGAAATCATGAGAGCAGGCAAATAACTCAAGTGTATGGCTTCTATGATGAATGCTTACG599
P1              AGGAAATCATGAGAGCAGGCAAATAACTCAAGTGTATGGCTTCTATGATGAATGCTTACG548
                ************************************************************
```

```
Old             GAAGTATGGAAATGCAAATGTATGGAAACACTTCCTCGATTTGTTTGATTATTCGGCCTC659
P1              GAAGTATGGAAATGCAAATGTATGGAAGCACTTCACCGATTTGTTTGATTATTTG-CCTC607
                ************************** ******  ***************** * ****

Old             TCAACCAGCTCTTATAGAAAAACCAGTATCTTCTGGCGCTACATGGGTGGGTCTCTCTCC719
P1              TCA--CAGCTCTTATAGAAAA-CCAG-ATCTTCTGCC--TACATGGT--GGTCTCTCTCC659
                ***  ***************** **** ******** *  *******   ***********

Old             ATCACTGGGAACACATTTGAGGATAATGATTTCCGTGCCCTGTTGATCGCATACAAGAAG779
P1              ATCACTGG--ACACATT---GGATAATA---TTCGTGCCCT--TGATCGCATACAAGAGG709
                ********  *******   *******    * ********  *************** *

Old             AGTTCCCTCA----------------------------------------------------789
P1              TCCCTCATGAAGACCTATGTGTGATCTTTTGTGGTCTGACCCGGATGACCGATGTGGATG769
                     * * *

Old             ------------------------------------------------------------
P1              GGGAATTTCACCAAGAGGAGCAGGTACACATTTGGGCAAGATATTGCACACAATTCAACC829


Old             ------------------------------------------------------------
P1              ATACCACGGGCTTAGTCTTATTTCAAGGGCTCATCAGCTTGTAATGGAAGGTTCAATTGG889


Old             --
P1              TG 891
```

16. Clone no. (OP) 254 (seven transmembrane helix receptor)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             TACCTAGTTATCATTTTTTTCTTACGGGCTGGTCAGTTCCGGTAAGCGTTTTTTTTATAA 60
new             ------------------------------------------------------------


Old             GAAAATAACTAAGATCGTCTCAACCAAAGCTAAGTCTAAAAACAAGTAGCGCAAGGATCA120
new             ------------------------------------------------------------


Old             GATCTAAAGTAGTAGCATTAGCAGGCTTTAGGTATTAAATAGGTAGAACCCCTTTTGAAT180
new             ---------------------------------GCACCAAGTCTGCTGATGCACT-----AT 24
                                                 * *  ** *  *  **  * **     **

Old             ATGCATAAGTAGAAGAAGCAGTGTTGAGAAGCGTCTCTGGAC-TGTTCTACTCCCCAAGT239
new             CTCGGAAAGTTGACGAAGC--TATCGAGAAGGATATATCGCTATGTGCCAATAGACTAGT 82
                 *    **** ** *****  * * ******  * * * *   *** * * *   * ***

Old             TCATGTTATTTGTG--GATGGAAGAATGTATTGGGA-GCATTGCAGGAGGTAGAATATTA296
new             CCATTTTGTTGAGACTATCAGCAAACCATACCAGGATGCATGTCAACAAAAGGTAGATTG 142
                 *** ** **          * * *   **   *** ****  **  *    * * ***

Old             TATACAATATTTGGTGTCGTAGAGAGGGTGATGGGCACGTCGTTGTATCTATGATCTTGT356
new             GTTACAG-----GGTATTCAGGGTGAGTTGTCCAGTGTTGAACGTAAGCTACAAACTTTG197
                  ****      *** *    *    * **    *           * ***  * ***

Old             AGGGTAGTAGTATATGTTGTATCTCAAGTACATTTCCAGCCCGTTGTTCGGCCGTTTTAC416
new             AAAGTTGAAATCCAAAATATTCATGGATAACTAGGTTTCACTTTTGTACATACTTTTGAT257
                *  ** * * *  *   * *   *  *  **         *  **** *   * *** *

Old             TGCAGTTAGTTTGATTTCTCGGA-GGTCTGCAGAACAGTAGTTGTTCCTGGGACGTAGGT475
new             TAT-GTTCACTTAATTCTTTGGATGATAGGTCCTACAATAATTACATCTTATTCTTCAAA316
                *   ***   ** ***  * *** * *  *    *** ** **    **    * *

Old             ATTGAATTTTTCAGTTTGTGCCATACTGTGCTACTAATAGTATGTCGCCACGAGTAAACT535
new             ------------------------------------------------------------


Old             GATTGCGGGTAATGTTGTCCTTCGCTATAACCCATGAAATGAATGTGCCCCTAAAGCAAC595
new             ------------------------------------------------------------


Old             ATTTGTTGCTGCAATGTATCAGCTGCTGGGTGTGGGGCGTCTGTATATTGCTCGCGCTCT655
new             ------------------------------------------------------------
```

```
Old             CTGTTTTAGTGTCTATTCTGTGTGCACGTCTTTGCTGTTCGTGCGCACCGTGTAGCAGTG715
new             ------------------------------------------------------------

Old             TAGCTGATTACGATAGATTATATGATGATGATGAGTGAGGTATGTGTGAGTGCGACGGAG775
new             ------------------------------------------------------------

Old             ATGAGCTACTCTTGTATTATTGTTTGTAGATTGAGTGCATACGTGTCGTTCGATGGGCCG835
new             ------------------------------------------------------------

Old             CGCGCAGTGCGTAGCTGCGTGTCAGTCGTTGCGATCTTGTCGTGTTATAGTGTGATTGTG895
new             ------------------------------------------------------------

Old             ACTACGTCCGTGTGTAGCGGTACGCTGTATACGCTAAGTGCTGTATTTGCTTGTGTATTA955
new             ------------------------------------------------------------

Old             GATTGTTGTCAGCATAGTGTCTGTGTGCTAGTCGATTGCGTGAGTATAGTGTGTTGTTTC1015
new             ------------------------------------------------------------

Old             AGACTAGTGATGTACTATACGCTATCTACTTACGCTTGTGATGTGCATGCGTAGCTGTCG1075
new             ------------------------------------------------------------

Old             CGGCCCGTGTCGTGTCTTTGTGTATGTGAGATGTCGTATACGGAGCGTGTTGCAGTCGCG1135
new             ------------------------------------------------------------

Old             CGTACGAGAGCGTGTGTGCTTTTGTTAGCAGCTGTGTTGAGATAGTAGTGTGGCGAGTAC1195
new             ------------------------------------------------------------

Old             TGTGTGTGATTTTGATCCTACTCCTCTATCGTCTGACTACCTACTGATTGGTGTATGCTG1255
new             ------------------------------------------------------------

Old             AGTTGTTAGTTGCTAGCTAGTGCGCTAGTGTAGCCGCTAGATATTGTGTTAGTGCGCGTG1315
new             ------------------------------------------------------------

Old             TACTGAGAGAGCGTGTGAGATACGAGCGAGTGCGTGAGTAGTGAGAGCAGAGTGCGTGCG1375
new             ------------------------------------------------------------

Old             TGTGTGTGATAAGTACATATATCTGCATGTGAGTATTGTGAGTGAGTAGTTGTGCTTGTG1435
new             ------------------------------------------------------------

Old             TGTATAGTGTGTGTGACGTGTTTGTATCGTGTGAGTGTGTGTATGCGCTGCGTGTGAGCG1495
new             ------------------------------------------------------------

Old             CGGAACGACGATGAGTATGTGCCGCGCGTAGCGTGTGTGTGTGTGTATGAGATGTTAGTG1555
new             ------------------------------------------------------------

Old             TGTGTGAGTGCGCTACGGATTTGTGATGTGCGTGTGTGCGCTGTTGCACGTACGTGTACT1615
new             ------------------------------------------------------------

Old             GTGTGATTGCGCGCTTGTGTGCTCAGTTTGCGAGACTAAAGCGAGATGAGAGACGTCACC1675
new             ------------------------------------------------------------

Old             CCGAGCGACATTAGTTGTAGTGTAGGCTCAGTGAGTACGATCGTTATGAGTGTGTTGTTG1735
new             ------------------------------------------------------------

Old             AGTGTGAATGAG 1747
new             ------------
```

17. Clone no. (OP) 53 (Low molecular mass early light-inducible protein HV90,)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             CGTCGCTGGTGCCACTGCTCCAGGGGGAGACCGCCGAGGCCAAGACCAGCGGACTCATGA 60
new             CGTCGCTGGTGCCACTGCTCCAGGGGGAGACCGCCGAGGCCAAGACCAGCGGACTCATGA 60
                ************************************************************

Old             ACGCCAATGCGGAGCTCTGGAATGGCCGCCTCGCCATGCTCGGGCTTGTTGCGCTCGCAG120
new             ACGCCAATGCGGAGCTCTGGAATGGCCGCCTCGCCATGCTCGGGCTTGTTGCGCTCGCAG120
                ************************************************************

Old             TGACCGAATACCTCACCGGCGCGCCCTTCATCAACGCGTAATTAATCATGTCTTATGTGT180
new             TGACCGAATACCTCACCGGCGCGCCCTTCATCAACGCGTAATTAATCATGTCTTATGTGT180
                ************************************************************

Old             ACATAGCTTAGGCATTCTCCGTATCAAATTTCCCGGGGCACGTGTATGACGTATATCAGG240
new             ACATAGCTTAGGCATTCTCCGTATCAAATTTCCCGGGGCACGTGTATGACGTATATCAGG240
                ************************************************************

Old             CGGTGTAATTGTACTAAGTGTACCATCCTGGAAACATTAAAAATAGACTGTTTTATGGCC300
new             CGGTGTAATTGTACTAAGTGTACCATCCTGGAAACATTAAAAATAGACTGTTTTATGGCC300
                ************************************************************

Old             TTCAATTTGCGCTGCAAAAAACATGTTACACTAGTTAAAAAAAAAAAAAAAAAAACTCGG360
new             TTCAATTTGCGCTGCAAAAAACATGTTACACTAGTTA-----------------------337
                ************************************

Old             GGGGGGGCCCGGTCCCCATTTCCCCCTATGGGGGGCCGAATTACAATTCCGGCCCCGGGG420
new             ------------------------------------------------------------

Old             GTTTTATTCAAAAAGGGGGGGGGGGGCGCCGAAGCGGGAGGGGAGGGGGGGGGGGGGGG480
new             ------------------------------------------------------------

Old             GGGGCCGCCGGCCGGCGGGCCCACGCCGCCGCCAGTCCCCGGGAAGCAGGAGGGGGGGAG540
new             ------------------------------------------------------------

Old             AGAAAGAACGACAAAAAAAAAAAAACAAACAAAAAAACAAAAACAAAAAAAACCAAACCC600
new             ------------------------------------------------------------

Old             TCAAATAATTTATCTTTTATAGAAGTAGGGGGGGGGCGGGCGCGAGCGGAAAAAATCAAA660
new             ------------------------------------------------------------

Old             TATCTAAATAAACAGGGTGAATTAAAATAGGAAAAAGATAGTGAGTGAGTATGCAGCTAG720
new             ------------------------------------------------------------

Old             GTGATCATACTGAGTCAGCGTGATG 745
new             -------------------------
```

18. Clone no. (OP) 61 (Farnesyl pyrophosphate synthetase)

```
CLUSTAL 2.0.12 multiple sequence alignment


Old             AAAGGACGGAAAGAGGCAAACACAGAACGAGAAACACAGAGAGAGAGGGAGAGAGAGAAG 60
new             ------------------------------------------------------------

Old             GCCGGCCGCCCGCTACATCCCTTGAGTGATGTTTACGTTCTTGTTTCAAGATAAACGACC120
new             ------------------------------------------------------------

Old             TTTCCCAAAAAAACCTACAGACGATAAGAAACGGGCGCGCCTGCCCCCATACTAAAGTGG180
new             ------------------------------------------------------------

Old             GATCCCGCGCGCGCGCGTGCTAGGCAATTCTTGCACGAAGGCACCCTGCTTCCGCGCCTG240
new             ----------------------------------------GACCCT--CTCCGCGCCTG 17
                                                        *****   **********
```

```
Old             CGCGGTACCTGTACGTGGGGAATGAGCTCCGAGGAGTGCGCAATGGCGACTGTGATGGCG300
new             CGGG--ACCTG-ACGTGGGGA-CGAGCTCCGAGGAGTGCGCAATGGCGACTGTGATGGCG 73
                ** *  ***** *********  ************************************

Old             GCGGCCAACGGCTCGGGCGGCGACACCAAGGCCGCCTTCGCGGAGATATACGACAAGCTC360
new             GCGGCCAACGGCTCGGGCGGCGACACCAAGGCCGCCTTCGCGGAGATATACGACAAGCTC133
                ************************************************************

Old             AAGCAGGAGATGCTCGAGGACCCCTCTTTCGAGTTCACCGACGAGTCGCTCCAGGGGATC420
new             AAGCAGGAGATGCTCGAGGACCCCTCTTTCGAGTTCACCGACGAGTCGCTCCAGTGGATC193
                ************************************************************ *****

Old             GACCGCATGGTGGATTACAATGTGCTTGGAGGAAAGTGCAATCGCGGGCTCTCTGTTATT480
new             GACCGCATGGTGGATTACAATGTGCTTGGAGGAAAGTGCAATCGCGGGCTCTCTGTTATT253
                ************************************************************

Old             GATAGCTACAAGATACTGAAGGGTGTTGATGTCCTGAGCAAGGAAGAGACATTTCTTGCT540
new             GATAGCTACAAGATACTGAAGGGTGTTGATGTCCTGAGCAAGGAGGAGACATTTCTTGCT313
                ******************************************** ***************
Old             TGCACCCTTGGATTGGTGTATTGAATGGATTCAAAGCTTATTTTCTTGGCGCTTGATGAT600
new             TGCACCCTTGG-TTGGTGTATTGAATGGCTTCAA-GCTTATTTTCTTG-TGCTTGATGAT370
                *********** ********************** ***** *************  **********

Old             ATTATGGACAACTCTCAAACAAGGCGTGGAACAACCTTGCTGGAACAAGGTGCCTCGGTG660
new             ATTATGGACAACTCTCAAACAAGGCGTGGT-CAGCCTTGCTGGTTCAGGGTGCCTCAG--427
                ****************************  ** *********  ** ******** *

Old             ATTGGAAGCTCAATTGCGGTAAATGCATAGAGAATTATCCTCTCCCAACCTTTATATCTC720
new             GTTGG---TCTCATTGCTGTAAATG---ATGGGATTATCCT-TCGCAACCAT-ATATCTC479
                 ****       ***** *******   *  * ******** ** ***** * *******

Old             GAAATCCTA---------TCTAGACGACCAGCATATGCTGAGAAGGGAATAGCACCAGT-A771
new             GGATCCTTCAACGCCACTTCAAAGGAAAGCCATATTATGTTGATCTCATTGATTTGTTCA539
                * *  * *          *  * * **    *****  **   *    ** *      * *

Old             GTGAGATGGAGCATGAA------TCGTGCCGAT--CTGAAG--TACTAGTGGATAGCAAG821
new             ATGAGGTTGAGTTCAAGACAGCGTCAGGGCAGTTGCTGGATCTTATCACTACTCATGAGG599
                 **** * ***    *       **  * *  *  *** *   **  * *    *  * *

Old             AGGAGAGAGTGCCAAGAGAGTGCTGTGAGACCATCAGCTATCA-GTAGGTGAGTGAGCAG880
new             GTGAAAAGGACCTATCTAAATATAACTTGACAGTTCACCGCCGCATTGTTCAATACAAGA659
                  ** *  *  * *    * *       ***  *   *   *   * * * * *

Old             CGACTCGCGAG--ATCCTGTATCGACAGTAGAC-----CAGCGAGTAGATGGATAGAAGA933
new             CAGCCTACTATTCATTCTATCTTCCTGTTGCATGTGCACTGCTGCTTTCTGGCGAGAACT719
                *  *   * *   ** ** * *      *  *      * **   *   ***  ****

Old             CAGAGGACCAGAACGCAACAGAAGAAGAGAGATATGTGAGAGAGAGATGAGAGACCGAGC993
new             TGGATAACTTTGGTG-ATGTCAAGAACATTCTTGT-TGACATGGGGACATACTTTCAAGT777
                  **  **      * *    ***** *    * * *** *  * **        * **

Old             ---GAGAGTGATGTGGTATAGCATAGCTA--CTGAGTG----AGCAGTCCAGAAGAGAGT1044
new             TCAGATGATTATCTAGATTGTTTTGGTGATCCTGAATTCATTGGCAGATCGGAACTGACA837
                   **   * ** * *  *    * *  *  **** *      ****  * ***  **

Old             AGCAGAGTGAC--GTAAACCTACACACTGAGATAGAAGATAGAGTGTAGAG------- 1093
new             TTGAAGATAATAAGTGTTCCTGGTTAGTCGTGCAAGCTCTTGAGCATGCTGATGAGAG 895
                   *   * *   **   ***    * *     *     * ***  *   *
```

19. Clone no. (OP) 96 (DRP4 protein)

Reaction failed

20. Clone no. (OP) 126 (monodehydroascorbate reductase)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             GGGATCCCAAGCTGGCTCTTACGCTGGCAGTGGGCATGGCAAGCCGTGAGCGAGCTTGCC 60
new             GGGATCCCAAGCTGGCTCTTACGCTGGCAGTGGGCATGGCAAGCCGTGAGCGAGCTTGCC 60
                ************************************************************

Old             AAGAGAAGGTGCTCGAGACTTTACCAAGATTATTATCTCGCATTATTGTGTACTGATTGT120
new             AAGAGAAGGTGCTCGAGACTTTACCAAGATTATTATCTCGCATTATTGTGTACTGATTGT120
                ************************************************************
```

```
Old             ATTGCTTCGTCCATGTGTTTCAGCCACCCAGCATCATTTGTGTTTATAACCGCAAATAAG180
new             ATTGCTTCGTCCATGTGTTTCAGCCACCCAGCATCATTTGTGTTTATAACCGCAAATAAG180
                ************************************************************

Old             GCTACAGCCGGCAGATTAAGAATAAGGTGCTAGCATATGTCTGTCCATGTAAACATTTTA240
new             GCTACAGCCGGCAGATTAAGAATAAGGTGCTAGCATATGTCTGTCCATGTAAACATTTTA240
                ************************************************************

Old             TATGGGTGTGTGTCTCACATTTATATGTGACCATGTACGATTGCTTATGTACATTTTGCC300
new             TATGGGTGTGTGTCTCACATTTATATGTGACCATGTACGATTGCTTATGTACATTTTGCC300
                ************************************************************

Old             TAAATAGTAGTATTCAGTTAAGAGAGAAAAAAAAAAAAAAAAAAACCCCGGGGGGGGGCC360
new             TAAATAGTAGTATTCAGTTAAGAGAGAA---------------------------------- 328
                ****************************

Old             CGGACCCCATTTCCCCCTAAGGGGGGCCGAATTAAAATTCCGGCCGGCCCGGTTTTTAAA420
new             ------------------------------------------------------------

Old             AAAAGGGGGGGGGGGGCGCGAGCGACACGGGGGGGGCTGCCACTTGCTTCTCTCGTCTCA480
new             ------------------------------------------------------------
Old             TTCCCCTACGTATATAATTGGGCGGGCAGGCGCGTATACAAATAAAGACAAAAATAAATA540
new             ------------------------------------------------------------

Old             TACAAAAAAAAATCAATTCCGGGGCGCGAAATTTGGGGGAGAGTGAGGGGGAAAGAGAAAA600
new             ------------------------------------------------------------

Old             AAAAAAAAAATAAAGGGGGGGCGACCAATGTACGGAGAGGGGGGGGGGTAACCATAATAG660
new             ------------------------------------------------------------

Old             AAGCAGAGAGCGAGAGAAGAGTAATACCAGCGATGATAGAAGATGACAACAAGAGCGCGA720
new             ------------------------------------------------------------

Old             AGTAGCTGAGTGAGAG 736
new             ----------------
```

21. Clone no. (NP)175 (RAB17 protein)

CLUSTAL 2.0.12 multiple sequence alignment

```
Old             CCTAGTGATTTGAGTGTTTGTGAGAGCAGCTTACAACTAGTCACCATGGAGCACCAGTGG 60
new             ------------------------------CTTGTATTGAGTTCTGTTGTTGATCTCTCGA 31
                                              ***  *   ***     **  *  *    *

Old             ACAGTACGGCCACGTGCACCACTCGGCCGCGTCGACGAGTACGTGCAACCCGGTTGCCGG120
new             TCTGAAAGATGTCGTGC----TGCGGAGGCAACTGCGGGTGCG---GCTCTGGCTGCAAG 84
                 * * * *    *****      ***  **  *  ** ** **      * ** ***  *

Old             AC--ACGGCACCACCGGAACCGGCGGCATGGGCATGGGGACACACGGGACCACCGGCACC178
new             TGCGGCAGCGGCTGCGGAGGGTGCAAGATGTACCCGGAGATGGCCGAGGAGGTGAGCACC144
                     * **  *  ****    **   ***  *  ** **    ** *       *****

Old             GGTACCGGCATGGGCATGGGAACTCACGGAACCGGTACCGGCGCCGGTATGGGTGGGCAG238
new             ACCACCCAGACCGTCATCATGGGTGTTTCGCCTTCCAAGGGGCACGCTGAGGACGGG---201
                   ***   *  * ***      *       *    *  **   ** *  **  ***

Old             TTCCAGCCCACCAGGGAGGAGCACAAGACTGGGGGCATACTGCACCGCTCTGCCAGTTCC298
new             TTCGAGGCTGCCGGAGCTGAGAACGACGGTTGCAA-GTGCGGCCCCAA-CTGC-AGCTGC258
                *** ** *  ** * *  *** ** *   * *     * * ** **   **** ** * *

Old             AGCTCCAGCTCGTCTGAGGACGACGGCTTGGGTGGCCGAAGGGA-AGAAGGGCATCAAAG357
new             AACCCCTGCAAGTGCGGCAAGTGAGCCCTGGCCGCCGCCGGAGATGGAAGGACTTGATAG318
                * * ** **  **  *   *    * * ***  * *    * **  ***** * * * **

Old             ACAATATCAAGGATTATCTCCCCGGTGGCCACAAGGTATGGACCAGCTGCATATGGGTTC417
new             TCGGAACACCTAATGGCGTCTATCATGTTGATCATACTTGTTTAAGC-GTGTGTAGCAAA377
                 *   *      **    **     **   *  *    **    *** *  * * *

Old             TGGGCGGAACCTTCTGCCTTTTAAGGACACACTCTGCTTGACCGGTACCTCTGGCACCTG477
new             TGCGTGGCA-----TGCTTAATAAC-ATGCATGCTGCCAG--TGGTAGTAAT--TAAGTA427
                ** * ** *     *** *  ***  *  **  ****  *   ****    *   *  *
```

```
Old             TTCCGGTTTCCACTGGATACCTACGGGCACCTTTACTTTGTGTAGTAATTAACGTCTATC537
new             CTGAGAATAATAAGGCAAAGCAGAGAGTA--------TGGCGTGGCCGT----GTCCTTC475
                 *  *  *   *  * * * *   * * *        * * ** *   *    ***  **

Old             ATTGTCCCAATATCAAGTGTAAATCTTCCCCCGTGCCTAATACTTTATCTGTTCTCTCCC597
new             CTCTCC-----AGCTAGAGTGAGGCTCTAGCTGTTTGTGGTCTAATATCTGCAGTGTGCC530
                 *   *     * * ** ** *  **    * **   *  *    ******   * * **

Old             CCTTTCTATTCCTA-GCTATTCTTTG-GCCCTTACCTGGTTTAACTTATCTTGGTTATCT655
new             CTGTGCATGTAGTATGGTGTCTCATGTATGGTCACTGTGTTTAAGTTGTTCTAAAAAAAA590
                *  * *   *  ** * * *    **     * **   ****** ** *  *    *

Old             ACTGTTATAGTATCGGAGCCTGCACCGTATCTGTGTATAATTATTTATTCCCTTTGATAT715
new             AA----AAAAAAAAAAAAAAAAAAAAAAAACCAAAAAAAAAAAAAAAAAAAAAAAAAA646
                *     * *  *    *      *    * *     * **  *   *         * *


Old             CCACTGGTCCACCATAGTCGTTGTCATTCGTATATCGTATGAGTATGTACATGCCTGGTC775
new             AAAAAAAACAAAAAAAAAAAAAAANCCCTCGGGGGGGGGCCCCGGNCCCCANTTCCCCCNA706
                  *     * *  * *        *  ***      *     *        * **
Old             TTGTTCGGGCTATGTTGTGTGAATACTGTCTCTGTTGCTTGTAT 819
new             TAGGGGAGTCTAT------------------------------- 719
                * *    * ****
```

22. Clone no. (C) 311 (trehalose-6-phosphate synthase/phosphatase)

Reaction failed

D. Alignment files of partial length of finger millet with full length complement from maize for TAF6 and signal recognition particle receptor.

1. TAF6

CLUSTAL 2.0.12 multiple sequence alignment

```
Zm              CTCCGGCGCCCGCGCGAGCTCTTCTTCCTTTCCGACGCCAGCGCGGTATCCTCGTCGTCG 60
Ec              ------------------------------------------------------------

Zm              CCAGTGCGCGAGCGTCTCCTCCTCCTCCTTGCGGCCTCCCGCGCCTCCGCCCCCCCCCCC 120
Ec              ------------------------------------------------------------

Zm              CCCCCCCCCCCCCACCCACACACACACATCCGGGAGAAGAGTAGTGGATGCGGAGCGTCTC 180
Ec              ------------------------------------------------------------

Zm              CTGTAAGCGGTCTCCCAAACGCCACGCTTGTACTCACTTTGTTTCTAATAAAAATTGCTG 240
Ec              ------------------------------------------------------------

Zm              TGGTGGAGGCCATAGATTCCAAGTGGTGGTTTAGAGAAGTCCCAAACATACCCTTTGTTT 300
Ec              ------------------------------------------------------------

Zm              GGGGAGAAGAATCGAGAGGGAGATAGCTAAACTAGGGTTAAGGGAATTAGGGGCCGAGAG 360
Ec              ------------------------------------------------------------

Zm              GAGCGCAATGAGCATCGTGCCCAAGGAGACGATCGAGGTCATCGCTCAGAGCGTCGGCAT 420
Ec              ------------------------------------------------------------

Zm              CCCCAGCCTCGGTGCCGATGTCGCCGTCGCGCTCGCCCCTGACGTCGAATACCGCCTCCG 480
Ec              ------------------------------------------------------------

Zm              CGAGATCATGCAGGAATCCATTAAATGTATGCGGCATGCAAAGCGGACTGTTCTGACGGC 540
Ec              ------------------------------------------------------------

Zm              GGATGATGTTGACAGTGCCCTGGGCCTCAGGAATGTTGAGCCCGTATATGGATTTGCCTC 600
Ec              ------------------------------------------------------------
```

```
Zm              TGGTGATCCATTACGTTTTAAGAGAGCTGTTGGACATAAGGATCTCTTCTATCTTGATGA 660
Ec              ------------------------------------------------------------

Zm              TCGGGAGGTAGACTTCAAAGAGATTATTGACTGTCCTCTGCCTAAAGCCCCACTGGACAC 720
Ec              ------------------------------------------------------------

Zm              ATCAGTTGTAGCCCATTGGCTAGCGATTGAGGGGGTACAGCCTGCAATACCGGAGAATCC 780
Ec              ------------------------------------------------------------

Zm              TGCTATTGACGCAATTGTACCACCAACTGAAAATAAAAGGTCTGAGCATGGGAAAGATGA 840
Ec              ------------------------------------------------------------

Zm              TGGACTTCCAGCTGATGTCAAGCTTCCTGTTAAGCATGTATTATCTAGAGAACTACAGAT 900
Ec              ------------------------------------------------------------

Zm              GTACTTTGATAAAATAGCTGAACTTACTATGAGTAGATCTGATACGTCCCTTTTTAAAGA 960
Ec              ------------------------------------------------------------

Zm              AGCTTTAGTGAGCTTGGCAAAAGATTCTGGTCTTCATCCTTTAGTTCCTTACTTTTCCTA 1020
Ec              ------------------------------------------------------------

Zm              CTTCATTGCAGACGAGGTTACCCGGAGTTTGGGTGATCTTCCTGTTTTACTTGCTCTAAT 1080
Ec              ------------------------------------------------------------

Zm              GAGAGTAGTCCAAAGTCTTCTCCGTAATCCACATATTCATATTGAACCTTATTTGCACCA 1140
Ec              ------------------------------------------------------------

Zm              GCTGATGCCATCAATGATCACTTGCATTGTTGCGAAAAGGCTAGGGCATAGGCTCTCAGA 1200
Ec              ------------------------------------------------------------

Zm              TAACCACTGGGAACTTAGGGACTTCTCTGCCAATTTGGTTGCCTTGGTATGTCAGAGGTT 1260
Ec              ------------------------------------------------------------

Zm              TGGTCATGTCTATCACAACCTTCAAAATCGGTTAACTAAAACATTGATCCATGCATTTCT 1320
Ec              ------------------------------------------------------------

Zm              TGATCCTGCCAAATCACTGACACAACATTATGGTGCTGTCCAAGGGATATCTGCACTGGG 1380
Ec              ------------------------------------------------------------

Zm              GCCCAGCGCGATTAGACTCCTTCTTTTGCCCAACCTTGTGACATATATGCAACTTCTGGA 1440
Ec              ------------------------------------------------------------

Zm              GCCGGAACTGCAGCTTGAGAAGCAGAAAAACGAGATGAAAAGGAAGGAAGCTTGGCGTGT 1500
Ec              ------------------------------------------------------------

Zm              TTATGGTGCCTTGCTGTGTGCTGCAGGCAAATGCTTGTATGATCGGTTGAAACTATTTCC 1560
Ec              ------------------------------------------------------------

Zm              TGGTTTGCTTTCTCCGTCGATGCGGCCTCTTTTACAGAGCAACAAAAGGGTCTTAACCAA 1620
Ec              ------------------------------------------------------------

Zm              CAACCCAAATAAGCGAAAATCTAGCACAGATCTCTCTGCAACTCAGCCACCTCTGAAGAA 1680
Ec              ------------------------------------------------------------

Zm              GATGGCAACGGACGCAACAGCCAACTCCATGGCCTCAGCTTCCATGGGAGGAAACATGCA 1740
Ec              ------------------------------CTGCGTCTTCTATGGCAGGAAACATGCA 28
                                              *  *  **** **** ************
```

```
Zm              AGGTGCCATGGATGGCTTTCCCAACCAGTTAGGCAACCCTGGCATGATGCAGGCCTCGTC 1800
Ec              AGGAACCATGGATGGGTTCTCCACCCAGTTAGCCAACCCTGGAATGATGCAAGCTTCATC 88
                ***  ********** **  *** ******** ********* ******** ** ** **

Zm              CTCTGGGCAGACGGTGGAGAGCATCCCATCAGCTGCGGTCCGGAGGGATCAGGGCAGTGA 1860
Ec              CACTTCACAAATGGTGGAAAACATTGCA---GTTGCGGTTCGGAGAGAACAGGGCAGTGA 145
                * **   ** * ****** * ***  **   * ****** ***** ** ***********

Zm              TCTCGCGCAGAGAGTTTCTGCAGTGCTTAGGCAAGCCTGGAAGGAGGACCAGGACACCGG 1920
Ec              TCTTGCGCAGAGGGTTTCCACGGTGCTGAGGCAAGCATGGAAGGAGGATCAAGATGCTGG 205
                *** ******** *****  * ***** ******** ********** ** **  * **

Zm              TCATCTCTTGGGGTCTCTGTACGAGGTCTTCGGGGAAGCCATCTTCTCCTTTGTCCAACC 1980
Ec              CCATCTCCTCGGGTCCTTGTACGAGGTCTTTGGTGAAGCCATCTTCTCCTTTGTTCAACC 265
                 ****** * *****  ************* ** ******************** *****

Zm              ACCAGAGATATCCTTGTTTGTGTAGAGAGAGAGAGAGAGGAAATGCATAGACTCTAGG 2040
Ec              ACCAGAGATATCCATCTTTGTGTAGAAAG------------GAAATGCGTAGGCTTTAGA 313
                ************* * ********** **            ******* *** ** ***


Zm              TTTTATTATTCTACGCTGTTGTACAACTAGGGATGTT-TTGAA----------GAGTTCA 2089
Ec              GGATTTTAATCTGCACTGTTGTATAATTAAGGGTATTCTTGAAATGGTAGACTGACATCA 373
                   * *** *** * ******** ** ** ** * ** *****          **  ***

Zm              CGAAACCCG--GTTTGACATTG---TCAAAGTAATGTGAC---TGAAGTTGTAGTTTACT 2141
Ec              GAAAACAAGTAATGTAATGCTAGGTTTATATTAATGTGAGGAATGAACATGTGTGTTGTA 433
                  ****  *   * * *   *    * * * ********    ****  ***   **

Zm              C------------------------------------------------------------ 2142
Ec              CACTTGTACTATTATTACTTTACGTGACTGATTATCTTAGCCATTTGGCTACTAATGTAT 493
                *

Zm              ----------------
Ec              ACGTACATCGTTTAAA 509
```

2. Signal recognition particle receptor

```
CLUSTAL 2.0.12 multiple sequence alignment


Zm              GGATTCATCTTCCTCTCGTCTTCCCTCTTCCTCTCCGAGGATACAAGGCGCCGTCCCAGG 60
Ec              ------------------------------------------------------------


Zm              TAGAGGTAGGGTCGCCACCATGCTGGAGGAGCTGCTCATCTTCACCCGGGGCGGCCTCAT 120
Ec              ------------------------------------------------------------


Zm              CCTGTGGTCGTCCTGCCGGGCGCTCGGCGCCGCAGCCCTCAAGGGCTCCCCCATCGACGC 180
Ec              ------------------------------------------------------------


Zm              CCTCATCCGCTCCTGTCTCCTCGAGGAGCGATCCGCCGACGCCAGCTTCTCCCAGGACAA 240
Ec              ------------------------------------------------------------


Zm              TTACGCCCTCAAGTGGACCTTCCACAACGACCTCGGCCTTGTCTTCGTCGCCGTATACCA 300
Ec              ------------------------------------------------------------


Zm              GAAGATACTCCACCTCCTCTACGTCGACGATCTCCTCGCCGCCGTCCGCAAGGAGTTCTC 360
Ec              ------------------------------------------------------------


Zm              CCAGATCTACGACCCCAAGCGCACCTCCTACGATGACTTTGGGGACATCTTCCGCCAGCT 420
Ec              ------------------------------------------------------------


Zm              CCACCGCGAGGCCGAGGCGCGCACCGAGGAAATGAAGAAGTCCAAGCAGGCCCTCTCGTC 480
Ec              ------------------------------------------------------------

Zm              CCGTCCGCAGCCTGCTCTCACCCAGAAGAAGCCTGGTCCCAAGGGCCGCGGCGGTGATTC 540
Ec              ------------------------------------------------------------
```

```
Zm              CCGCGGCGGCGCTGCTAACAAGGGCGGCTCTGGTGCCAAGGATGACTCTGATGGCGACGG 600
Ec              ------------------------------------------------------------

Zm              CTATGGCAACCAGGACCACCGTACTCTGGCAAATGGTGCCTTCAAGGGCCAAGAGAATGG 660
Ec              ------------------------------------------------------------GGA 3
                                                                             *

Zm              CGTCCGCAAGGAGAATTCTTCTCATGGCCGCACTGTTGTTGTCAAGGGAAAGGAGAATGG 720
Ec              CAAGGGACAGGAGAACTATCAGCCCCGTTCTGCTCTTCTTGTCAAAGGAAAGGAGAATGG 63
                *    *  ******* * *   *   *     ** ** ******* **************

Zm              GGCCCCCAAGGATGAA------GGGGCCTTCGATGTAAGTAAACTACAGAAGCTACGGGC 774
Ec              AGGTCCCAAGGACAACAACAACGGAGCCTTTGATGTAAATAAGTTACAGAAGCTGCGGAA 123
                 *  ********  *       ** ***** ******* ***  ********** ***


Zm              CAACAAGAAGAA---TGGTGCCACTGACAATGGAACCAAGAAGC---TGACCAAACCTGA 828
Ec              CAACAAGAAGAAAACTGTTCCCGCTGAAATCGGAACCAAGAAGAACACTACTAAACCTGA 183
                ************   ** * ** **** *  ************      ** ********
Zm              CATAAAGAAGAAAGGGAAGCAAGATAGGGTGTGGGACGACAAACCTTCC---AACAAGAA 885
Ec              CACTAAGAAGAAGGTCAA------CAGGGTCTGGGATGACAAACCTACCGACAACAAGAA 237
                **  ******** *  **       ***** ***** ********* **   ********

Zm              GCTGGATTTTACAGACCCGGCAGACGAGAGAGGGGACGA---GGTGGCAGACAATGTGGC 942
Ec              GCTGGATTTTACTGACCCAGCAGACAAGAGAGGGGATGATGAGGTGATCGACCATGCGCT 297
                ************ ***** ****** ********** **   ****   *** *** *

Zm              GGTCATCGAGGGAGAAAGCATGATGGATAAGGATGAGGATCTGAGCAGCGACAGCGAGGA 1002
Ec              CGTCAACCAGGGTGAGAGCATGATGGACAAGGATGAGAACGTCAGCAGCGACAGCGAGGA 357
                 **** * **** ** *********** ********* *  * *****************

Zm              CGAAGA---------GGTGGAAGATGGACCCAAAAAGAAAGGATGGTTCTCATCAATGTT 1053
Ec              TGAAGAAGAAGAACAGGTGCAAGGTGTCCCCAAAAAGAAAGGATGGTTCTCCTCAATGTT 417
                 *****         **** *** **  ********************** ********

Zm              CCAGAGTATTTCTGGTAACAATGTATTGGAGAAGTCTGATCTGCAGCCTGCGCTCAAAGC 1113
Ec              TCAGAGTATTTCAGGTAACAATGTGCTGGAGAAGTCTGATCTGCAACCTGCGCTTAAAGC 477
                 *********** ***********  ****************** ******** *****

Zm              TCTCAAAGATAGGCTGATGGCAAAGAACGTGGCAGAGGAAATTGCAGAGAAGCTTTGCGA 1173
Ec              TCTTAAAGATAGGCTGATGACCAAGAATGTGGCAGAGGAAATTGCAGAAAAGCTTTGTGA 537
                *** *************** * ***** ******************** ******** **

Zm              ATCTGTCGCAGCTAGCCTTGAGGGCAAGAAGCTAGGATCTTTCACTAGAATATCATCAAC 1233
Ec              ATCTGTTGCAGCTAGCCTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATCATCTAC 597
                ****** ***************** ** ******** ******** *********** **

Zm              AGTCCAGACAGCTATGGA-GGATGCGCTTCTTCGTATTTTGACCCCAAGGCGCTCTATTG 1292
Ec              TGTACAGACAGCAATGGATGGAGGCTCTTGTTCGCATTTTGACCCCAAGGCGATCAATTG 657
                 ** ******** ***** *** ** *** **** ***************** ** ****
Zm              ACATACTGAGGGACGTGCACGCTGCCAAGGAGCGTGGGAGGCCATATGTTATTGTTTTTG 1352
Ec              ACATACTGAGAGATGTGCAGTCTGCCAGGGAGCGTGGGAGGCCTTATGTCATTGTTTTCG 717
                ********** ** *****  ****** *************** ***** ******** *

Zm              TTGGGGTGAATGGAGTTGGCAAATCCACCAATCTTGCAAAGGTTGCTTATTGGCTTCTTC 1412
Ec              TTGGG-TCAATGGAGTTGGCACGTCCAC-AAT---------------------------- 747
                ***** * *************  ***** ***

Zm              AGCATAACCTCAGTGTTACACTTGCAGCATGTGACACCTTCAGATCTGGTGCTGTTGAGC 1472
Ec              ------------------------------------------------------------

Zm              AGCTGAGAACTCATGCAAGGAGGCTTCAGATACCTATATTTGAGAAGGGTTATGAGAAAG 1532
Ec              ------------------------------------------------------------

Zm              ATCCGGCAGTTGTAGCGAGGGAGGCTATTCAGGAAGCAACCCGGAACAAATCAGATGTTG 1592
Ec              ------------------------------------------------------------

Zm              TTCTGGTTGACACTGCTGGACGTATGCAGGACAACGAGCCACTCATGAGGGCACTCTCCA 1652
Ec              ------------------------------------------------------------

Zm              AGCTAATCAATCTCAATAGCCCAGACCTGGTTTTATTTGTTGGAGAAGCATTGGTTGGAA 1712
Ec              ------------------------------------------------------------

Zm              ATGATGCTGTGGATCAGCTCACTAAATTTAACCAGAAACTAGCAGACCTATCTGCTGTTC 1772
Ec              ------------------------------------------------------------
```

```
Zm              CCACTGCTAGATTGATCGATGGTATCCTGCTCACCAAGTTTGACACCATCGACGACAAGG 1832
Ec              ------------------------------------------------------------

Zm              TTGGAGCAGCGCTTTCCATGGTCTACATATCTGGAGCTCCAGTCATGTTCGTTGGCTGTG 1892
Ec              ------------------------------------------------------------

Zm              GTCAGTCCTACACGGACCTGAAGAAGCTGAACGTGAAATCCATTGTGAAGACCCTCCTGA 1952
Ec              ------------------------------------------------------------

Zm              AGTGATTGCACCACCGTGTTGAAGTGCAGTCTGTTATCTATTGTATCCCCAGTGACCTTT 2012
Ec              ------------------------------------------------------------

Zm              GTCATCCTCATGTTGGACGTCGTCGTGGATTTGTAACTGTCGGAGTTGCAACAAAACTTG 2072
Ec              ------------------------------------------------------------

Zm              TGTGCTTGCGTCAACGTGTGCCGTCTTGCTCTAATAATGCACTATCTGCATTTTGAATGG 2132
Ec              ------------------------------------------------------------

Zm              GTAAAAAATGGATGAGTATGCGAGTTTGCTATCATCTGGATGTTACCAGATT 2184
Ec              -----------------------------------------------------
```

APPENDIX IV

RESTRICTION ENZYMES AND SEQUENCES USED IN THE STUDY

Clone no. (C) 163 (MT-like protein)

GTTGCAAGTGCGGCTATATGTATCCTGAT^ATCCAGGAGAAAACCACCGGCGCCCCGTC
CACCGTCATCTTCGGNGTTGCAACGGATCTCAACGTCCCGCTCTTTGCTTCTACCNAAACA
TGCTAGATCGTCTTCTACTGCCGCTGCGGCTCCTTATGTTTTGCTATCCTGCACCTGCCTAT
TAGCTCAGAAC^CATGGGTGCTTGCTGCTTC

Comment [MSS1]: Eco R V site 29 + 182

Comment [MSS2]: Nco1 site 192 + 19

Clone no. (C) 172 (RISBZ4)

CCCCGACGGAGCATATGCAGAACTCGCCGTTGCAGAGCATCGCCACCAGCC^T
CGAGAGCCTCGACAACCGCATGGCCAGCGAGGTGACGAGCTGTGGCCGCGGCGG
CGCCGGCGTCGACGTCTGGCCCTGGGACTCTCACTCCAACGGTGGCTTGTCCAA
GTGAAGAGATGAAGC

Comment [MSS3]: Xho I site 51 + 124

Clone no. (C) 237 (sinapyl alcohol dehydrogenase) hypothetical

GCGTCGAAGACTGTGCTACTGAGGCGTGCAAAACTTGCTACGCGAATTGCCGG
ATTTGGTGCCGCTTTGTTCCGTCATGTTATCCCGACTGCGCCGAGCAATTTGAATG
CGATTCGAAGTGAGCATG^CATGTTGCCAGCGAGTGTCCCGCTGCTGTTAGGATA
CCTCGCCCTGCAACACAGTTGGGCGTTGCTTCATGTGTG

Comment [MSS4]: Pae I / Sph I site 127 + 75

Clone no. (C) 246 (putative TAF6)

CCAACCCTGGAATGATGCAAGCTTCATCCACTTCACAAATGGTGGAAAACATTG
CAGTTGCGGTTCGGAGAGAACAGGGCAGT^GATCTTGCGCAGAGGGTTTCCACGG
TGCTGAGGCAAGCATGGAAGGAG^GATCAAGATGCTGGCCATCTCCTCGGGTCCT
TGTACGAGGTCTTTGGTGAAGCCATCTTCTCCTTTGTTCAACCACC

Clone no. (C) 295 (farnesylated protein 1)

GTGCCGCACCCTTACGTGCCGGGTGCCTACGACAAGAAGGCGCCGCCC^GGGT
ACGTGCGCAACGTGCTCGACGACCCTGACGTCGGCCCGCTCGCGCGGGCAAACTC
CATGGAGGAGAAGTACACGTCCGCCTTCAGTGACGAGAATCCCAACGCGTGCGT
CGTCATGTGATCGGCTCGTGGCACTGTCTCGTTGCCTGTCG

Comment [MSS5]: Sma I site 48 + 154

Clone no. (C) 321 (putative signal recognition particle receptor)

GATAGGCTGATGACCAAGAATGTGGCAGAGGAAATTGCAGAAA^AGCTTTGTG
AATCTGTTGCAGCTAGCCTTGAGGGAAAAAAGCTAGGGTCTTTCACAAGAATATC
ATCTACTGTACAGACAGCAATGGAGGAGGCTCTTGTTCGCATTTTGACCCCAAGG
CGATCAATTGACATACTGAGAGATGTGCAGTCTGCCAAGGAGCGTGGGAGGCC
TTATGTC

Comment [MSS6]: Hind III site 43 + 179

Clone no. (OP) 168 (putative transcriptional regulator)

GAACGTCAGACTCGTCGCCCGTTTCCCATCATGGCCGATTCCGCCGCGCTCTCC
ACCGCCTCCTCCTCCACCCTTCCCCGCTCGCTGGCGTCCAGGAGGCCCACCCCTCT
CTACGCCGGGGCC^CCGCACACGCGCTGCTGCCGGTTCGGGCCGCGGTCGCTCCA
CCGGTTGCAGGCTCGGGCACGGAAGGATGATCCGGAGGACCTCTACGGCCCCTA
CCCCTGGGAACAGCCCCTGGATCGCACCACAGACTTCG

Comment [MSS7]: Apa I site 123 + 133

Clone no. (OP) 175 (protein phosphatase)

GGCGGCTACGGCTTCGGCGTTACAGGGGACTAGGTAAGGGAAGGGAGGGTGA
CGTCGAGCGGCAGCGGCGATGTCGTCGCACGCGGATCTGGACCGGCA^GATCTCG
CAGCTGCGGGAGTGCAAGTTCCTCCCCGAGGCAGAGGTGAAAGCACTCTGCGAG
AAGGCGAAGGCGATCCTCATGGAGGAATGGAACGTGCAGCCCGTGCGCTGCC

Comment [MSS8]: Bgl II site 99 + 113

Clone no. (OP) 53 (Low molecular mass early light-inducible protein HV90,)

GGTGCCACTGCTCCAGGGGGAGACCGCCGAGGCCAAGACCAGCGGACTCATGA
ACGCCAATGCGGAG^CTCTGGAATGGCCGCCTCGCCATGCTCGGGCTTGTTGCGC
TCGCAGTGACCGAATACCTCACCGGCGCGCCCTTCATCAACGCGTAATTAATCAT
GTCTTATGTGTACATAGCTTAGGCATTCTCCGTATCAAATTTCCCGGGGCA**CGTG**
TATGACGTATATCAGGCGG

Comment [MSS9]: Ecl136II site 67 + 169