DIFFERENTIAL TRANSCRIPT PROFILING AND CLONING OF CANDIDATE GENES IN *FUSARIUM*-WILT
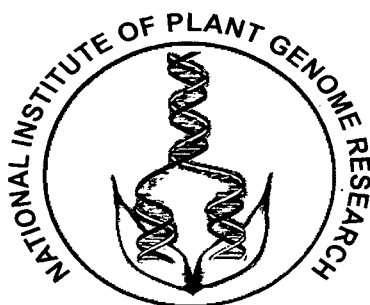
THESIS SUBMITTED TO
JAWAHARLAL NEHRU UNIVERSITY, NEW DELHI
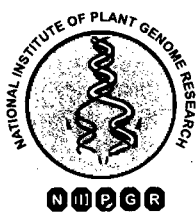FOR THE AWARD OF THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

NASHEEMAN ASHRAF

NATIONAL INSTITUTE OF PLANT GENOME RESEARCH
NEW DELHI-110067, INDIA

2009

CERTIFICATE

The research work embodied in this thesis entitled **"Differential Transcript Profiling and Cloning of Candidate Genes in *Fusarium*-Wilt"** has been carried out at the National Institute of Plant Genome Research, New Delhi. The work is original and has not been submitted so far in part or in full for any degree or diploma of any university.

Nasheeman Ashraf
(Student)

Prof. Akhilesh Kumar Tyagi
(Director)

Prof. Akhilesh Kumar Tyagi
Director,
National Institute of Plant Genome Research,
(Department of Biotechnology,
Ministry of Science & Technology, Govt. of India)
Aruna Asaf Ali Marg, New Delhi-110067

Dr. Subhra Chakraborty
(Supervisor)

सर्व शिक्षा अभियान

सब पढ़ें सब बढ़ें

Dedicated to the Victims of Terrorism...

Acknowledgements

يَا أَيُّهَا الَّذِينَ آمَنُوا اسْتَعِينُوا بِالصَّبْرِ وَالصَّلَاةِ إِنَّ اللّهَ مَعَ الصَّابِرِينَ

Strengthen yourselves by patience and prayer, and know that Almighty supports the patient people.

It is this verse from Holy Qur'an which kept me going through this tough journey.

As I sit back to write this little piece, I think of innumerable hands that steered me through this journey and put me on the right path of learning.

First and foremost I express my deep sense of gratitude to my supervisor Dr. Subhra Chakraborty for her encouragement, persistent attention and inspiring guidance that helped me to perform this task. I shall always be thankful to her for her belief and for giving me freedom of work.

I am thankful to Dr. Niranjan Chakraborty for his support and encouragement.

I am grateful to the founder director, NIPGR, Prof. Asis Datta, for being a source of inspiration and encouragement.

I extend my sincere thanks to Director NIPGR, Prof. Akhilesh Tyagi, for providing me an opportunity to work in this institute.

I am thankful to Prof. Sushil Kumar for helping me to understand the core genetics during my course work. I am also thankful to other faculty members of NIPGR family for their invaluable lessons during the course program which helped me in later part of my work.

I owe my thanks to my M.Sc teacher and supervisor, Prof. D.P Mishra. He has always been a source of inspiration and an icon of simplicity. He always wanted me to be at the top and his faith helped me to strive harder in order to achieve my goal.

My thanks are due to all my seniors Sarjeet da, Deepali di, Pranjan, Debaleena di, Arnab da, Deepti di Lalit sir, Ajay sir, Swagata di, Hanumanth, Vijay Nalla, Niroj sir, Rakesh sir, Kundu da and poornima di.

I owe my thanks to ever respected Vijay sir whose guidance helped me in developing scientific aptitude. Seeing him working sincerely from dawn to dusk has always inspired me to follow the same steps. I am also indebted to Nagu sir for all his help and encouragement. He is the person who made me to master all the computational analyses. Rajgourab, as a senior and friend was always there in good and bad times. I am thankful to him for always being supportive and helpful and for making life in lab cheerful.

I will always be indebted to my friend kishwer. Mere words won't suffice what she means to me. She was my family in hostel, always there to listen to my disappointments and dilemmas and always guided me to choose the right path.

I thank my batchmates Archana, Sreeram, Santosh, Shalu and Aarti, for their support and encouragement. Manikant and Kamal deserve special thanks for helping me in difficult situations. Manikant was always a friend in need and Kamal, the cloning expert, helped me learn many things. I m also thankful to Mihir for his help and even though he left me all alone in the midst of the journey, it helped me grow stronger.

I am grateful to all the juniors in THE GENOMICS group, Sudip, Manindra, Kanika, Rajul, Naval, Anand and Debarati. They have always helped me in one way or the other. And even though I used to scold them, I could always find them besides me to extend any help needed. During later part of my PhD when I was all stressed out and therefore in need of a helping hand, I found a great junior, a friend and a sister, Sushmita, who shared all my troubles and helped me to come out of tough situations. I am really very thankful to her. I am thankful to Swaraj for his help and suggestions and for giving new dimensions to my work. His scripts did the job of magic and made things easy for me. I am

TABLE OF CONTENTS

Chapter 1

Introduction

1. Transcriptional reprogramming-A means of survival

Biological phenomena at different organismic levels have revealed some sophisticated systematic architectures of cellular and physiological activities implicitly. All the cellular processes in living cells such as growth, development, morphogenesis, cellular differentiation and stress responses are controlled by gene expression programs involving complicated transcriptional regulation of several genes. The molecular machinery underlying such activities reveals that these processes are mostly interpreted as frameworks of connectivity between biochemical compounds and proteins which are synthesized from genes. These gene products function as transcription factors which bind to regulatory sites of other genes, as enzymes catalyzing metabolic reactions or as components of signal transduction pathways. This implies that in order to understand the molecular mechanism by which genes control cellular processes, the scope should be broadened from DNA sequences coding for proteins to the systems of genetic regulatory pathways determining which genes are expressed, when and where in the organism and to which extent. The existence of such regulatory switches makes sense evolutionarily because these mechanisms prevent the cell from devoting energy to producing the entire set of components (such as RNA, proteins, and metabolic intermediates) of a pathway when achieving the final effect is not beneficial to the cell at a particular point (Alberts *et al.*, 2002). In addition to on/off switches, living organisms have evolved mechanisms to quantitatively change the rate of cellular processes in response to internal and external cues. Indeed, the identification and dissection of such regulatory processes have come to occupy centre stage in molecular and cellular biology.

Most biological phenomena are directly or indirectly influenced by various environmental cues including various forms of stresses. For example, animal system responds to different diseases by activation of immune system which helps them fight such diseases. These immune responses are determined by programs of gene expression, which are strictly regulated spatio-temporally by a complex network of interacting molecular mechanisms that control this response. Similarly, plants as sessile organisms are constantly exposed to environmental stress conditions. The stresses can come in many forms depending on the crop production systems. In general, these stresses are categorized as biotic, which are caused by biological agents such as pathogens and insects and abiotic, which are brought about by the physical

environments, such as dehydration, salinity, cold, heat, ozone etc. (Bucahanan *et al.*, 2000). Many of these stresses are likely to intensify under the various scenarios of climate changes. During biotic stress, plants are subject to attack by a wide variety of microbial pathogens and insect herbivores and in response plants express numerous defense mechanisms. Pathogens deploy one of three main strategies to attack plants: necrotrophy, biotrophy, or hemibiotrophy. Necrotrophs first kill host cells and then metabolize their contents. These pathogens have a broad host range and cause cell death which is often induced by toxins and/or enzymes targeted to specific substrates. *Pythium*, *Botrytis* and *Fusarium* species are examples of fungal necrotrophs. Biotrophic and hemibiotrophic pathogens invade living cells and subvert metabolism to favour their growth and reproduction (Glazebrook, 2005). Biotrophs, for example, *Pseudomonas syringae*, tend to cause disease on only one or a few related plant species. Hemibiotrophic fungi such as *Phytophthora* and *Colletotrichum* kill surrounding host cells during the later stages of the infection (Hammond-Kosack and Jones, 1997).

Most plant-pathogen interactions are fierce battles of attack and counterattack. These battles are fought with highly sophisticated means for the survival of an individual. On the plant side, the most immediate defense response includes the reprogramming of cellular metabolism and highly dynamic structural rearrangements within and around the attacked cells. Whereas some of these defense mechanisms are preformed and provide physical and chemical barriers to hinder pathogen infection, others are induced only after pathogen attack. In animals these two forms of immune responses are called as innate and adaptive immunity respectively. Innate immunity forms a first line of defense against invading pathogens and is also a key element for the deployment of adaptive immunity. Plants lack adaptive immune system and depend entirely on innate immunity for defending themselves against pathogen attack.

Identification of the genetic regulatory networks is important to understand the diseases and work out the strategic control of such diseases. The rapid advances of genomics, transciptomics, proteomics and metabolomics make it possible to understand such networks. More particularly, the embedded time-course feature of microarray data would promote the system analysis of transcriptional and signal regulatory pathways. Various studies have been carried so far to understand plant

immune system and the molecular mechanism underlying this response. An overview of plant immunity is given below.

2. Plant immune system

Plant immune responses are governed by the combined genotypes of host and pathogen and depend on a complex exchange of signals and responses occurring under given environmental conditions. Both plants and animals have complex mechanisms to recognize and respond to attack by pathogenic microorganisms. Immune responses may be innate or adaptive. Plants, unlike mammals, lack mobile defender cells and a somatic adaptive immune system. Instead, they rely on the innate immunity of each cell and on systemic signals emanating from infection sites (Dangl and Jones, 2001). Plants respond to infections using a two-branched innate immune system. One is based on trans-membrane pattern recognition receptors (PRRs) that respond to slowly evolving microbial associated molecular patterns (MAMPS) or pathogen-associated molecular patterns (PAMPs) (Zipfel and Felix, 2005). PAMPs are evolutionarily conserved pathogen-derived molecules that distinguish hosts from pathogens and include lipopolysaccharide, peptidoglycan, bacterial flagellin and mannans of yeast. The nonpathogens also synthesize these molecules, therefore, 'microbe-associated molecular pattern' (MAMP) is a more precise term. The second branch of immune system acts largely inside the cell using protein products encoded by R genes (Dangl and Jones, 2001). Most of these R gene products belong to the class of NB-LRR proteins. The NB-LRR-mediated disease resistance is effective against obligate biotrophs or hemibiotrophic pathogens, but not against necrotrophs (Glazebrook *et al.*, 2005).

Jones and Dangl (2006), described plant immunity in the form of a four phased zig zag model (Figure 1.1). In phase 1, PAMPs or MAMPs are recognized by PRRs, resulting in PAMP-triggered immunity (PTI) that can halt further colonization. In phase 2, successful pathogens deploy effectors that interfere with PTI, contribute to pathogen virulence and these result in effector-triggered susceptibility (ETS). In phase 3, a given effector is specifically recognized by one of the NB-LRR proteins, resulting in effector-triggered immunity (ETI). Recognition is either indirect or through direct NB-LRR recognition of an effector. ETI is an accelerated and amplified PTI response, resulting in disease resistance and usually a hypersensitive cell death response (HR) at
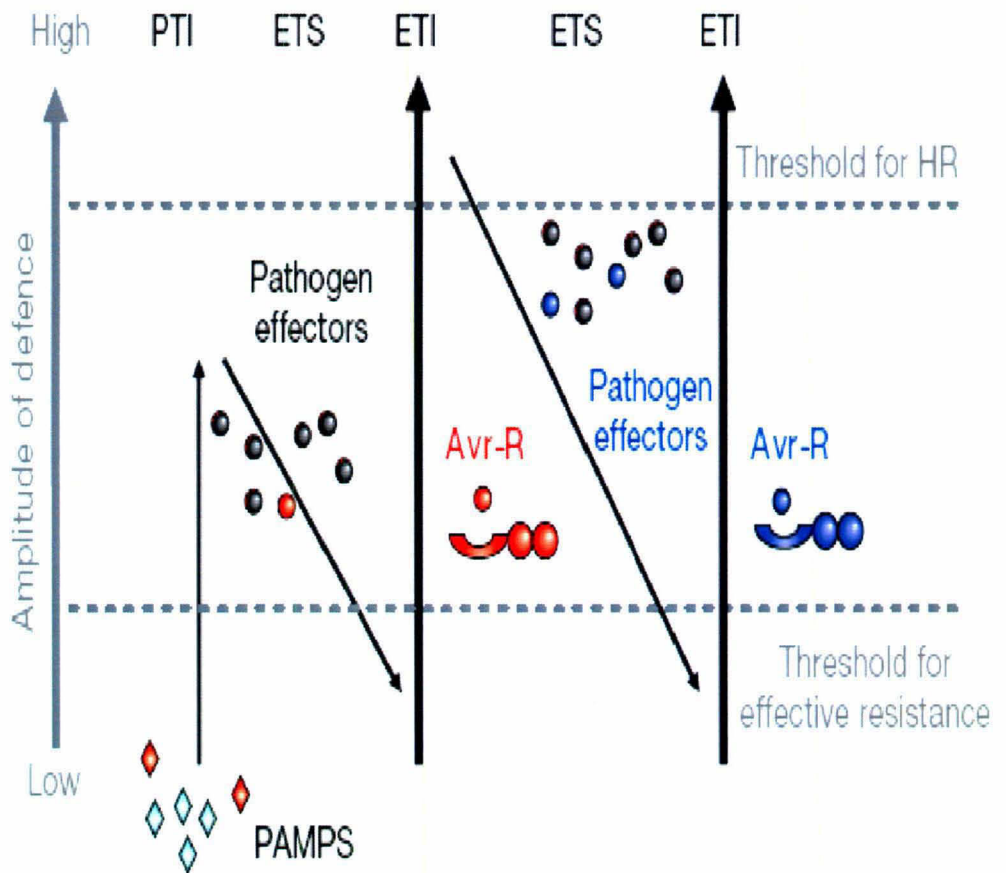
Figure 1.1: Zigzag Model showing quantitative output of plant immune system.

Source: Jones and Dangl. *Nature*, 2006, 444: 323-329

the infection site. In phase 4, natural selection drives pathogens to avoid ETI either by shedding or diversifying the recognized effector gene or by acquiring additional effectors that suppress ETI.

2.1 Components of two branches of plant immune system

The structure and function of the components of two branches of immune system, that is, PTI and ETI are discussed below:

2.1.1 Components of PTI

The structure and function of the components of PTI are discussed below:

2.1.1.1 Microbe-associated molecular patterns (MAMPs)

MAMPs are defense elicitors that form a core component of the microorganism and are evolutionarily stable. These cannot be sacrificed or altered much without seriously impairing viability of the microorganism. Microorganisms must reach a critical mass in the plant interior before the basal immune system is strongly activated. Smaller or primarily external/epiphytic microbial populations are usually less potent at inducing *PR* gene expression and other active defenses. MAMPs are known to induce PTI which is the first phase of overall defense response system in plants. The archetypal elicitor of PTI is bacterial flagellin, which triggers defence responses in various plants (Gomez-Gomez and Boller, 2002). A synthetic 22-amino-acid peptide (flg22) from a conserved flagellin domain is responsible for inducing cellular responses (Felix *et al.*, 1999). A genetic screen using flg22 led to the identification of *Arabidopsis* LRR-receptor kinase FLS2, which binds flg22 (Chinchilla *et al.*, 2006) and acts as flagellin receptor. Plant FLS2 and animal TLR5 both recognize flagellin domains that are highly conserved. Bacterial cold shock proteins and elongation factor Tu (EF-Tu) activate defence responses similar to flg22 (Felix and Boller, 2003; Zipfel *et al.*, 2006). Ef-Tu is recognized by an *Arabidopsis* LRR-kinase called EFR. Treatment with a conserved EF-Tu peptide induces expression of a gene-set nearly identical to that induced by flg22 (Zipfel *et al.*, 2006). Conversely, EFR transcription is induced by flg22. This suggests that the responses to MAMPs/PAMPs converge on a limited number of signaling pathways and lead to a common set of outputs that comprise PTI. Moreover, apart from the above mentioned MAMPs of bacterial origin, some fungal MAMPs have been identified, for example, chitin and ergosterol.

2.1.1.2 MAMP receptors

Most of the MAMP receptors, for example FLS2 and EFR1 encode transmembrane proteins with an extracellular LRR domain and an intracellular protein kinase domain. The first LRR kinase found to be involved in plant defense was an *R* gene product, Xa21 of rice, and other plant R proteins also have this structure (Song *et al.*, 1995). FLS2 itself functions in ways that resemble an R protein (Zipfel *et al.*, 2004), suggesting a further overlap between R proteins and MAMP receptors. However, MAMPs are directly bound to their receptors, while some R proteins do not directly bind pathogen effectors, but rather detect them indirectly via their perturbation of host proteins. FLS2 and EFR1 respond to entirely different MAMPs yet both activate very similar plant defense responses. The similarity of MAMP receptors between plants and animals has frequently been noted. For example, FLS2 and human TLR5 that perceive flagellin and activate innate immune responses are both LRR kinases (Ausubel *et al.*, 2005). However, the two proteins recognize different flagellin domains and do not exhibit common derivation. Recently, CEBiP, the high affinity-binding site for fungal chitin, has been identified in rice (Kaku *et al.*, 2006). However, it remains elusive whether there are CEBiP homologues of similar function present in other plants.

2.1.2 Components of ETI

The structure and function of the components of ETI are discussed below:

2.1.2.1 Pathogen effectors

Effectors, such as toxins and effector proteins, can be defined as pathogen-derived molecules intended to promote pathogen virulence by interacting with the host. In other words, effectors are virulence factors that usually do not have "housekeeping" function in microbial growth and development outside of the host. However, when an effector is recognized by a host defense receptor, the intended virulence function is often overshadowed by a dominant avirulence function. Many effectors were first identified on the basis of their avirulence activity. These were appropriately called *Avr* genes since their *R* gene-mediated activity induces defense that prevents virulence (Lucas, 1998). Bacteria deliver many type III effectors beyond those known to be Avr proteins. Recent work has demonstrated that oomycetes also secrete effectors into both the extracellular space and the inside of host cells (Kamoun, 2006). Thus,

pathogens can deliver a potpourri of effectors into cells of the host by a variety of mechanisms.

2.1.2.1 R proteins

Most *R* genes encode NB-LRR proteins and recognize the effectors which are termed as avirulence (Avr) proteins. The initial cloning and molecular characterization of *Avr* genes and *R* genes (Bent, 1996) were major landmarks in plant defense studies. A large number of plant *R* genes specifying resistance to bacterial, fungal, viral or nematode pathogens and aphids, have now been cloned. Despite the widely different modes of pathogen colonization, analysis of the structural features of R proteins reveals the existence of only a limited number of sequence motifs. These include putative protein interaction/recognition domains such as leucine-rich repeats (LRR) and leucine zippers (LZ), signaling functions such as a kinase domain or nucleotide-binding site (NB), and the TIR domain, defined by homology to the intracellular effector domains of the *Drosophila* Toll and human interleukin-1 receptors (Bent and Mackey, 2007). Recruitment of this rather limited repertoire of motifs suggests that the processes underlying R protein-specified recognition of unrelated pathogens are mechanistically highly conserved and probably engage a restricted number of downstream defense pathways.

LRRs make a highly adaptable structural platform on which very different binding specificities can evolve. Studies of a number of plant *R* gene families have shown that LRR domains can be under diversifying selection (Ellis *et al.*, 2000). In most other proteins the key domains are conserved across taxa, presumably due to natural selection to maintain function. But in some *R* gene families the predicted solvent-exposed residues along the concave face of the LRR, not only lack conservation, but are significantly more diverse than expected from random drift. This suggests selective pressure to adopt new function in this part of the R proteins. The prediction has been that the diversifying sites encode the pathogen-specificity domains of R proteins, and that their evolution permits recognition of different pathogen Avr proteins (Boller and Felix, 2009).

2.2 How do R proteins recognize the presence of effectors?

Three different mechanisms have been put forth for the molecular recognition of pathogen effectors by plant R proteins.

2.2.1 Direct interactions

This form of recognition occurs via direct physical association of the pathogen effector with the R immune receptor, similar to a ligand binding to its receptor (Figure 1.2A). The first example of this mode was shown between the Pita CC-NB-LRR immune receptor in rice and the AVR-Pita effector from the fungus *Magnaporthe grisea* (Jia *et al.*, 2000). The LRR domain of Pita directly interacts with the AVR-Pita effector, and a single amino acid substitution in the LRR can abolish this interaction, resulting in loss of resistance. Another example of recognition by direct association was discovered between the *Arabidopsis* RRS1-R immune receptor with the bacterial wilt effector PopP2 (Deslandes *et al.*, 2003). Also this type of interaction was studied in flax multigenic loci (*K, L, M, N,* and *P*) whose products recognize about 30 effector proteins from the flax rust fungus (Ellis *et al.*, 2007). The polymorphic L locus encodes a TIR-NB-LRR gene with 13 allelic variants (*L, L1-L11,* and *LH*). *L5, L6* and *L7* are alleles that show differing resistance specificities to corresponding fungal effectors encoded by the *AvrL567* multigenic locus of related genes. Yeast two-hybrid analysis showed a correlation between the direct physical interaction of AvrL567 effectors with their cognate L immune receptor and the activation of resistance. These detailed studies convincingly demonstrate that direct interactions can drive recognition specificity. Sequence analysis and domain swap experiments with L immune receptors have shown that the LRR domain is the major determinant for effector specificity (Caplan *et al.*, 2008).

2.2.2 Indirect interaction

R immune receptors can also recognize pathogen effectors in a more mechanistically complex and indirect way. Many effector proteins alter or modify certain host proteins during pathogen infection. The ''Guard Hypothesis'' suggests that R proteins monitor these host target proteins and activate defenses if they are perturbed (Dangl and Jones, 2001). There are two variations of this model. The R immune receptor can be constitutively bound to its guardee host factor (Figure 1.2B), or alternatively, the R immune receptor may bind to its guardee only after the guardee is bound by the pathogen effector (Figure 1.2C). The most extensively studied host target (guardee) protein is *Arabidopsis* RIN4 (RPM1 interacting protein 4), which constitutively associates with the CC-NB-LRR immune receptors RPM1 and RPS2. Three

structurally unrelated *Pseudomonas syringae* effector proteins modify RIN4, which leads to activation of RPM1 and RPS2. RPS2 is activated when AvrRpt2, a cysteine protease from *P. syringae* promotes cleavage of RIN4. Structure-function analysis of the RIN4-RPS2 association suggests that RIN4 keeps RPS2 in an inactive state until it is cleaved. RIN4 associates with two other *P. syringae* effectors, AvrRpm1 and AvrB, which promote the phosphorylation of RIN4 that then induces the activation of RPM1 immune response by an unknown mechanism (Mackey *et al.*, 2002).

2.2.3 Pathogen recognition through effector-mediated transcriptional activation of Non-NB-LRR resistance proteins

Recent work on four non-NB-LRR immune receptors has discovered a third novel mechanism of pathogen recognition (Figure 1.2D) that activates the transcription of resistance proteins by effector molecules that act as transcription factors. The AvrBs3 effector protein from *Xanthomonas campestris* contains a central repeat region, leucine zipper domain, an NLS (nuclear localization signal), and an acidic activation domain, which makes it similar to eukaryotic transcription factors. AvrBs3 localizes in the nucleus and is recognized by the pepper Bs3 resistance gene product. AvrBs3 directly binds to the promoter of Bs3 leading to Bs3 transcript accumulation, eventually resulting in HR. An AvrBs3-related protein, AvrXa27, from *Xanthomonas oryzae* is recognized by rice *Xa27*, which encodes a novel protein (Caplan *et al.*, 2008). Besides two predicted alpha helices, the protein's coding sequence does not contain homology to known domains to predict its structure and function. Resistant and susceptible *Xa27* alleles show near identical protein sequences with substantial differences only in their promoter regions. AvrXa27 specifically induces the transcription of the resistant allele of *Xa27*, indicating that the *Xa27* promoter determines its effector specificity. Direct binding of AvrXa27 to the promoter of *Xa27*, however, remains to be demonstrated. Similarly, the *Pantoea agglomerans* HsvG effector recognized in gypsophila and the HsvB effector recognized in beet have putative transcriptional activity in plants (Nissan *et al.*, 2006). Swapping of the transcriptionally active repeats of HsvG and HsvB swaps their specificity, suggesting that these repeats provide recognition specificity. Further studies on the targets of HsvG and HsvB will determine if they transcriptionally activate resistance proteins. Thus, these effector proteins act as transcription-factor mimics that activate synthesis of desired host proteins by binding to select promoter elements. Hence, the promoter
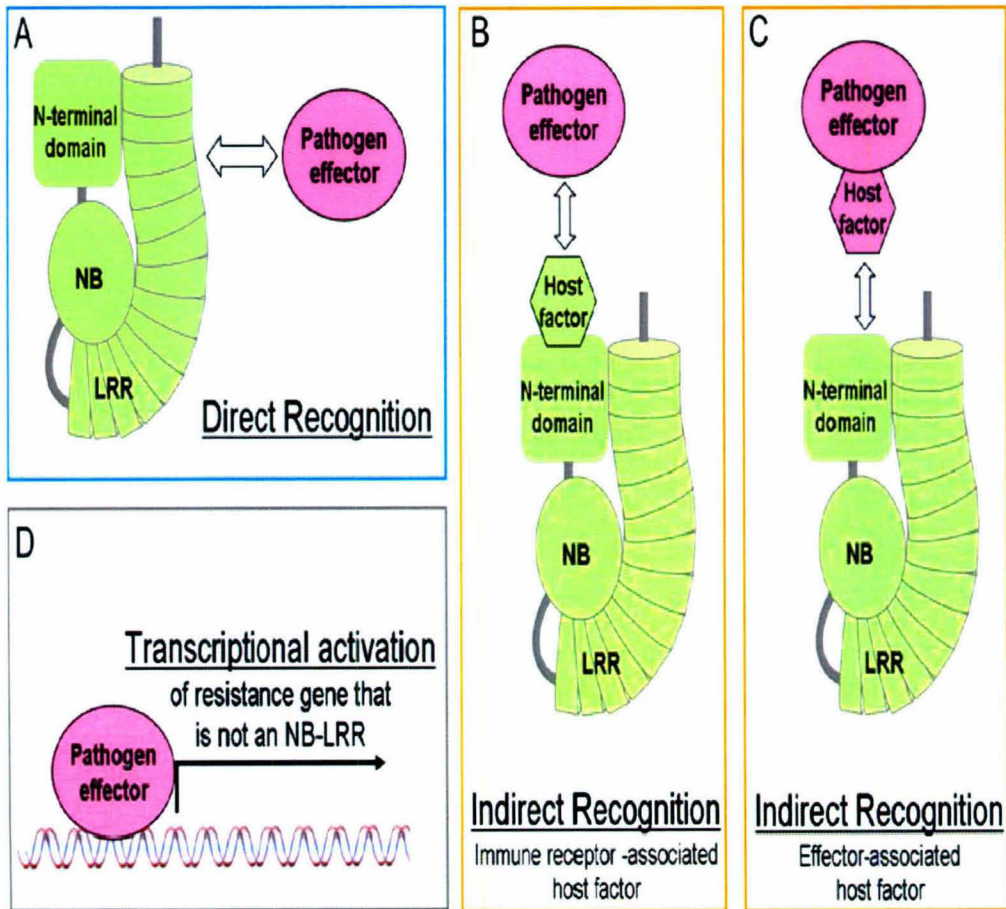
Figure 1.2: Various modes of pathogen recognition.

Source: Caplan *et al. Cell host and microbe*, 2008, 3: 126-135

is functioning as the immune receptor. The hosts, in turn, evolved to exploit this strategy to trigger expression of defense genes.

2.3 Induction of defense responses by R proteins

R proteins that are activated upon effector recognition fall into five main classes depending upon the structural motifs present in them. These classes include serine threonine kinases, NB-LRR proteins with an extra N-terminal coiled coil (CC) or leucine zipper (LZ) domain, NB-LRR proteins with N-terminal Toll and Interleukin 1 receptor like (TIR) domain, transmembrane (TM) LRR protein and TM-LRR kinase proteins (Martin *et al*., 2003). Among all these R proteins, NB-LRRs form the most predominant and investigated class. The mechanism of defense activation known so far is largely based on this particular class. Many NB-LRR R proteins are apparently maintained in an ADP-bound "off" state by interactions of LRR and NB domains. Elicitation disrupts these interactions and allows ADP release/ATP binding and opening of the protein for defense-signaling protein-protein interactions. Induction of defense responses by NB-LRRs occurs in three phases (Figure 1.3). In the first phase, the pathogen effector associates with the NB-LRR immune receptor. In the second phase, the NB-LRR immune receptor is activated by a conformational change and ATP binding to the NB domain. In third phase, activated NB-LRRs function in the nucleus to induce defense-related signaling. This can occur by inhibiting negative regulators of defense. Alternatively, an unknown transcription factor may bind to NB-LRRs to positively regulate and induce defense-related expression. The nuclear and cytoplasmic pools of NB-LRRs are most likely maintained by nucleocytoplasmic shuttling. To cross the nuclear pore NB-LRRs with a classical NLS require importin-a and importin-b for import and exportin for export. The inactive form of NB-LRRs is found in both the cytoplasm and nucleus in the absence of the pathogen. However, in presence of the pathogen effector, the activated form of an NB-LRR accumulates in the nucleus to initiate defense signaling (Caplan *et al*., 2008).

3. Defense signaling pathways

PRR or *R* gene mediated immune responses are very similar (Navarro *et al*., 2004). However, constitutive defense components and associated signaling events playing major roles in these two immunity barriers might differ (Zipfel *et al*., 2004). Overall, these responses involve ion fluxes across the plasma membrane, generation of
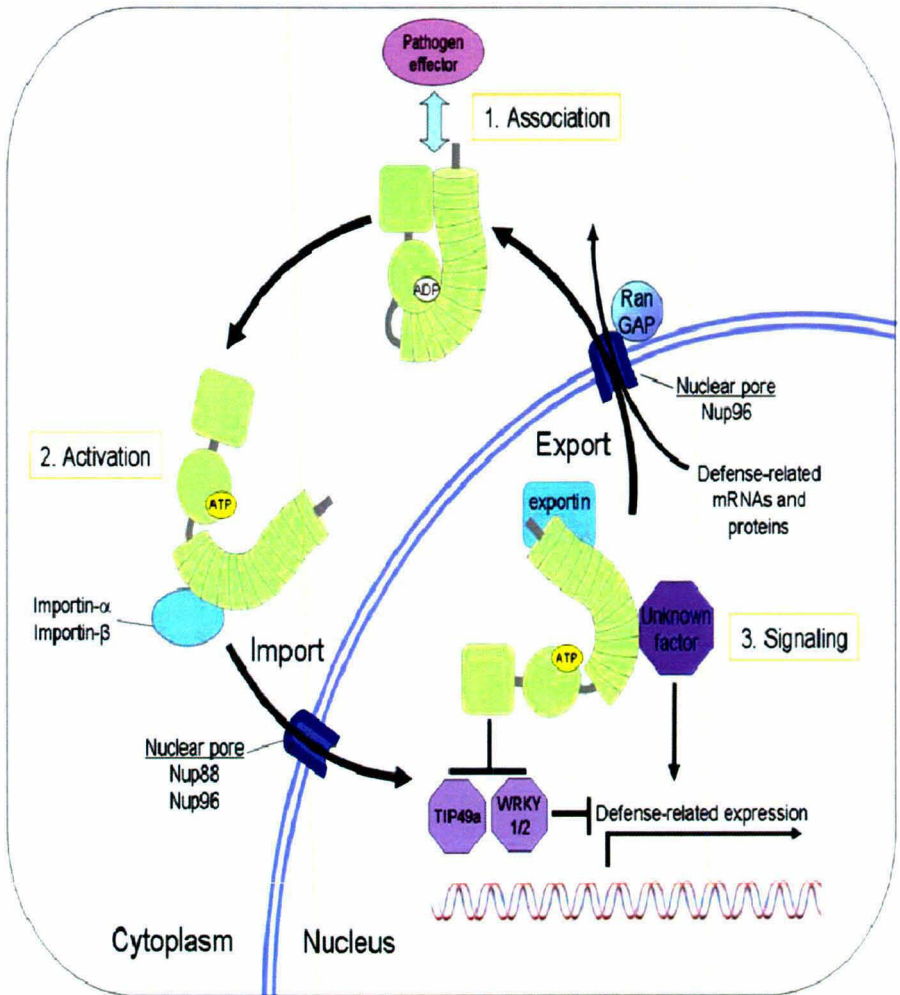
Figure 1.3: Model showing defense activation by NB – LRRs.

Source: Caplan *et al. Cell host and microbe*, 2008, 3: 126-135

reactive oxygen intermediates (ROI) and reactive nitric oxide (RNO), deposition of callose, activation of calcium-dependent and mitogen-activated protein kinases, and transcription of numerous defense genes. Many studies demonstrate that SA, JA and ET are the main molecules signaling the activation of defense genes (Thomma *et al.*, 2001). The SA pathway is primarily linked to resistance to biotrophic pathogens, whereas the JA and ET pathways mediate resistance mostly to necrotrophic pathogens, indicating that the activation of defense signaling pathways depends on the pathogen lifestyle and its mode of infection. Frequently, these pathways interact in an antagonistic manner (Glazebrook, 2005), a circumstance that can be used by pathogens to avoid plant defenses.

3.1 SA-dependent signaling

Occurrence of the HR or attack by certain pathogens triggers activation of SA dependent signaling. SA level rises, leading to activation of various presumed defense effector genes, including *PR-1*. Two genes, *PAD4* and *EDS1*, are required for activating SA accumulation in response to some, but not all, SA-inducing stimuli. These genes encode proteins similar to triacyl-glycerol lipases that interact with each other (Glazebrook, 2005). *SID2* is also required for SA signaling. It encodes isochorismate synthase, and SA production is drastically reduced in *sid2* plants, indicating that the majority of SA is produced from isochorismate. However, there is some SA present in *sid2* mutants, and this may be synthesized through the phenylalanine pathway. *EDS5* is also required for production of SA in response to pathogen attack. It encodes a MATE family transporter that may be involved in transport of intermediates for SA biosynthesis. Pathogen-induced expression of *EDS5* requires *EDS1* and *PAD4*, placing *EDS5* downstream from *PAD4* and *EDS1* in SA signaling. The transgene *NahG* encodes a bacterial salicylate hydroxylase that destroys SA by converting it to catechol (van Wees *et al.*, 2003). *NPR1* acts downstream of SA. When SA levels are low, NPR1 exists in an oligomeric form in the cytoplasm. When SA levels increase, the NPR1 oligomers dissociate into monomers owing to reduction of disulfide linkages holding the monomers together. The monomers then enter the nucleus, where they interact with TGA-type transcription factors which are required for the activation of *PR-1* expression by SA (Zhang *et al.*, 2003). The transcription factor WRKY70 is also required for full expression of *PR-1* in response to infection. Expression of *WRKY70* is SA-inducible

and *NPR1* dependent (Li *et al.*, 2004). Some SA-dependent defense responses are independent of *NPR1*, indicating the existence of another branch of the SA signaling pathway. SA levels are higher in infected *npr1* mutants than in infected wild-type plants, suggesting that *NPR1* functions in controlling SA levels as well as in responding to elevated SA (Dong, 2004). The occurrence of HR cell death leads to activation of SA signaling throughout the plant. Consequently, plants that have been challenged with an avirulent pathogen develop resistance to subsequent infections by pathogens that are sensitive to SA regulated defense responses. This phenomenon is known as systemic acquired resistance (SAR) (Glazebrook *et al.*, 2005).

3.2 JA and ET dependent signaling

JA-dependent signaling proceeds through increased JA synthesis in response to pathogen attack and consequent increases in expression of defense effector genes such as *PDF1.2*. Some JA regulated genes are also regulated by ET. In case of *PDF1.2*, induced expression requires both JA and ET. In contrast, ET is not required for expression of the JA inducible gene *VSP1* (Norman-Setterblad *et al.*, 2000). Cellulose synthases in the plant cell wall seem to be involved in regulation of JA levels, as the cellulose synthase mutant *cev1* displays constitutively high JA levels and JA dependent gene expression. Plants bearing *jar1* mutations are defective in responses to JA. *JAR1* encodes a JA-amino synthetase that can form conjugates between JA and several amino acids including isoleucine (Glazebrook, 2005). The isoleucine conjugate may be the active form of JA. All known activities of JA in *Arabidopsis* require the function of *COI1*. *COI1* encodes an F box protein that is presumed to act in proteolysis. It has been shown to form active SCFCOI1 complexes with cullin, Skp1-like proteins and AtRbx1 that are expected to function as E3 ubiquitin ligases (Xu *et al.*, 2002). Some responses to JA also require the function of a MAP kinase encoded by *MPK4*, as *mpk4* mutants fail to express *PDF1.2* and another JA regulated gene, *THI2.1*, in response to JA treatment (Petersen *et al.*, 2000). The transcription factors *ERF1*, *RAP2.6* (AP2 family), and *JIN1* (AtMYC2) are all inducible by JA. *ERF1* integrates signals from JA and ET at one hand and at the other, its expression requires both the hormones in addition to *COI1* and *EIN2*. Overexpression of *ERF1* results in activation of many defense-related genes and suppresses the defects in defense-related gene expression caused by *ein2* and *coi*. These results suggest that *ERF1* acts downstream from *COI1* to activate gene expression (Glazebrook, 2005).

JIN1 and *ERF1* regulate distinct sets of JA inducible genes. For genes that have been tested so far, those induced by *ERF1* are repressed by *JIN1*. A second group of genes is activated by *JIN1* and repressed by *ERF1*. The first group of genes seem to be important for disease resistance, as *ERF1* overexpression and *jin1* mutations enhance resistance to necrotrophic pathogens (Lorenzo *et al.*, 2004).

3.3 Hormone signaling pathways crosstalk to fine-tune defense

Plants often deal with simultaneous or subsequent invasion by multiple aggressors and beneficials which can influence the primary induced defense responses of the host plant. Activation of plant defense mechanism is associated with ecological fitness costs and hence plants need regulatory mechanisms to effectively and efficiently adapt to changes in their complex environment. Crosstalk between hormonal signaling pathways provides the plant with a powerful regulatory potential and allows the plant to tailor its defense response to the invaders encountered (Pieterse *et al.*, 2009). Global expression-profiling studies provided ample evidence that SA, JA, and ET pathways interact, either positively or negatively (Glazebrook *et al.*, 2003). Studies have demonstrated mutually antagonistic interaction between SA and JA dependent signaling, however, synergistic interactions may occur as well. For example in *Arabidopsis*, treatment with low concentrations of JA and SA resulted in a synergistic effect on the JA and SA responsive genes *PDF1.2* and *PR-1*, respectively. However, at higher concentrations the effects were antagonistic, demonstrating that the outcome of the SA-JA interaction is dependent on the relative concentration of each hormone (Mur *et al.*, 2006). As a result of negative crosstalk between SA and JA, activation of the SA response should render a plant more susceptible to attackers that are resisted via JA dependent defenses and vice versa. Many trade-offs have been reported between SA dependent resistance against biotrophic pathogens and JA dependent defense against insect herbivory and necrotrophic pathogens (Bostock, 2005). Besides SA/JA cross talk, interactions between SA and ET, JA and ABA, and JA and ET have been shown to function in the adaptive response of plants to herbivores and pathogens with different lifestyles (Figure 1.4). A number of key signaling nodes came to the light from various studies and were shown to play role in hormonal crosstalk.
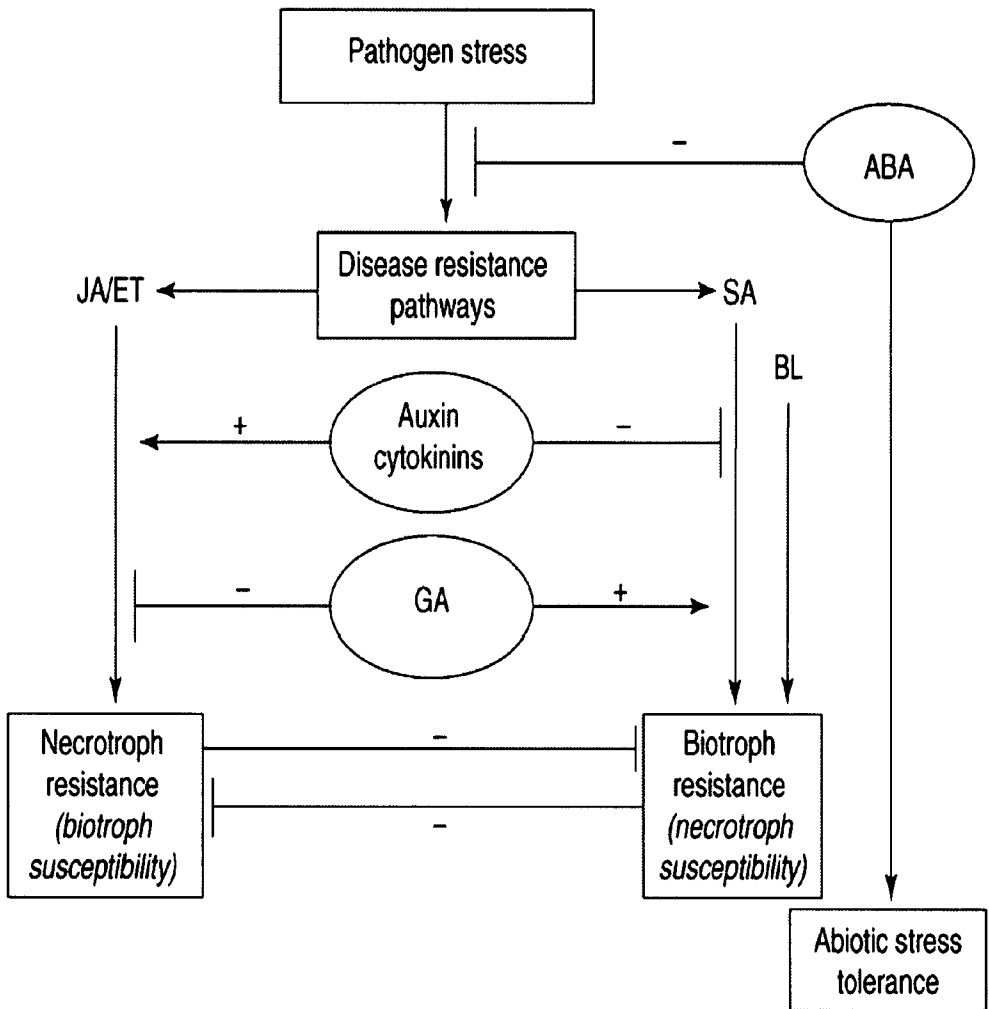
Figure 1.4: Model depicting interactions between various hormone signaling pathways.

Source: Robert-Seilaniantz *et al. Current Opinion in Plant Biology,* 2007, 10: 372–379

3.3.1 SA-JA

Several proteins with important regulatory roles in SA-JA crosstalk have been identified which include MPK4, EDS1, PAD4, NPR1, glutaredoxin (GRX480) and WRKY70 (Figure 1.5). The majority of the identified crosstalk regulators play pivotal roles in SA signal transduction in which NPR1 plays a central role. SA induced redox changes activate NPR1 by reducing inactive NPR1 oligomers to active monomers which are in turn translocated to nucleus where they interact with TGA transcription factors that activate SA responsive genes. But the nuclear localization of SA activated NPR1 is not required for suppression of JA responsive genes which indicates that the antagonistic effect of SA on JA signaling is modulated through a function of NPR1 in cytosol (Pieterse *et al.*, 2009). WRKY transcription factors are important regulators of SA-dependent defense responses and some of them have been implicated in SA/JA cross talk. *Arabidopsis* WRKY70 was identified as a node of convergence between SA and JA signaling when Li *et al.* (2004) showed that overexpression of WRKY70 caused enhanced expression of SA responsive *PR* genes and concomitantly suppressed methyl jasmonate induced expression of the JA-responsive marker gene *PDF1.2*. Hence, WRKY70 acts as a positive regulator of SA mediated defenses while repressing the JA response. Besides WRKY70, WRKY11 and WRKY17 of *Arabidopsis* have also been implicated in SA/JA cross talk. Glutaredoxin (GRX480) is a putative regulator in SA/JA cross talk. Overexpression of GRX480 was found to completely abolish MeJA induced *PDF1.2* expression, but did not affect the induction of the JA-responsive genes *LOX2* and *VSP2*. This suggests that GRX480 affects only a subset of the JA-responsive genes that are sensitive to SA-mediated suppression. The suppressive effect of GRX480 on PDF1.2 induction was abolished in *the tga2 tga5 tga6* triple mutant, indicating that the interaction between GRX480 and TGA transcription factors is essential for GRX480 dependent cross talk (Ndamukong *et al.*, 2007). These results suggest that SA activated NPR1 induces GRX480, which in turn interacts with TGA transcription factors to suppress JA responsive gene induction. MAP KINASE4 (MPK4) was identified as a negative regulator of SA signaling and a positive regulator of JA signaling in *Arabidopsis*. Inactivation of MPK4 in mutant *mpk4* plants resulted in elevated SA levels followed by constitutive expression of SA responsive *PR* genes, suppression of JA responsive genes, and enhanced susceptibility to the necrotroph *A. brassicicola* (Petersen *et al.*, 2000). Interestingly, the *mpk4*
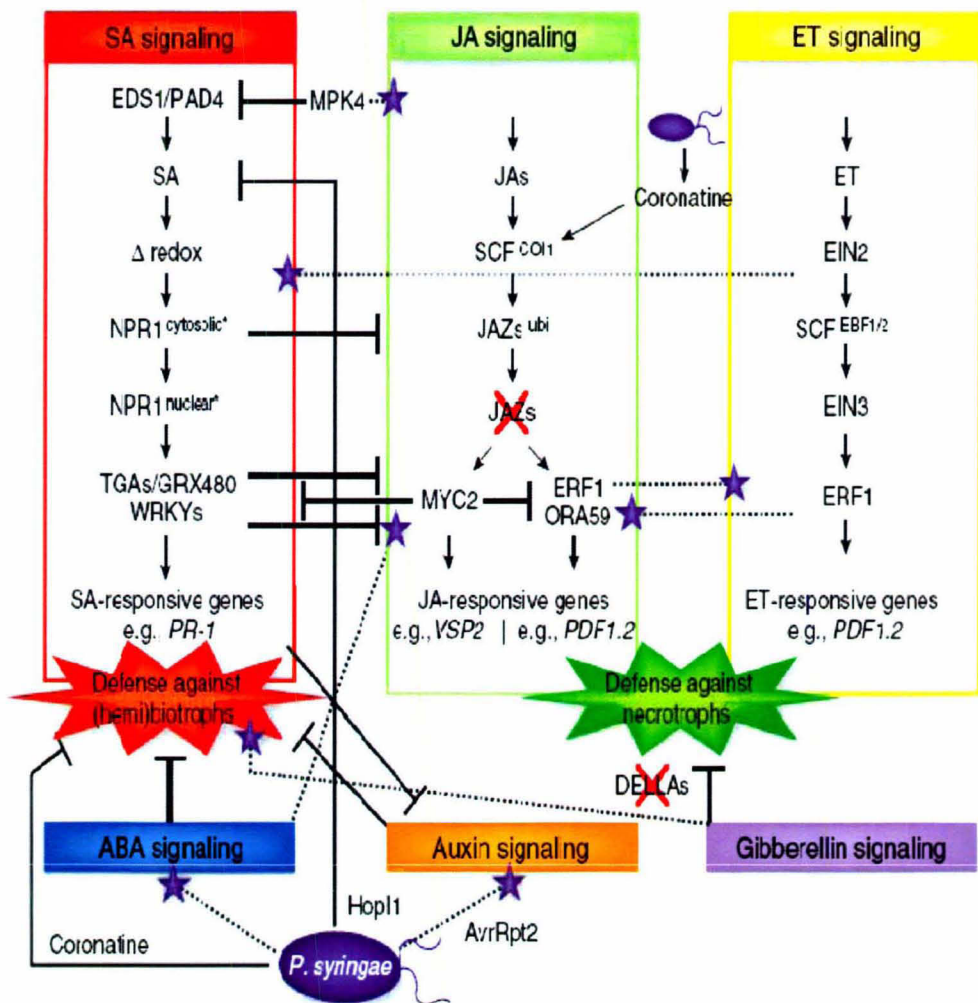
Figure 1.5: Networking by phytohormones in plant immune response.

Source: Pieterse *et al. Nature chemical biology*, 2009, 5: 308-316

mutation blocked JA-responsive gene expression independently of SA accumulation, as SA nonaccumulating *mpk4*/NahG transgenics still exhibited increased susceptibility to *A. brassicicola* and suppression of MeJA induced PDF1.2 expression (Petersen *et al.*, 2000). EDS1 and PAD4 were identified as downstream effectors of MPK4 function, having the opposite effect of MPK4 by behaving as activators of SA signaling and repressors of JA signaling (Brodersen *et al.*, 2006).

3.3.2 JA-ET

The interaction between JA and ET signaling has been found to be synergistic in most of the cases. For example, the regulation of the *Arabidopsis* plant defensin gene *PDF1.2* requires concomitant activation of the JA and ET response pathways (Penninckx *et al.*, 1998). ERFs are members of the large plant-specific AP2/ERF superfamily of transcription factors and two members of this superfamily, ERF1 and octadecanoid responsive Arabidopsis gene59 (*ORA59*), emerged as principal integrators of the JA and ET signaling pathways. The expression of both *ERF1* and *ORA59* is induced by JA and ET and can be activated synergistically by both hormones. Also the basic helix-loop-helix leucine zipper transcription factor MYC2 (originally called JIN1, for JASMONATEINSENSITIVE1) has been demonstrated to play an important role in the regulation of JA responsive genes (Figure 5). MYC2 functions as a positive regulator of JA responsive genes such as *VSP2* and *LOX2*, whereas it acts as a negative regulator of JA/ET responsive genes such as *PDF1.2* that are activated by ERFs. Hence, when the JA response is activated in combination with ET, the ERF branch of the JA response is activated, while the MYC2 branch of the JA response is activated when ET is absent. ABA was shown to play a role in favouring the MYC2 dependent branch of the JA response. Hence, the interplay between the ERFs and MYC2 may allow the plant to activate the set of JA responsive genes that is required for optimal defense against the attacker encountered. Moreover, MYC2 is involved in mediating the suppression of SA dependent defenses by coronatine, a phytotoxic virulence factor of *P. syringae* that mimics the action of JAs which makes this transcription factor an important node in the SA-JA-ET signaling network (Pieterse *et al.*, 2009).

3.3.3. ET-SA

ET has been demonstrated to be an important modulator of the plant's defense response to pathogen and insect attack (Van Loon *et al.*, 2006). ET was found to be essential for the onset of SA dependent SAR that is triggered upon infection by tobacco mosaic virus. Moreover, ET was shown to enhance the response of *Arabidopsis* to SA, resulting in a potentiated expression of the SA-responsive marker gene *PR-1*. This synergistic effect of ET on SA induced *PR-1* expression was blocked in the ET-insensitive mutant *ein2* which indicates that the modulation of the SA pathway by ET is EIN2 dependent and thus functions through the ET signaling pathway (Figure 1.5). The study of global expression profiles of *P. syringae* infected *Arabidopsis* wild-type and signalling defective mutant plants showed extensive crosstalk between the SA and ET signaling pathways, as evidenced by the fact that the expression of many SA-responsive genes was significantly affected in the *ein2* mutant background (Pieterse *et al.*, 2009).

3.4 Other hormone pathways connected to the SA-JA-ET backbone

The SA, JA and ET responsive pathways serve as the backbone of the induced defense signaling network, however, many other hormones also feed into this backbone (Figure 1.4, Figure 1.5). For example, ABA is commonly associated with plant development and abiotic stress, but in recent years its role in biotic stress has also become evident (Asselbergh *et al.*, 2008). ABA was shown to attenuate JA/ET dependent gene expression and to affect JA biosynthesis and resistance against JA inducing necrotrophic pathogens (Adie *et al.*, 2007). Moreover, ABA was demonstrated to antagonize the onset of SA dependent defenses and SAR (Yasuda *et al.*, 2008). Conversely, activation of SAR suppresses the expression of ABA related genes, which indicates that ABA serves as an important regulator that functions at the crossroad of abiotic and biotic stress responses.

Auxins play a role in virtually every stage of plant development and the auxin response pathway is connected to the SA-JA-ET signaling network in different ways. For instance, auxin has been demonstrated to affect JA biosynthesis and the expression of genes involved in JA production (Liu and Wang, 2006). Whole-genome expression profiling revealed that SA interferes with auxin responses by global repression of auxin-related genes, including the auxin receptor gene *TIR1*. The

inhibitory effect of SA on auxin responses stimulated effective defenses against the (hemi) biotrophic pathogens like *P. syringae*, resulting in heightened resistance to these pathogens (Wang *et al.*, 2007). Hence, the antagonistic effect of SA on auxin signaling seems to be an intrinsic part of SA-dependent resistance against (hemi)biotrophs.

Recently, gibberellins were shown to hook up to the SA-JA-ET network as well. Gibberellins are hormones that control plant growth by regulating the degradation of growth-repressing DELLA proteins. Navarro *et al.* (2008) demonstrated that DELLA proteins promote susceptibility to biotrophic pathogens and resistance to necrotrophic pathogens by modulating the relative strength of the SA and JA signaling pathways. Hence, it was postulated that by regulating the stability of DELLA proteins, gibberellins are able to modulate the SA-JA-ET network and affect the final outcome of the immune response.

Brassinosteroids play a key role in cell expansion and division, differentiation and reproductive development. Brassinosteroids are perceived by the receptor BRI1, which interacts with the receptor-like kinase BAK1 to initiate an intracellular signaling cascade that regulates growth and development related processes (Belkhadir and Chory, 2006). Interestingly, BAK1 also interacts with receptors that recognize PAMPs, such as bacterial flagellin, resulting in the initiation of innate immunity (Chinchilla *et al.*, 2007).

Cytokinins often work in concert with auxins in processes such as cell division and differentiation of plant tissues. They are linked to the response of plants to biotrophic pathogens that alter the host's physiology, such as *Plasmodiophora brassicae*, which causes aberrant root growth (club roots) in *Brassica* species (Siemens *et al.*, 2006). However, little is known about their connection with the SA-JA-ET network.

Past efforts to improve plant defense through breeding and mutant analysis have limited success owing to the genetic complexity of stress responses. Moreover, these methods are time taking. As the global population is increasing, science needs to move at a much faster rate. Therefore, it is the time to adapt faster methods of investigating plant genomes and their responses to various stresses. Genomics is one such tool and can prove to be useful in the modern era of research.

4. Need to pave way to genomics

Earlier knowledge about various cellular processes including stress was gained by working on individual genes in context of a particular process. But recent developments have shown that cellular processes are controlled by highly connected gene networks. Therefore, the cellular function of an individual gene cannot be understood at the level of isolated components alone, but needs to be studied in the context of its interplay with other gene products (Dittrich *et al.*, 2008). Also existence of cross talk between various processes and pathways has been revealed. For example many biotic and abiotic stress pathways have been shown to have some overlap. Further, many development related genes are shown to play role in other cellular processes like stress responses (Chung *et al.*, 2008). Therefore, in order to gain a broader understanding of biological processes, we cannot study genes in isolation from the larger context of other genes. We have to move beyond the single gene approach and pave way to the study of genome of organisms as a whole. This will allow developing a wide picture of gene characteristics. Further, more needs to be achieved in less period of time. Food and health related problems are appealing the scientific community to do better and come up with more comprehensive understanding of biology and provide solutions to the problems. Here genomics comes to our rescue as it has the potential to address many such problems.

In recent years genomics has fundamentally changed our ability to study the molecular basis of cells and tissues in health and diseases, giving a new comprehensive view. Genomics provides essential tools to fully understand the molecular basis of various agronomic traits and to manipulate them for human benefits (Harlizius *et al.*, 2004). The field of genomics allows the analysis of thousands of genes in parallel to understand the genetic architecture of genomes. Recent technological advances have changed the paradigm enabling the analysis of organisms in terms of genome organization, expression and interaction (Hocquette, 2005). The application of genomics has led to and will further extend substantial and rapid advances in our understanding of the molecular basis of various processes including stresses (Mathers, 2004). Genomics tools are essential to fully understand the synthesis of compounds that are at the basis of research community's focus on improvement of food and feed (Harlizius *et al.*, 2004). Such approaches may assist in illuminating the mechanism as it enables the simultaneous discovery and study of

many biological processes and genes involved in such processes. Another importance of genomics is that it helps to capture the structure-function relationships of genes. Also, the gene and protein expression studies, as well as metabolomics, will provide the link between the relatively static genome and the highly dynamic physiological processes.

ESTs and microarray are the two essential genomics tools and help immensely in discovering cellular processes. These genomic tools have helped in revolutionizing the biology by providing vast catalogues of gene sequences and the expression pattern of such genes. These tools have and will prove of immense use for further uncovering the cellular processes in diverse organisms including plants.

4.1 Expressed sequence tags (ESTs)

The manner in which various biological questions can be addressed has profoundly evolved in the last few years with the advent of genomics. The possibility to conduct large scale analyses in functional genomics opens way to the identification of large sets of coregulated genes involved in various biological processes. This is interesting to identify novel and possibly important molecular events, and also to investigate biological processes at the level of gene networks for a better global understanding of complex developmental programs by cataloguing the genes involved in such processes. One such approach involves the generation of ESTs.

ESTs are short cDNA sequences that serve to "tag" the gene from which the mRNA originated and that can serve multiple important uses. The EST collections for various organisms have been generated with the purpose of studying the genes underlying various functions and coding for various developmental plans and stress responses (White *et al.*, 2000; Ewing *et al.*, 1999; Jantasuriyarat *et al.*, 2005). EST data are generated by large-scale, single-pass, partial sequencing of cDNA clones usually from a large number of libraries representing diverse tissues, developmental stages or environmental conditions. ESTs have become an established method for rapidly developing gene databases (Adams *et al.*, 1995). Annotation of sequenced clones is made easier by the availability of rapidly growing sequence databases that allow for the detection of putative orthologs through which to assign potential function to cDNA clones. The value of EST data is not limited to the identification of orthologs of protein coding genes but can also be mined for non-coding sequences such as SSRs

micro RNAs (miRNAs) and small RNAs (si RNAs) (Ellis and Burke, 2007) and for gene expression data, which may be inferred based on the representation of any particular sequence in the data set (Ramirez *et al.*, 2005). By sequencing numbers of clones from cDNA libraries derived from RNA obtained from different source tissues, the total set of genes sampled from the genome can be maximized. Bioinformatic sorting and clustering of the resulting sequences yield databases of putative genes that form the basis of a functional genomics program. EST based information can be used for crop improvement in many ways as given below:

4.1.1 ETSs: towards gene inventory and function

ESTs obtained from different tissues and organs have been used as an effective method of gene discovery. cDNA libraries constructed from various tissues and the sequencing followed by annotation of the genes obtained thereof can lead to the discovery of genes associated with the specific tissues (Wellmer *et al.*, 2006). Also construction of subtracted libraries from plants subjected to various stresses has led to the identification of many stress related genes. There are many important agronomic traits for which no information is available regarding the genes controlling such traits. ESTs can help in identification of many such genes and hence pave way for improvement of such traits.

4.1.2 Generation of functional molecular markers

ESTs have facilitated the development of molecular markers from the transcribed regions of the genome. Among the important and popular molecular markers that can be developed from ESTs are single-nucleotide polymorphisms (SNPs) (Rafalski, 2002), simple sequence repeats (SSRs) (Varshney *et al.*, 2005) or conserved orthologous sets of markers (COSs) (Rudd *et al.*, 2005). Putative functions can be deduced for the markers derived from ESTs or genes using homology searches (BLASTX) with protein databases. Therefore, molecular markers generated from gene sequence data are known as 'functional markers' (FMs) (Gupta and Rustgi 2004). FMs have some advantages over random markers (RMs) that are generated from an anonymous region of the genome because they are completely linked to the desired trait allele. Such markers can be derived from the gene responsible for the trait of interest and target the functional polymorphism in the gene, thus allowing selection in different genetic backgrounds without revalidating the marker–quantitative-trait-locus

(QTL) allele relationship. These markers will be helpful to identify loci controlling important traits like disease resistance and quality and can be used for the development of improved cultivars.

4.1.3 Gene expression analysis

Comparisons of EST frequencies in different libraries can expose differential gene expression on a broad basis (Ewing *et al.*, 1999). In theory, the abundance of an EST is an exact digital representation of the number of copies of a transcript in the tissue. Large numbers of ESTs derived from diverse tissues produce quantitative estimates of gene expression. Theoretically, expression profiles could be derived for very weakly expressed genes if ESTs were sequenced in sufficient number. This has been performed with human EST libraries that contain tens of thousands of sequences (Adams *et al.*, 1995). In plants, Ewing *et al.* (1999) compared and analyzed 10 rice libraries containing between 1,000 and 5,000 ESTs and were able to identify statistically significant patterns of gene expression among several rice tissues. ESTs can also be used to assemble DNA arrays, a powerful tool for the analysis of the entire expression pattern of a tissue, a specific differentiation stage or mutants (DeRisi *et al.* 1997).

4.2 Microarray

The DNA microarray has produced a revolution in expression analysis. This technology has demonstrated the power of the high-throughput study of gene expression in unravelling key processes of biology (Brown and Botstein, 1999). Microarrays have become especially relevant for species where little genome information is available, and where intensive laboratory work is necessary to get insight into a particular biological process, as well as to identify candidate target genes for future. It is one of the most powerful techniques recently developed to bridge the gap between sequence information and functional genomics (Rensink and Buell, 2005). Two fundamentally different approaches are currently utilized in microarray fabrication. One involves deposition of minute quantities of DNA (mostly cDNA) on to the solid surface and in case of the other oligonucleotides are directly synthesized on the carrier.

4.2.1 Principle of the technology

Like other hybridization-based analysis methods in molecular biology, the specificity of microarray technology relies on the selective and differential hybridization of nucleic acids. It is based on a strategy where complex mixtures of labelled polynucleotides (such as cDNA derived from mRNA) are hybridized with large numbers of individual elements (e.g. unique PCR products in cDNA microarrays), attached to a solid surface (Aharoni and Vorst, 2001). In this way information on the abundance of many polynucleotide species is gained in parallel. Labelling with fluorescent dyes possessing different excitation and emission characteristics allows the simultaneous hybridisation of two samples on a single array. The strength of fluorescence emission at the two wavelengths represents the amount of a polynucleotide from each sample bound to the array. In this manner a single experiment provides quantitative hybridization data for hundreds to thousands of probes. For expression studies using microarrays, this approach of combining two differently labelled samples (reference and test sample) is a common practice. For each gene the corresponding amount of signal in both samples can then be quantified in parallel and expression ratios obtained (Richmond and Somerville, 2000). This strategy, to use expression ratios instead of absolute expression levels, for the analysis of changes in gene expression, has been shown to be a very powerful one and has helped to overcome a large source of experimental variation.

4.2.2 Applications of microarrays

Microarray technology is the most important tool for gene expression studies. It provides a global overview of biological mechanisms underlying various cellular processes. This technique finds immense use in biology in a number of ways as discussed below:

TH-17220

4.2.2.1 Functional discovery through microarray expression profiling

Access to a large collection of reference expression data from mutants, tissues or treatments can provide a tool for identifying the function of unknown genes. This assumption is based on numerous examples in which gene function was tightly connected to precise expression patterns under certain conditions. This concept of functional discovery via a compendium of expression profiles was first demonstrated in yeast (Hughes *et al.*, 2000). With the growing availability of microarray expression

data, similar approaches could also be used to analyze many other organisms including plants. A recent study used publicly available *Arabidopsis* arrays to identify genes required for cellulose synthesis through the analysis of co-regulation with cellulose synthase gene (Persson *et al.*, 2005). This illustrates the potential of mining many expression profiles for functional discovery via co-regulation. Microarray was used to study expression profile in ovarian cancer in order to gain an insight into the catalogue of genes and their expression pattern induced in response to this disease (Fehrmann *et al.*, 2007). It has also been used in plant systems in order to study transcriptional programs regulating various processes. Many studies were conducted in model plants like *Arabidopsis* to study transcript profile in response to stresses like oxidative stress (Desikan *et al.*, 2001); salt stress (Jiang and Deyholos, 2006); pathogen stress (Schenk *et al.*, 2000) etc.

4.2.2.2 Identification of cell-specific phenotypes

Specialized cells can differ dramatically in their morphology and function and discovering the function of genes that contribute to cell specialization is one goal of transcriptional profiling at the cell level. Global cell-profiling studies have shown that several hundred transcripts are unique to specialized cells, or at least highly localized to a few widely dispersed cell types. For example, a developmental profile of maturing *Arabidopsis* pollen showed that about 650 to 850 transcripts are specifically expressed at various stages of pollen (Honys and Twell, 2004). Similarly *Arabidopsis* root transcript analysis showed that about 200–400 transcripts are uniquely expressed in only one of the five tissues profiled in the root (Birnbaum *et al.*, 2003). Thus, there are dramatic differences in the number of cell-specific genes and function in bringing out cell type specificity.

4.2.2.3 Study of developmental programs

The developmental plans in plants are controlled by a repertoire of genes and their spatio-temporal expression. Microarray technology plays a great role in understanding the transcriptional control over various developmental processes. For example it helped in identification of distinct metabolic switches controlling seed development in *Arabidopsis* (Fait *et al.*, 2006). In *Medicago*, it helped in understanding the role played by various genes in nodulation (Lohar *et al.*, 2006). Further, transcriptional profiling using microarray revealed dynamics and complexity in gene expression

during embryo (Spencer *et al.*, 2007) and flower development in *Arabidopsis* (Wellmer *et al.*, 2006).

4.2.2.4 Study and comparison of diverse genotypes

Natural variation in gene expression levels between closely related plant varieties can be treated as a genetic polymorphism. Microarrays can be used to describe patterns of gene expression among individuals in a mapping population. Each pattern constitutes a molecular phenotype. Transcript abundance levels differing in the parents of a mapping population and segregating among the progeny can be mapped and characterized as quantitative traits (Cheung and Spielman, 2002). These expression profiles may be more easily interpreted or quantified than some visible phenotypes.

4.2.2.5 Analysis of DNA sequence variation

Oligonucleotide arrays are well suited for the detailed analysis of DNA variation as they allow the detection of single nucleotide mismatches during hybridization. These analyses can include both the discovery of novel DNA variants and the determination of known variants. Identification of sequence variations, such as single nucleotide polymorphisms (SNPs), can serve as genetic markers (Fan *et al.*, 2000).

4.2.2.6 Measurement of allele-specific differences

Beyond measuring expression level differences among homozygous inbred lines, an additional challenge for gene expression technologies is to characterize and quantify subtle allele-specific differences in expression at heterozygous loci. For example, hybrid vigour is a well-characterized but poorly understood trait that is important to modern agriculture. One possible explanation for hybrid vigour is transgressive variation in expression. Expression differences for a particular allele in a hybrid compared with the parental lines result either from imprinting, a *cis* effect, or *trans* acting regulatory elements encoded in the two genomes. Imprinting is generally associated with monoallelic expression (Oakey and Beechey, 2002). The biallelic nonparental expression is indicative of trans-acting regulation of expression. To put it differently, the promoter and other adjacent regulatory elements for a given allele are identical in the F1 hybrid and parental lines and therefore any differences in expression for a specific allele between an inbred parent and the F1 hybrid must result from the interchromosomal effects in the hybrid. Similar intergenomic effects may

alter gene expression patterns in polyploids (Osborn *et al.*, 2003). Thus microarray technology can help in revealing the molecular basis of hybrid vigour and other allele specific variations.

4.2.2.7 Identification of promoter elements and their targets

Data from genome-wide gene expression studies provide a novel approach to the identification of new *cis*-regulatory elements in promoter regions and to the classification of genes in similar regulatory circuits according to the elements identified (Bucher, 1999). For example, an 'evening element motif' conferring rhythmic gene expression in *Arabidopsis* was identified by microarray analysis (Harmer *et al.*, 2000). Also, DNA microarrays were used in combination with chromatin immunoprecipitation (ChIP) methods, for the identification of *in vivo* binding sites for yeast transcription factors. This study led to identification of novel target genes of *GAL4*, which is one of the most investigated yeast transcription factors (Ren *et al.*, 2000; Iyer *et al.*, 2001). More recently ChIP on chip assays are also used for the identification of various transcription factors which bind to a particular DNA sequence (Qin *et al.*, 2009).

4.2.2.8 Study of localized hormone responses

Global gene expression profiling can help to reveal the mechanism by which plants mediate broad hormone fluxes to elicit local responses that trigger cell fate, patterning, differentiation and tropisms. The meristematic region of the root near the QC is also an auxin source (Ljung *et al.*, 2005). The root cell profiling data revealed the presence of several genes known to be involved in auxin biosynthesis, including *CYP79B2*, *CYP79B3*, *ASA1*, and *ASA2*. However, the global profile of the root QC, which is patterned by the auxin maxima, showed enrichment of both *SUPERROOT1 (SUR1)* and *SUR2*, which negatively regulate auxin biosynthesis (Nawy *et al.*, 2005). Thus, global cell-specific transcript analysis provides evidence for both positive and negative regulation of auxin in the root tip. Thus gene expression profiling helps to gain insight into the hormone control of various tissue patterns.

4.2.2.9 Unraveling stress responsive genes and related regulatory pathways

Employment of microarray technology has rapidly produced vast catalogues of gene expression activities. In addition, various computational methods developed recently

can help in elucidating the pathways and interaction networks based on microarray profiles underlying disease or stress responses. For example, microarray data along with protein-protein interaction data has been used for the construction of cancer related networks (Chu and Chen, 2008). Microarray can be employed in concert with other interaction databases to build the pathways controlling stress responses in plants. For example, redox regulation of *Arabidopsis* transcriptome was analysed using this approach (Khandelwal *et al.*, 2008). Microarray-based expression profiling methods, together with the availability of genomic and/or EST sequence data allow significant progress in the characterization of plant pathogenesis-related responses (Wan *et al.*, 2002). The expression profiling studies conducted till date have already identified a large number of genes that had not previously been implicated in plant defense. Potential transcriptional *cis*-regulatory elements upstream of co-regulated genes can also be identified.

Many efforts have been made to study plant pathogen interaction and the associated defense pathways using microarray gene expression profiling. For example, gene expression changes were examined in *Arabidopsis* under 14 different SAR-inducing or repressing conditions, including a notable focus on plant mutants (Maleck *et al.* 2000). This led to the identification of around 413 ESTs that appeared to be associated with SAR. Similarly, 705 genes were identified by microarray analysis that were responsive to the fungal pathogen *Alternaria brassicicola* or to the defense activating signaling molecules SA, methyl jasmonate (MeJ) or ethylene treatments in *Arabidopsis* (Schenk *et al.*, 2000). Also 41 jasmonate-responsive *Arabidopsis* genes were identified of which 5 genes were JA biosynthesis genes, 3 genes were involved in other signaling pathways (ethylene, auxin, and SA), while others had some known defense association, but most were functionally unknown genes (Sasaki *et al.*, 2001).

Using custom cDNA microarray representing 13,000 randomly chosen ESTs, changes in *Arabidopsis* transcript levels after attempted infection with an avirulent *P. syringae* strain were monitored at different time points. This led to the identification of a massive shift in the expression pattern of around 2,000 genes representing many cellular processes (Scheideler *et al.* 2002). Apart from *Arabidopsis*, microarray gene expression analysis was also used in many other plants to reveal the genes or the pathways underlying defense responses.

4.3 Next generation sequencing

Genomics has performed a great role in exploring the world of plant and animal biology. Researchers are trying to improve the genomics tools so as to get more and more benefits. For example, over the past few years, new DNA sequencing platforms referred to as next generation sequencing are becoming widely available. This technology is rapidly evolving and includes the development of robust protocols for generating sequencing libraries and building effective approaches to data analysis. The important applications of this technology include full genome sequencing, discovery of mutations, mapping of structural rearrangements, detection of copy number variation, chromosomal inversions, analysis of gene expression and genome wide mapping of DNA-protein interactions. Thus this technology has the potential to dramatically accelerate biological research by enabling the comprehensive analysis of genomes, transcriptomes and interactomes to become inexpensive, routine and widespread (Shendure and Ji, 2008).

5. Transcriptional regulation of plant immunity

Genes are coordinately expressed under tight regulation by transcriptional factors to carry out complex and condition-specific biological functions in living cells. Stress responsive gene expression is also regulated primarily at the level of transcription. Transcriptional control of the expression of stress-responsive genes is a crucial part of the plant response to a range of abiotic and biotic stresses. Plants devote a large portion of their genome capacity to transcription, with the *Arabidopsis* genome coding around 1500 transcription factors (Riechmann *et al.*, 2000). These transcription factors often belong to large gene families, some of which are unique to plants. Members of various classes of transcription factors have been found to play role in plant defense responses. For example, ethylene-responsive-element-binding factors (ERF), basic-domain leucine-zipper (bZIP), Myb and WRKY proteins were shown to exert influence on stress responsive gene expression (Singh *et al.*, 2002).

ERF proteins are a subfamily of the APETALA2 (AP2)/ethylene-responsive-element-binding protein (EREBP) transcription factor family that is unique to plants. ERF proteins share a conserved 58–59 amino-acid domain (the ERF domain) that can bind to two similar *cis*-elements: the GCC box, which is found in several *PR* gene promoters where it confers ethylene responsiveness, and the C-repeat

26

(CRT)/dehydration-responsive element (DRE) motif, which is involved in the expression of dehydration and low temperature-responsive genes. Transcript levels of several ERF family members were found to increase after ethylene treatment, wounding and defense (Suzuki *et al.*, 1998; Onate-Sanchez and Singh, 2002). Overexpression of *Arabidopsis ERF1* as well as the tomato ERF gene *Pti4* in *Arabidopsis* enhanced resistance to several biotrophic and necrotrophic pathogens (Gu *et al.*, 2002; Lorenzo *et al.*, 2003). Microarray analyses also suggested that ERF1 directly or indirectly regulates numerous known defense-associated genes.

bZIPs are a large family of transcription factors in plants with 75 members present in *Arabidopsis* (Jakoby *et al.*, 2002). One class of bZIP proteins that is linked to stress responses comprises the TGA/*octopine synthase* (*ocs*)-element-binding factor (OBF) proteins. These bind to the *activation sequence-1* (*as-1*)/*ocs* element, which regulates the expression of some stress-responsive genes such as the *PR-1* and *GLUTATHIONES-TRANSFERASE6* (*GST6*) genes. In *Arabidopsis*, there are seven members of the TGA/OBF family, which play roles in plant defense, xenobiotic stress responses and development. A major advance was the discovery that TGA/OBF family members interact with NPR1, a key component in the SA defense signaling. Individual TGA/OBF proteins vary in their DNA-binding specificity, protein–protein interaction properties and expression patterns (Eulgem, 2005). Also different TGA/OBF members may be involved in different other stress responses.

WRKY proteins are a novel family of transcription factors that are unique to plants and form a large family with 74 members in *Arabidopsis*. WRKY proteins contain either one or two WRKY domains, a 60-amino-acid region that contains the amino-acid sequence WRKYGQK and a zinc-finger-like-motif. Specific WRKY family members show enhanced expression and/or DNA-binding activity following induction by a range of pathogens, defense signals and wounding (Eulgem *et al.*, 2000). Some WRKY proteins are induced during leaf senescence consistent with other studies that demonstrate an overlap between the gene expression profiles seen during senescence and stress. WRKY proteins bind to the W-box, which is found in the promoters of many plant defense genes. W-box or W-box-like sequences often occur in clusters within short promoter stretches suggesting that WRKY proteins may act synergistically with other WRKY proteins and/or other classes of transcription factors. When arranged as multimers, W-box sequences are sufficient to respond to a range of

pathogens, elicitors and wounding (Singh *et al.*, 2002). WRKY proteins also regulate the expression of regulatory genes such as receptor protein kinases and *NPR1*.

Several genes encoding Myb transcription factors are up-regulated by *P. syringae* infections and other defense related stimuli (Stracke *et al.*, 2001). Some Myb factors can bind to promoters of defense-associated genes. A T-DNA insertion in the gene encoding the *Arabidopsis* Myb factor BOS1 resulted in enhanced disease symptoms after infection with several biotrophic and necrotrophic pathogens. Myb factors bind to a variety of different DNA sequences. Two separate consensus sequences of Myb binding sites have been derived. No enrichment of these motifs in promoters of defense-related genes has been reported yet. However, one Myb-binding motif (type I, GG/TTA/TGG/TT) is generally conserved in promoters of *WRKY* genes, which further supports the suggestion that Myb factors have roles in defense regulation (Eulgem, 2005).

6. Role of small RNAs in plant immunity

Small RNAs constitute a family of regulatory non-coding RNAs of 19-28 nt in length, which are derived from double-stranded RNAs (dsRNAs) and can induce gene silencing through specific base-pairing with the target molecules. Small RNAs are derived from dsRNAs through the processing mediated by RNase III type enzymes. Two relatively well-defined classes of small RNAs are involved in RNA silencing: microRNAs (miRNAs) and small interfering RNAs (siRNAs). miRNAs are generated from the dsRNA region of the hairpin-shaped precursors while siRNAs are derived from long dsRNAs. Small RNAs mediate gene silencing through at least four different mechanisms: (1) endonucleolytic cleavage of the cognate mRNAs (2) translational repression (3) transcriptional repression through the modification of DNA and/or histone and (4) DNA elimination through the modification of histone. When small RNA targets mRNA, silencing occurs at the post-transcriptional level, which is accordingly known as post-transcriptional gene silencing (PTGS) in plants (Bartel, 2004).

In animals small RNAs are involved in many cellular processes such as establishment, maintenance, and function of hematopoietic lineages and regulation of organogenesis. They are also involved in metabolic processes and diseases (Tili *et al.*, 2008). Extensive genome-wide expression profiling of cells and tissues in different stages of

development or differentiation, metabolic conditions, and disease have revealed a broad range of functions associated with small RNAs (Bartel, 2004). More recently small RNAs have been shown to have functions in plants also. Studies have shown that gene silencing by small RNAs is widely adopted in plant immunity. First illustration of gene silencing in plants was when Tobacco mosaic virus (TMV) containing a stretch of phytoene desaturase (PDS) was found to silence the transcription of PDS mRNA of plant (Carr *et al*. 1992). Recently, an endogenous siRNA, nat-siRNAATGB2, has been proven to contribute to RPS2-mediated disease resistance. It repressed PPRL, a putative negative regulator of the RPS2 resistance pathway (Katiyar-Agarwal *et al*. 2006). Numerous miRNAs have been discovered in *Arabidopsis* and rice. More recently, miRNAs have also been identified in other plant species (Lu *et al*., 2008). About 71 plant miRNA families have been identified so far. Numerous miRNAs have been predicted or validated to be involved in plant defense. For example, plants activate defenses after perceiving pathogen-associated molecular patterns (PAMPs) such as bacterial flagellin. In *Arabidopsis*, flg22 triggers rapid changes in transcript levels, including down-regulation of a gene subset, potentially by posttranscriptional mechanisms involving siRNAs and microRNAs. It was found that these small RNAs negatively regulate mRNAs for auxin receptors TIR, AFB2 and AFB3, thereby repressing auxin signaling. Since auxin signaling is known to cause susceptibility to biotrophic pathogens, its repression leads to resistance (Navarro *et al*., 2006). Small RNAs have also been implicated in defense responses through transcriptional modifications. For example, chromatin-associated, 24-nt siRNAs was found to induce transcriptional gene silencing by guiding DNA methylation and histone modifications. The bacterial pathogen Pst DC3000 was shown to induce DNA hypomethylation in the absence of DNA replication at certain genomic loci, including some peri/centromeric repeats and retrotransposons, in *Arabidopsis* (Pavet *et al*., 2006). Both virulent Pst DC3000 and avirulent Pst DC3000 (avrRpt2) strains, but not the non-pathogenic strain could induce chromatin alteration, such as decondensation of chromocenters, in the infected tissue. This led to the speculation that Pst DC3000-induced DNA hypomethylation and heterochromatin decondensation may promote increased DNA recombination and genetic variability at some genomic loci, including centromeric regions harboring a group of resistance genes (Pavet *et al*., 2006). Thus small RNAs in plants play important role in gene

expression reprogramming by transcriptional or post transcriptional gene silencing during plant disease or defense responses.

7. Post translational regulation by ubiquitination during plant defense

Ubiquitin-mediated proteolysis is a central regulatory mechanism in the control of several cellular processes in yeast and animals. Ubiquitination has also been implicated in a growing number of plant signaling pathways, including those mediating responses to hormones, light, sucrose, developmental cues and pathogens (Zeng *et al.*, 2006). Two distinct yet coordinate steps are involved in the degradation of a protein by the ubiquitin system: (i) covalently attaching multiple ubiquitin molecules to the target protein; (ii) degrading the modified protein by the 26S proteasome. The biochemical process of ubiquitination is operated by a multienzymatic system that consists of ubiquitin-activating (E1), conjugating (E2), and ligating (E3) enzymes. Ubiquitin-mediated proteolysis performs various roles, for example, by removing abnormal proteins the ubiquitin/26S proteasome system performs an essential housekeeping role in the cell. Importantly, the system also serves as control/regulation mechanism by eliminating normal proteins, such as rate-limiting enzymes and key regulators, in different pathways. In the past several years, many ubiquitination-related components have been identified that are involved in plant-pathogen or plant-insect interactions. These genes were identified from different plant species and most of them are induced after pathogen or elicitor challenge suggesting that they might be involved in the response to biotic stresses.

7.1 Members of the ubiquitination system implicated in plant defense signaling

E3 ubiquitin ligases are the most abundant members of the ubiquitination system and they are generally considered to be most important in controlling substrate specificity. Therefore, most of the components of the ubiquitination system that have been shown to play a role in the regulation of various pathways belong to the E3 group. Many classes of E3s have been implicated in plant defense. The *ACRE132* (Avr9/Cf9 rapidly elicited), *ACRE189,* and *ACRE276* genes which encode different types of E3 components were identified in Cf9 tobacco cell cultures treated with Avr9-containing intercellular fluid. Silencing of *ACRE189* and *ACRE276* dramatically reduced *Avr4/Cf4* or *Avr9/Cf9* induced HR suggesting that they are involved in *Cf4* and *Cf9* mediated cell death and/or resistance (Rowland *et al.*, 2005). F-box-type E3 ligases

have also been implicated in the regulation of plant defense responses (Sullivan *et al.*, 2003). For example, *Coronatine Insensitive1* (*COI1*) encodes a protein with amino-terminal F-box motif and a leucine-rich repeats (LRR) domain and is required for plant response to jasmonic acids (JAs), which regulate root growth, pollen fertility, wound healing, and defense against insects and pathogens (Devoto *et al.*, 2003). The *coi1* mutant fails to express the defense-related gene *plant defensin 1.2* (*PDF1.2*) and is susceptible to insect herbivory and to pathogens. Thus, it is believed that COI1 plays a positive role in the control of plant defense. U-box-containing proteins are newly identified E3 ubiquitin ligases that are present in yeast, plants, animals, and humans (Zeng *et al.*, 2006). To date, several U-box proteins have been shown to be involved in disease resistance, however, the mechanism remains elusive.

7.2 Other components of the ubiquitination system involved in plant defense signaling

Evidence on the involvement of other components in the ubiquitination system in plant defense signaling came from the genetic analysis of the SCF complexes in *Arabidopsis*. Several components that potentially regulate E3 ligase activity were implicated in plant defenses in these studies. The RUB (related to ubiquitin, also known as Nedd8) is a ubiquitin-like protein that covalently attaches to cullins, the core subunit of the SCF E3 ligase complex. Rub1/Nedd8 and consequently its conjugation pathway are required for the SCF E3 activity (Gray *et al.*, 2003). Mutation of *RCE1* (RUB-conjugating enzyme1) resulted in defects in the expression of defense-related genes like basic chitinase and *PDF1.2*, inferring that the SCF E3 complex is required for plant defense. The plant *SGT1* (*Suppressor of G2 allele of skp1*) gene is known to be required by a diverse range of *R* genes from different plant species in response to a variety of pathogens (Peart *et al.*, 2002). In yeast, SGT1 is a key regulatory protein in centromere function and cell cycle transition and is associated with SKP1 to regulate the activity of the SCF E3 ligase complex (Kitagawa *et al.*, 1999). Thus it is speculated that SGT1 functions in plant defense by regulating the activity of the yet unknown SCF E3 ligase that regulates plant defense signal transduction. This is partly supported by the observation that the *Arabidopsis SGT1* gene could complement yeast *sgt1* mutations and that SGT1 is associated with SCF subunits SKP1 and CUL1 in barley and tobacco cell extracts.

8. Necrotrophs- killers set free

Plant immune responses differ depending upon the type of pathogen. As biotrophic pathogens require a living host, localized controlled cell death in the region of pathogen attack forms part of an effective defense strategy known as the hypersensitive response (HR). Necrotrophic pathogens are distinguished from biotrophs by their ability to actively kill host tissue; therefore programmed cell death initiated by the plant is not intuitively an effective strategy to limit necrotrophic pathogen growth. Thus gene for gene resistance, SA-dependent responses and SAR are not predicted to play a role, whereas responses mediated by JA and ET are expected to be involved in defense responses against necrotrophs (Glazebrook, 2005). Till so far not much has been studied in terms of molecular events that come into play when necrotrophs attack a plant. Further, apart from *Aternaria brassicola* and *Botrytis cineria* (Glazebrook, 2005), no other necrotrophs have been studied in detail in spite of the fact that a wide range of necrotrophs including *Fusarium*, *Ascochyta*, *Sclerotinia* etc. destroy crop fields. Therefore in order to explore immune pathways other than those mediated by gene for gene resistance, and know more about the signalling cascades involved in necrotroph-host interaction, strategies should be framed to work on necrotrophic pathogens.

9. Need to shift the research from model plants to the crop plants

Genomic research has and will continue to revolutionize plant biology. The adoption of *Arabidopsis* as model species has done much to speed the development of plant genomics and to hasten our increased understanding of basic plant biology. However, *Arabidopsis* is not an "omniscient" model because this plant does not encompass all of the diverse physiological, developmental, and environmental processes seen throughout the plant kingdom. For example, unlike *Arabidopsis*, legumes develop important and interesting symbioses with nitrogen (N)-fixing rhizobia and with mycorrhizal fungi. They also exhibit interesting differences in secondary metabolism, pod development, and other processes that cannot be adequately modeled with *Arabidopsis*. Thus, to study these processes and to bring the genomic revolution to crop plants, there is a need to develop genomic resources in these plants. Over the past several years, this realization has led to the adoption of the model species concept to the study of other plant families like cereals and legumes.

9.1 Why chickpea??

The *Fabacease* or legume family is second to cereal crops in agricultural importance based on area harvested and total production and is a corner stone in the biological nitrogen cycle. Legumes are cultivated on 180 million hectares and account for about 27% of world's primary crop production (Graham and Vance, 2003). The grain legumes are cultivated primarily for their seeds which are rich in energy and protein. On a worldwide basis, legumes contribute about one third of humankind's protein intake, while also serving as an important source of fodder and forage for animals and of edible and industrial oils. Seeds of grain legumes contain at least 20% to 40% of protein. In many places of the world, legumes complement cereals or root crops, the primary source of carbohydrates. Legumes also provide essential minerals required by humans and produce health promoting secondary compounds that can protect against human cancers (Grusak, 2002) and protect the plant against the onslaught of pathogens and pests (Dixon *et al.*, 2002). Grain legumes have blood cholesterol-reducing effect as well hypoglycaemic effect, reducing the increase in blood glucose after a meal.

Legumes are divided into five tribes and one representative member from each tribe except the tribe *ciceri* has been used for the generation of EST as a first step towards genomic prospecting (Asamizu *et al.*, 2000; Cheung *et al.*, 2006). The tribe *ciceri* has a single genus *Cicer* which has remained as the understudied legume. Thus there is a need to explore this legume as well so as to gain an insight of the properties and functions of the left out *ciceri* tribe.

10. Chickpea (*Cicer arietinum* L.)

Chickpea is a cool season food legume belonging to the *Leguminosae* or *Fabaceae* family. This crop is grown on >10 million ha in 45 countries of the world. It is the second most important, rainfed legume, grown mainly by small farmers in the semi-arid tropics (SAT), West Asia and North African (WANA) regions. This crop is also grown in southern and eastern Africa (particularly in Ethiopia), Europe, the Americas, and more recently in Australia. Average annual chickpea cultivation area is >16,000 hectares in each of the 23 most important chickpea-growing countries. At present worldwide production of chickpea is 8 million tonnes (FAO, 2004). Chickpea probably originated in southeastern Turkey and spread to other parts of the world. The

two most common types of chickpea are the white-seeded "*Kabuli*" and the small dark-seeded "*Desi*". In traditional production systems of Asia and Africa, it was grown as an intercrop, but in recent years, it is mostly cultivated as a single crop. In low input traditional production systems, chickpea has been a preferred crop because of its minimal dependence on monetary inputs of N and P containing fertilizers, irrigation, and agrochemicals in general. It is known to confer sustainability to cropping systems.

Chickpea is an annual crop with plant height ranging between 30-70 cm, but tall varieties can be >1.0 m in height. The foliage is covered with glandular hair, which secretes highly acidic exudates, and is considered important in conferring tolerance to insect pests, such as the pod borer. Leaves are compound, arranged in an alternate phyllotaxy, and generally inparipinnate with 11 to 13 leaflets. Flowers are axillary, solitary, or in inflorescence of two or three and they are white, pink, purplish, or blue in color. The plant has a deep root system and is considered a hardy crop. It produces nodules in common with other legumes, and is efficient in fixing atmospheric nitrogen (N) in a plant-usable form through biological nitrogen fixation (BNF). The crop is highly efficient in uptake of phosphorus (P) from soils containing low amounts of available P. Chickpea is a self-pollinated crop and cross-pollination is rare. Optimum conditions for chickpea cultivation include 18-26°C day and 21-29°C night temperatures and annual rainfall of 600-1000 mm. It is generally grown on heavy black or red soils (pH 5.5-8.6). Although spoken of as "day-neutral," chickpea is a quantitative long-day plant, but flowers in every photoperiod (Duke, 1981).

Chickpea has one of the highest nutritional compositions of any dry edible legume and does not contain any specific major antinutritional factors. It is a good source of carbohydrates and proteins, together contributing about 80% of the total seed dry weight. The chickpea grains are rich in minerals and vitamins, and also form a good source of livestock feed It is also known for its use in herbal medicine and cosmetics. Chickpeas are mostly consumed as a mature pulse (cooked whole, dehulled or as flour), but are also served as a vegetable (immature shoots and seeds). Seeds contain about 20% protein, 5% fat and 55% carbohydrate (Roy et al, 2001) and represent a basic food crop in many developing countries, especially India, where they have a high economic value. Chickpea is a low input crop that often completes its lifecycle in drought and heat stress. The average yield of chickpea is about 0.8t/Ha but it has an

estimated yield potential of 5t/Ha. In India, the area under chickpea cultivation is around 8.4mHa and the around production around 6.33mT.

Chickpea is affected by various biotic and abiotic factors which limit the yield of this crop. Among the biotic constraints to increased and stable yields of chickpea are *Fusarium* wilt (*Fusarium oxysporum* f.sp. *ciceri*), *Ascochyta* blight (*Ascochyta rabiei*), dry root rot (*Rhizactonia bataticola*), *Botrytis* gray mold (*Botrytis cinerea*), *Helicoverpa* pod borer etc. Among the abiotic factors, drought stands to be the number one problem in major chickpea growing regions because the crop is grown on residual moisture and the crop is eventually exposed to terminal drought (Johansen *et al.*, 1994). In west Asia and North African countries, low temperature causing freezing injury or death or delayed onset of podding reduces yield tremendously (Singh, 1987). Heat and salinity problems are relatively important following drought and cold stresses.

11. *Fusarium* – a ubiquitous pathogen

The genus *Fusarium* includes numerous plant pathogens that cause destructive diseases on some of the world's most agriculturally important plant species, including corn, wheat, potato, cassava, palm, banana, pine, and numerous vegetables. Some pathogenic *Fusarium* species also produce mycotoxins that pose health risks to humans and animals and that markedly reduce the value of the crops in which they occur. *Fusarium* mycotoxins span a structurally diverse array of compounds and are produced in infected plant tissue. They become a safety and economic concern when they accumulate in food or feeds to the levels toxic to ingesting organisms. *Fusarium* is distributed worldwide and is highly abundant in some environments. The wide distribution of *Fusarium* is likely due to their ability to colonize diverse plants and soils and to disperse widely (Brown *et al.*, 2006).

Fusarium species cause various diseases in animals including humans. For example, *Fusarium* mycotoxins are considered one of the five major mycotoxin groups affecting human health. *F. oxysporum* and *F. verticillioides* produce fumonisins that are cytotoxic to several mammalian cell lines. Their mode of action may involve competition with sphingosine in sphingolipid metabolism. The disruption of sphingolipid metabolism by fumonisins, as well as folate transport and neural tube development in embryo culture suggests that these mycotoxins play a role in diseases

affecting human neural tube development. Fumonisins are also implicated in esophageal cancer. Many *Fusarium* species produce toxins which upon inhalation lead to the development of debilitating diseases such as bronchial asthma, allergic alveolitis, allergic rhinitis, atopic conjunctivitis, organic dust toxic syndrome, and chronic fatigue-like syndrome. *Fusarium* species cause disseminated disease in severely immunocompromised patients. Symptoms of disseminated *Fusarium* infection include persistent fever refractory to antibiotics, skin lesions, and pneumonia. This is a highly fatal infection that merges fungemia with multiple organ injuries such as in the lung, liver, spleen, kidney, and heart. Mortality of immunocompromised patients having fusariosis ranges from 50% to 80% (De Lucca *et al.*, 2007).

Fusarium oxysporum (Fox) stands out in several ways. It has an apparent long history of predominant, perhaps exclusive, asexual reproduction (Gordon and Martyn, 1997). It invades roots and can cause wilt diseases through colonization of xylem tissue (Tjamos and Beckman, 1989), and it displays apparent gene-for-gene relationships with several hosts. Although *F. oxysporum* has been found to cause disease in a large number of plant species, the study of the molecular basis of the pathogenicity has necessarily involved a limited number of hosts, mainly tomato, melon, bean, banana, cotton, chickpea and *Arabidopsis* (Raabe *et al.*, 1981). Large scale gene expression studies were conducted on cotton wilt caused by *Fusarium oxysporum.*f.sp *vasinfectum* (Dowd *et al.*, 2004). The studies revealed that a large repertoire of genes showed differential expression in response to wilt in cotton. However, the molecular mechanism of this disease remains largely unknown. Very recently, in an effort to dissect the mechanism of disease development or defense, a detailed study was conducted on *Arabidopsis* defense response against *Fusarium* (Berrocal-Lobo and Molina, 2007). Six dominant resistance loci to *Fox f.sp. matthiolae* (*RFO*) were identified in the *Arabidopsis* Col-0 accession. Among these *RFO* loci, *RFO1* was the largest contributor controlling the resistance mediated by *RFO2*, *RFO4* and *RFO6* loci. *RFO1* encodes the cell wall-associated kinase-like 22 (WAK/WAKL), one of the 26 members of the *Arabidopsis* WAK/WAKL class, which belongs to the receptor-like kinase (RLK) protein family. Like the perception of a bacterial PAMP by the RLKs, FLS2 and EFR, ROF1 might be envisaged to play a role in the perception of a fungal PAMP (Berrocal-Lobo and Molina, 2007). In tomato, Six *I* loci (I for 'immunity to *Fusarium* wilt') were identified conferring resistance to different FOX races and some

of them have been found to encode resistance proteins of the NBS-LRR subclass (Simons *et al.*, 1998; Lim *et al.*, 2006). The results obtained in the analysis of the interaction between tomato, *Arabidopsis* and Fox illustrate the genetic complexity and variability of plant resistance to Fox that can be mediated either by recognition of elicitor/PAMP or effector/Avr proteins.

Analysis of the *Arabidopsis*–Fox interaction has lead to the identification of signaling pathways required for plant resistance to Fox, as well as key regulators of innate immunity against this type of vascular pathogen. Fox was shown to induce systemic acquired resistance (SAR) and pathogenesis-related proteins (PRs) in *Arabidopsis*, indicating that the SA pathway plays a role in plant resistance to Fox (Mauch-Mani and Slusarenko, 1994). Different mutant analyses conducted to explore the signaling pathways that come into play in *Arabidopsis* in response to *Fusarium* revealed that SA, ET and JA pathways influence the Fox disease outcome in *Arabidopsis* and these pathways interact in a positive way in the activation of *Arabidopsis* resistance to Fox. Similar cooperative effects have been described for *Arabidopsis* resistance to other pathogens, such as the necrotrophs *B. cinerea* and *P. cucumerina* or the vascular oomycete *P. irregulare* (Adie and Perez-Perez, 2007). ABA was shown to have negative effect on *Arabidopsis* resistance to Fox. Different sets of transcription factors have been implicated in the regulation of *Arabidopsis* resistance to Fox. One of these transcription factors is ATAF2, a member of NAC protein family. ATAF2 has been proposed to function as a repressor of Fox-inducible defense responses in *Arabidopsis* (Delessert *et al.*, 2005). Several ERFs have been implicated directly in the activation or inhibition of *Arabidopsis* defense response against Fox. Overexpression of ERF1, an integrator of ET and JA responses, enhanced resistance to Fox in *Arabidopsis* and also to necrotrophic fungi, such as *B. cinerea and P. cucumerina*. A similar function in resistance to Fox has been described for ERF14, as loss-of-function mutants in this gene showed increased susceptibility to Fox (Berrocal-Lobo and Molina, 2007).

11.1 *Fusarium* wilt of chickpea

Chickpea wilt caused by *Fusarium oxysporum* f. sp. *ciceri* is one of the major yield limiting factors in chickpea. The disease causes 10-90% yield losses annually in chickpea. Eight physiological races of the pathogen (0, 1A, 1B/ C, 2, 3, 4, 5 and 6) are reported so far whereas additional races are suspected from India. The distribution

pattern of these races in different parts of the world indicates regional specificity for their occurrence leading to the perception that *F. oxysporum* f. sp. *ciceri* evolved independently in different regions. Of the various races of the pathogen identified, four have been reported in India (Jiménez-Díaz *et al.*, 1993). Pathogen isolates also exhibit differences in disease symptoms. Races 0 and 1B/C cause yellowing syndrome whereas 1A, 2, 3, 4, 5 and 6 lead to wilting syndrome.

In general, *Fusarium* wilt first appears as slight vein clearing on the outer portion of the younger leaves, followed by epinasty (downward drooping) of the older leaves. At the seedling stage, plants infected by *F. oxysporum* may wilt and die soon after symptoms appear. In older plants, vein clearing and leaf epinasty are often followed by stunting, yellowing of the lower leaves, formation of adventitious roots, wilting of leaves and young stems, defoliation, marginal necrosis of remaining leaves, and finally death of the entire plant (Agrios, 1988). Browning of the vascular tissue is strong evidence of *Fusarium* wilt. In chickpea, wilt can be observed within 25 days of sowing into infected soil (Nene *et al* 1987). Affected seedlings show drooping of the leaves and are dull green in colour. Seedlings may collapse and lie flat on the ground and, when uprooted, may show uneven shrinkage around the collar at the base of the stem. The roots do not show any external rotting and look apparently healthy. When split vertically from the collar region downward, such roots show a brown discolouration of the internal tissues. Transverse sections of the infected root examined under the microscope show the presence of hyphae and spores of the fungus in the xylem (Agrios, 1988). Wilting may also occur in adult plants up until the reproductive and podding stage. Drooping of the petioles, rachis and leaflets in the upper part of the plant, together with the pale green colour of the foliage, are the most common symptoms. This is a diagnostic feature of *Fusarium* wilt. In certain chickpea cultivars typical symptoms may not develop. Instead, there is a yellowing and drying of the lower leaves, and a stunting of the plant. While the affected plant is alive the pathogen is confined to the vascular system and possibly a few surrounding cells. At plant death, the fungus moves to other tissues and sporulates at or near the plant surface. Plants grown from infected seed develop wilt faster than plants originating from clean seed.

Fusarium oxysporum ciceri belongs to the deuteromycetes. It is an imperfect fungus (one with no known sexual stage) and spreads by means of asexual spores. It produces

three types of asexual spores: microconidia, macroconidia, and chlamydospores (Agrios, 1988). Microconidia are one or two celled, and are most abundantly and frequently produced by the fungus under all conditions and within the vessels of infected plants. Macroconidia are three to five celled, gradually pointed and curved toward the ends. These spores are commonly found on the surface of plants killed by this pathogen. Chlamydospores are round, thick-walled spores, produced either terminally or intercalary on older mycelium or in macroconidia. These spores are either one or two celled (Agrios, 1988). The fungus can survive either as mycelium, or as any of its three different spore types (Agrios, 1988).

Healthy plants can become infected by *F. oxysporum* if the soil in which they are growing is contaminated with the fungus. The fungus can invade a plant either with its sporangial germ tube or mycelium by invading the plant's roots. The roots can be infected directly through the root tips, through wounds in the roots, or at the formation point of lateral roots (Agrios, 1988). Once inside the plant, the mycelium grows through the root cortex intercellulary. When the mycelium reaches the xylem, it invades the vessels through the xylem's pits. At this point, the mycelium remains in the vessels, where it usually advances upwards towards the stem and crown of the plant. As it grows the mycelium branches and produces microconidia, which are carried upward within the vessel by way of the plant's sap stream. When the microconidia germinate, the mycelium can penetrate the upper wall of the xylem vessel, enabling more microconidia to be produced in the next vessel. The fungus can also advance laterally as the mycelium penetrates the adjacent xylem vessels through the xylem pits (Agrios, 1988). Due to the growth of the fungus within the plant's vascular tissue, the plant's water supply is greatly affected. This lack of water induces the stomata to close, the leaves wilt, and the plant eventually dies. It is at this point that the fungus invades the plant's parenchymatous tissue, until it finally reaches the surface of the dead tissue, where it sporulates abundantly (Agrios, 1988). The resulting spores can then be used as new inoculum for further spread of the fungus.

12. Genomic prospecting of chickpea – so far and beyond

Chickpea like other legumes has a narrow genetic base, making it a difficult system for genetics and breeding studies. Thus genomics knowledge and tools can be used for comprehensive study of this important crop so as to provide an assurance for its future

progress and an insurance against unforeseen threats to the productivity. Various efforts have been made by different research groups to address various impediments to the chickpea production. Extensive marker based analysis was carried out for identifying genetically diverse chickpea germplasm with beneficial traits for use in crop improvement programs (Varshney *et al.*, 2007). The gene expression study using microarray was carried out in chickpea to identify the abiotic stress responsive genes (Mantri *et al.*, 2007). Also SuperSAGE, an improved version of serial analysis of gene expression technique was used to address the drought problem in chickpea (Molina *et al.*, 2008). Further, proteomics approach has been used to study the drought responses in chickpea (Bhushan *et al.*, 2007; Pandey et al, 2008). Among the biotic stresses, *Ascochyta* blight has been studied to a considerable extent. *Ascochyta* responsive chickpea ESTs were generated as a step towards this. Also expression profiling of chickpea genes differentially regulated during a resistance response to *Ascochyta rabiei* was conducted which led to the identification of several defense related genes (Coram and Pang, 2005). As far as the other very important disease of chickpea, *Fusarium* wilt, is concerned, not much has been done in this regard. cDNA-AFLF approach was used to identify the wilt responsive genes in chickpea. The study resulted in identification of 273 transcripts differentially expressed during wilt (Nimbalkar *et al.*, 2006).

Fusarium wilt being an important factor to limit chickpea production deserves greater attention. Since nothing much is known about the disease, the filed is left open to be explored. This led us to initiate our study on *Fusarium* wilt of chickpea. As discussed earlier in this section, understanding the immense potential of the genomics tools, we used EST and microarray tools to study this disease. Here we present first ever report on large scale gene expression profiling involved in *Fusarium* wilt of chickpea. We have also compared the response of susceptible and resistant genotypes of chickpea in response to wilt. Also few of the identified immune responsive genes were isolated and characterized. Future efforts would be to construct immune responsive regulatory network and cloning and characterization of candidate genes and their genetic manipulation for the development of resistant cultivars.

Development and analysis of Chickpea EST database

2.1 Introduction

Expressed sequence tags (ESTs) provide a very useful means of quickly accessing gene sequence information and bridge the gap between the genome sequence and gene function. ESTs provide an invaluable resource for analysis of gene expression associated with specific organs, growth conditions, developmental processes and responses to various environmental stresses (White *et al.*, 2000; Ewing *et al.*, 1999; Jantasuriyarat *et al.*, 2005). EST projects have been useful for intra and intergenomic comparisons, gene discovery and molecular marker generation, for example, SSRs and SNPs that can be further used for creating genetic maps (Morgante *et al.*, 2002; Rafalski., 2002). ESTs also find use in microarray development and transcript pattern characterization (Andrews *et al.*, 2000). One of the important advantages of ESTs is that the identification of putative genes by BLAST comparisons enables researchers to begin biological analyses prior to the completion of full genome sequence. EST based transcriptome analysis has become an indispensable source of information about gene expression. Previous EST projects have revealed the tremendous utility of this approach for gene discovery (Ramirez *et al.*, 2005; Jantasuriyarat, 2005; Lazo *et al.*, 2004; Udall, 2005). Large scale EST databases now offer new perspectives for understanding the molecular basis of important traits in plants of agricultural relevance. Gene mining of these databases, aided by techniques such as microarrays, can be used to select candidate genes that are implicated in a particular biological process or in any kind of stress (Andrews *et al.*, 2000; Laitinen, 2005). Furthermore, large-scale analysis in functional genomics opens up the avenue for the identification of molecular events and biological processes at the level of gene networks rather than the individual genes. This approach also helps in global understanding of complex developmental programs such as those activated during interactions between plants and the invading pathogens.

Legumes are valuable agricultural and commercial crops that serve as important nutrient sources for human diet and animal feed. Chickpea is the world's third most important legume and is grown in many drier regions of the world where it is an important protein rich food and an increasingly valuable traded commodity. It is also known for its use in herbal medicine and cosmetics. Currently chickpea is grown on about 10 mHa area worldwide and the global production exceeds 8 million tons (FAOSTAT). However, from past few years, there has been stagnancy in chickpea

production due to various biotic and abiotic stresses. Among biotic stresses, *Fusarium* wilt has been one of the major constraints in chickpea production (Haware *et al.*, 1990). During recent years, plant defense systems against the invading pathogens are being elucidated in numerous plant species. However, genome scale gene expression in response to defense signaling has not been studied much extensively in pulses in general and chickpea in particular. The annual wild *Cicer* species are known to possess unique sources of resistance to pests and diseases and tolerance to environmental stresses; however, there has been a limited utilization of these wild species by breeding programs due to interspecific crossing barriers (Croser *et al.*, 2003; Singh *et al.*, 1998). Previously, studies have been undertaken to elucidate the genetic basis of the disease resistance and tolerance to abiotic stresses and identify the differentially expressed genes. However, the overall coordinated defense response remains largely uncharacterized. A genomics approach may, therefore, assist in illuminating the chickpea resistance mechanism as it enables the simultaneous discovery and study of many biological processes and genes involved in such processes.

In the present study, we used large scale EST sequencing as a step towards studying molecular mechanism of *Fusarium* wilt in chickpea. Plasmid isolation from the two subtracted cDNA libraries which were earlier constructed in our laboratory (Ghai *et al.*, unpublished results) using mRNA isolated from *Fusarium* infected tissue of resistant and susceptible genotypes of chickpea was carried out. This was followed by sequencing, annotation, computational and comparative analyses of these clones. This study provides a genomic resource of chickpea, an understudied legume and also reveals information on the expression pattern of chickpea genes in response to vascular wilt

2.2 Material and methods

2.2.1 Plasmid isolation and quality check

The plasmids were isolated in a high throughput manner using Perfectprep Plasmid 96 Vac Direct bind kit (Eppendorf) following the protocol given below:

1. The plasmids were inoculated directly from the glycerol stock plates in 1.2ml of LB dispensed in the deep well culture plates (Nunc) and grown overnight.

2. The deepwell plates were centrifuged at 4000 rpm for 5 minutes for pelleting the bacterial culture.

3. The supernatant was poured off and 150µL of solution 1 (resuspension buffer) was added to all the wells using multi channel pipettes.

4. The plates were vortexed at maximum rpm for 5 minutes so as to resuspend all the bacterial pellelts.

5. 150µL of solution 2 (lysis buffer) was added to all the wells and the plates shaked at 800 rpm for 2 minutes.

6. 150µL of solution 3 (neutralization buffer) was added and the plates again shaked at 800 rpm for 2 min.

7. Filter plate A was placed in top and filter plate DB in bottom of a vacuum manifold in such a way that filter plate A filter tips fit into the filter plate DB wells.

8. Whole of the 450µL of deepwell culture plate contents were pipetted into the respective wells of filter plate A and the plate was sealed.

9. Vacuum was applied at -20 mm Hg for 5 minutes or until all the contents passed from filter plate A into the filter plate DB.

10. Then the filter plate DB was moved to the top of the manifold and 300µL of DNA binding buffer added into filter plate DB and vacuum at -20 mm Hg was again applied for 5 min.

11. 400µL of purification buffer was added into filter plate DB and vacuum applied at -20 mm Hg for 5 minutes.

12. The bottom of the filter plate DB was blotted on a paper towel. A fresh collection plate was placed in the manifold and filter plate DB placed over it. 50µl of sterile MQ water was added to the wells of DB plate.

13. Vacuum was applied for 5 minutes to elute the plasmid DNA and collection plate was recovered from the manifold. The eluted plasmid DNA was stored at -20C.

The quality of the plasmid DNA was checked by agarose gel electrophoresis as mentioned in the section 8 of appendix III.

2.2.2 Sequencing, processing and assembly of ESTs into contigs.

The individual plasmids were sequenced using the BigDye terminator cycle sequencing kit (Perkin Elmer, Applied Biosystems) with M13 forward and reverse primers for 5′ and 3′ single pass sequencing respectively, in ABI Prism 3700 sequencer (Applied biosystems, CA). Sequence base calls were made using Phred (Ewing *et al.*, 1998) with a quality cutoff of 15. Vector filtering was performed using the cross match program (P.Green, http://bozeman.mbt.Washington. edu/phrap.docs/phrap.html) followed by manual trimming of low quality sequences. The sequences were individually inspected for chimeras, short reads, *E.coli* and mitochondrial sequences which were subsequently removed. The processed ESTs with 100 bases or longer were assembled into contigs by CAP3 programme using standard parameters (Huang and Madan, 1999). The final assembly of contigs and singletons constituted the chickpea gene index.

2.2.3 Blast analysis and annotation

The EST contig and singleton sequences were annotated for homology using BLASTX and BLASTN algorithms against non redundant protein and nucleotide databases respectively. For BLASTX, an e-value cut-off of 10^{-15} was used and the sequences with e-value below this cutoff were then subjected to BLASTN analysis with e-value cutoff of 10^{-20}. In addition, BLASTN was used to compare the chickpea sequences from this study to a database of legume sequences. This database included sequences of *Lotus japonicus* (gene index release 3.0); *Medicago trunculata* (gene index release 8.0); *Glycine max* (gene index release 12.0) and *Phaseolus vulgaris* (gene index release 1.0) from The Institute for genomic Research (TIGR; Quackenbush *et al.*, 2001) and *Arachis hypogea, Cajanus cajan, Pisum sativum* and *Robinia pseudoacacia* available from NCBI taxonomy browser [http://www.ncbi. nlm.nih.gov/Taxonomy/taxonomyhome.html]. The following criteria were used in stand-alone BLASTN comparison: (1) exact match bp = 11; (2) e-value $\leq 10^{-5}$ and (3) DNA identity $\geq 80\%$ and 90%. Further, TBLASTX [with e-value cutoff of 10^{-15}] was used for comparing chickpea sequences to GenBank's EST_others database for the identification of chickpea specific sequences. Sequences were additionally functionally characterized in the context with the Swiss Prot 2 Gene Ontology, Metacyc and COG.

2.2.4 Identification of gene families using single-linkage clustering

In order to identify gene families, the chickpea contigs and singletons were combined into a single dataset. TBLASTX with e-value cutoff of 10^{-15} was used to compare the dataset against itself. Sequences with at least one sequence in common in their BLAST reports were combined into a putative gene family as described in Graham *et al.*, 2004.

2.2.5 Identification of SNPs

For the identification of SNPs, the ace file output of CAP3 programme was used as input to the PolyBayes SNP detection program along with the base values assigned by Phred for each of the contigged sequences. Perl scripts were used to parse the PolyBayes output file. The cutoff value for the probability of SNP was put at 0.99. Only base change mutations were considered in order to avoid any discrepancy in the results due to any error in the alignment. The SNPs were further divided into high-quality if minimum of two sequences from each genotype showed the same base change and low-quality if minimum of two sequences from one genotype and one from the other showed the same base change.

2.3 Results and discussion

2.3.1 Development of Chickpea EST database

In order to construct a functional EST dataset for discovering wilt responsive genes expressed in chickpea, two suppression-subtracted cDNA libraries, one from vascular wilt susceptible genotype (JG-62) and the other from resistant genotype (WR-315) were present in the lab. The source of RNA for each library was root and collar tissue from the 25-d-old chickpea seedlings challenged with wilt pathogen, *Fusarium oxysporum ciceri* race 1, vs control tissue.

The present study was started with plasmid isolation and sequencing from two thousand subtracted clones from susceptible (JG-62) and one thousand clones from resistant (WR-315) library. The quality of the plasmids was checked by agarose gel electrophoresis and the high quality plasmid DNA for all the clones was given for sequencing (Figure 2.1). In parallel to the work done by other lab members on this project, in total, 6955 ESTs were generated and sequenced from the two libraries out of which 2908 were from susceptible and 4047 from resistant genotypes. Further
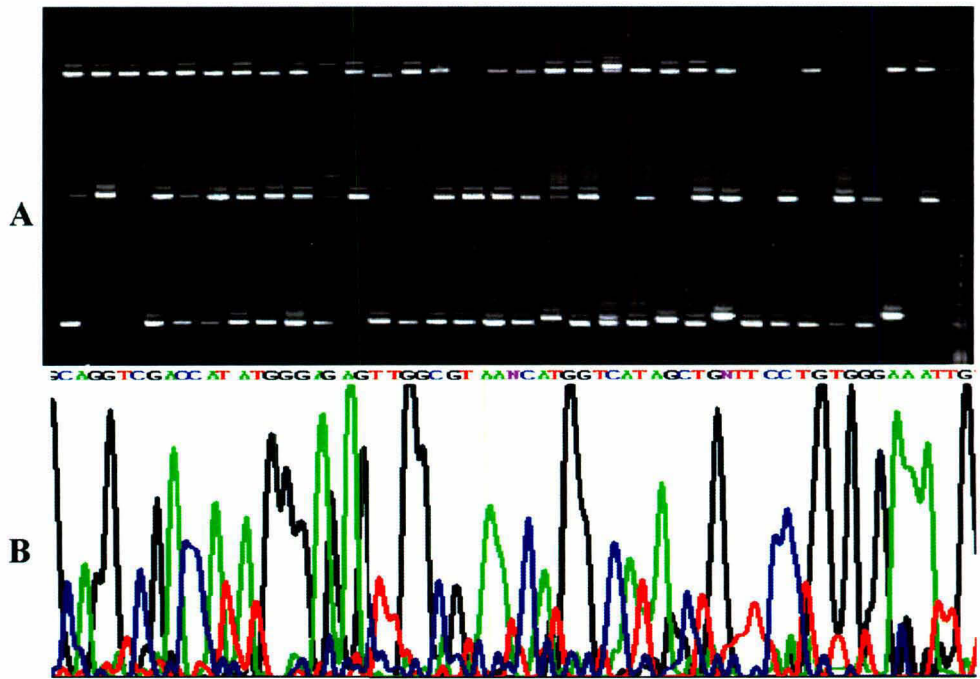
Figure 2.1: (A) 0.8% agarose/EtBr gel showing plasmid DNA (B) chromatogram as obtained from the sequencer.

work was proceeded with the whole set of 6955 sequences which was analyzed for chimeric sequences, low quality sequences, short sequence length, sequences from other organellar origin or the absence of the insert. On the basis of this analysis, 9.8% of the sequences were discarded and the remaining high quality 6272 ESTs were considered for sequence assembly (Table 2.1). The high quality ESTs were assembled into contigs using CAP3 program (Huang and Madan, 1999). Of the total 6272 ESTs, 1040 were found to be classified as singletons whereas 5232 assembled into 973 contigs (Table 2.2). The total contigs and singletons comprised a non redundant chickpea unigene (*Ca*Unigene) set of 2013 different transcripts. The sequences reported in this study have enriched the collection of ESTs in this food legume which would form a basis in our understanding of plant metabolism, development and adaptation to pathogen stress. The name of unigenes and the corresponding BLAST homology is given in table 2.3.

Of the total ESTs considered for assembly, 1202 were found to be specific to susceptible genotype whereas 2168 were specific to resistant genotype and 2902 were common to both (Figure 2.2). Thus although a large percentage of ESTs (46.27%) was common to both the susceptible and resistant genotypes, a significant percentage of ESTs was specific to each of them indicating the possible divergence in immune signaling pathways (Ashraf *et al.*, 2009).

2.3.2 Functional annotation and classification of chickpea ESTs

In order to understand the function of wilt-responsive chickpea ESTs, the entire set of 2013 *Ca*Unigenes was annotated on the basis of similarities to the known or putative ESTs in the NCBI database. Using the best hits found by BLAST, an inferred putative function was assigned to the sequences and were sorted into various functional categories as shown in figure 2.3. We were able to assign function to 60% of the genes while our data analysis revealed that 18.28 % of the *Ca*Unigenes belong to no significant homology (NSH) class. In addition, 18.33% and 3.47% of the unigenes matched with hypothetical and unknown proteins respectively. These ESTs may represent unique chickpea sequences or might suggest that the *Fusarium* infection of chickpea causes substantial changes in the expression of numerous unknown transcripts yet to be assigned a cellular function (Ashraf *et al.*, 2009). Genes were assigned to the functional classes according to their biochemical function using gene

Table 1. Chickpea library and EST characterization

	Total	Genotype	
		WR	JG
Sequence reads	6955	4047	2908
Failed base calling QC	8	6	2
Low quality sequence	214	121	93
Short insert sequence	422	246	176
No insert	30	30	0
E. coli	0	0	0
Mitochondrial	9	1	8
Total high quality	6272	3643	2629

Table 2. Sequencing and contigging statistics of chickpea ESTs.

Genotype	Total No. of ESTs sequenced	Sequencing success percentage	Good-quality ESTs used for contigging	ESTs in contigs	EST singletons
JG-62	2908	90.4	2629	2316	313
WR-315	4047	90.0	3643	2916	727
Total ESTs	6955	90.2	6272	5232	1040
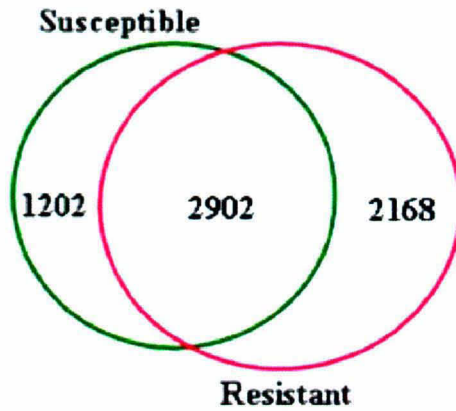
Figure 2.2: Venn diagram depicting exclusive and overlapping *Ca***ESTs.** The numbers signify ESTs specific to the susceptible and resistant genotypes and common between them.

Translation, ribosomal structure and biogenesis 4.57%

Secondary metabolism 1.29%

RNA processing and modification 1.09%

Signaling 4.12%

Transcription 3.52%

Replication, recombination and repair 0.20%

Stress 1.84%

Unknown function 3.43%

Post translational modification, protein turn over and chaperons 4.27%

Photosynthesis 0.60%

Nucleotide binding proteins 0.94%

Cell cycle control and cell division 0.60%

Cellular redox state 2.98%

Cellular Transport/inorganic ion transport and metabolism 2.29%

Cytoskeleton 0.79%

Defense mechanism 1.34%

Development/storage dormancy and senescence 0.45%

Energy production and conversion 0.79%

Hormone responsive 0.50%

Hypothetical protein 18.33%

No significant homology 18.22%

Miscellaneous 13.16%

Metabolism 14.56%

Figure 2.3: Functional classification of *Ca*Unigenes. The genes identified were grouped into 23 functional classes as shown in the piechart and the values represent the percentage of unigenes assigned to a particular functional class.

Table 3. Functional annotation and assignment of chickpea ESTs to different functional classes

Cell cycle control and cell division

CaF1_WIE_54_D_07	gb|ABE81376.1|	Cell division protein FtsZ [Medicago truncatula]
CaF1_WIE_54_E_01	emb|CAA65982.1|	cdc2MsF [Medicago sativa]
Contig281	gb|ABO81258.1|	Targeting for Xklp2 [Medicago truncatula]
Contig282	emb|CAB87834.1|	putative kinetochore protein [Vicia faba var. minor]
Contig733	gb|ABO84368.1|	Cell division protein FtsZ [Medicago truncatula]
Contig876	gb|ABO84368.1|	Cell division protein FtsZ [Medicago truncatula]
Contig895	gb|AAO72990.1|	cyclin D [Populus alba]
Contig965	gb|ABF66654.1|	EBP1 [Ammopiptanthus mongolicus]

cellular redox state

CaF1_JIE_23_C_09	emb|CAB41490.1|	cytochrome P450 monooxygenase [Cicer arietinum]
CaF1_JIE_24_D_02	gb|ABE93515.2|	Cytochrome b5 [Medicago truncatula]
CaF1_JIE_24_F_04	emb|AJ243804.1|	Cicer arietinum mRNA for cytochrome P450 (cyp93C3 gene)
CaF1_JIE_25_A_06	gb|ABE79228.1|	2OG-Fe(II) oxygenase [Medicago truncatula]
CaF1_JIE_40_E_05	dbj|BAD97435.1|	peroxidase [Pisum sativum]
CaF1_WIE_47_B_08	gb|ABO83980.1|	Thioredoxin domain 2; Thioredoxin fold [Medicago truncatula]
CaF1_WIE_50_B_08	gb|ABG90381.1|	glutathione S-transferase [Caragana korshinskii]
CaF1_WIE_54_E_07	emb|CAD31718.1|	putative cytochrome P450 monooxygenase [Cicer arietinum]
Contig104	gb|ABC68399.1|	cytochrome P450 monooxygenase CYP76O2 [Glycine max]
Contig143	dbj|BAC81649.1|	glutathione S-transferase [Pisum sativum]
Contig190	emb|CAB50768.1|	cytochrome P450 [Cicer arietinum]
Contig193	gb|ABN09112.1|	Haem peroxidase, plant/fungal/bacterial [Medicago truncatula]
Contig197	emb|CAB50768.1|	cytochrome P450 [Cicer arietinum]
Contig254	ref|NP_567919.1|	peroxidase, putative [Arabidopsis thaliana]
Contig275	emb|CAD31838.1|	putative quinone oxidoreductase [Cicer arietinum]
Contig278	gb|ABE77501.1|	Glutathione S-transferase
Contig299	gb|AAD37427.1|	peroxidase 1 precursor [Phaseolus vulgaris]
Contig313	emb|CAB56741.1|	cytochrome P450 monooxygenase [Cicer arietinum]
Contig315	emb|CAB50768.1|	cytochrome P450 [Cicer arietinum]
Contig317	gb|ABC59084.1|	cytochrome P450 monooxygenase CYP83G1 [Medicago truncatula]
Contig379	sp|P32110|GSTX6_SOYBN	Probable glutathione S-transferase
Contig43	emb|CAA62082.1|	cytochrome p450 [Arabidopsis thaliana]
Contig486	emb|CAA10067.1|	cytochrome P450 [Cicer arietinum]
Contig594	gb|AAZ32865.1|	thioredoxin h [Medicago sativa]
Contig613	emb|CAA10067.1|	cytochrome P450 [Cicer arietinum]
Contig637	gb|AAD37428.1|	peroxidase 3 precursor [Phaseolus vulgaris]
Contig674	gb|AAD37376.1	peroxidase [Glycine max]
Contig710	gb|ABE89853.1|	Thioredoxin-related; Thioredoxin domain 2 [Medicago truncatula]
Contig732	dbj|BAD97439.1|	peroxidase [Pisum sativum]
Contig830	emb|AJ487465.1|	Cicer arietinum mRNA for putative quinone oxidoreductase (qor gene)
Contig854	emb|CAA10132.1|	superoxide dismutase [Cicer arietinum]
Contig868	gb|ABC59084.1|	cytochrome P450 monooxygenase CYP83G1 [Medicago truncatula]
Contig909	gb|ABC59078.1|	cytochrome P450 monooxygenase CYP72A59 [Medicago truncatula]
Contig917	gb|ABE92135.1|	Glutathione S-transferase

Cellular Transport/inorganic ion transport and metabolism

CaF1_JIE_15_G_09	gb|AAA66200.1|	signal recognition particle 54 kDa subunit
CaF1_JIE_19_A_02	gb|AAQ87023.1|	VDAC3.1 [Lotus corniculatus var. japonicus]
CaF1_JIE_24_F_11	ref|NP_179680.1|	secretory carrier membrane protein (SCAMP) family protein [Arabidopsis thaliana]
CaF1_JIE_26_G_09	sp|Q41009|TOC34_PEA	Translocase of chloroplast 34
CaF1_JIE_27_E_08	emb|CAJ29291.1|	putative polyol transporter protein 4 [Lotus japonicus]
CaF1_JIE_36_B_11	ref|NP_194186.1|	clathrin adaptor complexes medium subunit family protein [Arabidopsis thaliana]
CaF1_JIE_36_H_10	gb|ABN05714.1|	metal ion transporter , putative [Medicago truncatula]
CaF1_JIE_37_C_10	gb|ABN08360.1|	Cation transporting ATPase, C-terminal [Medicago truncatula]
CaF1_WIE_45_A_04	gb|AAF22842.1|	vacuolar sorting receptor protein [Prunus dulcis]
CaF1_WIE_47_C_04	gb|ABE82410.1|	SBP [Medicago truncatula]
CaF1_WIE_47_D_07	ref|NP_199656.1|	SKS3 (SKU5 Similar 3); copper ion binding [Arabidopsis thaliana]
CaF1_WIE_48_B_10	gb|ABE93939.2|	Nuclear pore complex protein Nup205
CaF1_WIE_48_F_05	gb|ABD32921.2|	F5O11.19, related [Medicago truncatula]
CaF1_WIE_50_E_05	emb|CAC67501.1|	selenium binding protein [Medicago sativa]
CaF1_WIE_51_F_04	gb|ABN08184.1|	General substrate transporter [Medicago truncatula]

CaF1_WIE_55_G_06	gb|ABE90826.1|	von Willebrand factor, type C; Heavy metal transport/detoxification protein [Medicago truncatula]
CaF1_WIE_56_G_07	emb|CAC67501.1|	selenium binding protein [Medicago sativa]
Contig171	gb|AAC64109.1|	signal recognition particle 54 kDa subunit precursor [Pisum sativum]
Contig202	gb|ABE79078.1|	Mitochondrial carrier protein [Medicago truncatula]
Contig219	gb|ABP02565.1|	General substrate transporter [Medicago truncatula]
Contig305	gb|ABO82182.1|	Zinc finger, Sec23/Sec24-type; Sec23/Sec24 trunk region;
Contig331	gb|AAM91533.1|	membrane related protein CP5,
Contig441	emb|CAC67501.1|	selenium binding protein [Medicago sativa]
Contig472	ref|NP_190919.1|	ABC transporter family protein [Arabidopsis thaliana]
Contig484	gb|ABN08635.1|	Mitochondrial import inner membrane translocase, subunit Tim17/22 [Medicago truncatula]
Contig597	ref|NP_563985.1|	GOS11 (GOLGI SNARE 11)
Contig626	sp|Q41009|TOC34_PEA	Translocase of chloroplast 34
Contig648	gb|ABC01895.1|	transporter-like protein [Solanum tuberosum]
Contig683	gb|AAL17949.1|	type IIB calcium ATPase [Medicago truncatula]
Contig736	dbj|BAC65212.1|	phosphate permease [Fusarium oxysporum]
Contig749	gb|AAO39834.1|	ferrous ion membrane transport protein DMT1 [Glycine max]
Contig827	gb|ABE86679.1|	Intracellular chloride channel [Medicago truncatula]
Contig828	gb|ABM69111.1|	phosphate transporter 5 [Medicago truncatula]
Contig862	ref|NP_564367.1|	integral membrane HRF1 family protein
Contig938	gb|ABE93220.1|	General substrate transporter [Medicago truncatula]

Cytoskeleton

CaF1_JIE_09_H_01	gb|AAX86048.1|	tubulin B4 [Glycine max]
CaF1_JIE_20_C_03	gb|EDN04544.1|	actin [Ajellomyces capsulatus NAm1]
CaF1_JIE_40_B_07	sp|P28551|TBB3_SOYBN	Tubulin beta chain (Beta tubulin) emb|CAA42777.1| beta-tubulin [Glycine max]
CaF1_WIE_41_E_04	sp|Q39445|TBB_CICAR	Tubulin beta chain (Beta tubulin) emb|CAA67056.1| beta-tubulin [Cicer arietinum]
CaF1_WIE_46_D_04	gb|ABI51282.1|	actin [Hibiscus cannabinus]
CaF1_WIE_48_E_08	sp|Q8SAG3|ADF_VITVI	Actin-depolymerizing factor (ADF)
CaF1_WIE_53_E_05	gb|ABF59516.1|	putative spindle disassembly related protein CDC48 [Nicotiana tabacum]
Contig413	gb|AAZ94896.1|	actin-like protein [Catharanthus roseus]
Contig416	dbj|BAB40710.1|	BY-2 kinesin-like protein 10 [Nicotiana tabacum]
Contig846	dbj|BAB88648.1|	microtubule bundling polypeptide TMBP200 [Nicotiana tabacum]
Contig855	gb|ABE90729.1|	Myosin II heavy chain-like [Medicago truncatula]
Contig912	emb|X68649.1|PSACTG	P.sativum mRNA for actin

Defense mechanism

CaF1_JIE_20_A_02	emb|CAA10189.1|	class I chitinase [Cicer arietinum]
CaF1_JIE_25_B_07	gb|AAP03880.2|	Avr9/Cf-9 induced kinase 1 [Nicotiana tabacum]
CaF1_WIE_41_E_01	gb|ABO84738.1|	Harpin-induced 1 [Medicago truncatula]
CaF1_WIE_48_F_10	gb|ABL98074.1|	chitinase-related agglutinin [Robinia pseudoacacia]
Contig264	gb|AAC49370.1|	non-specific lipid transfer-like protein
Contig411	emb|CAA10189.1|	class I chitinase [Cicer arietinum]
Contig559	sp|P27047|DRR4_PEA	Disease resistance response protein DRRG49-C gb|AAA33663.1|
Contig725	gb|AAP37978.1|	class 10 pathogenesis-related protein [Lupinus luteus]
Contig773	gb|ABE84488.1|	Plant lipid transfer/seed storage/trypsin-alpha amylase inhibitor [Medicago truncatula]
Contig834	emb|CAA56142.1|	pathogenesis related protein [Cicer arietinum]
Contig863	gb|AAP03880.2|	Avr9/Cf-9 induced kinase 1 [Nicotiana tabacum]
Contig892	gb|AAV92899.1|	Avr9/Cf-9 rapidly elicited protein 140 [Nicotiana tabacum]
Contig941	gb|ABE79318.1|	Plant lipid transfer/seed storage/trypsin-alpha amylase inhibitor [Medicago truncatula]
Contig942	gb|AAC77929.1|	similar to Nicotiana HR lesion-inducing ORF [Medicago sativa]
Contig953	sp|P22196|PER2_ARÁHY	Cationic peroxidase 2 precursor (PNPC2) gb|AAA32676.1| cationic peroxidase

Development/storage/dormancy and senescence

CaF1_JIE_25_A_02	gb|AAB84193.1|	dormancy-associated protein [Pisum sativum]
CaF1_JIE_42_A_04	sp|P08688|ALB2_PEA	Albumin-2
CaF1_WIE_53_E_03	dbj|AB049721.2|	Pisum sativum ssa-11 mRNA for putative senescence-associated protein
Contig266	emb|CAD31716.1|	putative ripening related protein [Cicer arietinum]
Contig532	sp|P08688|ALB2_PEA	Albumin-2
Contig575	dbj|BAB33421.1|	putative senescence-associated protein [Pisum sativum]

Contig654	emb|CAD31716.1|	putative ripening related protein [Cicer arietinum]
Contig962	dbj|BAB33421.1|	putative senescence-associated protein [Pisum sativum]

Energy production and conversion

Contig165	gb|ABP02830.1|	H+-transporting two-sector ATPase, alpha/beta subunit,
Contig231	dbj|BAE71236.1|	putative ADP,ATP carrier-like protein [Trifolium pratense]
Contig348	ref|NP_180560.1|	NDA2 (ALTERNATIVE NAD(P)H DEHYDROGENASE 2)
Contig356	gb|AAY56795.1|	vacuolar H+-ATPase subunit A [Vigna unguiculata]
Contig651	gb|AAY44152.1|	ATP synthase beta subunit [Ecballium elaterium]
Contig656	gb|AAM65274.1|	NADH dehydrogenase [Arabidopsis thaliana]
Contig709	gb|AAZ23107.1|	plastid ATP/ADP transport protein 2 [Manihot esculenta]
Contig739	ref|XP_384488.1|	ATPB_NEUCR ATP synthase beta chain, mitochondrial precursor [Gibberella zeae PH-1]
Contig884	sp|O23948|VATE_GOSHI	Vacuolar ATP synthase subunit E (V-ATPase subunit E)

Hormone responsive

CaF1_JIE_20_C_08	emb|CAJ13711.1|	putative ethylene response protein [Capsicum chinense]
Contig116	gb|AAL32037.2|	ethylene-responsive transciptional coactivator-like protein [Retama raetam]
Contig203	gb|AAV63565.1|	auxin-induced putative aldo/keto reductase family protein [Arachis hypogaea]
Contig240	gb|AAM65588.1|	putative auxin-induced protein, IAA12 [Arabidopsis thaliana]
Contig261	emb|CAC84710.1|	aux/IAA protein [Populus tremula x Populus tremuloides]
Contig558	gb|AAZ66745.1|	coronatine-insensitive 1 [Glycine max]
Contig605	gb|ABO82441.1|	Auxin responsive SAUR protein [Medicago truncatula]
Contig758	sp|P35694|BRU1_SOYBN	Brassinosteroid-regulated protein BRU1 precursor gb|AAA81350.1| brassinosteroid-regulated protein

Metabolism

CaF1_JIE_12_H_08	dbj|BAD94487.1|	phosphoenolpyruvate carboxykinase-like protein [Arabidopsis thaliana]
CaF1_JIE_13_E_03	ref|NP_194214.2|	arginosuccinate synthase family [Arabidopsis thaliana]
CaF1_JIE_14_B_09	gb|AAL74418.2|	ATP sulfurylase [Glycine max]
CaF1_JIE_14_H_09	ref|NP_200706.1|	ceramidase family protein [Arabidopsis thaliana]
CaF1_JIE_15_A_10	ref|XP_389647.1|	GR78_NEUCR 78 KDA GLUCOSE-REGULATED PROTEIN
CaF1_JIE_15_B_02	gb|ABE81529.1|	Pyridoxal-5-phosphate-dependent enzyme, beta subunit [Medicago truncatula]
CaF1_JIE_15_H_01	gb|ABE78289.1|	UDP-glucuronosyl/UDP-glucosyltransferase [Medicago truncatula]
CaF1_JIE_16_B_05	sp|P12886|ADH1_PEA	Alcohol dehydrogenase 1 emb|CAA29609.1| alcohol dehydrogenase [Pisum sativum]
CaF1_JIE_17_F_04	gb|ABD28734.1|	UDP-N-acetylglucosamine transferase subunit ALG14, related [Medicago truncatula]
CaF1_JIE_17_G_04	dbj|BAF44219.1|	polyketide reductase [Lotus japonicus]
CaF1_JIE_18_C_10	ref|NP_172147.2|	2-oxoglutarate-dependent dioxygenase, putative [Arabidopsis thaliana]
CaF1_JIE_18_F_02	sp|O04866|ARGD_ALNGL	Acetylornithine aminotransferase
CaF1_JIE_19_F_11	sp|P51850|PDC1_PEA	Pyruvate decarboxylase isozyme 1 (PDC) emb|CAA91444.1| pyruvate decarboxylase [Pisum sativum]
CaF1_JIE_20_D_06	dbj|BAE99337.1|	putative beta-amylase [Arabidopsis thaliana]
CaF1_JIE_22_A_06	gb|AAF70823.1|AF154422_1	beta-galactosidase [Lycopersicon esculentum]
CaF1_JIE_22_A_09	emb|CAK54360.1|	putative desaturase-like protein [Trifolium repens]
CaF1_JIE_26_C_07	gb|ABE91874.1|	SAM (and some other nucleotide) binding motif [Medicago truncatula]
CaF1_JIE_26_G_07	dbj|BAF49052.1|	phytoene synthase [Prunus mume]
CaF1_JIE_27_D_10	ref|NP_563662.1|	pectinesterase family protein [Arabidopsis thaliana]
CaF1_JIE_28_B_05	gb|ABD28403.1|	Formylmethionine deformylase [Medicago truncatula]
CaF1_JIE_28_D_10	ref|NP_563662.1|	pectinesterase family protein [Arabidopsis thaliana]
CaF1_JIE_31_A_04	gb|AAT36331.1|	nitrilase 4A [Lupinus angustifolius]
CaF1_JIE_34_D_01	gb|ABN08040.1|	Acyl-coA-binding protein, ACBP; Serine/threonine protein phosphatase, BSU1 [Medicago truncatula]
CaF1_JIE_37_B_02	gb|AAX63898.1|	geranylgeranyl reductase [Medicago truncatula]
CaF1_JIE_38_F_06	gb|AAK13318.1|	ATP:citrate lyase [Capsicum annuum]
CaF1_JIE_39_C_03	gb|ABE77854.2|	Phospholipid/glycerol acyltransferase [Medicago truncatula]
CaF1_JIE_41_B_07	sp|P12886|ADH1_PEA	Alcohol dehydrogenase 1 emb|CAA29609.1| alcohol dehydrogenase [Pisum sativum]
CaF1_JIE_41_E_01	gb|ABE89691.1|	Pectinesterase [Medicago truncatula]
CaF1_JIE_41_G_04	pir||S47243	starch phosphorylase (EC 2.4.1.1) isoform L precursor, chloroplast - fava bean
CaF1_JIE_42_A_01	ref|NP_181224.1|	xyloglucan:xyloglucosyl transferase

CaF1_WIE_42_A_02	gb|ABE84980.1|	NAD-binding site; Nucleotide sugar epimerase [Medicago truncatula]
CaF1_WIE_42_E_04	gb|AAB40396.1|	glycolate oxidase [Mesembryanthemum crystallinum]
CaF1_WIE_43_C_10	emb|CAC43237.1|	lipoxygenase [Sesbania rostrata]
CaF1_WIE_43_H_02	gb|ABE93021.1|	Protein prenyltransferase [Medicago truncatula]
CaF1_WIE_44_D_02	dbj|BAE93460.1|	diacylglycerolacyltransferase-1a [Glycine max]
CaF1_WIE_45_C_09	gb|AAL66290.1|AF452450_1	adenosine 5'-phosphosulfate reductase [Glycine max]
CaF1_WIE_45_E_10	ref|NP_001077888.1|	ALDH6B2 (Aldehyde dehydrogenase 6B2) [Arabidopsis thaliana]
CaF1_WIE_46_D_11	sp|O24301|SUS2_PEA	Sucrose synthase 2 (Sucrose-UDP glucosyltransferase 2) emb|CAA04512.1| second sucrose synthase [Pisum sativum]
CaF1_WIE_46_F_09	gb|ABK55756.1|	P450 monooxygenase-like protein [Stylosanthes guianensis]
CaF1_WIE_47_A_05	gb|ABO81445.1|	Glycoside hydrolase, family 1 [Medicago truncatula]
CaF1_WIE_47_A_11	gb|ABE78454.1|	Phosphoesterase, DHHA1 [Medicago truncatula]
CaF1_WIE_48_C_04	dbj|BAA36291.1|	HMG-CoA reductase [Cucumis melo]
CaF1_WIE_48_D_02	dbj|BAA36291.1|	HMG-CoA reductase [Cucumis melo]
CaF1_WIE_49_A_10	gb|AF461200.1|	Medicago truncatula nodule-enhanced malate dehydrogenase gene
CaF1_WIE_49_B_08	emb|CAA09588.1|	phosphoenolpyruvate-carboxylase [Vicia faba]
CaF1_WIE_49_D_11	ref|NP_178980.1|	malate oxidoreductase, putative [Arabidopsis thaliana]
CaF1_WIE_50_C_09	gb|ABE87035.1|	Orn/DAP/Arg decarboxylase 2; Protease-associated PA; Proteinase inhibitor I9, subtilisin propeptide [Medicago truncatula]
CaF1_WIE_50_D_11	sp|P52904|ODPB_PEA	Pyruvate dehydrogenase E1 component subunit beta
CaF1_WIE_50_E_10	gb|ABP65665.1|	VTC2-like protein [Actinidia chinensis]
CaF1_WIE_50_G_05	gb|ABO84404.1|	SAE2 , putative [Medicago truncatula]
CaF1_WIE_51_B_08	sp|Q39366|LGUL_BRAOG	Putative lactoylglutathione lyase (Methylglyoxalase)
CaF1_WIE_51_B_10	emb|CAA63598.1|	glyoxysomal beta-ketoacyl-thiolase [Brassica napus]
CaF1_WIE_51_F_07	gb|ABM91070.1|	xyloglucan endotransglycosylase/hydrolase precursor XTH-3 [Populus tremula x Populus tremuloides]
CaF1_WIE_52_B_01	gb|ABP02258.1|	Glycoside hydrolase, family 3, N-terminal; Glycoside hydrolase, family 3, C-terminal [Medicago truncatula]
CaF1_WIE_53_A_01	gb|ABR15094.1|	ATP citrate lyase alpha subunit [Glycyrrhiza uralensis]
CaF1_WIE_53_C_10	sp|P31239|ACCO_PEA	1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase)
CaF1_WIE_54_A_05	gb|AAG31076.1	sucrose-phosphatase [Medicago truncatula]
CaF1_WIE_54_C_10	emb|CAJ15149.1|	sialyltransferase-like protein [Medicago truncatula]
CaF1_WIE_54_E_05	sp|O24301|SUS2_PEA	Sucrose synthase 2 (Sucrose-UDP glucosyltransferase 2)
CaF1_WIE_55_F_07	gb|ABN08775.1|	Glycoside hydrolase, family 19 [Medicago truncatula]
CaF1_WIE_55_G_09	emb|CAJ38375.1|	nucleoside-diphopshate-sugar dehydratase [Plantago major]
CaF1_WIE_55_H_09	gb|ABO81708.1|	2-oxoglutarate dehydrogenase, E1 component [Medicago truncatula]
CaF1_WIE_56_D_11	gb|ABE82917.1|	Aldo/keto reductase [Medicago truncatula]
CaF1_WIE_56_F_05	gb|ABE86452.1|	Aminotransferase, class I and II [Medicago truncatula]
CaF1_WIE_56_G_06	gb|ABI34092.1|	cystathionine gamma-synthase [Medicago sativa]
Contig10	emb|CAA09040.1|	glyceraldehyde 3-phosphate dehydrogenase, cytosolic [Cicer arietinum]
Contig134	gb|AAS18240.1|	enolase [Glycine max]
Contig135	gb|ABE77912.1|	SAM (and some other nucleotide) binding motif [Medicago truncatula]
Contig148	gb|ABE81243.1|	Malic oxidoreductase [Medicago truncatula]
Contig152	gb|AAS46232.1|	methionine sulfoxide reductase A [Populus trichocarpa x Populus deltoides]
Contig153	emb|CAA10287.2|	glucan-endo-1,3-beta-glucosidase [Cicer arietinum]
Contig154	gb|AAL11502.1	NAD-dependent malate dehydrogenase [Prunus persica]
Contig159	gb|ABE86430.2|	Glycosyl transferase, family 20; Trehalose-phosphatase [Medicago truncatula]
Contig163	gb|ABE89232.1|	Inositol polyphosphate related phosphatase [Medicago truncatula]
Contig164	ref|NP_566081.1|	pyridoxine 5'-phosphate oxidase-related [Arabidopsis thaliana]
Contig179	gb|ABE93770.1|	Adenylosuccinate synthetase [Medicago truncatula]
Contig180	gb|ABP49577.1|	oleate desaturase [Caragana korshinskii var. intermedia]
Contig182	ref|NP_193961.2|	carboxylic ester hydrolase [Arabidopsis thaliana]
Contig186	gb|ABC87913.1|	polygalacturonase precursor [Glycine max]
Contig192	ref|NP_568756.2|	ceramide kinase-related [Arabidopsis thaliana]
Contig208	emb|CAM35498.1|	spermine synthase [Lotus japonicus]
Contig21	gb|AAR30118.1|	putative histidine phosphotransferase HPT1p [Gibberella moniliformis]
Contig213	emb|CAI91291.1|	deoxyhypusine synthase [Crotalaria juncea]
Contig224	sp|O65735|ALF_CICAR	Fructose-bisphosphate aldolase
Contig239	gb|ABE83505.2|	TPP-binding enzymes [Medicago truncatula]
Contig24	gb|ABO80586.1|	Amidase [Medicago truncatula]
Contig243	emb|CAH05011.1|	alpha-dioxygenase [Pisum sativum]
Contig249	gb|AAZ20293.1|	isomerase-like protein [Arachis hypogaea]

Contig25	gb|AAC49902.1|	diadenosine 5',5'''-P1,P4-tetraphosphate hydrolase [Lupinus angustifolius]
Contig250	gb|ABO84301.1|	Alpha-isopropylmalate/homocitrate synthase [Medicago truncatula]
Contig252	gb|ABE88265.1|	Generic methyltransferase [Medicago truncatula] gb|ABN08826.1| Generic methyltransferase [Medicago truncatula]
Contig255	sp|Q43070|GALE1_PEA	UDP-glucose 4-epimerase (Galactowaldenase) (UDP-galactose 4-epimerase) gb|AAA86532.1| UDP-galactose-4-epimerase
Contig263	gb|ABE82022.1|	Adenylate kinase [Medicago truncatula]
Contig265	gb|ABE82974.1|	Pyridoxal-5-phosphate-dependent enzyme, beta subunit [Medicago truncatula]
Contig267	gb|AAO72533.1|	pyruvate decarboxylase 1 [Lotus corniculatus]
Contig286	sp|P12886|ADH1_PEA	Alcohol dehydrogenase 1 emb|CAA29609.1| alcohol dehydrogenase [Pisum sativum]
Contig289	gb|ABE94198.1|	NAD-dependent epimerase/dehydratase [Medicago truncatula]
Contig296	gb|AAD55298.1|	F25A4.24 [Arabidopsis thaliana]
Contig3	gb|ABE93138.1|	Amino acid/polyamine transporter II [Medicago truncatula]
Contig30	sp|Q03460|GLSN_MEDSA	Glutamate synthase [NADH], chloroplast precursor (NADH-GOGAT)
Contig304	gb|AAC50039.1|	polynucleotide phosphorylase [Pisum sativum]
Contig308	sp|P21727|TPT_PEA	Triose phosphate/phosphate translocator
Contig327	gb|ABN08532.1|	Prephenate dehydratase with ACT region [Medicago truncatula]
Contig328	emb|CAA05979.1|	adenine nucleotide translocator [Lupinus albus]
Contig337	ref|XP_387704.1|	KPYK_TRIRE Pyruvate kinase [Gibberella zeae PH-1]
Contig339	gb|ABD33414.1|	Pectinesterase; Pectinesterase inhibitor [Medicago truncatula]
Contig346	gb|AAL85977.1|	putative uridine kinase [Arabidopsis thaliana]
Contig352	gb|ABG73467.1|	6-phosphogluconolactonase [Oryza brachyantha]
Contig36	gb|ABE89691.1|	Pectinesterase [Medicago truncatula]
Contig381	emb|CAA08855.1|	copper amine oxidase [Cicer arietinum]
Contig382	sp|Q9AT63|PDX1_GINBI	Probable pyridoxal biosynthesis protein PDX1
Contig39	gb|AAS67005.1|	Phosphoenolpyruvate carboxylase [Glycine max]
Contig403	gb|ABD33005.1|	Membrane bound O-acyl transferase, MBOAT [Medicago truncatula]
Contig430	dbj|BAB43909.1|	phosphoenolpyruvate carboxykinase [Flaveria pringlei]
Contig435	sp|P12886|ADH1_PEA	Alcohol dehydrogenase 1 emb|CAA29609.1| alcohol dehydrogenase [Pisum sativum]
Contig440	gb|ABE84165.2|	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase
Contig450	gb|AAM65998.1|	putative dTDP-glucose 4-6-dehydratase [Arabidopsis thaliana]
Contig453	emb|CAC06095.1|	ferredoxin-nitrite reductase [Lotus japonicus]
Contig459	gb|ABB69781.1|	beta-glucan-binding protein 1 [Medicago truncatula]
Contig488	sp|P13603|ADH1_TRIRP	Alcohol dehydrogenase 1
Contig49	emb|CAD70620.1|	branched-chain amino acid aminotransferase-like protein [Cicer arietinum]
Contig515	gb|AAD19957.1|	thiosulfate sulfurtransferase [Datisca glomerata]
Contig53	ref|NP_173390.1|	pfkB-type carbohydrate kinase family protein [Arabidopsis thaliana]
Contig531	gb|ABB89021.1|	CXE carboxylesterase [Actinidia deliciosa]
Contig539	gb|ABE89020.1|	Adenosine kinase [Medicago truncatula]
Contig54	dbj|BAA13032.1|	phosphoribosylanthranilate transferase [Pisum sativum]
Contig543	gb|AAP33475.1|	polygalacturonase-like protein [Fragaria x ananassa]
Contig545	gb|ABN08040.1|	Acyl-coA-binding protein, ACBP; Serine/threonine protein phosphatase, BSU1 [Medicago truncatula]
Contig555	gb|AAS46231.1|	methionine sulfoxide reductase A [Populus trichocarpa x Populus deltoides]
Contig574	sp|O24301|SUS2_PEA	Sucrose synthase 2 (Sucrose-UDP glucosyltransferase 2)
Contig577	gb|AAB19212.1|	polygalacturonase-inhibiting protein [Malus x domestica]
Contig595	gb|ABE80304.1|	Extradiol ring-cleavage dioxygenase, class III enzyme, subunit B [Medicago truncatula]
Contig600	gb|AAZ32910.1|	polygalacturonase-like protein [Medicago sativa]
Contig603	emb|CAA10287.2|	glucan-endo-1,3-beta-glucosidase [Cicer arietinum]
Contig606	gb|AAG17666.1|	S-adenosylmethionine synthetase [Brassica juncea]
Contig608	gb|ABE92923.1|	C2; Peptidase, cysteine peptidase active site [Medicago truncatula]
Contig665	emb|CAH55772.1|	putative His-Asp phosphotransfer protein [Pisum sativum]
Contig678	dbj|BAF31848.1|	nitrite reductase [Fusarium oxysporum]
Contig694	gb|AAP33475.1|	polygalacturonase-like protein [Fragaria x ananassa]
Contig697	gb|ABE83264.1|	Isocitrate dehydrogenase NADP-dependent, plant [Medicago truncatula]
Contig704	gb|ABA86966.1|	triosephosphate isomerase [Glycine max]
Contig724	gb|AAB99755.1|	malate dehydrogenase precursor [Medicago sativa]

Contig734	gb|AAL16064.1|	S-adenosyl-L-methionine synthetase [Dendrobium crumenatum]
Contig740	gb|ABO15567.1|	cysteine synthase [Glycine max]
Contig741	gb|ABE84165.2|	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase
Contig746	gb|ABE92907.2|	Peptidase S24, S26A and S26B [Medicago truncatula]
Contig755	gb|ABN07990.1|	SAM (and some other nucleotide) binding motif [Medicago truncatula]
Contig793	gb|ABE84165.2|	5-methyltetrahydropteroyltriglutamate--homocysteine S-methyltransferase
Contig795	gb|ABE82378.1|	Aldehyde dehydrogenase [Medicago truncatula]
Contig800	gb|ABO78563.1|	Glycine cleavage system P-protein [Medicago truncatula]
Contig804	gb|ABE81243.1|	Malic oxidoreductase [Medicago truncatula]
Contig813	emb|CAD29733.1|	pectin methylesterase [Sesbania rostrata]
Contig831	gb|AAT45084.1|	proline dehydrogenase [Medicago sativa]
Contig844	sp|P34921|G3PC_DIACA	Glyceraldehyde-3-phosphate dehydrogenase, cytosolic
Contig85	dbj|BAE98181.1|	putative pyruvate decarboxylase [Fusarium oxysporum f. sp. lycopersici]
Contig858	sp|Q9M4D8|DCAM_VICFA	S-adenosylmethionine decarboxylase proenzyme
Contig864	gb|ABO77440.1|	S-adenosylmethionine decarboxylase [Medicago sativa subsp. falcata]
Contig865	ref|NP_568117.1|	phosphoadenosine phosphosulfate (PAPS) reductase family protein [Arabidopsis thaliana]
Contig867	gb|AAG23130.1|	diacylglycerol kinase variant A [Lycopersicon esculentum]
Contig870	gb|ABE85996.1|	Glycoside hydrolase, family 1 [Medicago truncatula]
Contig877	ref|NP_850271.1|	ATGSL08 (GLUCAN SYNTHASE-LIKE 8)
Contig879	gb|ABE82897.1|	Acetohydroxy acid isomeroreductase [Medicago truncatula]
Contig880	gb|AAP83926.1|	transaldolase [Lycopersicon esculentum]
Contig890	gb|AAL01888.1|	acyl-CoA oxidase [Glycine max]
Contig899	ref|NP_187341.1|	DIN3/LTA1 (DARK INDUCIBLE 3)
Contig900	gb|ABE78689.2|	AMP-dependent synthetase and ligase [Medicago truncatula]
Contig907	gb|ABE78365.1|	Metallophosphoesterase
Contig91	dbj|BAF31848.1|	nitrite reductase [Fusarium oxysporum]
Contig911	pdb|2P4H|X	Chain X, Crystal Structure Of Vestitone Reductase From Alfalfa (Medicago Sativa L.)
Contig918	gb|ABE81243.1|	Malic oxidoreductase [Medicago truncatula]
Contig923	gb|ABB72820.1|	oligouridylate binding protein-like protein [Solanum tuberosum]
Contig927	gb|ABO80948.1|	S-adenosylmethionine synthetase [Medicago truncatula]
Contig933	gb|ABE89160.1|	O-methyltransferase, family 2; Dimerisation [Medicago truncatula]
Contig935	gb|ABA03227.1|	glyceraldehyde-3-phosphate dehydrogenase [Populus maximowiczii x Populus nigra]
Contig936	sp|Q9FVL0|HBL1_MEDSA	Non-symbiotic hemoglobin 1
Contig95	dbj|BAA76430.1|	fructose-bisphosphate aldolase [Cicer arietinum]
Contig951	emb|CAD91338.1|	beta-fructofuranosidase [Glycine max]
Contig955	gb|AAA33358.1|	3-hydroxy-3-methylglutaryl-coenzyme A reductase
Contig967	gb|AAU07997.1|	phosphoenolpyruvate carboxylase 2; LaPEPC2 [Lupinus albus]
Contig973	emb|CAA52800.1|	T-protein of the glycine decarboxylase complex [Pisum sativum]

Miscellaneous

CaF1_JIE_09_H_09	ref|NP_194420.1|	AT-HF (Arabidopsis thaliana HisF protein) sp|Q9SZ30|HIS5_ARATH Imidazole glycerol phosphate synthase hisHF
CaF1_JIE_13_C_01	gb|ABD32291.1|	Uncharacterized Cys-rich domain [Medicago truncatula]
CaF1_JIE_14_C_09	gb|ABO84327.1|	Putative non-LTR retroelement reverse transcriptase, related [Medicago truncatula]
CaF1_JIE_15_D_07	gb|AAU05467.1|	At5g22850 [Arabidopsis thaliana] gb|AAV59285.1| At5g22850 [Arabidopsis thaliana]
CaF1_JIE_16_H_02	dbj|BAD94495.1|	sigma-like factor [Arabidopsis thaliana]
CaF1_JIE_16_H_10	gb|AC166093.12|	Medicago truncatula clone mth2-17k6, complete sequence
CaF1_JIE_17_C_11	gb|AC146573.21|	Medicago truncatula clone mth2-145j1, complete sequence
CaF1_JIE_19_A_01	gb|ABG21940.1|	AGR_C_5039p, putative, expressed [Oryza sativa (japonica cultivar-group)]
CaF1_JIE_19_F_04	gb|AC151956.5|	Medicago truncatula clone mth2-52p17, complete sequence
CaF1_JIE_19_G_04	gb|ABE79642.1|	Cellular retinaldehyde-binding/triple function, N-terminal [Medicago truncatula]
CaF1_JIE_20_D_05	gb|AC147430.9|	Medicago truncatula clone mth2-71e6, complete sequence
CaF1_JIE_22_E_07	gb|AAK82520.1|	AT5g24810/F6A4_20 [Arabidopsis thaliana]
CaF1_JIE_23_A_03	sp|Q850K7|EXLB1_ORYSJ	Expansin-like B1 precursor (OsEXLB1)
CaF1_JIE_26_E_04	emb|AJ411814.1|	Cicer arietinum Ty3-gypsy like Retrotransposon CaRep

CaF1_JIE_29_F_09	gb|AC157490.18|	Medicago truncatula clone mth2-123f23, complete sequence
CaF1_JIE_30_D_06	gb|ABE77505.1|	DECOY (exp=-1; , putative [Medicago truncatula]
CaF1_JIE_31_F_03	gb|AC167403.2|	Medicago truncatula chromosome 7 BAC clone mte1-14f12
CaF1_JIE_31_F_11	gb|AAM49801.1|	GFA2 [Arabidopsis thaliana]
CaF1_JIE_32_B_05	gb|AC137822.30|	Medicago truncatula clone mth2-31e20, complete sequence
CaF1_JIE_32_G_08	gb|AY232722.1|	Fusarium oxysporum f. sp. vasinfectum strain Ag149 Foxy transposable element, partial sequence
CaF1_JIE_32_H_02	gb|AC148397.13|	Medicago truncatula clone mth2-22h4, complete sequence
CaF1_JIE_34_B_06	gb|AC151621.20|	Medicago truncatula clone mth2-14p3, complete sequence
CaF1_JIE_34_B_07	gb|K03313.1|RIATL	Integrated Ri plasmid agropine (A. rhizogenes strain A4) complete TL-DNA and flanking plant (Convolvulus arvensis) DNA
CaF1_JIE_35_B_03	gb|ABL59986.1|	brittle stalk-2-like protein 6 [Zea mays]
CaF1_JIE_35_F_10	gb|AC175685.3|	Medicago truncatula chromosome 2 BAC clone mte1-55k6
CaF1_JIE_36_B_02	gb|AAW33880.1|	RING-H2 subgroup RHE protein [Populus alba x Populus tremula]
CaF1_JIE_36_E_08	dbj|BAB86895.1|	syringolide-induced protein B15-3-5 [Glycine max]
CaF1_JIE_37_A_07	gb|AC146746.15|	Medicago truncatula clone mth2-108p9, complete sequence
CaF1_JIE_40_E_01	gb|AAM19795.1|	At2g04030/F3C11.14 [Arabidopsis thaliana]
CaF1_JIE_41_G_02	gb|ABP03273.1|	Uncharacterized Cys-rich domain [Medicago truncatula]
CaF1_WIE_41_D_01	gb|ABF06706.1|	UP-9A [Nicotiana tabacum]
CaF1_WIE_41_H_04	dbj|AP004505.1|	Lotus japonicus genomic DNA, chromosome 3, clone:LjT10E18, TM0035
CaF1_WIE_43_H_03	emb|CAH66506.1|	OSIGBa0111I14.1 [Oryza sativa (indica cultivar-group)]
CaF1_WIE_44_C_04	ref|NP_001050957.1|	putative HIPL1 protein
CaF1_WIE_45_B_03	gb|ABE84233.1|	Arf GTPase activating protein [Medicago truncatula]
CaF1_WIE_45_C_07	gb|AY032742.1|	Fusarium sporotrichoides guanine nucleotide-binding protein mRNA
CaF1_WIE_45_H_01	gb|EF025129.1|	Medicago truncatula PHD6 mRNA, complete cds
CaF1_WIE_46_A_06	gb|ABQ11262.1|	mago nashi-like protein 1 [Physalis pubescens]
CaF1_WIE_46_A_09	gb|AC150798.3|	Medicago truncatula chromosome 2 clone mth2-33b11, complete sequence
CaF1_WIE_46_D_05	gb|ABE92992.1|	Exostosin-like [Medicago truncatula]
CaF1_WIE_46_G_07	emb|AJ489609.1	Cicer arietinum mRNA for alpha-expansin 4 (expa4 gene)
CaF1_WIE_47_A_02	gb|AC148171.16|	Medicago truncatula clone mth2-30a2, complete sequence
CaF1_WIE_47_D_01	gb|ABP03277.1|	UBX; PUG; Zinc finger, C2H2-type [Medicago truncatula]
CaF1_WIE_47_E_07	gb|AC149579.11|	Medicago truncatula clone mth2-99d8, complete sequence
CaF1_WIE_48_A_04	ref|NP_196983.1|	putative protein [Arabidopsis thaliana]
CaF1_WIE_48_D_10	gb|AC148237.7|	Medicago truncatula clone mth2-1o14, complete sequence
CaF1_WIE_49_B_02	emb|CT954252.6|	M.truncatula DNA sequence from clone MTH2-60M21 on chromosome 3
CaF1_WIE_49_D_04	ref|NP_851260.1|	ATHDH (HISTIDINOL DEHYDROGENASE) [Arabidopsis thaliana]
CaF1_WIE_49_D_06	emb|CR962137.2|	Medicago truncatula chromosome 5 clone mte1-10p15
CaF1_WIE_49_G_05	gb|AC144340.30|	Medicago truncatula clone mth2-7k2,
CaF1_WIE_49_H_08	dbj|AP004534.1|	Lotus japonicus genomic DNA, chromosome 4, clone:LjT14P20
CaF1_WIE_51_C_06	gb|DQ117568.1|	Phaseolus vulgaris clone PvD34, mRNA sequence
CaF1_WIE_52_A_11	dbj|AP006629.1|	Lotus japonicus genomic DNA, chromosome 2, clone:LjT10B11, TM0008
CaF1_WIE_52_F_05	emb|CR940305.16|	M.truncatula DNA sequence from clone MTH2-28N4 on chromosome 3
CaF1_WIE_54_F_05	emb|CAJ31277.1|	autophagy protein 5 [Glycine max]
CaF1_WIE_54_F_07	gb|AAM65872.1|	ferritin subunit, putative [Arabidopsis thaliana]
CaF1_WIE_55_B_04	emb|CR962124.2|	Medicago truncatula chromosome 5 clone mth2-44c15
CaF1_WIE_55_B_09	gb|AAP22955.1|	Potyvirus VPg interacting protein [Pisum sativum]
CaF1_WIE_56_B_03	ref|NP_001046176.1|	KH domain-containing protein-like [Oryza sativa (japonica cultivar-group)]
CaF1_WIE_56_D_10	ref|NP_001058575.1|	putative spastin protein [Oryza sativa (japonica cultivar-group)]
Contig119	gb|AAT08648.1|	ADP-ribosylation factor [Hyacinthus orientalis]
Contig123	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene
Contig124	emb|AJ250814.1|	Fusarium oxysporum f. sp. lycopersici insertion sequence Foxy
Contig129	emb|CAC43238.1|	calcium binding protein [Sesbania rostrata]
Contig141	gb|ABE77505.1|	DECOY (exp=-1; , putative [Medicago truncatula]
Contig144	emb|CAA93759.1|	putative transposase [Tolypocladium inflatum]
Contig157	ref|NP_001045159.1|	putative ribophorin II precursor [Oryza sativa (japonica cultivar-group)]
Contig162	gb|ABE92863.1|	F-actin capping protein, alpha subunit [Medicago truncatula]
Contig167	gb|AAL06496.1|	AT5g04420/T32M21_20 [Arabidopsis thaliana]
Contig169	emb|CR962130.3|	Medicago truncatula chromosome 5 clone mte1-63h20
Contig175	gb|ABE93060.2|	CD9/CD37/CD63 antigen [Medicago truncatula]

Contig185	gb|AC134322.25|	Medicago truncatula clone mth2-17i21, complete sequence
Contig198	gb|ABE80069.1|	Root cap [Medicago truncatula]
Contig204	gb|ABE78019.1|	Concanavalin A-like lectin/glucanase [Medicago truncatula]
Contig205	ref|NP_001042188.1|	putative aspartate aminotransferase [Oryza sativa (japonica cultivar-group)]
Contig216	gb|AAL06505.1|	At2g43970/F6E13.10 [Arabidopsis thaliana]
Contig225	gb|ABE87946.1|	CAP protein [Medicago truncatula]
Contig253	gb|BT013115.1|	Lycopersicon esculentum clone 114402R, mRNA sequence
Contig276	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene
Contig279	dbj|AK247755.1|	Solanum lycopersicum cDNA, clone: LEFL2005N13, HTC in fruit
Contig300	gb|EF416175.1|	Sesbania rostrata cDNA-AFLP fragment 043BT43M21-641.8 genomic sequence
Contig307	emb|CR932967.2|	Medicago truncatula chromosome 5 clone mth2-49e21
Contig318	gb|AY232720.1|	Fusarium oxysporum f. sp. vasinfectum strain X515-II Foxy transposable element, partial sequence
Contig334	emb|CAE55867.1|	Fot5 transposase [Fusarium oxysporum f. sp. lycopersici]
Contig336	ref|NP_001048662.1|	Os03g0102400 [Oryza sativa (japonica cultivar-group)] dbj|BAF10576.1| Os03g0102400 [Oryza sativa (japonica cultivar-group)]
Contig349	gb|AY232723.1|	Fusarium oxysporum f. sp. vasinfectum strain Ag149-1 Foxy transposable element, partial sequence
Contig351	gb|AY232720.1|	Fusarium oxysporum f. sp. vasinfectum strain X515-II Foxy transposable element, partial sequence
Contig355	gb|ABE80340.1|	TonB box, N-terminal; Tetratricopeptide-like helical [Medicago truncatula]
Contig357	dbj|BAE71282.1|	putative receptor-like GPI-anchored protein 2 [Trifolium pratense]
Contig358	gb|ABE77790.2|	Patatin [Medicago truncatula]
Contig361	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene, shh1 gene and ORF2 (partial)
Contig368	gb|AC145156.61|	Medicago truncatula clone mth2-7h6, complete sequence
Contig370	ref|NP_001054396.1|	putative ankyrin protein [Oryza sativa (japonica cultivar-group)]
Contig386	gb|AC140850.20|	Medicago truncatula clone mth2-11i23, complete sequence
Contig390	gb|ABD32881.1|	Nascent polypeptide-associated complex NAC; UBA-like [Medicago truncatula]
Contig398	gb|ABE88379.1|	Nascent polypeptide-associated complex NAC [Medicago truncatula]
Contig399	gb|ABE82627.1|	Argonaute and Dicer protein, PAZ; Stem cell self-renewal protein Piwi [Medicago truncatula]
Contig407	gb|AC140914.20|	Medicago truncatula clone mth2-18h17, complete sequence
Contig408	ref|XM_381275.1|	Gibberella zeae PH-1 chromosome 1 RAN_BRUMA GTP-binding nuclear protein RAN/TC4 (FG01099.1) partial mRNA
Contig419	ref|NP_850995.1|	binding [Arabidopsis thaliana] gb|AAG51417.1|AC009465_17 unknown protein
Contig420	emb|CR940305.16|	M.truncatula DNA sequence from clone MTH2-28N4 on chromosome 3
Contig424	gb|ABR25719.1|	ADP-ribosylation factor 1 [Oryza sativa (indica cultivar-group)]
Contig468	gb|AC148528.15|	Medicago truncatula clone mth2-53h4, complete sequence
Contig487	dbj|BAB43813.1|	CaNAG2 [Candida albicans] dbj|BAB43820.1| CaNAG2 [Candida albicans]
Contig509	gb|ABE93756.1|	Longin-like [Medicago truncatula]
Contig52	gb|AAD32141.1|	Nt-gh3 deduced protein [Nicotiana tabacum]
Contig524	gb|ABE93901.1|	Nonaspanin (TM9SF) [Medicago truncatula]
Contig525	dbj|BAA25187.1|	ARG10 [Vigna radiata]
Contig528	gb|AC124218.18|	Medicago truncatula clone mth2-30b20, complete sequence
Contig547	gb|AC139745.35|	Medicago truncatula clone mth2-17d15, complete sequence
Contig548	emb|CR936368.12|	M.truncatula DNA sequence from clone MTH2-4O11 on chromosome 3
Contig565	gb|AC134822.19|	Medicago truncatula clone mth2-15j20, complete sequence
Contig581	emb|CAA66108.1|	specific tissue protein 1 [Cicer arietinum]
Contig604	gb|AC166897.12|	Medicago truncatula clone mth2-64l14, complete sequence
Contig607	ref|NM_123209.3|	Arabidopsis thaliana GRF3 (GENERAL REGULATORY FACTOR 3)
Contig633	ref|NP_001047515.1|	Os02g0634800 [Oryza sativa (japonica cultivar-group)] dbj|BAD25096.1| putative ubiquitin-conjugating enzyme E2
Contig636	emb|AJ250814.1	Fusarium oxysporum f. sp. lycopersici insertion sequence Foxy
Contig639	ref|NP_001060360.1|	Transcription elongation factor 1 homolog
Contig647	gb|ABE87918.1|	Cupin region [Medicago truncatula]
Contig67	gb|ABN08080.1|	ATA15 protein, putative [Medicago truncatula]

Contig671	ref|XM_381601.1|	Gibberella zeae PH-1 chromosome 1 PMA1_NEUCR Plasma membrane ATPase (Proton pump)
Contig690	emb|CAA78515.1|	dehydrin-cognate [Pisum sativum]
Contig70	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene
Contig716	emb|CU075908.9|	M.truncatula DNA sequence from clone MTH2-31B23 on chromosome 3
Contig717	gb|ABE77920.1|	Rhodanese-like [Medicago truncatula] gb|ABE92275.1| Rhodanese-like [Medicago truncatula]
Contig722	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene
Contig73	emb|AJ608703.3|	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene, six2 gene, shh1 gene and ORF2 (partial)
Contig753	gb|AC187294.1|	Glycine max clone gmw1-105h23, complete sequence
Contig805	gb|AY874423.1|	Fusarium oxysporum voucher VPRI 19292 mitochondrion, partial genome
Contig818	gb|ABI34274.1|	IS10 transposase, putative [Lycopersicon esculentum]
Contig820	emb|CT027663.2|	Medicago truncatula chromosome 5 clone mth2-139g23
Contig822	ref|NP_001065513.1|	putative urea active transport protein [
Contig848	gb|AC139854.21|	Medicago truncatula clone mth2-16e16, complete sequence
Contig850	gb|ABE85139.1|	TPR repeat [Medicago truncatula]
Contig859	ref|NP_564405.1|	LOL1 (LSD ONE LIKE 1) [Arabidopsis thaliana]
Contig866	emb|CAA80983.1|	narbonin [Vicia narbonensis]
Contig882	gb|ABE84580.1|	fiber protein Fb15 [Medicago truncatula]
Contig893	gb|AY874423.1|	Fusarium oxysporum voucher VPRI 19292 mitochondrion, partial genome
Contig897	gb|AAR14273.1|	predicted protein [Populus alba x Populus tremula]
Contig902	ref|XP_001257448.1|	cupin domain protein [Neosartorya fischeri NRRL 181]
Contig905	sp|Q9ZSW9|TCTP_HEVBR	Translationally-controlled tumor protein homolog (TCTP)
Contig910	gb|ABO79239.1|	KOB1 , putative [Medicago truncatula]
Contig922	emb|CT009479.4|	M.truncatula DNA sequence from clone MTH2-164E3 on chromosome
Contig939	ref|XP_001216717.1|	predicted protein [Aspergillus terreus NIH2624]
Contig956	gb|AC174326.12|	Medicago truncatula chromosome 8 clone mth2-107i14, complete sequence
Contig961	gb|AAD32880.1|	F14N23.18 [Arabidopsis thaliana]
Contig972	gb|ABO82310.1|	Dimerisation [Medicago truncatula]

Nucleotide binding proteins

CaF1_JIE_27_C_04	gb|AAC32610.1|	ras-like small monomeric GTP-binding protein [Avena fatua]
CaF1_JIE_32_G_03	emb|CAA06731.1|	GDP dissociation inhibitor [Cicer arietinum]
CaF1_WIE_46_D_06	gb|ABP03326.1|	Calcium-binding EF-hand; Ras small GTPase, Rho type [Medicago truncatula]
CaF1_WIE_52_C_06	gb|ABE89416.1|	TGS; Small GTP-binding protein domain [Medicago truncatula]
Contig287	emb|CAA98163.1|	RAB1X [Lotus japonicus]
Contig290	dbj|BAA02118.1|	GTP-binding protein [Pisum sativum]
Contig5	dbj|BAA02115.1|	GTP-binding protein [Pisum sativum] prf|2001457G GTP-binding protein
Contig751	gb|AAF65513.1|	GTP-binding protein [Capsicum annuum]
Contig790	gb|ABE91892.1|	Ras small GTPase, Ras type; Small GTP-binding protein domain [Medicago truncatula]
Contig871	gb|ABE78982.1|	Ras small GTPase, Rab type [Medicago truncatula]
Contig885	gb|AAL49957.1|	GTP cyclohydrolase I [Lycopersicon esculentum]

Photosynthesis

CaF1_JIE_29_E_03	gb|AAG26305.1|	photosystem II CP47 protein [Trochodendron aralioides]
CaF1_JIE_36_B_01	sp|P06452|ATPI_PEA	Chloroplast ATP synthase a chain precursor (ATPase subunit IV)
CaF1_WIE_54_B_02	emb|CAA10290.1|	ribulose 1,5-bisphosphate carboxylase small subunit [Cicer arietinum]
Contig174	gb|ABE84246.1|	Chlorophyll A-B binding protein [Medicago truncatula]
Contig227	emb|AJ404642.1|	Cicer arietinum partial ORF for NAD-dependent malic enzyme
Contig785	gb|ABO87610.1|	chloroplast ferredoxin-NADP+ reductase [Pisum sativum]
Contig81	emb|CAA10284.1|	chlorophyll a/b binding protein [Cicer arietinum]
Contig857	dbj|BAC22609.1|	41 kD chloroplast nucleoid DNA binding protein (CND41) [Nicotiana sylvestris]
Contig904	dbj|BAE71227.1|	putative rubisco subunit binding-protein alpha subunit [Trifolium pratense]
Contig916	gb|ABE80903.2|	Light chain 3 (LC3) [Medicago truncatula]
Contig98	ref|NP_563815.1|	PSAO (photosystem I subunit O) [Arabidopsis thaliana] gb|AAK93637.1| unknown protein

Post translational modification, protein turn over and chaperons

CaF1_JIE_14_D_03	ref|NP_197749.2|	DNAJ heat shock N-terminal domain-containing protein [Arabidopsis thaliana]
CaF1_JIE_15_A_03	gb|ABO83982.1|	protein binding , related [Medicago truncatula]
CaF1_JIE_29_H_02	gb|ABD28395.1|	Nucleoporin interacting component; Protein prenyltransferase [Medicago truncatula]
CaF1_JIE_35_D_08	gb|ABG22120.1|	polyprotein [Cynara scolymus]
CaF1_WIE_41_A_05	gb|ABE85043.1|	HSP20-like chaperone [Medicago truncatula]
CaF1_WIE_42_H_11	gb|ABE81079.1|	PpiC-type peptidyl-prolyl cis-trans isomerase [Medicago truncatula]
CaF1_WIE_44_D_03	gb|ABP02242.1|	Cyclin-like F-box [Medicago truncatula]
CaF1_WIE_44_D_07	emb|AJ009878.1|CAR9878	Cicer arietinum mRNA for cysteine proteinase
CaF1_WIE_46_B_03	gb|ABE80335.1|	Kunitz inhibitor ST1-like [Medicago truncatula]
CaF1_WIE_47_F_03	ref|NP_174675.2|	STT3B (STAUROSPORIN AND TEMPERATURE SENSITIVE 3-LIKE B); oligosaccharyl transferase [Arabidopsis thaliana]
CaF1_WIE_47_H_03	gb|ABE85038.2|	Peptidase S1 and S6, chymotrypsin/Hap; Immunoglobulin/major histocompatibility complex; AAA ATPase, central region; SMAD/FHA [Medicago truncatula]
CaF1_WIE_49_F_03	sp|Q9SXU1|PSA7_CICAR	Proteasome subunit alpha type 7 (20S proteasome alpha subunit D)
CaF1_WIE_51_E_07	gb|ABD32889.1|	AAA ATPase; 26S proteasome subunit P45 [Medicago truncatula]
CaF1_WIE_52_B_09	gb|ABD32628.1|	Granulin; Peptidase C1A, papain [Medicago truncatula]
CaF1_WIE_52_C_05	gb|ABP03389.1|	20S proteasome, A and B subunits [Medicago truncatula]
CaF1_WIE_52_G_04	ref|NP_197500.1|	RPT6A (regulatory particle triple-A 6A); ATPase [Arabidopsis thaliana]
CaF1_WIE_53_D_07	ref|NP_001077535.1|	UBX domain-containing protein [Arabidopsis thaliana]
CaF1_WIE_55_D_03	gb|ABE78927.1|	Peptidase T1A, proteasome beta-subunit [Medicago truncatula]
Contig103	gb|AAQ18141.1|	poly(A)-binding protein C-terminal interacting protein 6 [Cucumis sativus]
Contig130	gb|ABO82668.1|	Proteinase inhibitor I25, cystatin [Medicago truncatula]
Contig150	gb|AAM22748.1|	polyubiquitin 2 [Deschampsia antarctica]
Contig156	gb|ABE78557.1|	Oligosaccharyl transferase, STT3 subunit [Medicago truncatula]
Contig199	gb|AAM64316.1|	multicatalytic endopeptidase complex, proteasome precursor, beta subunit [Arabidopsis thaliana]
Contig226	gb|AAC49013.1|	polyubiquitin containing 7 ubiquitin monomers
Contig244	gb|ABE84531.1|	Polyadenylate binding protein, human types 1, 2, 3, 4 [Medicago truncatula]
Contig325	gb|ABE77684.1|	PpiC-type peptidyl-prolyl cis-trans isomerase; Rhodanese-like [Medicago truncatula]
Contig344	emb|CAA51821.1|	ubiquitin conjugating enzyme E2 [Solanum lycopersicum] gb|ABB02644.1| ubiquitin conjugating enzyme E2-like [Solanum tuberosum]
Contig363	gb|AAZ98791.1|	cystatin [Medicago sativa]
Contig392	emb|CAA83548.1|	PsHSC71.0 [Pisum sativum]
Contig421	gb|ABE85685.1|	Cyclin-like F-box [Medicago truncatula]
Contig491	ref|NP_563657.1|	CLPP5 (NUCLEAR ENCODED CLP PROTEASE 1); endopeptidase Clp [Arabidopsis thaliana]
Contig634	dbj|BAA25755.1|	vcCyP [Vicia faba]
Contig655	gb|AAY54007.1|	subtilisin-like protease [Arachis hypogaea]
Contig663	gb|ABE78148.1|	Heat shock protein DnaJ [Medicago truncatula]
Contig693	gb|ABK42077.1|	ubiquitin extension protein [Capsicum annuum]
Contig702	gb|ABE86297.1|	Ubiquitin; Apoptosis regulator Bcl-2 protein, BAG [Medicago truncatula]
Contig721	emb|CAA04447.1|	DnaJ-like protein [Medicago sativa] gb|AAC19391.1| DnaJ-like protein MsJ1 [Medicago sativa]
Contig735	emb|CAA48140.1|	ubiquitin [Antirrhinum majus]
Contig807	emb|X93220.2|CACG2	Cicer arietinum partial mRNA for cysteine proteinase (cacG2 gene)
Contig851	gb|ABE90991.1|	Heat shock protein DnaJ, N-terminal; Tetratricopeptide-like helical [Medicago truncatula]
Contig874	gb|ABD32628.1|	Granulin; Peptidase C1A, papain [Medicago truncatula]
Contig925	dbj|BAD24713.1|	protein disufide isomerase-like protein [Glycine max]
Contig931	emb|CAA08906.1|	cysteine proteinase [Cicer arietinum]
Contig968	sp|Q41649|FKB15_VICFA	FK506-binding protein 2 precursor (Peptidyl-prolyl cis-trans isomerase)
Contig971	gb|AAT74554.1|	QM family protein [Caragana jubata]

Replication, recombination and repair

CaF1_JIE_42_G_01	gb|AAT38758.1|	Putative gag-pol polyprotein, identical [Solanum demissum]

Contig189   gb|AAT58770.1| putative polyprotein [Oryza sativa (japonica cultivar-group)]

Contig310   gb|AAK54302.1| putative helicase [Oryza sativa (japonica cultivar-group)]

Contig391   gb|ABE94111.1| DNA repair protein RadA; Peptidase M41, FtsH [Medicago truncatula]

RNA processing and modification

Clone	Accession	Description
CaF1_JIE_12_D_06	gb|ABE79750.1|	AAA ATPase, central region; DEAD/DEAH box helicase, N-terminal [Medicago truncatula]
CaF1_JIE_28_F_10	gb|AAF75791.1|	DEAD box protein P68 [Pisum sativum]
CaF1_JIE_40_H_08	gb|ABE88390.1|	Pre-mRNA processing ribonucleoprotein, binding region; NOSIC [Medicago truncatula]
CaF1_WIE_43_F_10	emb|Y16672.1|MSY16672	Medicago sativa mRNA for putative arginine/serine-rich splicing factor
CaF1_WIE_48_H_11	gb|ABE89510.1|	U2 auxiliary factor small subunit [Medicago truncatula]
Contig17	gb|ABE93398.1|	RNA-binding region RNP-1 (RNA recognition motif); HMG-I and HMG-Y, DNA-binding [Medicago truncatula]
Contig176	ref|NP_851141.1|	RNA recognition motif (RRM)-containing protein [Arabidopsis thaliana]
Contig206	gb|ABE82737.1|	RNA-binding region RNP-1 (RNA recognition motif) [Medicago truncatula]
Contig235	gb|AAN74635.1|	DEAD box RNA helicase [Pisum sativum]
Contig321	gb|ABE90931.1|	Sm-like protein [imported] - Arabidopsis thaliana [Medicago truncatula]
Contig47	gb|AAN74635.1|	DEAD box RNA helicase [Pisum sativum]
Contig779	gb|AAN74636.1|	DEAD box RNA helicase [Pisum sativum]
Contig883	gb|ABE82737.1|	RNA-binding region RNP-1 (RNA recognition motif) [Medicago truncatula]

Secondary metabolism

Clone	Accession	Description
CaF1_JIE_27_B_02	gb|AAN31890.1|	putative sterol-C-methyltransferase [Arabidopsis thaliana]
CaF1_JIE_37_F_08	emb|CAA10131.1|	chalcone synthase [Cicer arietinum]
CaF1_JIE_38_C_11	gb|AAY86360.1|	cinnamoyl-CoA reductase [Acacia mangium x Acacia auriculiformis]
CaF1_JIE_42_G_06	sp|Q00016|IFR_CICAR	Isoflavone reductase (IFR) (2'-hydroxyisoflavone reductase)
CaF1_WIE_43_D_11	gb|AAT94364.1|	chalcone isomerase 1B2 [Glycine max]
CaF1_WIE_53_B_09	dbj|BAF34844.1|	pterocarpan reductase [Lotus japonicus]
Contig178	gb|ABC94943.1|	squalene epoxidase [Medicago sativa]
Contig215	gb|L46857.1|ALFCAD1B	Medicago sativa cinnamyl-alcohol dehydrogenase (cad1) mRNA
Contig232	gb|ABO93014.1|	putative sterol desaturase [Solanum tuberosum]
Contig323	gb|AAZ29733.1|	phenylalanine ammonia lyase [Trifolium pratense]
Contig365	gb|ABC94943.1|	squalene epoxidase [Medicago sativa]
Contig393	gb|ABC94943.1|	squalene epoxidase [Medicago sativa]
Contig429	sp|O81928|TCMO_CICAR	Trans-cinnamate 4-monooxygenase (Cinnamic acid 4-hydroxylase)
Contig482	sp|Q9SML4|CHS1_CICAR	Chalcone synthase 1 (Naringenin-chalcone synthase 1)
Contig483	sp|Q96423|TCMO_GLYEC	Trans-cinnamate 4-monooxygenase (Cinnamic acid 4-hydroxylase)
Contig550	gb|AAM65672.1|	4-coumarate-CoA ligase-like protein [Arabidopsis thaliana]
Contig616	gb|AAZ29733.1|	phenylalanine ammonia lyase [Trifolium pratense]
Contig719	gb|AAQ20041.1|	isoflavone 3'-hydroxylase [Medicago truncatula]
Contig738	dbj|BAA76417.1|	chalcone reductase [Cicer arietinum]
Contig780	emb|CAA10131.1|	chalcone synthase [Cicer arietinum]
Contig957	sp|P28012|CFI1_MEDSA	Chalcone--flavonone isomerase 1 (Chalcone isomerase 1)

Signaling

Clone	Accession	Description
CaF1_JIE_12_E_05	dbj|BAB09853.1|	ER66 protein-like [Arabidopsis thaliana]
CaF1_JIE_12_E_08	dbj|BAC07504.2|	receptor-like protein kinase [Nicotiana tabacum]
CaF1_JIE_14_A_04	gb|AAT68475.1|	calcium/calmodulin-regulated receptor-like kinase [Medicago sativa]
CaF1_JIE_15_B_11	gb|ABE92129.1|	Protein kinase [Medicago truncatula]
CaF1_JIE_19_F_10	gb|ABO81298.1|	Protein kinase; TonB box, N-terminal [Medicago truncatula]
CaF1_JIE_20_A_11	ref|NP_195483.1|	HHP4 (heptahelical protein 4); receptor [Arabidopsis thaliana]
CaF1_JIE_29_F_07	gb|ABO78866.1|	WD40-like [Medicago truncatula]
CaF1_JIE_29_G_07	gb|ABN09164.1|	Protein phosphatase 2C-like [Medicago truncatula]
CaF1_JIE_33_A_07	gb|AAF78397.1|	Contains similarity to a putative protein T2J13.100 gi|6522560 from Arabidopsis thaliana BAC T2J13 gb|AL132967
CaF1_JIE_33_D_01	gb|ABE93018.1|	cAMP response element binding (CREB) protein; Prefoldin [Medicago truncatula]
CaF1_JIE_34_H_11	gb|ABE77486.1|	Protein kinase [Medicago truncatula]
CaF1_JIE_35_E_02	gb|ABN08649.1|	C2 [Medicago truncatula]
CaF1_JIE_38_B_08	dbj|BAD95892.1|	Ser/Thr protein kinase [Lotus japonicus]
CaF1_WIE_41_C_08	gb|ABE93295.1|	Leucine-rich repeat; Leucine-rich repeat, cysteine-containing subtype [Medicago truncatula]

CaF1_WIE_41_G_03	gb|ABK06434.1|	flag-tagged protein kinase domain of putative mitogen-activated protein kinase kinase kinase [synthetic construct]
CaF1_WIE_41_H_01	emb|Y17329.1|PSA17329	Pisum sativum mRNA for calnexin
CaF1_WIE_46_H_04	gb|ABO78866.1|	WD40-like [Medicago truncatula]
CaF1_WIE_47_C_03	gb|ABE93337.1|	Protein kinase [Medicago truncatula]
CaF1_WIE_49_B_05	gb|ABG73621.1|	leucine-rich repeat receptor-like kinase [Populus tomentosa]
CaF1_WIE_50_B_03	gb|AAO49473.1|	putative serine/threonine kinase [Vitis vinifera]
CaF1_WIE_50_G_06	ref|XP_001267009.1|	calmodulin [Neosartorya fischeri NRRL 181]
CaF1_WIE_51_C_11	gb|ABE78903.1|	Protein kinase [Medicago truncatula]
CaF1_WIE_53_A_02	gb|AAG29593.1	Ser/Thr specific protein phosphatase 2A A regulatory subunit alpha isoform [Medicago sativa subsp. x varia]
CaF1_WIE_53_H_05	gb|ABE84183.2|	Protein kinase [Medicago truncatula]
CaF1_WIE_54_C_09	ref|NP_568217.1|	BolA-like family protein [Arabidopsis thaliana]
CaF1_WIE_55_G_01	gb|ABP02712.1|	14-3-3 protein [Medicago truncatula]
Contig112	gb|ABD32712.1|	Response regulator receiver; CCT [Medicago truncatula]
Contig114	gb|ABE94388.1|	WD-40 repeat [Medicago truncatula]
Contig133	gb|AAF64040.1	14-3-3-like protein [Glycine max]
Contig161	gb|ABE86996.1|	Protein kinase [Medicago truncatula]
Contig191	gb|AAT08753.1|	LRR [Hyacinthus orientalis]
Contig194	gb|AAK11734.1|	serine/threonine/tyrosine kinase [Arachis hypogaea]
Contig20	dbj|BAE71262.1|	putative protein kinase APK1A [Trifolium pratense]
Contig228	dbj|BAE71265.1|	putative serine/threonine protein kinase-like protein [Trifolium pratense]
Contig234	gb|ABP02851.1|	Calcium-binding EF-hand [Medicago truncatula]
Contig270	gb|ABE85490.2|	Response regulator receiver; CCT [Medicago truncatula]
Contig280	gb|ABO78021.1|	Curculin-like (mannose-binding) lectin [Medicago truncatula]
Contig302	ref|NP_201198.1|	leucine-rich repeat transmembrane protein kinase, putative [Arabidopsis thaliana]
Contig332	sp|Q53IP3|MBF1_GIBFU	Multiprotein-bridging factor 1 emb|CAG28684.1| multiprotein bridging factor [Gibberella fujikuroi]
Contig353	emb|CAB43932.1|	putative serine/threonine-specific receptor protein kinase [Arabidopsis thaliana]
Contig367	gb|ABO80441.1|	EPS15 homology (EH) [Medicago truncatula]
Contig372	sp|P04353|CALM_SPIOL	Calmodulin (CaM)
Contig377	gb|AAW31901.1|	calcium-dependent/calmodulin-independent protein kinase isoform 3 [Cicer arietinum]
Contig396	sp|P15001|PHYA_PEA	Phytochrome A gb|AAA33682.1| phytochrome [Pisum sativum]
Contig405	gb|AAK68074.1	somatic embryogenesis receptor-like kinase 3 [Arabidopsis thaliana]
Contig406	ref|NP_200932.2|	protein binding [Arabidopsis thaliana] dbj|BAB08479.1| leucine-rich repeat disease resistance protein-like [Arabidopsis thaliana]
Contig425	sp|P42654|1433B_VICFA	14-3-3-like protein B (VFA-1433B) emb|CAA88416.1| 14-3-3 brain protein homolog [Vicia faba]
Contig527	gb|ABQ95992.1|	14-3-3-like protein [Cicer arietinum] gb|ABQ95994.1| 14-3-3-like protein [Cicer arietinum]
Contig568	dbj|BAB10271.1|	ankyrin-like protein [Arabidopsis thaliana]
Contig645	gb|ABE81462.1|	Protein kinase [Medicago truncatula]
Contig668	gb|AAK11734.1|	serine/threonine/tyrosine kinase [Arachis hypogaea]
Contig670	gb|AAL17948.1|	phosphoinositide-specific phospholipase C [Medicago truncatula]
Contig752	ref|NP_568466.1|	CIPK25 (CBL-INTERACTING PROTEIN KINASE 25); kinase [Arabidopsis thaliana]
Contig815	gb|ABE90033.1|	Protein kinase [Medicago truncatula]
Contig86	gb|ABE83899.1|	Serine/threonine protein kinase, active site [Medicago truncatula]
Contig89	emb|CAC36428.1|	mitogen activated protein kinase [Gibberella fujikuroi]
Contig896	gb|AAT37529.1|	purple acid phosphatase 1 [Solanum tuberosum]
Contig903	gb|ABD28527.1|	Protein kinase [Medicago truncatula]
Contig926	gb|ABE92804.1|	Curculin-like (mannose-binding) lectin [Medicago truncatula]
Contig928	ref|NP_565408.1|	protein kinase, putative [Arabidopsis thaliana]
Contig947	gb|ABQ95992.1|	14-3-3-like protein [Cicer arietinum]

Stress

CaF1_JIE_18_D_08	gb|AAV66464.1|	drought responsive element binding protein [Glycine soja]
CaF1_JIE_36_F_10	ref|NP_181934.1|	late embryogenesis abundant family protein / LEA family protein [Arabidopsis thaliana]
CaF1_WIE_42_G_01	emb|AJ299396.1|	Cicer arietinum partial mRNA for putative extensin (ORF)
CaF1_WIE_47_D_03	gb|AAY22204.1|	putative aquaporin [Phaseolus vulgaris]

CaF1_WIE_48_F_09	gb|ABE90758.1|	Annexin [Medicago truncatula]
CaF1_WIE_51_D_06	gb|AAB71830.1|	annexin [Lavatera thuringiaca]
CaF1_WIE_52_B_07	gb|AF397032.2|	Pisum sativum clone PsEXT3.28 root nodule extensin mRNA, partial cds
Contig27	emb|CAB61749.1|	putative water channel protein [Cicer arietinum]
Contig319	emb|CAG14983.1|	putative universal stress protein [Cicer arietinum]
Contig345	ref|NP_566991.2|	universal stress protein (USP) family protein [Arabidopsis thaliana]
Contig350	sp|Q39458|MT1_CICAR	Metallothionein-like protein 1 (MT-1) emb|CAA65008.1| metallothionein [Cicer arietinum]
Contig431	sp|Q9FY14|TIP1_MEDTR	Probable aquaporin TIP-type (MtAQP1) emb|CAC01618.1| aquaporin [Medicago truncatula]
Contig436	gb|AAB86380.1|	aquaporin-like transmembrane channel protein [Medicago sativa]
Contig649	emb|AJ299396.1	Cicer arietinum partial mRNA for putative extensin (ORF)
Contig707	gb|AF155232.1|AF155232	Pisum sativum extensin (Ext) mRNA, complete cds
Contig720	gb|ABE84254.1|	Universal stress protein (Usp) [Medicago truncatula] gb|ABE88281.1| Universal stress protein (Usp) [Medicago truncatula]
Contig745	gb|AAK66766.1|	aquaporin protein PIP1;1 [Medicago truncatula]
Contig759	gb|AAQ21120.1|	early light inducible protein [Trifolium pratense]
Contig908	gb|ABE85050.1|	Universal stress protein (Usp) [Medicago truncatula]
Contig945	emb|AJ006770.1|	Cicer arietinum mRNA for extensin, partial
Contig958	gb|AAX18706.1|	cold-related protein Cor413 [Gossypium barbadense]
Contig959	emb|CAC08564.1|	wound-induced GSK-3-like protein [Medicago sativa subsp. x varia]

Transcription

CaF1_JIE_22_G_08	gb|ABE86660.1|	Zinc finger, C2H2-type [Medicago truncatula]
CaF1_JIE_41_D_04	emb|CAA87075.1|	heat shock transcription factor 29 [Glycine max]
CaF1_WIE_45_E_07	gb|ABH02845.1|	MYB transcription factor MYB93 [Glycine max]
CaF1_WIE_49_C_09	gb|ABE91004.1|	Zinc finger, CCCH-type; Zinc finger, RING-type [Medicago truncatula]
CaF1_WIE_52_E_05	ref|NP_181843.2|	DNA binding / transcription factor [Arabidopsis thaliana]
CaF1_WIE_53_F_03	gb|ABG90380.1|	bZIP transcription factor [Caragana korshinskii]
CaF1_WIE_54_A_01	dbj|BAB32793.1|	110 kDa 4SNc-Tudor domain protein [Pisum sativum]
Contig111	gb|ABE84970.1|	Pathogenesis-related transcriptional factor and ERF [Medicago truncatula]
Contig147	gb|ABE83604.1|	PUG; Zinc finger, C2H2-type; UBA-like [Medicago truncatula]
Contig160	dbj|BAE71188.1|	BEL1-like homeodomain transcription factor [Trifolium pratense]
Contig196	ref|NP_568725.1|	ELO3 (ELONGATA 3); N-acetyltransferase/ catalytic/ hydrogen ion transporting ATP synthase
Contig212	sp|Q8LJS2|HDT1_SOYBN	Histone deacetylase HDT1 (Histone deacetylase 2a)
Contig217	ref|NP_974856.1|	zinc finger (ZPR1-type) family protein [Arabidopsis thaliana]
Contig22	gb|ABE83191.2|	Zinc finger, CCHC-type [Medicago truncatula]
Contig223	gb|ABE93791.1|	NOT2/NOT3/NOT5 [Medicago truncatula]
Contig23	gb|AAC28907.1|	phaseolin G-box binding protein PG2 [Phaseolus vulgaris]
Contig236	gb|ABD32320.1|	DNA-directed RNA polymerase, subunit C11/M/9 [Medicago truncatula]
Contig241	emb|CAA54168.1|	HMG 1 protein [Pisum sativum]
Contig247	gb|ABD32869.1|	Zinc finger, RING-type; RINGv [Medicago truncatula] gb|ABE79868.2| Zinc finger, RING-type [Medicago truncatula]
Contig293	gb|AAK84887.1	homeodomain leucine zipper protein HDZ3 [Phaseolus vulgaris]
Contig31	ref|NP_565451.1|	nucleic acid binding / zinc ion binding [Arabidopsis thaliana] gb|AAM14872.1| Expressed protein [Arabidopsis thaliana] gb|ABD19677.1| At2g19385 [Arabidopsis thaliana]
Contig316	ref|NP_196487.1|	KIWI; DNA binding / transcription coactivator [Arabidopsis thaliana]
Contig360	gb|ABE84268.2|	Zinc finger, CCHC-type; Putative 5-3 exonuclease [Medicago truncatula]
Contig362	ref|NP_198099.1|	seryl-tRNA synthetase / serine--tRNA ligase [Arabidopsis thaliana]
Contig374	ref|NP_200914.2|	HDA05 (HISTONE DEACETYLASE5); histone deacetylase [Arabidopsis thaliana]
Contig378	emb|CAA66478.1|	transcription factor [Vicia faba var. minor]
Contig387	gb|ABO84100.1|	Something about silencing protein 10 , related [Medicago truncatula]
Contig395	gb|ABE85988.1|	Zinc finger, RING-type; Zinc finger, RanBP2-type; Zinc finger, C6HC-type [Medicago truncatula]
Contig4	gb|ABP03302.1|	Zinc finger, C2H2-type [Medicago truncatula]
Contig434	gb|AAZ38969.1|	GAMYB-binding protein [Glycine max]
Contig452	gb|AAK84885.1|	homeodomain leucine zipper protein HDZ1 [Phaseolus vulgaris]
Contig458	gb|ABP03503.1|	Pathogenesis-related transcriptional factor and ERF [Medicago truncatula]
Contig502	gb|AAX47170.1|	SHORT VEGETATIVE PHASE [Pisum sativum]

Contig571	ref|XP_391802.1|	H2B_NEUCR Histone H2B [Gibberella zeae PH-1] sp|Q4HTT2|H2B_GIBZE Histone H2B
Contig586	gb|ABH02865.1|	MYB transcription factor MYB176 [Glycine max]
Contig587	dbj|BAA19156.1|	HMG-1 [Canavalia gladiata]
Contig650	gb|ABO81948.2|	Zinc finger, GATA-type [Medicago truncatula]
Contig691	gb|ABE84071.1|	Nucleic acid-binding, OB-fold, subgroup [Medicago truncatula]
Contig756	emb|AJ006767.1|CAR6767	Cicer arietinum mRNA for histone H1
Contig83	gb|AAZ14831.1|	putative AP2-binding protein [Jatropha curcas]
Contig833	dbj|BAB32793.1|	110 kDa 4SNc-Tudor domain protein [Pisum sativum]
Contig875	gb|ABH02852.1|	MYB transcription factor MYB112 [Glycine max]
Contig891	gb|ABE86663.2|	Zinc finger, RING-type [Medicago truncatula]
Contig929	gb|ABD33016.1|	Transcription Factor IIF, Rap30/Rap74, interaction [Medicago truncatula]
Contig934	ref|NP_181843.2|	DNA binding / transcription factor [Arabidopsis thaliana]
Contig966	ref|NP_001078516.1|	histone H3.2 [Arabidopsis thaliana]
Contig970	gb|ABD32383.1|	Zinc finger, RING-type [Medicago truncatula]

Translation, ribosomal structure and biogenesis

CaF1_JIE_10_B_07	gb|ABE85391.1|	S25 ribosomal protein [Medicago truncatula]
CaF1_JIE_11_E_07	gb|ABE88774.1|	Translation factor; Elongation factor G, III and V [Medicago truncatula]
CaF1_JIE_16_G_09	ref|XP_380571.1|	RS15_PODAN 40S RIBOSOMAL PROTEIN S15 (S12) [Gibberella zeae PH-1]
CaF1_JIE_21_B_05	gb|AAK92832.1|	putative glycyl tRNA synthetase [Arabidopsis thaliana]
CaF1_JIE_25_C_01	gb|ABE94142.2|	Translation initiation factor IF5 [Medicago truncatula]
CaF1_JIE_36_H_11	gb|ABE83633.1|	Ribosomal protein L10; Ribosomal protein 60S [Medicago truncatula]
CaF1_WIE_41_H_09	gb|ABP02866.1|	Ribosomal protein L30e [Medicago truncatula]
CaF1_WIE_43_G_02	ref|XP_359800.2|	ribosomal protein L39 [Magnaporthe grisea 70-15] gb|EDJ95441.1| ribosomal protein L39 [Magnaporthe grisea 70-15] gb|EDN06404.1| ribosomal protein L39 [Ajellomyces capsulatus NAm1]
CaF1_WIE_48_E_04	emb|CAI48073.1|	60S ribosomal protein L37a [Capsicum chinense]
CaF1_WIE_49_B_04	sp|Q9MAV7|RL31_PANGI	60S ribosomal protein L31
CaF1_WIE_50_E_03	gb|AAS47511.1|	ribosomal protein S6 [Glycine max]
CaF1_WIE_51_C_09	sp|Q945F4|IF5A2_MEDSA	Eukaryotic translation initiation factor 5A-2 (eIF-5A-2)
CaF1_WIE_51_H_05	gb|AF071889.1	Prunus armeniaca 40S ribosomal protein S8 (RPS8) mRNA
CaF1_WIE_52_A_07	gb|AAP80667.1	ribosomal Pr 117 [Triticum aestivum]
CaF1_WIE_53_A_10	sp|O65731|RS5_CICAR	40S ribosomal protein S5 emb|CAA06491.1| 40S ribosomal protein S5 [Cicer arietinum]
CaF1_WIE_53_G_02	sp|P55844|RL14_PEA	Probable 60 ribosomal protein L14 (Hydroxyproline-rich glycoprotein HRGP1)
CaF1_WIE_54_D_03	dbj|BAB86847.1|	elongation factor EF-2 [Pisum sativum]
CaF1_WIE_56_H_11	gb|AAD47346.1	ribosomal protein S26 [Pisum sativum]
Contig127	gb|AAA34366.1|	ribosomal protein L41
Contig14	gb|AAT01416.1|	translation initiation factor 5A [Tamarix androssowii]
Contig218	gb|ABO78621.1|	Translation factor [Medicago truncatula]
Contig230	gb|ABO78621.1|	Translation factor [Medicago truncatula]
Contig260	emb|CAA06245.1|	elongation factor 1-alpha (EF1-a) [Cicer arietinum]
Contig285	gb|ABE81204.1|	Ribosomal protein S13 [Medicago truncatula] gb|ABE81211.1| Ribosomal protein S13 [Medicago truncatula]
Contig383	emb|CAA71881.1|	Tyrosyl-tRNA synthetase [Nicotiana tabacum]
Contig40	emb|CAI48073.1|	60S ribosomal protein L37a [Capsicum chinense]
Contig400	gb|EF672342.1|	Hypocrea sp. Z28 18S ribosomal RNA gene, partial sequence
Contig412	emb|AJ577394.1	Cicer arietinum 18S rRNA gene, 5.8S rRNA gene, IGS, ITS1 and ITS2
Contig457	gb|AAP80667.1|	ribosomal Pr 117 [Triticum aestivum]
Contig496	gb|ABE82951.1|	Translation protein SH3-like [Medicago truncatula]
Contig500	dbj|AK226272.1|	Arabidopsis thaliana mRNA for ribosomal protein S30 homolog, complete cds, clone: RAFL05-08-E11
Contig517	sp|O81361|RS8_PRUAR	40S ribosomal protein S8 gb|AAC24583.1|
Contig588	gb|AAG17879.1	60S ribosomal protein L10A [Phaseolus coccineus]
Contig589	gb|ABE82951.1|	Translation protein SH3-like [Medicago truncatula]
Contig6	sp|P62302|RS13_SOYBN	40S ribosomal protein S13 gb|AAS47510.1| ribosomal protein S13 [Glycine max]
Contig609	emb|CAD56219.1|	ribosomal protein S3a [Cicer arietinum]
Contig641	ref|NP_191771.1|	tRNA synthetase class II (G, H, P and S) family protein [Arabidopsis thaliana]

Contig66	sp|O81361|RS8_PRUAR	40S ribosomal protein S8 gb|AAC24583.1| 40S ribosomal protein S8 [Prunus armeniaca]
Contig715	gb|AAS47511.1|	ribosomal protein S6 [Glycine max]
Contig796	gb|ABE84969.1|	Ribosomal protein L19e [Medicago truncatula]
Contig80	gb|ABB29934.1|	acidic ribosomal protein P1a-like [Solanum tuberosum]
Contig802	gb|ABO79321.1|	Ribosomal protein S24e [Medicago truncatula]
Contig810	gb|ABO81713.1|	Ribosomal protein L36E [Medicago truncatula]
Contig840	gb|ABE88774.1|	Translation factor; Elongation factor G, III and V [Medicago truncatula]
Contig847	gb|ABE85391.1|	S25 ribosomal protein [Medicago truncatula] gb|ABE85967.1| S25 ribosomal protein [Medicago truncatula]
Contig849	gb|ABE82912.1|	Ribosomal protein S4, bacterial and organelle form [Medicago truncatula]
Contig889	gb|ABE90867.1|	Ribosomal L18ae protein [Medicago truncatula]
Contig90	gb|ABE79479.1|	Ribosomal protein L10E [Medicago truncatula]
Contig906	gb|ABE87516.2|	Ribosomal protein L7Ae/L30e/S12e/Gadd45 [Medicago truncatula]
Contig921	gb|ABE88774.1|	Translation factor; Elongation factor G, III and V [Medicago truncatula]
Contig943	sp|P49163|RK22_MEDSA	50S ribosomal protein L22, chloroplast precursor (CL22)
Contig946	gb|ABE92854.1|	Ribosomal L22e protein [Medicago truncatula]
Contig97	emb|CAB76914.1|	60S ribosomal protein L6 [Cicer arietinum]
Contig99	gb|AAK95391.1|	ribosomal protein L2 [Gossypium arboreum]

Source: Ashraf *et al. BMC genomics*, 2009, in press

function databases like Swiss–Prot 2 GO (www.geneontology.org), metacyc (www.metacyc.org) and COG (www.ncbi.nlm.nih.gov/COG). However, the classification of transcripts is only tentative, since the biological function of many genes identified has not yet been established experimentally. The important functional classes are described below:

Genes involved in metabolism

The functional class of metabolism comprised of about 14.56% of the unigenes and represented the largest class apart from NSH and hypothetical proteins. The genes present in this class represented several biochemical pathways such as carbohydrate, fatty acid, energy and protein metabolism besides nitrogen and nucleotide metabolism. The genes involved in carbohydrate metabolism comprised alcohol dehydrogenase (ADH), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), enolase, fructose-bisphosphate aldolase, malate dehydrogenase, phosphoenol pyruvate (PEP) carboxylase and triosephosphate isomerase. Most of these enzymes are known to show altered expression in response to stress (Laxalt *et al.*, 1996; Cushman *et al.*, 1989). Though very little is known about the direct role of these proteins in defense, many such enzymes are involved in maintaining metabolite pool that may drive the metabolic processes to overcome such stresses. Polygalactouronase inhibiting proteins (PGIP) were present in both the susceptible and resistant genotypes. The PGIPs are cell wall proteins, which act against fungal polygalactouronases that are important pathogenecity factors (De Lorenzo and Ferrari, 2002).

The genes involved in lipid biosynthesis, for example, acyl Co-A synthetase and acyl Co-A binding proteins were found to be present in both the genotypes. The enzyme acyl Co-A synthetase has been shown to be induced in response to compatible and incompatible interactions between *Xanthomonas* and *Capsicum annum* (Lee *et al.*, 2001), while acyl Co-A binding proteins play an important role in intracellular transport and formation of acyl Co-A pools (Knudsen *et al.*, 1999). Acyl Co-A binding proteins also interact with ethylene-responsive element binding proteins suggesting their possible role in defense signaling (Li and Chye, 2004). ESTs encoding desaturases with varied substrate specificity, for example, oleate desaturase, sterol desaturase and putative desaturase-like protein were found in the *Ca*Unigene set. Desaturases catalyze the formation of double bonds in lipids thereby increasing

the lipid fluidity, which might be required as an adjustment associated with membrane stability under stress conditions. ESTs corresponding to other lipid metabolism genes such as serine C-palmitoyltransferase, were also found to be present in stressed tissues. Identification of sterol metabolism related enzymes such as sterol methyl transferases and oxysterol binding proteins suggest the occurrence of lipid modifications during pathogen stress.

Several amino acids act as precursors for the synthesis of specialized metabolites during varied cellular adaptations. ESTs encoding enzymes involved in amino acid metabolism, for example, arginine decarboxylase, aspartate aminotransferase, and cysteine synthase were identified in both the genotypes. Recent evidence suggests that arginine decarboxylase, involved in putrescine biosynthesis, is induced in response to various environmental stresses (Liu *et al.*, 2006). Proline dehydrogenase and many proline-rich proteins were identified in resistant genotype. Proline dehydrogenase is the rate-limiting enzyme in proline degradation and serves important functions in stress responses (Weltmeier *et al.*, 2006). Furthermore, we found accumulation of transcripts encoding branched chain amino acid (BCAA) aminotransferase, which serves in detoxification mechanism that maintains the pool of free branched chain amino acids at low and non toxic levels under dehydration. *Fusarium* grows in the vascular system of the plant and the mycelial growth of the fungus may cause blockage of vascular tissue, which mimics the symptoms of dehydration and therefore, may lead to the accumulation of BCAA aminotransferase.

Nucleotide metabolism related genes such as nuleoside diphosphate kinase (NDK), adenine nucleotide translocator and polynucleotide phosphorylase were also identified in both the genotypes. Previously, it has been shown that the overexpressing NDK provides higher ability to eliminate H_2O_2, indicating its potential role in the management of reactive oxygen species under stress (Fukamatsu *et al.*, 2001).

Genes involved in cell signaling

Genes involved in cell signaling are known to regulate many cellular responses in an organism. Our data showed the presence of different types of protein kinases in both susceptible and tolerant chickpea genotypes. These include serine/threonine protein kinase, somatic embryogenesis receptor-related kinase, and putative serine/threonine-specific receptor protein kinase. Protein kinases are involved in disease response via a

signaling cascade presumably conserved in plants, insects and mammals (Zhang and Klessig, 2001). Somatic embryogenesis related kinases are associated with brassinosteroid signaling pathway that plays a key role in plant defense (Karlova *et al.*, 2006). The EST data revealed presence of many components of calcium signaling such as calmodulin, calcium dependent calmoludin independent kinase, CIP kinase and calcium/calmodulin-regulated receptor-like kinase indicating an elaborate role of calcium signaling in immune response. The subsets of 14-3-3 proteins were also identified in both the susceptible and resistant genotypes. The interaction of fusicoccin, the fungal toxin, secreted by *Fusarium,* with 14-3-3s causes membrane hyperpolarization through activation of plasma membrane H^+ATPase (Roberts and Bowles, 1999).

Genes involved in transcription

The functional class of transcription associated genes, in this study, constituted about 3.52 % of the total ESTs that comprised families of transcription factors including zinc finger, MYB, BEL1-like homeodomain transcription factor, homeodomain leucine zipper HDZ3, G-box binding PG2, Zinc finger GATA-type, putative AP2-binding protein, bHLH and WRKY. This data suggests that interplay of a broad spectrum of transcription factors possibly regulates multiple signaling cascades during immune-response. There are reports of involvement of many of the members of these transcription factor families in plant stress. For example, some members of Myb family have been found to act as positive regulators of HR response that is commonly associated with disease resistance in plants (Vailleau *et al.*, 2002). Members of bZip family of transcription factors have also been shown to play role in plant defense response (Thurow *et al.*, 2005). WRKY group of transcription factors are very well known to be involved in transcriptional reprogramming during plant immune responses (Eulgem *et al.*, 2006). Cys2/His2-type zinc-finger transcription factor encoding gene, CAZFP1, from pepper has been found to function in enhancing disease resistance and drought tolerance (Kim *et al.*, 2004).

Genes involved in translation, ribosome structure and biogenesis

In the present study, this class constituted about 4.57% of ESTs and comprised predominantly the ribosomal proteins apart from some translation initiation factors. Our results showed that genes encoding a diverse range of ribosomal proteins are

present in both susceptible and resistant libraries. However, there is not much evidence available as yet, regarding the role of these genes in plant defense. Genes encoding major ribosomal proteins in susceptible and resistant genotypes are S6, S8, S9, and L24/L26. Translation factor SUI1 homolog, and eukaryotic elongation factors, EF-1α and EF-2 play a pivotal role in protein biosynthesis. Their presence in *Ca*Unigene set is indicative of their role in immune response.

Genes involved in post translational modification, protein turn over and chaperons

4.27% of the *Ca*Unigenes correspond to proteins involved in protein modifications and turnover. The presence of genes involved in protein modifications like cysteine proteinases and cyclophilin-type peptidyl-prolyl cis-trans isomerases indicate their role in pathostress responses. Cyclophilin-type peptidyl-prolyl cis-trans isomerase is a component of photosynthetic membranes and is reported to be light responsive (Romano *et al.*, 2004). The presence of heat shock protein, DnaJ is interesting as it has been shown to be involved in different environmental stresses including high temperature and salinity treatment (Futamura *et al.*, 1999). Our results revealed the presence of a wide range of proteins involved in protein turnover that include ubiquitin conjugating enzyme E2, ubiquitin-protein ligase, 20S proteasome alpha 6 subunit, F-box protein and protein disulphide. Earlier studies have shown that genes encoding ubiquitin pathway components like E2 or ring E3 ligase are upregulated in response to pathogen derived elicitors (Dahan *et al.*, 2001). The presence of these genes suggests that there might be a need of proteome reorganization during defense responses.

Genes involved in RNA processing and modification

Various RNA binding proteins have potential to modulate gene expression and might be involved in processes like RNA metabolism, mRNA splicing, ribosome biogenesis, transport and translation. We observed the presence of dead box RNA helicase, pre-mRNA processing ribonucleoprotein binding region, splicing factor 3B subunit 10 and RNA-binding region RNP-1 during vascular wilt of chickpea. While dead box RNA helicase was reported to be involved in development and stress responses (Gong *et al.*, 2005), the role of splicing factor and RNA-binding region RNP-1 in pathostress is yet to be established.

Genes involved in cellular transport and homeostasis

Stress-induced reorganization and spatial distribution of many key metabolites in plants require efficient transport machinery. Various transport associated genes were identified in this study that included general substrate transporter, transporter-like protein, SecY protein, ABC transporter related proteins, intracellular chloride channel, phosphate transporter 5, putative polyol transporter protein 4, multidrug resistance protein and metal ion transporter. While ABC transporters and multidrug resistance genes are known to function in plant defense (Kobae *et al*, 2006; Simmons *et al*, 2003), role of other genes in plant defense needs to be explored.

Genes involved in cellular redox

The production of reactive oxygen species (ROS) via consumption of oxygen in a so called oxidative burst is one of the earliest cellular responses following successful pathogen recognition. ROS is produced as part of a complex network of signals that responds to pathogen attack and mediates multiple responses in different contexts or in response to different types of pathogens (Mittler *et al.*, 2004). The expression of several ROS associated enzymes was induced in response to chickpea wilt which included quinone oxidoreductase, glutathione S transferase, superoxide dismutase etc. These enzymes restrict the ROS dependent damage or finely tune ROS dependent signal transduction and may lead to the activation of defense response following pathogen infection.

Genes involved in energy metabolism

Energy production has an impact on the overall metabolic state and the energy supply is the key factor for the maintenance of cell intactness under various stress conditions. We observed the presence of many genes encoding different proteins related to ATP synthesis like ABC-2, AAA ATPase, putative ADP, and ATP carrier-like proteins. Alternative NAD(P)H dehydrogenase 2 and oxidase 2b were also identified. While the involvement of these proteins in abiotic stress, energy conservation and maintenance of redox potential is well established (Arnholdt-Schmitt *et al*, 2006), their exact role in plant immune response is yet to be elucidated.

Genes involved in secondary metabolism

Most secondary metabolites of phenyl propanoid pathway, including lignins, isoflavonoid-phytoalexins and other phenolic compounds are instrumental in plant's ability to enforce successful defenses against invading pathogens. In this study, several genes were identified that are associated with biosynthesis of secondary metabolites. The important enzymes in this category include phenyl ammonia lyase (PAL), chalcone synthase, chalcone isomerase, and chalcone-flavonone isomerase-1. These enzymes are known to modulate plant defense response against invading pathogens and insects (Dixon *et al.*, 2002; Zabala *et al.*, 2006). Other members of this class are caffeic acid O-methyltransferase II, isoflavone 3'-hydroxylase, squalene monooxygenase 2, trans-cinnamate 4-monooxygenase and dihydroflavonol reductase. Earlier studies have shown that transgenic rice plants overexpressing dihydroflavonol reductase can provide tolerance to biotic and abiotic stresses (Takahashi *et al.*, 2006). Caffeic acid O-methyltransferase catalyzes a key step in lignin biosynthesis (Gowri *et al.*, 1991) and thereby gives protection against pathogen attack.

Defense responsive genes

Pathogen attack is often accompanied by the accumulation of elevated levels of transcripts of disease related proteins. The ESTs encoding proteins implicated in defense responses accounted for 1.34% of the total unigene set. Most dominant candidates among the defense responsive genes were disease resistance response protein DRRG49C and PR10. Other genes identified include those encoding chitinase, non-specific lipid transfer proteins, thaumatin, and pathogenesis-related protein RH2. The involvement of these proteins in plant defense responses is well known (Kader *et al.*, 1996). Also, we found the presence of ESTs encoding proteins like dirigent and harpin-induced 1 in the resistant genotype, implying their role in pathostress response. Dirigent protein has recently been shown to be involved in lignification and hence imparting disease resistance (Ralph *et al.*, 2007).

Stress responsive genes

Many other stress induced proteins were identified and constituted around 1.84% of the total unigenes. These included extensin, universal stress protein, aquaporin, annexin, cold acclimation responsive protein and metallothionien. Many of these genes have been earlier reported to be involved in various stresses. For example,

extensin was shown to be involved in development as well as stress response in *Arabidopsis* (Merkouropoulos and Shirsat, 2003). Water channel proteins, aquaporins, play fundamental roles in transmembrane water movements in plants and are involved in stress responses (Zhang *et al.*, 2007; Jang *et al.*, 2007). Metallothioneins are a group of low molecular mass and cysteine-rich metal-binding proteins, ubiquitously found in most living organisms and play an important role in maintaining intracellular metal homeostasis (Jin *et al.*, 2006). Annexins are a multigene family in most plant species and are known to play a role in a wide variety of essential cellular processes including stress (Cantero *et al.*, 2006).

Genes involved in development/storage/dormancy and senescence

The genes involved in development/storage/dormancy and senescence accounted for 0.45% of the total *Ca*Unigene set. The candidate genes were those encoding enzymes associated with fruit ripening and senescence, and several storage proteins like albumin and agglutinin. While there have been reports on association of the processes of plant defense and senescence (Barth *et al.*, 2004), the involvement of ripening related proteins has never been implicated in immune responses. Seed storage proteins like germin and albumin have been shown to be involved in stress responses (Agizzio *et al.*, 2003), however, role of agglutinin in plant immunity is not known.

Genes involved in cytoskeletal organization

Genes belonging to this class constituted around 0.79% of the unigene set. Cytoskeleton is thought to contribute to the establishment of effective barriers at the cell periphery against pathogen ingression. Pathogen attack leads to reorganization of cytoskeleton which leads to pivotal role in defense response (Miklis *et al.*, 2007). Substantiating this phenomenon, several structural proteins were identified that include actin, microtubule bundling polypeptide, and beta tubulin, besides genes associated with cytoskeletal reorganization like actin-depolymerizing factor and putative spindle disassembly related genes.

Genes involved in hormone responses

Salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) play key roles in developmental regulation and stress responses through cross communicating signal transduction pathways. These hormones accumulate in response to pathogen infection

53

and in turn lead to the activation of distinct sets of defense related genes (Bostock, 2005). We observed presence of ethylene and JA signaling pathway genes, for example, ethylene receptor, ethylene responsive transcoactivator and coronatine-insensitive 1 in the both resistant and susceptible genotypes. In addition, we found auxin signaling related proteins such as, auxin-responsive SAUR, auxin-regulated dual specificity cytosolic kinase, aux/IAA protein and putative auxin-induced protein, IAA12. Brassinosteroid pathway associated gene BRU1 precursor was also present.

Genes encoding nucleotide binding proteins

In our study, GTP binding proteins were found to be the major class of nucleotide binding proteins in susceptible as well as resistant genotypes. The stress-responsive predominant GTP binding proteins were of Ras, Rab and Ran type. GTP binding proteins are known to play diverse roles including stress (Fujiwara *et al.*, 2006).

Genes involved in cell cycle control and cell division

Cell division and cell cycle progression in plants is often altered in response to various environmental stresses (Potters *et al.*, 2007). Many cell division and cell-cycle related proteins, for example, putative kinetochore, cell division protein FtsZ and cdc2MsF were identified in this study suggesting pathostress responsive alterations of cell cycle in chickpea. Genes involved in DNA replication and repair were also identified that included putative helicase and DNA repair protein, RadA. These findings are interesting because very little is known about the role of these genes in plant immune responses.

Photosynthesis related genes

The genes involved in photosynthesis made 0.6% of the unigenes and included genes coding for chlorophyll a/b-binding protein, chloroplast ferredoxin-NADP+ reductase and NAD-dependent malic enzyme. There is not any strong evidence regarding the role of these genes in biotic stress. Therefore, it would be quite interesting to explore their role in stress.

Genes belonging to miscellaneous class

The genes for which no sufficient information was available regarding the cellular process they are involved in, were grouped in the miscellaneous class. This class

constituted 13.16% of the total unigenes and included many interesting genes. For example, dehydrin-cognate, macrophage migration inhibitory factor family protein, tumor protein, xylogen protein, pollen-specific protein, membrane attack complex component etc. were present in this class. This class also harboured genes encoding exostosin-like protein, nonaspanin, ADP-ribosylation factor 1, expansin-like B1 precursor and autophagy protein 5. Exostosin gene is known to be involved in animal diseases (Schrage *et al.*, 2009) but there is no report regarding its role in plants. ADP-ribosylation factor plays role in vesicular transport (Wei *et al.*, 2009) and recently many vesicular transport related genes have been found to be regulated during plant defense. Expansins are cell wall loosening enzymes (Kwon *et al.*, 2008) and might be involved in pathogen stress. This class also included many fungal genes like foxy transposable element, fot5 gene, putative transposase etc. Presence of these genes and their abundance indicates that these genes have an important role to play in plant-pathogen interactions.

2.3.3 Relative abundance of ESTs in contigs

In order to determine the rate of redundancy of ESTs in the contigs, we calculated the number of ESTs present in each contig. We found that maximum number of contigs had only upto five ESTs (Figure 2.4). The number of contigs with higher number of ESTs was considerably less. Furthermore, we observed that many contigs with higher number of ESTs coded for metabolic enzymes like gyleraldehyde-3-phosphate dehydrogenase and alcohol dehydrogenase suggesting that there is a strong alteration in metabolic pathways in response to stress. Many contigs with high number of ESTs coded for hypothetical proteins. Some stress related genes like extension were also very redundant. The high rate of redundancy indicates that such genes play an important role during defense responses. We also observed that ESTs present in very highly abundant contigs coded for some fungal genes like Fot and Foxy. Moreover, such contigs had more number of ESTs from susceptible genotype indicating that these genes might have a role to play in pathogenesis. Thus the relative abundance of ESTs in a contig gives an idea about their expression pattern.

2.3.4 Comparison with other legume genome and EST databases

In order to investigate how many of the chickpea ESTs from this study are highly homologous to other legume ESTs, we performed a comparative analysis of our
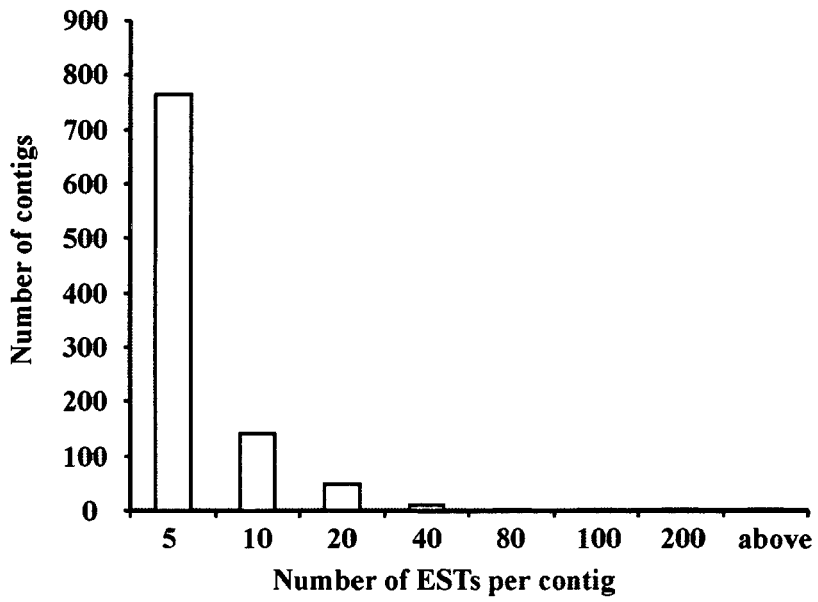
Figure 2.4: Bar diagram representing the rate of redundancy of ESTs in the contigs. Each bar represents the number of contigs with a specific number of ESTs.

CaEST dataset to publicly available dataset comprising ESTs from *Arachis hypogea, Cajanus cajan, Pisum sativum, Robinia pseudoacacia, Lotus japonicus, Medicago trunculata, Glycine max* and *Phaseolus vulgaris.* The highest number of common orthologous ESTs was found to be present in *Pisum sativum*, although, *Cajanus cajan* and *Arachis hypogea* also showed higher similarity with *Ca*Unigenes. The percentage of chickpea ESTs that matched to *Pisum sativum* was 88.8% and 79.43% at DNA sequence identity of ≥ 80% and 90%, respectively. A total of 87.6% and 83% *Cajanus cajan* ESTs and 85.89% and 55.54% of *Arachis* ESTs were common with *Ca*Unigenes. Chickpea and *Pisum* show close phylogenetic affliations, both being representatives of galegoid, a group of cool season legumes (Zhu *et al.*, 2005) and is reflected in their sequence similarities. However, high sequence similarity with *Arachis* and *Cajanus* came as a surprise since these legumes are phylogenetically distant relatives of *Cicer* and belong to tropical season legumes (phaseoloid). The comparative analysis of chickpea ESTs with that of *Glycine max* showed homology of 80.32% and 31.54%. Interestingly, *Medicago* and *Lotus* showed more divergent trend with only 58.02% and 30.50% and 43.31% and 14.25% ESTs respectively, having counterparts in *Ca*Unigene set. A recent phylogenetic study based on penalized likelihood analysis (Choi *et al.*, 2004) indicated that *Medicago, Lotus* and *Cicer* are closer and fall in the same galegoid clade, however, the divergence/difference in the pattern of ESTs may be attributed to the fact that gene expression is shaped by cellular environment besides ecological niche of the corresponding organism. Chickpea ESTs showed lesser homology with *Robinia pseudoacacia* and was about 12% and 3.92% at DNA sequence identity of 80% and 90%, respectively. This can be due to the fact that *Robinia* is a perennial tree while chickpea is an annual herb. Thus the current study documents substantial conservation as well as genome divergence amongst legume crops and in future can facilitate cross species analysis of gene function (Ashraf *et al.*, 2009). Detailed results of comparative analysis are shown in table 2.4.

2.3.5 Identification of chickpea specific transcripts and enrichment of *Ca*ESTs

In an attempt to identify the genotype-specific transcript signatures in chickpea, we performed a blast analysis of *Ca*Unigene set. Of the total 2013 *Ca*Unigenes, about 18.22% sequences that belong to NSH group represented potential chickpea specific sequences. To verify whether these sequences were indeed chickpea specific,

Table 4. Comparative matching of the chickpea ESTs to the ESTs of other legume databases.

Gene indices	Identity >80%	Identity>90%
Arachis hypogea	1686 (83.75)	1118 (55.54)
Cajanus cajan	1764 (87.6)	1729 (85.89)
Pisum sativum	1788 (88.8)	1599 (79.43)
Robinia pseudoacacia	245 (12.17)	79 (3.92)
Lotus japonicus	872 (43.31)	287 (14.25)
Medicago trunculata	1168 (58.02)	614 (30.50)
Glycine max	1617 (80.32)	635 (31.54)
Phaseolus vulgaris	928 (46.10)	234 (11.62)

The chickpea ESTs isolated in this study were compared to other legume databases like *Arachis hypogea, Cajanus cajan, Pisum sativa, Robinia psedoacacia, Lotus japonicus, Medicago trunculata, Glycine max and Phaseolus vulgaris* collected from NCBI and TIGR gene indices. The criteria for stand alone BLASTn were (1) exact match = 11; (2) E-value cutoff 1E5; and (3) identity > 80% and 90% at DNA sequence level.

TBLASTX was used for comparing them to the EST_others database. Nearly 64% of the NSH class of *Ca*EST dataset did not show any significant match thereby confirming that these ESTs represent unique chickpea sequences. Thus this dataset represents 234 novel chickpea specific ESTs, which were never reported earlier (Ashraf *et al.*, 2009).

2.3.6 Comparative stress responsive transcriptome reveals canonical and non canonical genes

We compared the *Ca*Unigenes with previously reported stress- responsive genes from other organisms. We classified genes known to be involved in multi-stress responses as ubiquitous, while those found to be specific to *Fusarium* infection were categorized as canonical. All other genes, which have never been implicated in any stress, were designated as noncanonical (Figure 2.5A). We found that around 516 genes are ubiquitous, suggesting broad similarities in stress responses across most of the organisms and confirming cross communicating pathways playing role in different kinds of stresses. However, only 41 genes were found to be canonical all of which showed an overlap with ubiquitous class. A significantly large number of genes (649) were found to be noncanonical (Ashraf *et al.*, 2009). This difference in the pattern of immune-responsive root transcriptome may be attributed to the fact that the gene expression in an organism is shaped by the cellular environment and the epigenetic factors. Metabolism was one of the most abundant functional class in all the three groups (Figure 2.5B), suggesting that any stress response results in the alteration of the metabolic pathways of an organism. A few of the metabolic enzymes found to be expressed in response to multiple stresses represented in the ubiquitous category included S-adenosylmethionine decarboxylase, glyceraldehydes-3-phosphate dehydrogenase, methionine sulfoxide reductase, and copper amine oxidase. Cell signaling related genes also formed an important class that included leucine-rich repeat receptor-like kinase and multiple bridging factor. As expected, many canonical genes belonged to functional class of cellular redox, defense and signaling and included Cationic peroxidase 2, chitinase and 14-3-3 protein. Our data on immune responsive transcriptome and comparative analysis thereof provide evidence for molecular diversity vs commanality in gene expression profile at organismal level. The comparative analysis revealed few canonical genes, several ubiquitous genes across different stresses while most of the genes were found to be noncannonical. An
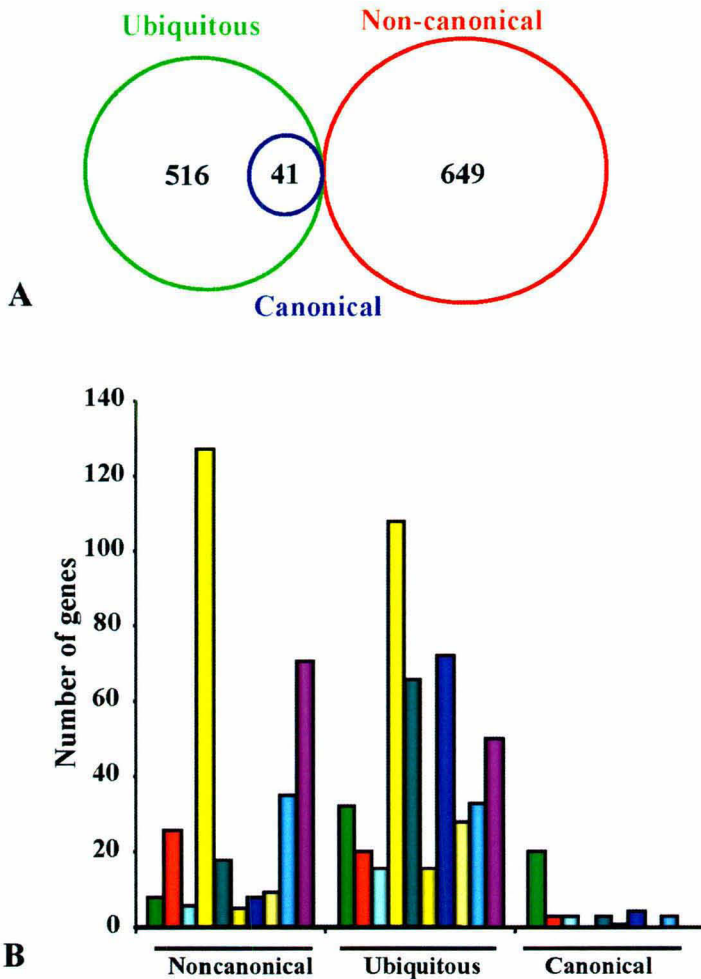
Figure 2.5: Comparative analysis of chickpea stress responsive genes with earlier known stress related genes. (A) Venn represents the overlap between ubiquitous, canonical and non canonical genes and the numbers signify unique and common stress responsive genes. **(B)** Prevalence of functional classes within the three groups. Each bar represents the number of genes in the respective functional class.

interesting observation was that apart from metabolism, most of the noncanonical genes belonged to the functional classes of translation, posttranslational modifications, transcription, and signaling suggesting the occurrence of new, yet undiscovered immune responsive pathways. This study thus provides a comprehensive catalogue of noncanonical immune responsive genes or might suggest their species specificity with new insight into their identity and function

2.3.7 Identification of gene families

To assign *Ca*Unigenes from this study to putative gene families, we used single linkage clustering. The individual contigs and singletons were combined into a single dataset which was then compared to itself using TBLASTX with an e - value cutoff of 10^{-15}. Sequences with overlapping BLAST reports were assigned to a putative gene family. We identified 209 gene families ranging in size from 2 to 29 members (Table 2.5) (Ashraf *et al.*, 2009). This analysis gives insights into following three areas: firstly, it identifies the genes that are likely to cross hybridize during the microarray hybridizations. Secondly, it helps in assigning possible function to genes that had no significant homology to known proteins or belonged to class of UF but clustered with proteins of known function. Thirdly, the identification of gene families provides a base for uncovering and understanding the biological rationale of functional novelty and partitioning following gene duplications. Some of the gene families identified were 14-3-3 proteins, alcohol dehydrogenase, copper amine oxidase, DnaJ-like protein etc.

2.3.8 Analysis of genotype-specific SNPs of chickpea

SNPs between genotypes/haplotypes once discovered are extensively used for many applications for instance generation of very dense genetic maps, to construct the specific genotypes required for quantitative genetic studies, to enhance understanding of genome organization and function and to address fundamental questions related to evolution. SNPs can also be used for genome-wide linkage disequilibrium and association studies that assign genes to specific functions or traits. Furthermore, transcript-associated SNPs can be used to develop allele-specific assays for the examination of *cis* regulatory variation within a species (Barbazuk *et al.*, 2007). In the present study, SNPs were identified between JG-62, a susceptible and WR-315, a resistant genotype of chickpea to vascular wilt. A total of 279 contigs (28.67% of the

Table 5: Identification of chickpea gene families by single linkage clustering

Group	Top BLAST HIT
1	110 kDa 4SNc-Tudor domain protein
2	14-3-3-like protein
3	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
4	2OG-Fe(II) oxygenase
5	2-oxoglutarate deHydrogenase, E1 component
6	40S ribosomal protein S8
7	5-metHyltetraHydropteroyltriglutamate--homocysteine S-metHyltransferase
8	60S ribosomal protein
9	60S ribosomal protein L37a
10	Acyl-coA-binding protein, ACBP
11	Agaricus bisporus partial mRNA for putative myosin heavy chain kinase
12	Albumin-2 (PA2)
13	Alcohol deHydrogenase 1
14	Aldo/keto reductase
15	Amino acid/polyamine transporter II
16	AMP-dependent synthetase and ligase
17	Avr9/Cf-9 induced kinase 1
18	Belgica antarctica clone Ba-U40 CG32816-like mRNA
19	Brassica napus isolate mutant Cr3529 clone Bncr10 unknown mRNA
20	C.arietinum mRNA for metallothionein (clone: CanMT-1)
21	Calcium-dependent calmodulin-independent protein kinase isoform 2
22	cAMP response element binding (CREB) protein
23	CAP protein
24	Catharanthus roseus clone CrP15 T-DNA sequence
25	Catharanthus trichopHyllus genotype CtN58 microsatellite CATR10 sequence
26	Cationic peroxidase 2 precursor (PNPC2)
27	Cell division protein FtsZ
28	Chalcone synthase
29	Chaperone DnaK
30	Chitinase-related agglutinin
31	Cicer arietinum partial mRNA for putative extensin (ORF), clone CanEXT-1
32	Crassostrea gigas BAT1 homolog mRNA, complete cds
33	Class I chitinase
34	Concanavalin A-like lectin/glucanase
35	Copper amine oxidase
36	Crassostrea gigas BAT1 homolog mRNA, complete cds
37	Curculin-like (mannose-binding) lectin
38	Cyclic peptide transporter
39	Cyclin-like F-box
40	Cytochrome P450
41	Cytochrome P450 monooxygenase
42	Cytochrome P450 monooxygenase CYP83G1
43	Cytosolic malate deHydrogenase
44	Danio rerio tf2a mRNA, 3'UTR
45	DEAD box RNA helicase
46	DECOY, putative
47	DiHydrolipoyl deHydrogenase
48	DNA binding / transcription factor
49	DnaJ-like protein
50	Elongation factor EF-2
51	Enolase

52	Eristalis tenax partial mRNA for Hypothetical protein (ORF1), isolate 3
53	Flag-tagged protein kinase domain of putative mitogen-activated protein kinase
54	Fructose-bisphosphate aldolase, cytoplasmic isozyme
55	Fusarium oxysporum f. sp. lycopersici insertion sequence Foxy
56	Fusarium oxysporum f. sp. lycopersici six1 gene, fot5 gene,
57	Fusarium oxysporum f. sp. vasinfectum strain Ag149 Foxy transposable element
58	Fusarium oxysporum f. sp. vasinfectum strain Ag149-III Foxy transposable element,
59	Fusarium oxysporum f. sp. vasinfectum strain X515-II Foxy transposable element
60	Fusarium oxysporum voucher VPRI 19292 mitochondrion, partial genome
61	GDP dissociation inhibitor
62	General substrate transporter
63	Glucan-endo-1,3-beta-glucosidase
64	Glutathione S-transferase, C-terminal-like; Thioredoxin fold
65	Glycoside Hydrolase, family 1
66	Granulin; Peptidase C1A, papain
67	GTP-binding protein
68	Haem peroxidase, plant/fungal/bacterial
69	Heat shock protein DnaJ
70	Histone deacetylase HDT1
71	HMG-CoA reductase
72	Homeodomain leucine zipper protein HDZ1
73	Hypothetical protein
74	C2 Peptidase, cysteine peptidase active site
75	Calcium-binding EF-hand
76	Hypothetical protein [Trifolium pratense]
77	Hypothetical protein [Nicotiana tabacum]
78	Hypothetical protein [Cicer arietinum]
79	Hypothetical protein [Gibberella fujikuroi]
80	Hypothetical protein [Vitis vinifera]
81	Hypothetical protein [Vitis vinifera]
82	Hypothetical protein [Vitis vinifera]
83	Hypothetical protein [Vitis vinifera]
84	Hypothetical protein [Vitis vinifera]
85	Hypothetical protein [Vitis vinifera]
86	Hypothetical protein [Picea mariana]
87	Hypothetical protein OsI_018180
88	Hypothetical protein OsJ_009000
89	Hypothetical protein OsJ_025206
90	Hypothetical protein FG02001.1
91	Hypothetical protein FG04176.1
92	Hypothetical protein FG07152.1
93	Hypothetical protein FG09053.1
94	Isoflavone 3'-Hydroxylase
95	Ixodes ricinus cytochrome oxidase subunit I mRNA, partial cds
96	LEM3 (ligand-effect modulator 3) family protein / CDC50 family protein
97	Lotus japonicus genomic DNA, chromosome 3, clone:LjT10E18, TM0035
98	M.truncatula DNA sequence from clone MTH2-28N4 on chromosome 3
99	Malic oxidoreductase
100	Medicago truncatula chromosome 5 clone mth4-20m5
101	Medicago truncatula clone mth2-17d15
102	Medicago truncatula clone mth2-18h17
103	Medicago truncatula clone mth2-1o14
104	Medicago truncatula clone mth2-69j4
105	Medicago truncatula clone mth2-85g12

106	Medicago truncatula mRNA for MtN4 gene, partial
107	Methionine sulfoxide reductase A
108	NAD-dependent epimerase/deHydratase
109	Nascent polypeptide-associated complex NAC; UBA-like
110	Nerium oleander microsatellite CATR25 sequence
111	Nicotiana tabacum serine/threonine kinase mRNA, partial cds
112	Nitrite reductase
113	O-metHyltransferase, family 2
114	Orn/DAP/Arg decarboxylase 2
115	Pathogenesis-related transcriptional factor and ERF
116	Pectinesterase
117	Pectinesterase family protein
118	Peptidase S10, serine carboxypeptidase
119	Peptidase S24, S26A and S26B
120	Peptidase T1A, proteasome beta-subunit
121	Peroxidase
122	Peroxidase, putative
123	Phenylalanine ammonia lyase
124	Human DNA sequence from clone RP11-305D15 on chromosome 13
125	Photobacterium damselae subsp. piscicida partial coi genes for putative cytochrome C oxidase
126	Photobacterium damselae subsp. piscicida partial ORF1 DNA for Hypothetical protein
127	Photobacterium damselae subsp. piscicida trpA gene for putative transposase
128	Photobacterium damselae subsp. piscicida trpB gene for putative transposase
129	Plant lipid transfer/seed storage/trypsin-alpha amylase inhibitor
130	Plasmodiophora brassicae 16S ribosomal RNA gene
131	Platynereis dumerilii mRNA for Hypothetical protein (ORF1)
132	PMA1_NEUCR Plasma membrane ATPase (Proton pump)
133	Polygalacturonase-like protein
134	Poplar cDNA sequences
135	Poplar cDNA sequences
136	PpiC-type peptidyl-prolyl cis-trans isomerase; Rhodanese-like
137	Prefoldin
138	Hypothetical protein
139	Protein kinase
140	Protein kinase
141	Protein of unknown function DUF676, Hydrolase-like
142	Protein of unknown function DUF707
143	Protein phosphatase 2C-like
144	Protein prenyltransferase
145	Solanum lycopersicum cDNA, clone: FC06CB10, HTC in fruit
146	Pseudomonas aeruginosa gene for 16S rRNA
147	Putative ADP,ATP carrier-like protein
148	Putative beta-amylase
149	Putative desaturase-like protein
150	Putative imbibition protein
151	Putative protein
152	Putative receptor-like GPI-anchored protein 2
153	Putative ripening related protein
154	Putative senescence-associated protein
155	Putative spindle disassembly related protein CDC48
156	Pyridoxal-5-phosphate-dependent enzyme, beta subunit
157	Pyruvate decarboxylase isozyme 1 (PDC)
158	Pyruvate deHydrogenase E1 component subunit beta
159	Pyruvate kinase

160	Rana pirica mRNA for trypsinogen
161	Ras small GTPase, Rab type
162	Rattus norvegicus obese protein gene, 5' flanking region
163	Response regulator receiver; CCT
164	Ribosomal Pr 117
165	Ribosomal protein L19e
166	Ribosomal protein L7Ae/L30e/S12e/Gadd45
167	Ribosomal protein S6
168	RNA-binding region RNP-1 (RNA recognition motif)
169	S25 ribosomal protein
170	S-adenosyl-L-homocysteine Hydrolase
171	S-adenosylmethionine synthetase 2
172	SAM (and some other nucleotide) binding motif
173	Secretory peroxidase
174	Selenium binding protein
175	Serine C-palmitoyltransferase like protein
176	Serine/threonine protein kinase, active site
177	Serine/threonine/tyrosine kinase
178	Serine/threonine-protein phosphatase PP1
179	Sesbania drummondii clone SSH-1_01_F12_T3 mRNA sequence
180	Sesbania drummondii clone SSH-36_01_A09_T3 mRNA sequence
181	Signal recognition particle 54 kDa subunit precursor
182	Siniperca chuatsi transposase mRNA, partial cds
183	Protein binding |leucine-rich repeat disease resistance protein-like
184	Squalene epoxidase
185	Sucrose synthase 2 (Sucrose-UDP glucosyltransferase 2)
186	Synthetic construct arsenic-like protein gene, complete cds
187	Unknown protein
188	Tonoplast intrinsic protein
189	Translation factor
190	Translation factor; Elongation factor G, III and V
191	Translation protein SH3-like
192	Translocase of chloroplast 34
193	Triosephosphate isomerase
194	Triticum aestivum clone wlsu2.pk0001.h3:fis
195	tRNA synthetase class II (G, H, P and S) family protein
196	Tubulin beta chain (Beta tubulin)
197	Ubiquitin
198	Ubiquitin-conjugating enzyme, E2
199	Uncharacterized Cys-rich domain
200	Universal stress protein (Usp)
201	Unknown protein [Vitis pseudoreticulata]
202	Vigna unguiculata partial mRNA for putative ATP synthase
203	Vigna unguiculata partial mRNA for putative CBL-interacting protein kinase
204	WD40-like
205	Wound-induced GSK-3-like protein
206	Zinc finger, RING-type
207	Zinc finger, RING-type; RINGv
208	Heat shock protein DnaJ
209	Thioredoxin domain 2; Thioredoxin fold

total) contained at least one sequence from both the genotypes and were mined for potential SNPs. We identified 262 SNPs which were further classified into high and low quality SNPs depending upon the number of sequences from each genotype showing the same base change. High-quality SNPs were confirmed by two or more sequences from each genotype showing the same base change, while low-quality SNPs were confirmed by one sequence from one genotype and at least two from the other. Thus, we identified 136 high-quality and 126 low-quality SNPs (Ashraf *et al.*, 2009). In the present analysis only base pair mutations were taken into consideration and among these transitions (73.3%) were more common than transversions (26.7%). Within the transitions, occurrence of both adenine to guanine and cytosine to thymidine base changes were found to be almost equal.

2.4 Conclusion

The present study was directed towards the development of chickpea EST database and provides an initial platform for functional genomics of this food legume. The ESTs identified in this study represent a major attempt so far to define chickpea transcriptome and offers a base for expression profiling thereby enabling the identification of genes involved in specific physiological and biochemical processes. This will pave way for creation of other resources such as microarray chips that can help provide a view of global gene expression and construct network models for various stress pathways. Identification and sequencing of genotype specific ESTs would help to understand the host specific responses against stress. The classification of ESTs into various functional classes suggests that there is alteration of many cellular processes during stress. The redundancy of ESTs in contigs provides a clue about the relative expression of such genes without going further into the expression studies. Study of comparison between chickpea and other legume ESTs showed substantial conservation as well as genome divergence amongst legume crops. Identification of a catalogue of non canonical genes in this study provides a new dimension to stress genomics. This may lead to the discovery of new defense regulators and provide a clue to the existence of new pathways and also missing links in already known defense signaling pathways. Our study led to the identification of many gene families which helped to assign function to the genes which otherwise

were grouped with genes of unknown function. A very important aspect of the study was identification of SNPs between two chickpea genotypes. The SNPs would be of great help to study genotype specific response towards a particular cellular process.

Differential transcript profiling during compatible and incompatible Chickpea-Fusarium interaction

3.1. Introduction

Many biological processes are controlled by intricate regulatory networks of gene expression. Identifying these networks and hierarchical relationship between them are vital to the understanding of biological systems (Wang *et al.*, 2006). Reprogramming of cellular functions in response to external stimuli involves complex changes in gene expression. Unravelling the way in which such stimuli are perceived and transduced to downstream machinery is of strategic importance to the understanding of cellular responses triggered by various environmental cues. DNA microarray technology has opened up the possibility to study gene expression on a genome wide scale and rapid advances are being made towards the understanding of transcriptional programs of various biological processes. This approach may provide clues for elucidating the functions of genes underlying specific processes including stress and identifying the candidate genes predicted to regulate such processes. The employment of microarray technology has rapidly produced vast catalogues of dynamic expression patterns of genes activated in response to various stresses (Chen *et al.*, 2008). Understanding these changes is of interest for both pure and applied sciences. Further, function of an individual gene in a particular biological process cannot be understood on the level of isolated components alone, but needs to be studied in the context of its interplay with other gene products (Dittrich *et al.*, 2008). Microarray technique helps in studying the co-regulatory expression patterns of number of genes and hence helps in revealing the concerted effort of such genes to regulate a developmental plan or a stress response. Also information from many other databases like protein-protein interaction databases can be integrated to construct networks regulating various biological processes (Beyer *et al.*, 2007).

Plant diseases have major effects on agricultural production and the food supply. Development of new strategies based on plant's own defense mechanisms for disease control is critical for sustained agricultural production and improving our environment and health (Yang *et al.*, 1997). During the process of host pathogen coevolution, plants have developed elaborate mechanisms to resist the pathogen attack. These defense mechanisms include both preformed barriers as well as those induced after pathogen attack. Such induced mechanisms involve a network of signal transduction and activation of gene expression. Plant disease resistance and susceptibility are governed by combined genotypes of host and pathogen and depend on a complex

exchange of signals and responses occurring under given environmental conditions (Yang *et al.*, 1997). A key difference between resistance and susceptibility is the timely recognition of the invading pathogen and the rapid and effective activation of host defense mechanism in case of resistance in contrast to susceptibility where plant exhibits much weaker and slower response and fails to restrict pathogen growth and/or spread (McDowell and Dangl, 2000). At genomic level, plant defense responses are complex and diverse, and every gene involved from recognition to signal transduction to the direct involvement forms a part of the coordinated response network. Transcriptional profiling may, therefore, provide new tools for identifying the key genes that govern host responses against pathogens and the variations in host gene expression that may be associated with resistance or susceptibility to pathogen stress (Delahaye *et al.*, 2007).

Chickpea is the third most important pulse crop in the world but its production is limited due to severe abiotic and biotic stresses, *Fusarium* wilt being one of them. *Fusarium* wilt caused by *Fusarium oxysporum ciceri* is an important and widespread soil borne disease in many chickpea producing regions. Various breeding programmes are being carried out aiming at producing resistant cultivars, however, the mechanism controlling effective resistance remains largely unknown. Therefore, transcript profiling of chickpea-*Fusarium* interaction may generate a wealth of information and improve overall understanding of the coordinated defense response to this disease at the molecular level.

In the present study, we carried out gene expression profiling in susceptible and resistant cultivars of chickpea in response to *Fusarium* wilt using 12K cDNA microarray consisting of 12116 probes representing 2013 unigenes in duplicates. The aim was to create a signal transduction catalogue for chickpea defense signaling as well as to compare the expression pattern in compatible and incompatible interactions. We also studied the co-expression patterns of differentially regulated genes using K-means clustering. The study provides an insight into the molecular mechanism of defense pathways. More importantly, we identified a subset of novel regulators of plant immunity. Also genes with significant differences in regulation between the two cultivars were identified which might have role to play in displaying disease resistance or susceptibility.

3.2. Material and methods

3.2.1. Plant material and *in planta* infection

The chickpea (*Cicer arietinum .L*) seeds were sterilized with 70% ethanol and 0.1% $HgCl_2$ followed by repetitive washing with autoclaved water. The seedlings were grown in sterile conditions in glass tubes containing germination medium solidified with 0.6% agar (appendix I) in an environmentally controlled growth room and maintained at $25 \pm 2°C$ temperature, $50\pm5\%$ relative humidity under 16 h photoperiod (270 μmol m-2 s-1 light intensity). *Fusarium oxysporum ciceri race 1* was grown in potato dextrose broth (PDB). For this, the fungus was inoculated into PDB and grown at 28°C and 200 rpm shaking for 3 days. The fungal suspension was then filtered through autoclaved double layered cheese cloth to remove the mycelia. The concentration of the recovered spore suspension was checked by haemocytometer. The spore suspension was diluted to a final concentration of $1x10^6$/ml. The 3-week old seedlings were inoculated with *Fusarium* spore suspension while the control plants were treated with water. Root and collar tissue was sampled as experimental material, harvested at 6h, 12h, 24h, 48h and 5d after inoculation and stored at -80 °C after quick-freezing in liquid nitrogen.

3.2.2. Screening of chickpea varieties

Five chickpea varieties were screened based on the some physiological tests for their susceptibility or resistance response to *Fusarium* wilt. The various tests and their procedure are given below:

3.2.2.1. Determination of relative water content

Chickpea plants were grown as mentioned in section 3.2.1. The leaf tissue samples were collected and immediately weighed (fresh weight, FW). The samples were rehydrated in water for 24 hours until fully turgid, surface dried, and re-weighed (turgid weight, TW) followed by oven drying at 80°C for 48 hours and again weighed (dry weight, DW). The RWC was calculated by the following formula: RWC (%) = (FW-DW/TW-DW) x 100. The experiment was carried out in triplicates.

3.2.2.2. Proline estimation

For measuring the proline content, the tissues were homogenized in 3% aqueous sulphosalicylic acid and the homogenate was centrifuged at 9,000 x g. The reaction mixture consisting of 2ml supernatant, 2ml acid ninhydrin and 2ml glacial acetic acid was boiled at 100°C for 1h. After termination of the reaction on ice, the reaction mixture was extracted with 4ml of toluene, and the absorbance was read at 520 nm. The assays were done in triplicates using corrected weight calculated for the actual moisture content of tissue at each time point. Proline was calculated as per the formula: proline (mg/g) = 36.6 x A520 x volume/2 x fresh weight.

3.2.2.3. Pigment estimation

Tissue harvested at different time points was ground in 80% chilled acetone. The supernatant was taken for determination of photosynthetic pigments: chlorophyll a (mg/g) = [(12.7 x A663 – 2.69 x A) v/w], chlorophyll b (mg/g) = [(22.9 x A645 – 4.68 x A663) v/w], and carotenoid (mg/g) = [{(1000 x A470) – (3.27 x chlorophyll a + 1.04 x chlorophyll b)} /227 v/w]. The experiments were done in triplicates using corrected tissue weights calculated for actual moisture content of the tissue at the respective time points.

3.2.2.4. Determination of lipid peroxidation

Lipid peroxidation was estimated in terms of malonaldehyde (MDA) production. Fresh chickpea tissues were ground in 5% TCA. One part of the supernatant was mixed with four parts of 0.5% thiobarbituric acid (TBA) prepared in 20% TCA and heated at 95°C for 30 min. Absorbance at 532 nm was measured and corrected for nonspecific absorbance at 600 nm. The pellet was washed in 100% chilled acetone and suspended in 0.1N NaOH to estimate total protein using Bradford's reagent. Oxidized MDA was calculated as per the formula: MDA (g/mg) = {(A532 – A600) x volume x 100}/ {155 x total protein (mg)}. The experiments were carried out in triplicates using corrected weights calculated for actual moisture content of the seedlings at each time point.

3.2.2.5. Electrolyte leakage assay

Electrolyte leakage was assayed by estimating the ions leaching from the leaflets into MQ water. Plant materials were taken in 20ml of MQ in two sets. The first set was

kept at room temperature for 4h and its conductivity (C1) was recorded using a conductivity meter. The second set was autoclaved and its conductivity was also recorded (C2) and electrolyte leakage [1-(C1/C2)] x 100 was calculated. The experiments were carried out in triplicates.

3.2.3. Construction of cDNA microarray

The cDNA microarray setup consisted of 6058 cDNA clones corresponding to 2013 unigenes. The microarray construction was done in four steps as given below:

1. PCR amplification of cDNA clones

The cDNA clones from both susceptible and resistant libraries were amplified by performing colony PCR. For template preparation, 5µl of the culture from the glycerol stocks plates was directly inoculated into 50µl of sterile water dispensed into each well of a 96-well plate. The plate was sealed and heated to 100°C for 10 minutes in a thermocycler to lyse the cells and release the plasmid DNA. In order to remove the cell debris, the plate was centrifuged at 4,000 x g for 5 minutes. 5µl of the supernatant was used as template for the PCR. The PCR reaction for each clone was carried out in 100µl volume and was set as follows:

10X PCR buffer	10.0µl
25mM MgCl$_2$	6.0µl
10mM dNTP mix	2.0µl
10µM M13 forward primer	2.0 µl
10µM M13 reverse primer	2.0µl
Template	5.0µl
DNA polymerase (5 units/µl)	0.5µl
Sterile MQ water	72.5µl

PCR was performed using the following program

1. Initial denaturation	94°C for 2 min
2. Denaturation	94°C for 30 sec
3. Annealing	52°C for 30 sec

| 4. Primer extension | 72°C for 1 min |

(steps 2 to 4 cycled 30 times)

| 5. Final extension | 72°C for 10 min |

| 6. Indefinite hold | 4°C |

2. Purification of PCR products

The PCR amplified products were purified using Perfectprep PCR clean up kit (Eppendorf) as per the following protocol:

1. The entire volume of 100μl of the PCR reactions was transferred to the wells of the filter plate provided with the kit and placed on the vacuum manifold.

2. Vacuum was applied for 10 minutes at -20 mm Hg.

3. Vacuum was turned off and 50μl of molecular biology grade water (Eppendorf) was added to the center of the filter plate membrane and incubated at room temperature for 5 minutes.

4. The contents of each well were mixed by pipetting up and down 5 times.

5. The purified PCR product from each well was pipetted off the membrane and placed in a collection plate.

6. The purified PCR products (50μl) were stored at -20°C till further use.

3. Quality check of purified PCR products

The quality of the PCR products (2μl) was checked on 1.0% agarose gels as mentioned in section 8 of appendix III. The PCR products which showed distinct and sharp bands were further used for the array preparation. For the clones which did not amplify or which showed more than one band were excluded and again amplified as per the reaction mentioned in 3.2.3.1.

4. Printing of cDNA clones on slides

The purified PCR products were transferred from 96-well plates to 384-well plates for long term storage. For the printing purpose, five microlitres of the PCR products from 384-well storage plates were reorganized on the fresh 384-well printing plates with 5μl of 100% DMSO. The printing was done in duplicates on the poly-L-lysine coated

slides (Sigma) using a high throughput arrayer (Arrayer ESI, SDDC2) followed by UV cross linking. A total of 6058 clones were printed in duplicates on the slides. 5 positive and 14 negative controls (see appendix I) were also printed at three different locations in the microarray.

3.2.4. Target preparation and labelling

In order to find out the differential gene expression during chickpea-*Fusarium* interaction, root tissue samples were collected from *Fusarium* infected WR-315 and JG-62 chickpea seedlings after 6h, 12h, 24h, 48h and 5d of infection. Water treated plants were taken as control. Two biological replicates were done for each time point. Total RNA was isolated from all the tissue samples using TRIzol reagent (Invitrogen, CA). 6μg of the total RNA from each sample was used for cDNA synthesis using indirect TSA labeling and detection kit Micromax (PerkinElmer). The RNA isolated from uninfected and infected tissue was labelled with flourscein and biotin, respectively.

The steps involved in cDNA synthesis are:

1. 6μg of control and treated RNA were added in two separate tubes in RNase free water and the reaction mixture was prepared as following

REAGENTS	VOLUME
6μg total RNA in RNase free water	X μl
Reaction mix concentrate	1μl
Fluorescein or biotin Nucleotide	1μl
RNase free water	18-Xμl
TOTAL VOLUME	20μl

2. To denature any secondary structure, RNA samples were incubated at 65°C for 10 minutes.

3. In order to anneal the primer to the RNA template, the reactions were cooled to room temperature (25°C) for 5 minutes.

4. The samples were then warmed to 42°C for 3 minutes and the following reagents were added to each sample containing tube separately in the order indicated. After each addition the reagents were mixed well by pipetting up and down.

REAGENTS	VOLUME (initial volume 20ul)
10X RT reaction buffer	2.5µl
AMV RT\RNase inhibitor mix	2.0µl
TOTAL VOLUME	24.5µl

5. The whole reaction mixture was incubated continuously at 42ºC for 60 minutes.

6. The labelling reaction was cooled to 4ºC for 5minutes and the following reagents were added in the order indicated.

REAGENTS	VOUME (initial volume 24.5ml)
0.5M EDTA, pH 8	2.5µl- Stops the reaction
1.0N NaOH	2.5µl- initiates hydrolysis
TOTAL VOLUME	29.5µl

7. Then it was again incubated at 65ºC for 5 minutes.

8. The reaction mixtures were cooled to 4ºC for 5 minutes

Purification of labelled cDNAs

The labelled cDNAs were purified using microcon YM 100 columns (Millipore, Bedford, MA) as per the following procedure:

1. 200µl of 10mM Tris HCl (pH 7.5) was added to a microcon YM 100 filter cartridge. Then the Fl labelled and biotin labelled cDNA was added to the filter catridges.

2. The samples were centrifuged at 500 x g for 30 minutes. Centrifugation was continued until the sample volume reached about 20µl.

3. The samples were washed by adding 400µl of 10mM Tris HCl (pH 7.5) by pipetting up and down.

4. The centrifugation step was followed as in step 2.

5. The labelled cDNAs were recovered by inverting the membrane cartridge in clean 1.5mL tubes

6. The labelled cDNAs were taken down to near dryness in a savant speed vacuum and then redissolved in 200µl of hybridization buffer.

3.2.5. Microarray hybridization and TSA detection

For hybridization, total 200µl each of purified flourscein and biotin labeled cDNAs were mixed and hybridized to the microarray slides in hybridization chambers (Corning) at 65ºC incubator. The slides were washed for 10 minutes in 30ml of 0.5X SSC and 0.01% SDS, 10 minutes in 30ml of 0.06X SSC and 0.01%SDS and for 5 minutes in 30ml of 0.06X SSC. The slides were subsequently processed for replacement of flourscein with Cy3 and biotin with Cy5 as given below:

1. Microarray slides were incubated with 600µl of TNB-10% goat serum buffer for 10 minutes which was followed by rinsing once in 30ml TNT buffer for 1 minute with agitation.

2. The slides were then incubated with 300µl of Anti-FL-HRP conjugate solution for 10 minutes followed by rinsing of slides in 30ml of TNT buffer with agitation thrice for one minute each.

3. Slides were again incubated with 300µl of cyanine 3 tyramide solution for 10 minutes followed by rinsing in 30ml of TNT buffer three times each for five minutes.

4. The slides were incubated with 300µl of HRP inactivation solution for 10 minutes and were then rinsed, 3 x 1 minute, in 30ml of TNT buffer with agitation.

5. This was followed by incubation with 300µl streptavidin-HRP conjugate solution for 10 minutes. Then the slides were washed 3 x 1 minute, in 30ml of TNT buffer with agitation.

6. Slides were further incubated with 300µl cyanine 5 tyramide solution for 10 minutes and then washed 3 x 5 minutes, in 30ml of TNT buffer with agitation.

7. The slides were lastly rinsed for one minute in 30ml of 0.6X SSC with agitation.

8. In order to dry the slides, these were given a spin in a swinging bucket rotor at $\leq$ 3,000 x g for 2 minutes, in a 50ml conical tube.

3.2.6. Scanning and data analysis

Micraoarrays were scanned using Scan array 5000 scanner (PerkinElmer) to produce two separate tiff images. Spot finding and quantification of the spots were done by using scan array express software (PerkinElmer). Spots appearing bad due to poor

morphology, high local background and bubbles were flagged off and were excluded from further analysis. Spots with both channel intensities less than 500 were also filtered out. Spots were quantified using an adaptive method. Avadis software (PerkinElmer) was used for further data transformation which consisted of background correction and normalization. For background correction, local background intensity of each spot was subtracted from its foreground intensity value. Due to non linearity of the data, intensity dependent Lowess normalization was applied. Cy5/Cy3 signal ratio was also calculated for each spot on the array. Cross slide one class t - test with Benjamini and Horchberg FDR multiple correction was performed on the four replicate data points for each probe. P value of 0.05 and fold induction of 2.5 was used as a limit for statistically significant differences in the expression.

3.2.7. Gene clustering

The clustering analysis was performed using MEV software (TIGR). The analysis included 1280 genes which had altered expression during at least one time point in one of the genotypes. Expression values for the same gene measured in each of the two genotypes were entered into analysis as independent sets of values in a single clustering analysis to detect the pattern of similarity and differences between the two cultivars. K-means clustering method was employed.

3.2.8. Real time PCR

The microarray expression results were validated by performing quantitative real time PCR. For each genotype and time point, total RNA was isolated and quality of RNA checked by agarose gel electrophoresis. The total RNA samples were quantified using a NanoDrop Spectrophotometer (Nanodrop Technologies) and cDNA was prepared using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). For the preparation of cDNA, 2μl of 10X RT buffer, 0.8μl of 25X dNTP mix (100mM each), 1μl of 10X random primer, 1μl of multiscribe RT, 1μl of RNase inhibitor and 3.2μl of MQ were mixed in 0.2ml PCR tubes and then 10μl of total RNA dissolved in DEPC treated water was added (final concentration 2μg). The tubes were incubated for reverse transcription at 25°C for 10 min; 37°C for 2h and 85°C for 5 sec. The cDNA was diluted 10 times and qRT-PCR was performed for each clone of interest in triplicates in ABI 7500 sequence detection system. Reactions were performed in a

20μl volume containing 10μl SYBR Green Master Mix (Applied Biosystems), 2μl of diluted cDNA and 0.5μl of each gene-specific primer (10μM each). PCR cycles were performed as: 95°C for 10 min; 40 cycles of 95°C for 15 sec and 60°C for 1 min. Primer titration and dissociation experiments were performed to ensure that no primer dimers are formed. The sequences of the primers used for the real time PCR are given in table 1 of appendix II. The relative quantification method (ΔΔ-CT) was used to evaluate quantitative variation between the replicates examined. The amplification of 18S RNA was used as an endogenous control to normalize all data.

3.3. Results and discussion

3.3.1. Screening of chickpea varieties for *Fusarium* resistance and susceptibility

Fusarium wilt is the most important root disease of chickpea in the semi arid tropics (Pande *et al.*, 2007). The fungus grows in the water conducting vessels of the roots and stems which are subsequently plugged and collapse thereby blocking the water supply. Due to this reason, the wilt usually mimics the dehydration stress symptoms. The screening of germplasm has led to the identification of stable sources of resistance to race 1 of *Fusarium*. While WR-315 is known to be resistant to *Fusarium* wilt, JG-62 is highly susceptible and early wilting genotype (Brinda and Ravikumar, 2005). However, no detailed physiological and biochemical study has been conducted so far. Thus in the present study, we performed some physiological tests in order to complement to the earlier genetics based data. We screened five commercial varieties of chickpea for the relative water content (RWC) at various time points after fungal infection. All the varieties showed a gradual decline in the RWC, however, WR-315 was able to maintain it at a higher level than all other varieties (Figure 3.1A). JG-62 showed a much lower value and can be attributed to the inability of the plant to take up water due to the maximum fungal growth. Stress in plants is often followed by the accumulation of solutes like proline. We estimated proline content in chickpea varieties and found its maximum accumulation in WR-315 followed by CPS1 (Figure 3.1B). This solute helps the plants to resist the dehydration stress. The fungal attack often leads to membrane damage in plants and is required for the entry of the pathogen into cell. The ability of a plant to avoid membrane damage is a measure of its resistance against the invasion. To determine the level of membrane integrity, the lipid peroxidation and status of electrolyte leakage was monitored in all the five
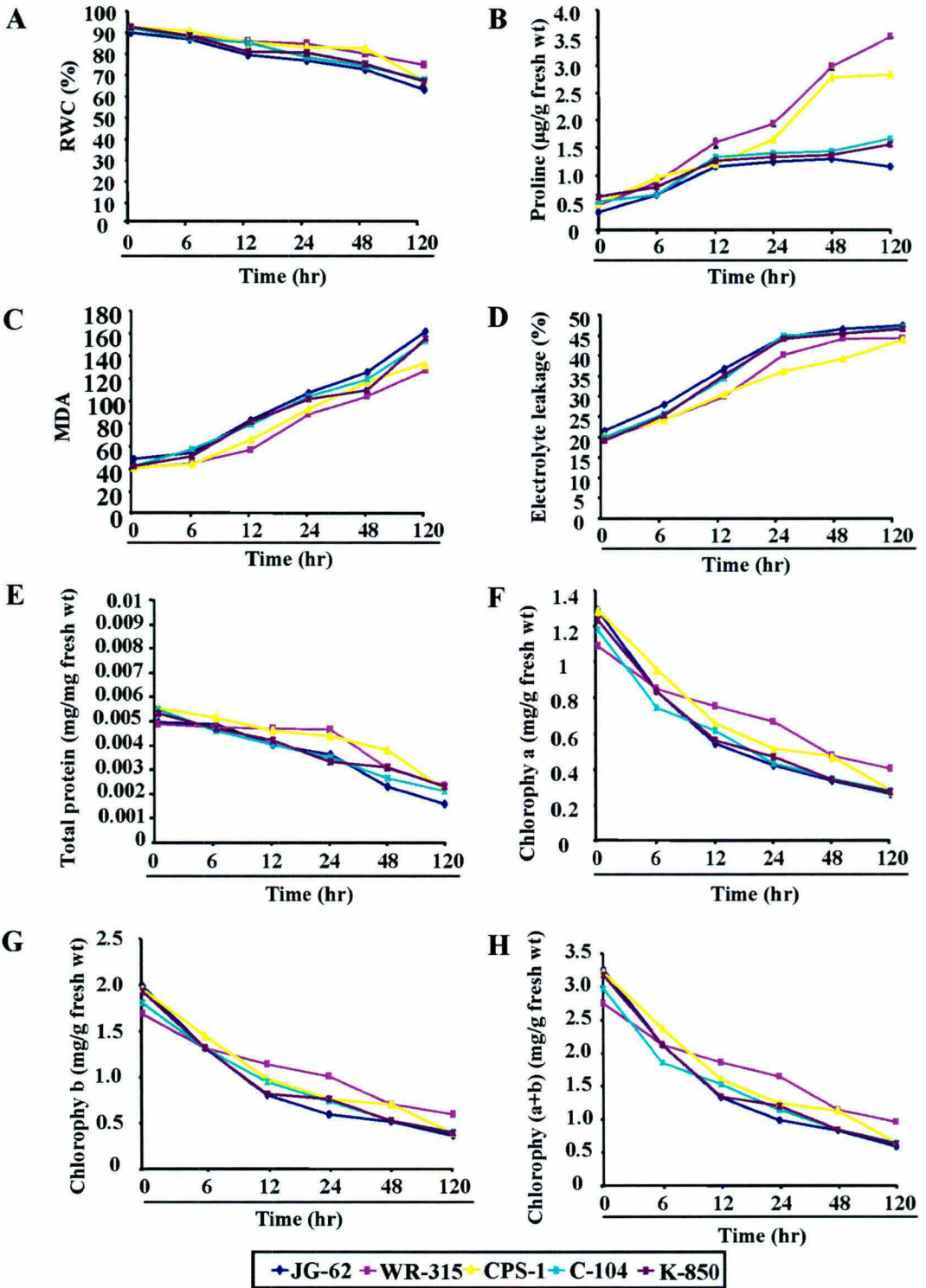
Figure 3.1: Comparative analysis of effect of *Fusarium* wilt on (A) Relative water content (RWC) **(B)** Proline accumulation **(C)** MDA **(D)** Electrolyte leakage **(E)** Total protein content **(F)** Chlorophyll a **(G)** Chlorophyll b **(H)** Chlorophyll (a+b) among five varieties of chickpea. All experiments were performed in triplicates and average values were plotted against time.

varieties. All the varieties showed increase in lipid peroxidation, however, it was maximum in JG-62 while as WR-315 and CPS1 showed lesser values (Figure 3.1C) indicating lesser extent of membrane damage in these genotypes. Electrolyte leakage also showed a similar trend (Figure 3.1D). This together suggested that WR-315 displayed maximum membrane stability in response to fungal attack. The amount of total protein was estimated in all these varieties and results showed higher level in WR-315, though it showed a decreasing trend in all genotypes at various time points (Figure 3.1E). In order to understand the effect of pathogen stress on photosynthetic apparatus, status of photosynthetic pigments was determined at various time points in all the varieties. It was observed that chlorophyll a and b followed a decline in all genotypes in the progressive time points after pathogen stress, WR-315, however, maintained the photosynthetic apparatus better than other genotypes (Figure 3.1F, G, H). Taken together, these results confirm that WR-315 is most resistant to *Fusarium* wilt while JG-62 is the most susceptible one. These results are in confirmation with earlier microscopic data from our laboratory where it has been shown that hyphal progression and accumulation were maximum in JG-62 and minimum in WR-315 (Ghai *et al*., unpublished results).

3.3.2. Construction of chickpea cDNA microarray and experimental design

We constructed chickpea cDNA microarray comprising of 6058 cDNA clones representing 2013 unigenes from the two subtracted cDNA libraries which were previously synthesized in our laboratory and characterized as mentioned in the previous chapter. Towards this, the present study was started with amplification of 3000 cDNA clones by colony PCR in 96 well formats as mentioned in material and methods. The quality of the PCR products was checked by agarose gel electrophoresis. A high percentage of clones showed single sharp bands and the clones which gave multiple bands were rejected and reamplication was done for such clones. All the amplified products were then purified in 96 well format using PCR purification kit (Eppendorf) following manufacturer's instructions. The amplicons were then reorganized on to the 384 well plates and given for printing. All the clones are present twice on the array. Also many positive and negative controls were included on the array at different places as mentioned in section 3.2.3.

In an effort to dissect the pathways that come into play during plant immune responses, we directed our analyses to the detection of genes that are differentially regulated in two genotypes of chickpea in response to *Fusarium* wilt. Gene expression was measured in JG-62 and WR-315 genotypes which are respectively the susceptible and resistant cultivars of chickpea, at different time points post pathogen inoculation. To minimize experimental variability and ensure accurate representation of changes in mRNA abundance, a standardized regime for plant growth and subsequent fungal infection was used for all the experiments. The experimental design utilizes RNA isolation from the two chickpea genotypes at 6h, 12h, 24h, 48h and 5d post fungal inoculation. A schematic representation of the experimental design for the hybridization is given in figure 3.2. For each hybridization, labelled control and target cDNA samples were mixed and added to a microarray slide which was then kept in a hybridization chamber (corning) for 16 hours at 65°C. Two replicate slides were done for each time point. After the hybridization procedure, the slides were scanned and the images were captured for the hybridized slides of all the time points for both the susceptible and the resistant genotypes (Figure 3.3 and 3.4). These images were then used to calculate the intensity values for each spot on the array.

3.3.3. Gene expression dynamics during time course of compatible and incompatible plant pathogen interaction

The transcript level for each cDNA probe was calculated as average intensity of four data points from duplicate spots of the two replicate slides for each of the time points. We used a fold cutoff of 2.5 and students t-test with $P<0.05$ ranking with FDR multiple testing correction to identify the differentially expressed genes. The reproducibility of the microarray data was evaluated by comparing the signal ratios of duplicate spots for each probe within an array as well as in the biological replicate slides and also by checking the reciprocity of the gene signals in dye swap experiments. Our results indicated that the comparison of the intensity ratios between duplicate spots within an array exhibited a linear relationship with R^2 consistently >0.9 (Figure 3.5A). The average intensity ratios of the replicate slides also showed R^2 >0.8 (Figure 3.5B). Furthermore, Cy5 channel intensity of the original slide showed high correlation to the Cy3 intensity of the dye swap slide (Figure 3.5C). Taken together these results confirmed high reproducibility of the microarray data.
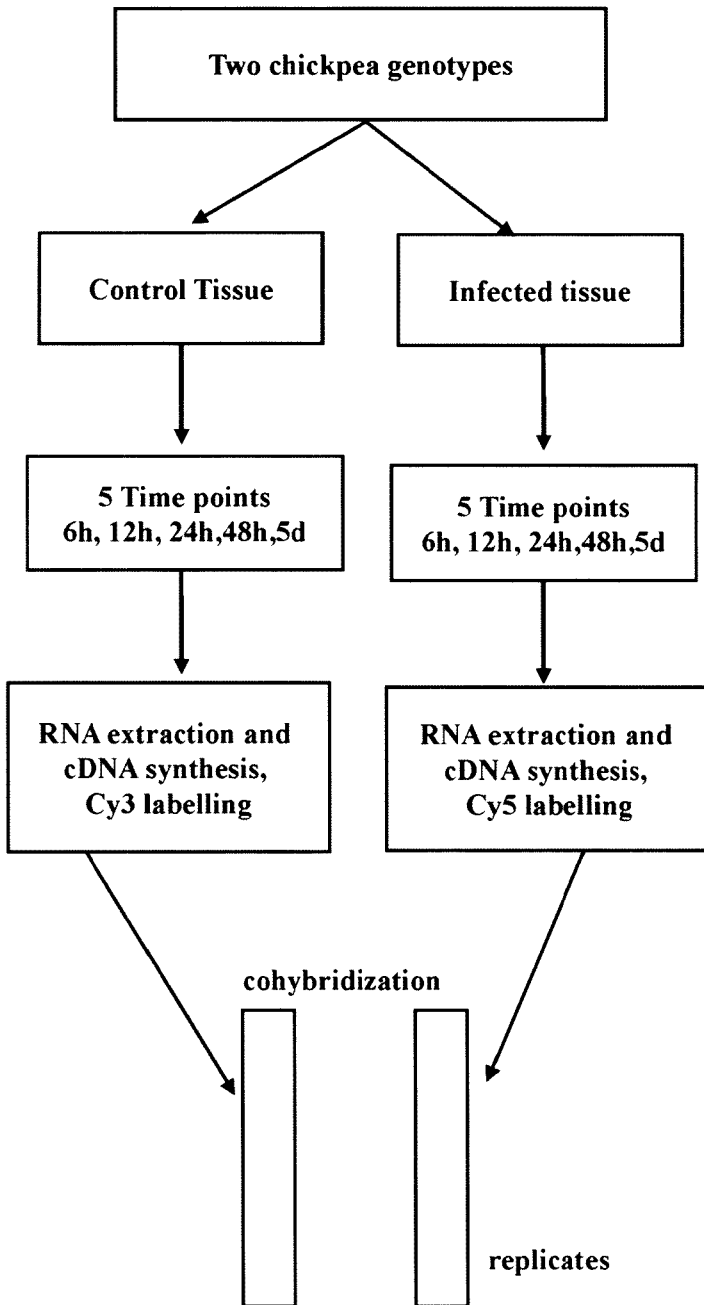
Figure 3.2: Schematic diagram depicting the microarray experimental design.
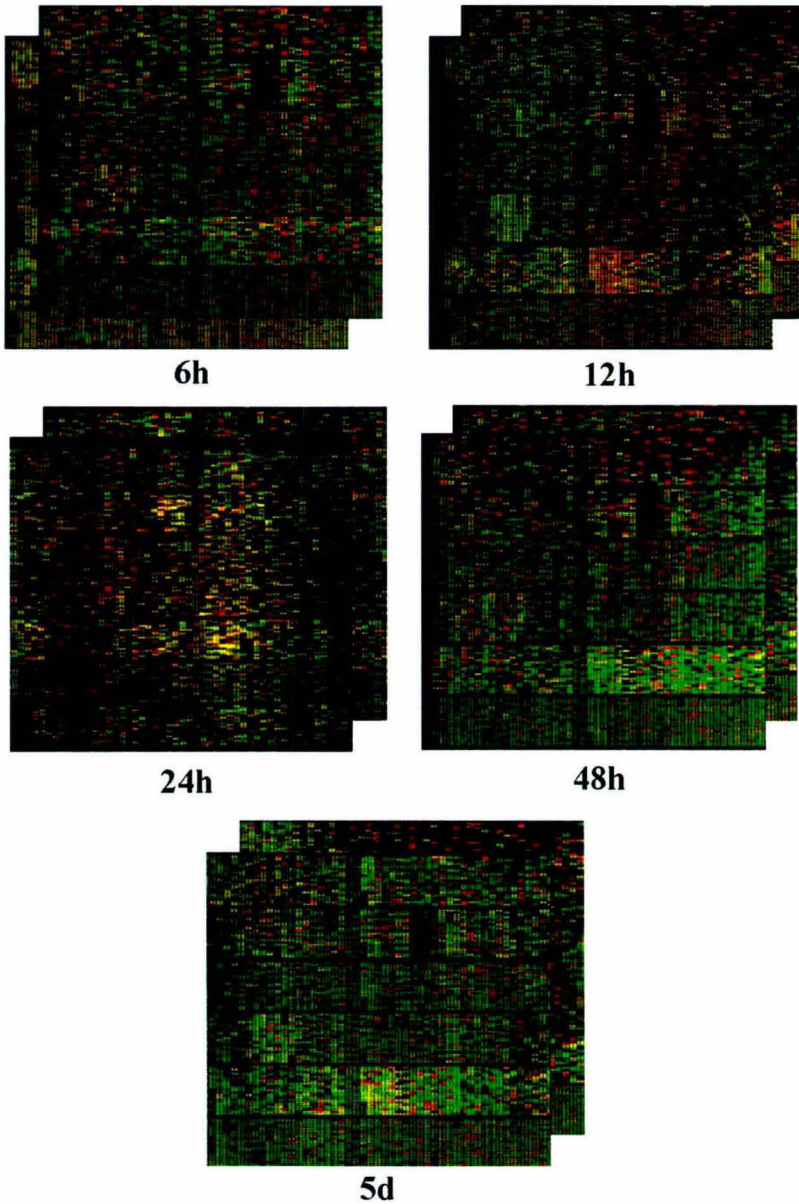
6h

12h

24h

48h

5d

Figure 3.3: Images of microarray slides hybridized with cDNA targets prepared from root tissue of susceptible variety of chickpea harvested at various time points after *Fusarium* infection.
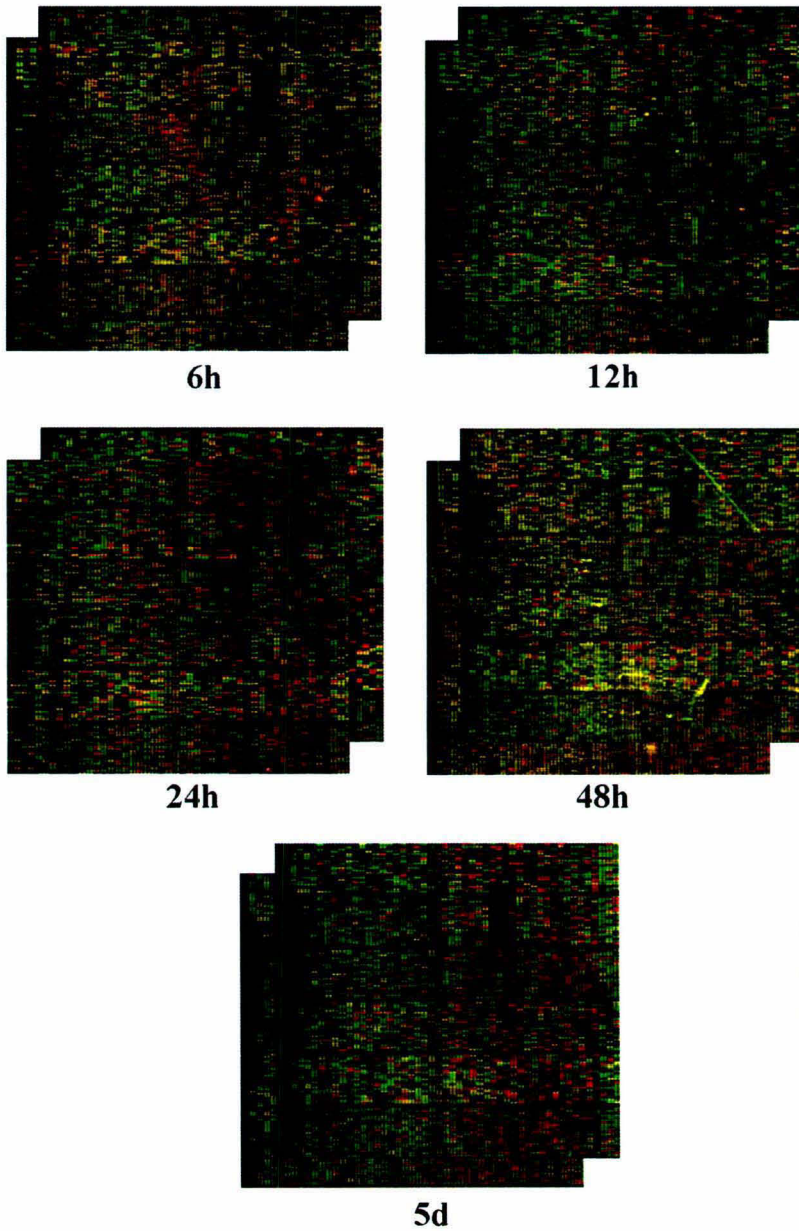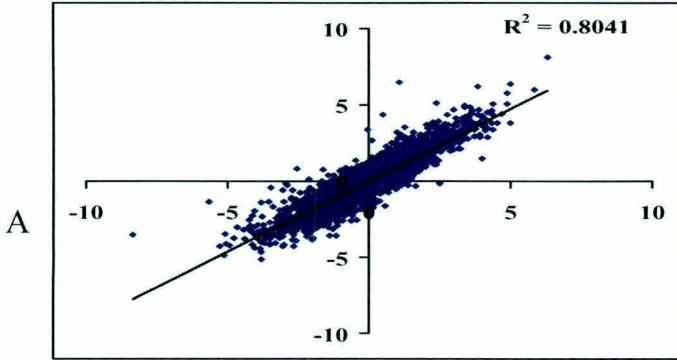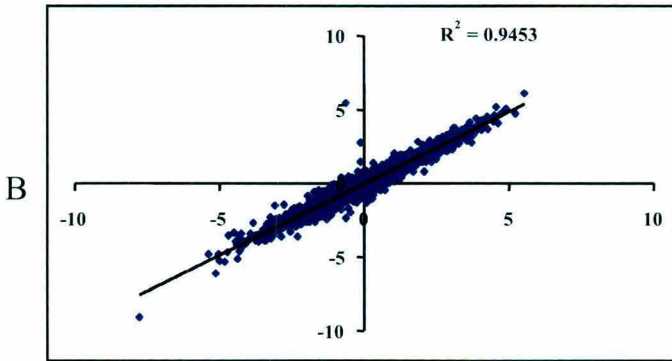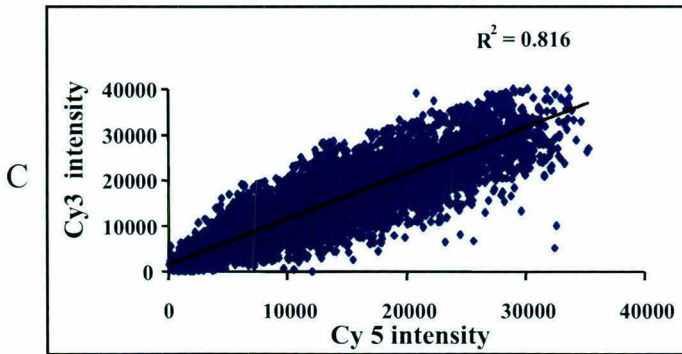
6h

12h

24h

48h

5d

Figure 3.4: Images of microarray slides hybridized with cDNA targets prepared from root tissue of resistant variety of chickpea harvested at various time points after *Fusarium* infection.

A Log of Cy5/Cy3 ratio for spots on two replicate slides

B Log of Cy5/Cy3 ratio for duplicate spots in the same microarray

C Dye swap experiment

Figure 3.5: Scatter plot of log transformed cy5/cy3 ratios of (A) each spot from two replicate slides **(B)** duplicate spots on the same slide. **(C)** Scatter plot of cy3 signal intensity values of a slide versus cy5 signal intensity values of dye swap slide.

Using this approach we found that out of 6058 probe sets representing 2013 unigenes present on the microarray, a total of 1280 unigenes were significantly regulated in at least one of the time points in one of the genotypes. Among these significantly expressed genes, 416 were unique to susceptible; 224 unique to resistant and 640 common in both the genotypes (Figure 3.6A). Furthermore, around 549 genes showed up and 450 downregulation in the susceptible genotype while 57 genes were upregulated in some time points and downregulated in others (Figure 3.6B). Likewise, 456 genes were up and 352 downregulated in the resistant genotype and 56 genes showed a mixed trend (Figure 3.6C).

The time course data shows that each stage of disease development or for that matter plant defense is represented by distinct transcriptional profile in both compatible and incompatible plant pathogen interactions. We detected considerable differences in the gene expression dynamics at different time points after fungal infection in both susceptible and resistant genotypes, the percentage of genes up and down regulated at each time point being different in each genotype. In case of the susceptible genotype, more number of genes showed altered expression at 6h post infection than its resistant counterpart. This might be due the fact that pathogen does not enter the resistant plants at such an early stage. At 12h post infection, both the genotypes showed similar kind of expression pattern. More number of genes was expressed differentially at 24h and 48h after pathogen infection in the susceptible genotype, however, the proportion of upregulated genes was more in resistant genotype at 24h post infection. Approximately same number of genes showed upregulation at 5 days after infection in both the genotypes but more number of genes were downregulated in the resistant one (Figure 3.7).

3.3.4. Specific functional gene groups are regulated during biotic stress in chickpea

The pathogen stress regulated genes identified in this study were categorized into different functional classes in order to gain an insight into their putative role during biotic stress. A large proportion (20.5%) of differentially regulated genes belonged to NSH class. Further 17% of genes matched with hypothetical and 3% with unknown proteins. Rest of the genes belonged to various functional categories as shown in figure 3.8. The name of the genes and their expression pattern at various time points
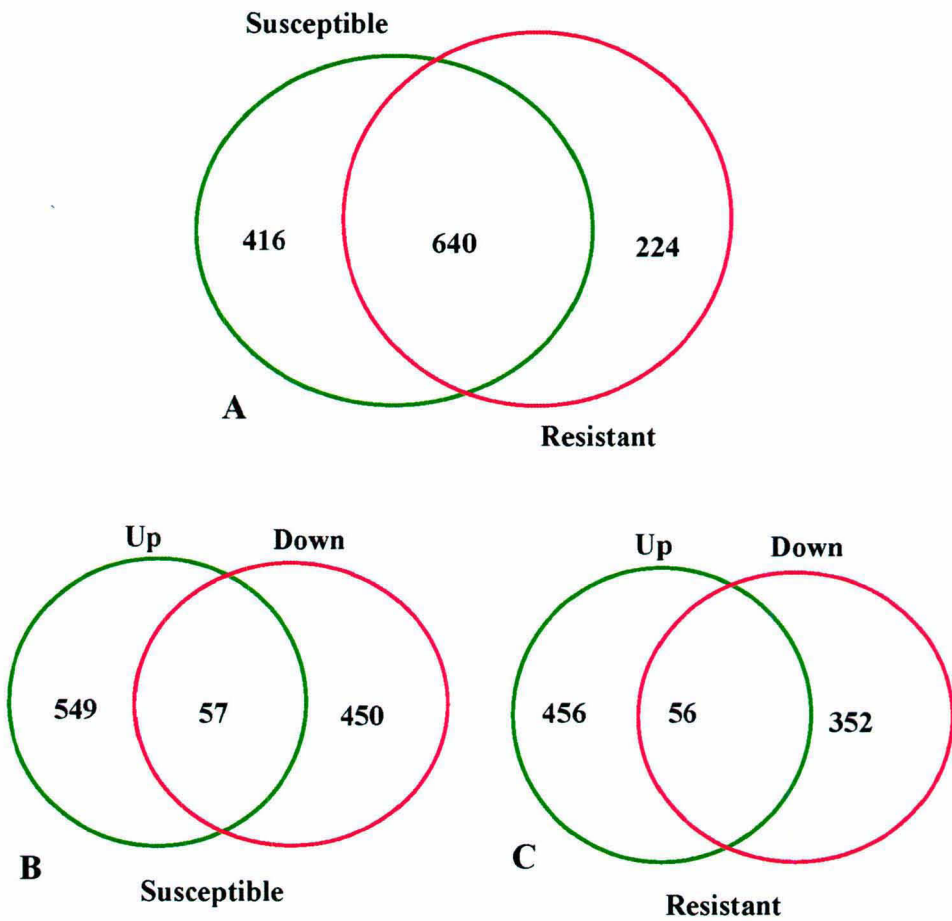
Figure 3.6: Venn diagrams depicting (A) exclusive and overlapping differential genes between susceptible and resistant chickpea genotypes **(B)** genes showing up, down or mixed regulation in susceptible genotype and **(C)** genes showing up, down or mixed regulation in resistant genotype.
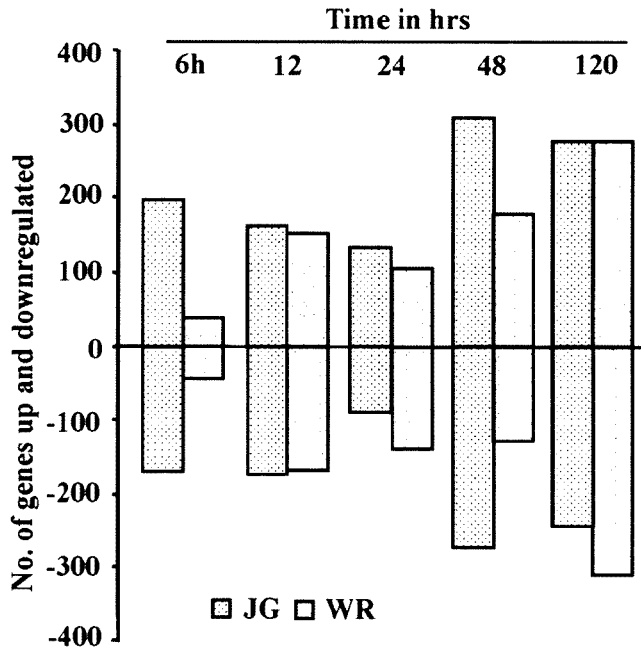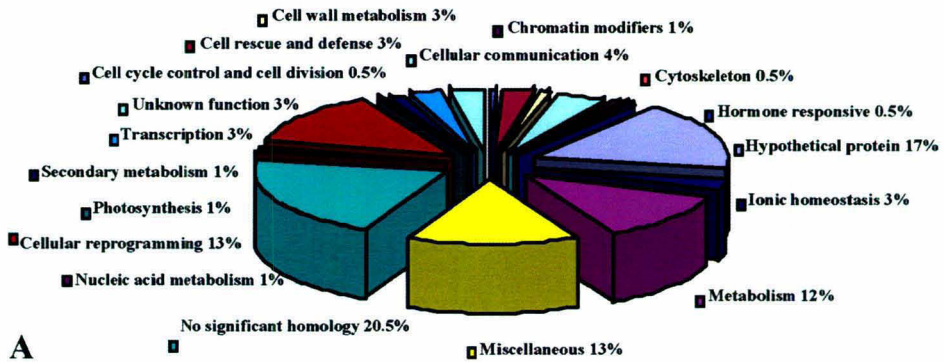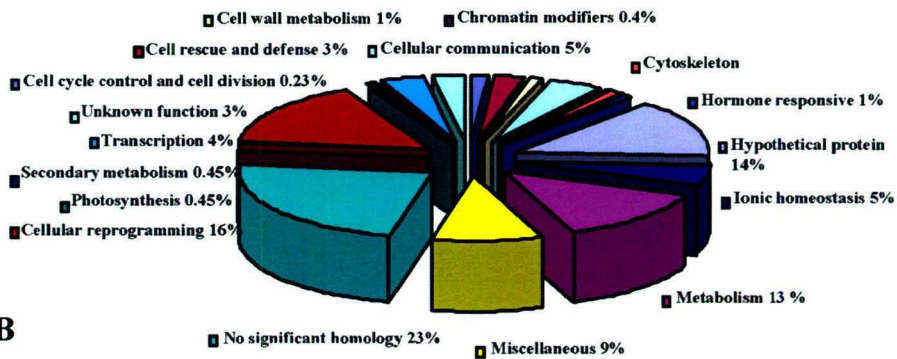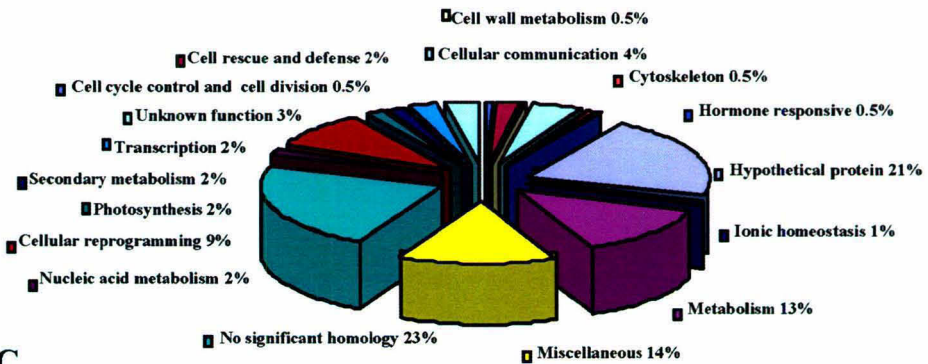
Figure 3.7: Kinetic trends of differential gene expression for susceptible (JG-62) and resistant (WR-315) genotypes of chickpea over the time course after inoculation with *Fusarium.*

Figure 3.8: Functional classification of wilt responsive chickpea genes. The genes identified were grouped into functional classes as shown in the piechart based on metacyc, KOG and GO databases and the values represent the percentage of **(A)** total unigenes **(B)** resistant genotype specific **(C)** susceptible genotype specific genes assigned to various functional classes.

in susceptible and resistant genotypes is given in figure 3.9. Some of the important functional classes are described below.

Genes involved in Cell wall metabolism

Any kind of stress in plants is associated with cell wall reorganization which involves role of many cell wall modifying enzymes. Our results showed upregulation of pectin esterase and xyloglucan:xyloglucosyl transferase, an enzyme involved in cell wall restructuring, in both susceptible and resistant genotypes, suggesting that the pathogen attack may be followed by cell wall restructuring resulting in softening and loosening of primary cell wall that leads to subsequent entry of pathogen into the plant system. Earlier also, these enzymes have been reported to be involved in plant stress (Vercauteren *et al*., 2002; Kagan-Zur *et al*., 1995; Albert *et al*., 2004).

Cell rescue and defense

Plants have evolved a variety of complex array of chemical and enzymatic defenses, both constitutive and inducible and whose effectiveness influences pathogenesis and disease resistance. In confirmation with these reports, we observed strong upregulation of genes coding for class I chitinase and disease response related protein in both susceptible and resistant genotypes, however, the induction was early in the latter case. Genes encoding nsLTP, pathogenesis related protein and putative universal stress related protein showed higher expression in susceptible genotype as compared to the resistant one. The role played by these proteins during plant stress is a well documented phenomenon (Wu and Bradford, 2003). Many other stress related genes for which not much evidence has been collected so far regarding their role in defense responses were also found to be differentially expressed. Among these, aquaporin, extensin and putative Bet v I family protein were found to be downregulated in both the genotypes.

Cytoskeleton

Cytoskeleton helps a plant to prevent itself against pathogen ingress and its dynamics and organization also contribute to the transmission of signals to downstream targets (Lipka and Panstruga, 2005). Our results showed downregulation of actin gene in both the susceptible and resistant genotypes. However, gene coding for actin depolymerising factor ADF-6 was found to be upregulated in resistant genotype.
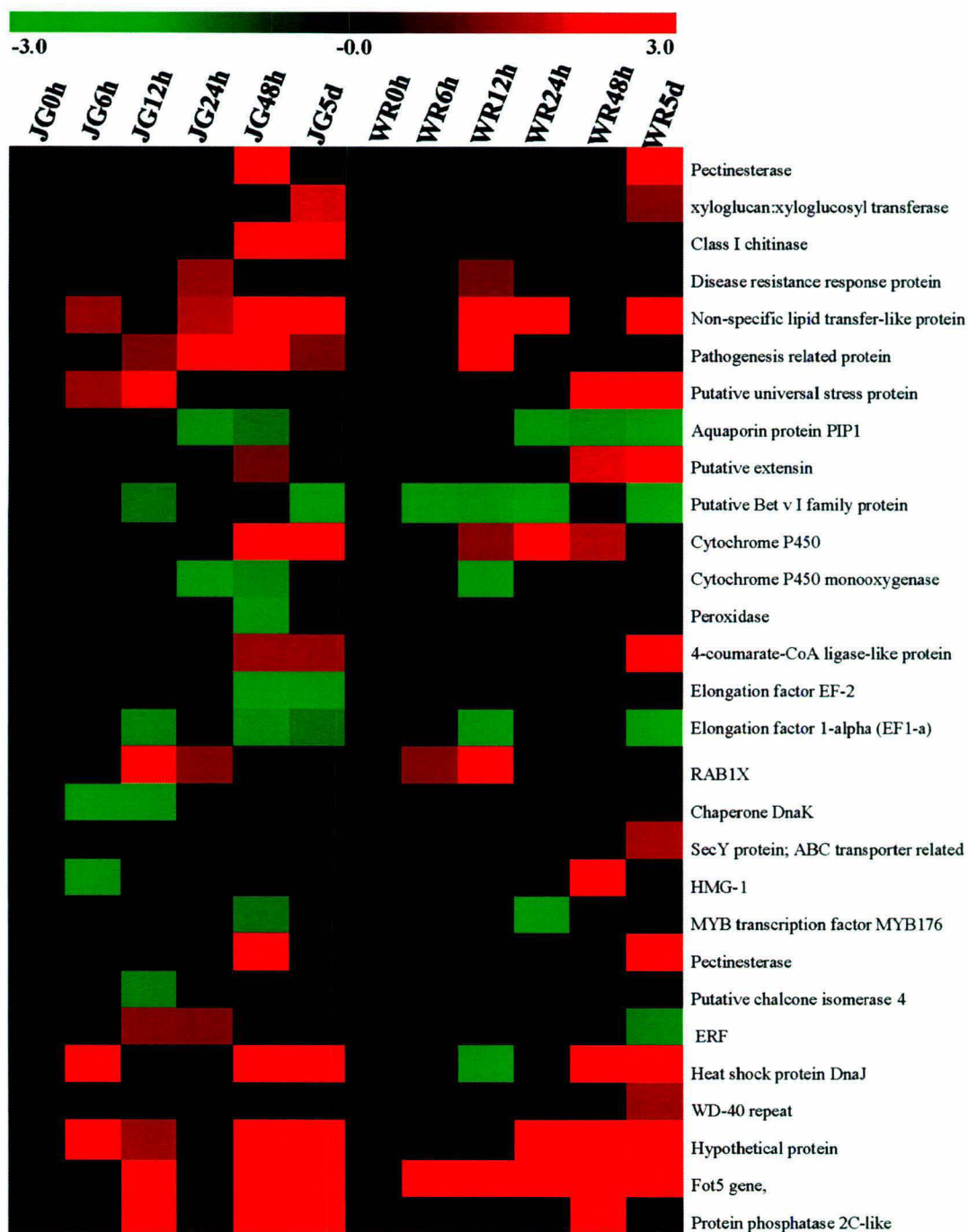
Figure 3.9: Heatmap representing the expression pattern of some of the genes differentially expressed during *Fusarium* wilt in chickpea

Profilin like protein was also induced in both the genotypes. Profilin has several cellular functions and is best known for its ability to promote the exchange of nucleotide in actin monomers released from filaments (Dos Remedios *et al.,* 2002).

Ion homeostasis

ROS have taken a centre stage in signal transduction network of stress-inducible genes (Davletova *et al* 2005). Confirming these reports, our data showed the altered expression of a number of ROS metabolism related genes including cytochrome P450, cytochrome P450 monooxygenase and peroxidase. As a general trend, most of them showed upregulation in both the susceptible and the resistant genotypes. Many other ROS metabolism genes like putative quinone oxidoreductase and cytochrome b5 were found to be downregulated in both the genotypes. These results suggest that different enzymes might play diverse roles in stress responsive pathways.

Genes involved in secondary metabolism

Many secondary metabolites derived from multiple branches of phenylpropanoid pathway are instrumental in plant's ability to mount successful defenses to invading pathogens (Deavours and Dixon, 2005). The enzymes which catalyze the initial steps of this pathway including 4-coumarate-CoA ligase-like protein (CL) were found to be upregulated in susceptible and resistant genotypes. Legumes have some unique enzymatic machinery including cytochrome P450 monooxygenases which helps them in the production of isoflavones (Zabala, 2006). We observed upregulation of many members of this gene family in both the genotypes. Other enzymes involved in isoflavone synthesis like chalcone reductase and chalcone isomerase also showed upregulation in both the genotypes. The expression profile of these genes suggests that there is a strong bias in chickpea towards the synthesis of isoflavonoids rather than flavones or flavonoids in response to pathogen stress.

Pathogen stress induces regulation of genes involved in cell reprogramming

Genes involved in translation, protein processing, cellular trafficking and those encoding GTP binding proteins demonstrate significant changes in their expression during compatible and incompatible plant pathogen interaction indicating that there is a strong reprogramming of cellular processes during plant defense. Several components of translation machinery like ribosomal proteins S6, L10 and S3a and

translation factors EF1A-2 and EF2 were found to be downregulated in both susceptible and resistant genotypes. There are some earlier evidences regarding the effect of stress on ribosome biogenesis but not much is known about this phenomenon. We also observed downregulation of SAR DNA binding protein, the domain analysis of which showed that it is involved in ribosome biogenesis. Interestingly these proteins were found to be downregulated in early time points in the susceptible genotype than in the resistant one. The temporal expression studies of these proteins suggest that stress leads to the suppression of translation machinery and this happens earlier in compatible interaction due to which the susceptible genotype succumbs to pathogen attack easily than its resistant counterpart.

GTP binding proteins are an integral component in a plethora of signal transduction pathways in plants. More recently GTP binding proteins have been implicated in plant defense (Trusov *et al.*, 2006). We observed upregulation of a GTP-binding protein during both compatible and incompatible interactions, however, the fold induction was higher in the former case. The domain analysis showed that the members of this family are involved in translation indicating its possible role in synthesis of proteins related to disease development. The overexpression of a family of rab proteins has been observed in many types of cancers in the mammalian system. Our data showed induction of a Rab protein which belongs to this family. This protein was found to be expressed during early time points in the susceptible genotype while in case of resistant genotype; its induction was limited to later stages after pathogen stress. This observation was indicative of the possible involvement of this protein in plant disease development, congruent with its role in mammalian cells. Another gene encoding a RAB1X protein, which is known to be involved in vesicular transport, was induced in both the genotypes. This suggests the possible involvement of this gene in mediating vesicular trafficking during defense mechanisms.

Selective protein processing and degradation by ubiquitin/proteosome pathway has recently emerged as a powerful regulatory mechanism in a wide variety of cellular processes. It was quite interesting to see that genes encoding various components of ubiquitin/proteosome pathway are differentially expressed during chickpea-*Fusarium* interaction. Further, these proteins showed varied expression trends in the two genotypes. For example, we observed downregulation of 26S protease regulatory subunit 6, DnaK, cyclin-like F-box, HSP20-like chaperone and ubiquitin-conjugating

enzyme, E2 in susceptible genotype. These proteins might have a role to play in defense signaling and their downregulation may render a plant susceptible to pathogen attack. Some other components of the proteosome pathway like putative ubiquitin-conjugating enzyme were found to be downregulated in the resistant genotype. This indicates that the regulated processing of endogenous proteins can contribute to multiple levels of plant defense.

Several important components of vesicle trafficking pathway in eukaryotic cells have been found to play a variety of roles in plant physiology, development and stress (Collins et al., 2003). We observed induction of many such genes in chickpea in response to wilt. A gene encoding GOLGI SNARE 11 (GOS11) showed induction in both genotypes suggesting its role in pathogen stress. Many other transport associated proteins like SecY protein; ABC transporter related protein was found to be induced during resistance pathway. Such components may be involved in the transport of pathogenesis related proteins from the site of their synthesis to the site of their action.

Regulation of chromatin modifiers, Transcription factors and other signalling mediators during pathogen stress

Any kind of stress can have restructuring effect on the genome of an organism. Similarly, pathogen attack can induce remodelling and regulatory changes in the genome of a plant. In the present study, concomitant differential expression of many chromatin modifiers and transcription factors in concert with signalling components illustrates the interplay of sub-networks of transcription modulation. Altered expression of proteins involved in different combinations of histone modifications for example, deacetylases, methylases, HMG proteins and histone variants may result in distinct outcomes in terms of chromatin-regulated functions. For example, downregulation of histone deacetylase HDT1 in both susceptible and resistant genotypes likely indicates chromatin decondensation for loading of appropriate factors, driving new transcription patterns. Earlier, histone deacetylase has been shown to be involved in jasmonic acid and ethylene signalling during defense response (Zhou et al., 2005). Further, HMG proteins act as versatile modulators of chromatin function (Stros et al., 2007). In the present study HMG-1 was found to be downregulated in susceptible and upregulated in resistant genotypes suggesting increase in the structural flexibility of DNA, promoting the assembly of nucleoprotein

complexes that control DNA-dependent processes including transcription of resistance related genes. A histone variant, histone H3.2 was found to be downregulated in both genotypes. We also observed downregulation of histone 2 and histone 3 in both the genotypes, however, this repression in expression was during early time points in susceptible genotype. Many other broad spectrum as well as stress responsive transcription factors showed differential expression during chickpea-*Fusarium* interaction. Broad spectrum transcription factors represented many classes like Myb, leucine zipper, bHLH, Zinc finger, etc. and were found to display diverse patterns of expression during the time course of defense response in the susceptible and the resistant genotypes. In this context, MYB176 was repressed in both the genotypes. Further, putative AP2-binding protein was repressed and homeodomain leucine zipper protein HDZ1 induced in susceptible genotype. The members of these families of transcription factors are known to play role in plant stress (Vailleau, 2002; Kim *et al.*, 2004).

Plants are under constant threat of pathogen attack and have in turn developed sophisticated detection and response systems that decipher pathogen signals and induce appropriate defense responses. In the present study, many cell signalling components were found to be induced in both susceptible and resistant genotypes and included serine/threonine/tyrosine kinase, leucine-rich repeat transmembrane protein kinase and putative protein kinase. Also putative mitogen-activated protein kinase kinase kinase was found to be upregulated during compatible interaction. Protein phosphatase 2C-like protein was found to be upregulated in both susceptible and resistant genotypes, however, expression was higher and more consistent in susceptible one. Many families of 14-3-3 proteins also showed significant changes in expression during chickpea - *Fusarium* interaction. To our surprise, our study showed downregulation of many calcium related proteins like calmodulin and calcium-dependent calmodulin-independent protein kinase isoform 2 during defense response. These results suggest that a range of signalling components are differentially regulated and may in turn have a role to play in resistance or susceptibility.

3.3.5. Cluster analysis of genes differentially regulated during pathogen stress

In order to further evaluate the impact of *Fusarium* wilt on gene expression in chickpea, we used K-means clustering. To facilitate the detection of differences and

similarities of gene expression in the susceptible and resistant cultivars, a combined data set comprising of 1280 unigenes that showed detectable expression in at least one time point in one of the genotypes was used for cluster analysis. The genes in this data set were grouped into 10 clusters (Figure 3.10). We also analysed these groups of co-expressed genes with respect to their distribution into functional categories. Clusters 2, 6 and 10 consisted of genes co-regulated during compatible and incompatible interactions while as rest of the clusters mostly consisted of genes expressed in either of the two genotypes. Cluster 2 consisted of 69 genes upregulated at 24h and 12h post pathogen infection in susceptible and resistant genotypes, respectively. The genes in this cluster included those coding for disease resistance response protein, polygalacturonase-inhibiting protein and peroxidase1C which are well known for their role in defense (Di Matteo *et al.*, 2003; Kotchoni and Gachomo, 2006). Their early expression in resistant genotype may lead to resistance response. Cluster 6 consisted of 109 genes upregulated mostly at 48h and 5d after pathogen attack in both the genotypes. Some genes also showed upregulation at 6, 12 and 24 hpi in susceptible and 12 and 24 hpi in resistant genotype. Putative Pi starvation-induced protein, putative kinetochore protein, putative protein kinase, 14-3-3-like protein were few of the members of this cluster. Early expression of such genes in susceptible genotype indicates their involvement in disease development. While 14-3-3 has been implicated in plant defense earlier also (Chen *et al.*, 2006), there are no reports of the involvement of other above mentioned genes in plant immune responses. A total of 88 genes were grouped into Cluster 10 and included those upregulated at 48h and 5d after fungal infection. In all these three clusters, metabolism and cellular reprogramming were the abundant functional classes. Cluster 1 and 3 consisted of genes downregulated mostly in the susceptible genotype. Cluster 3 was the largest group and apart from metabolism and cellular reprogramming related genes, majority of transcription factors belonged to this cluster. Cluster 4 and 7 included genes downregulated in resistant genotype. The genes which were upregulated during susceptible genotype were grouped into cluster 5 and 8. Many important genes like ERD, Cationic peroxidase 2 precursor and transcription Factor IIF, Rap30/Rap74 were present in this cluster. Although cationic peroxidase is known to be involved in plant defense (Reimers *et al.*, 1992), role of ERD and Rap30/Rap74 is yet to be understood. Cluster 9 consisted of genes which showed mixed trend in expression, however, harboured many genes which were regulated in the resistant genotype only
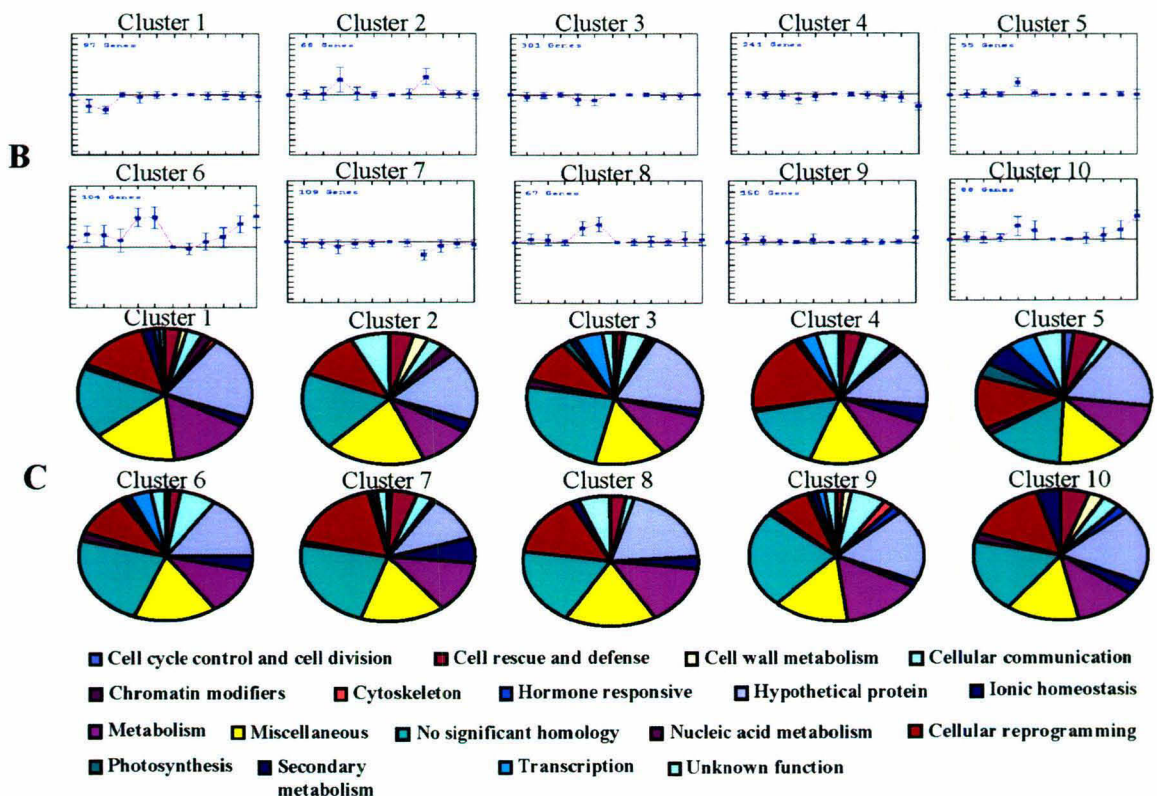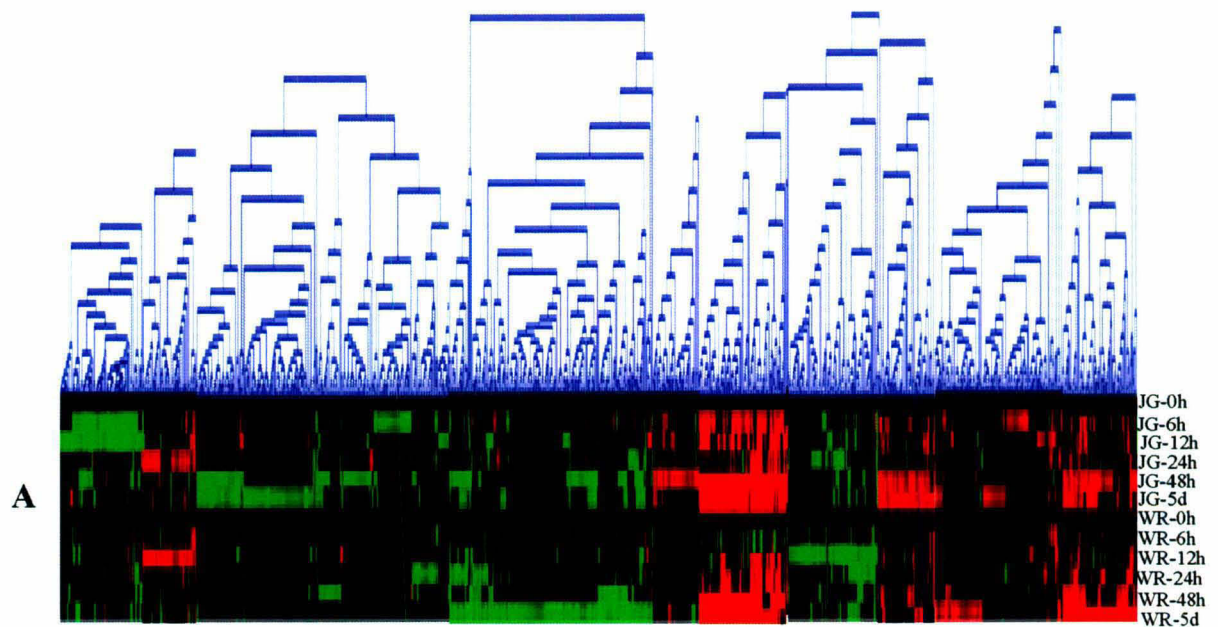
Figure 3.10: Tree view, expression profiles and the functional categories for clustered genes. (A) K-means clustering analysis on 1280 genes which showed significant regulation (fold ≥ 2.5, P < 0.05) in at least one time point in at least one genotype. Genes were clustered into10 clusters. Red color represents upregulation; green, downregulation and black no change and the intensity of the color represents the value of the ratio. **(B)** Expression profile of each individual cluster and **(C)** pie charts showing distribution of genes to major functional categories.

like ras-like small monomeric GTP-binding protein, WD40-like, LRR strongly suggesting their role in defense.

### 3.3.6. Validation of microarray data by quantitative real time PCR

To validate the gene expression profiles obtained from microarray data, quantitative RT-PCR was performed on 12 candidate genes belonging to different functional groups using gene specific primer pairs. The reference gene chosen for the normalization was 18S RNA. The comparative Ct method was used. The overall patterns of up and down regulation were conserved for all the genes at each time point studied, although the ratios were higher in case of quantitative RT PCR. The results showed that PR10 which is marker gene for plant response to necrotrophic pathogens was upregulated during all time points in the resistant genotype. However, in case of susceptible one, it was downregulated at two time points. PR5 showed upregulation in both the genotypes at all time points. Same was the case with chalcone isomerase and water channel protein. Ethylene response factor 5 (ERF5) which is involved in ethylene mediated signalling during plant defense was found to be upregulated in all time points in resistant genotype while in case of susceptible one, it showed down regulation in a few time points while at other time points it was upregulated. DnaJ, cystain and pectin esterase showed upregulation in both the genotypes, however, expression was more in case of resistant one. In case of extensin and WD40, transcript abundance was more in resistant genotype during all time points except at 6h post infection. Interestingly Fot 5, a fungal gene also was upregulated during time course in both the genotypes and expression was more in the resistant genotype. A hypothetical protein showed higher expression in susceptible genotype as compared to the resistant one (Figure 3.11). The results showed that the expression patterns detected by chickpea array were in good correlation with those obtained by qRT-PCR.

3.3.7 Plant immune responsive pathways

DNA microarray technology has opened up the possibility to study gene expression on a wide scale and advances are being made towards understanding of the transcriptional programs during various biological processes and in response to various forms of stresses. In the present study, we have made an attempt to get an overview of the regulatory network involved in the fungal stress. Our EST data as described in the chapter 2 showed that although a large number of ESTs are common
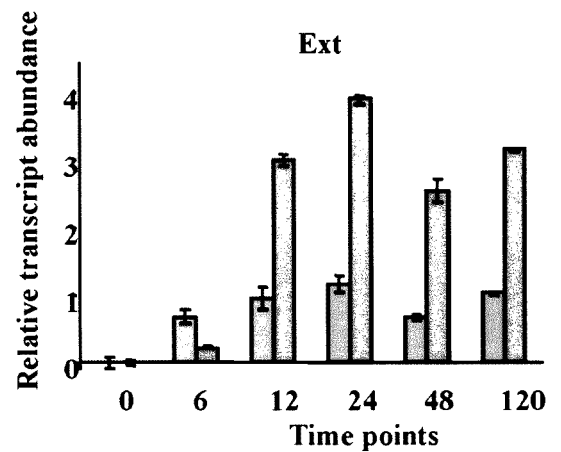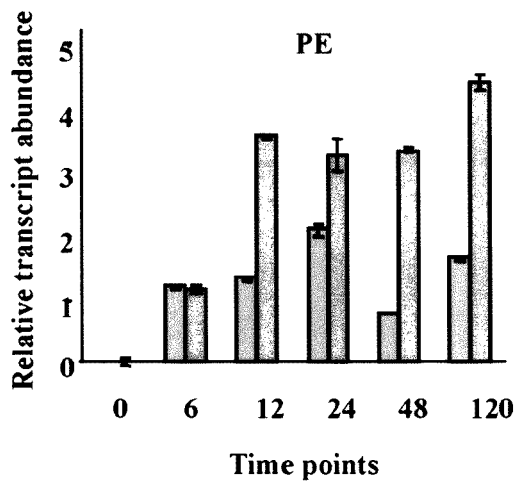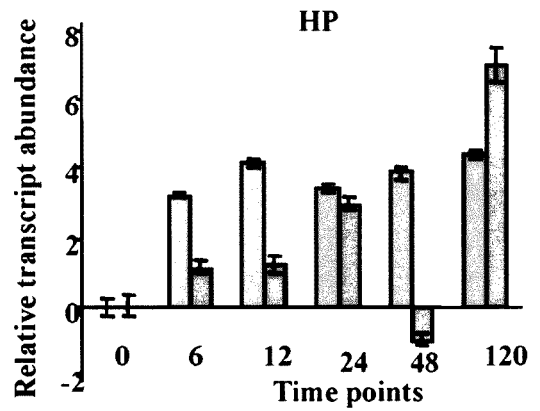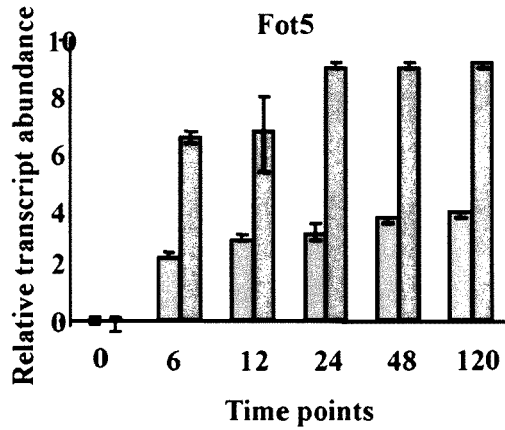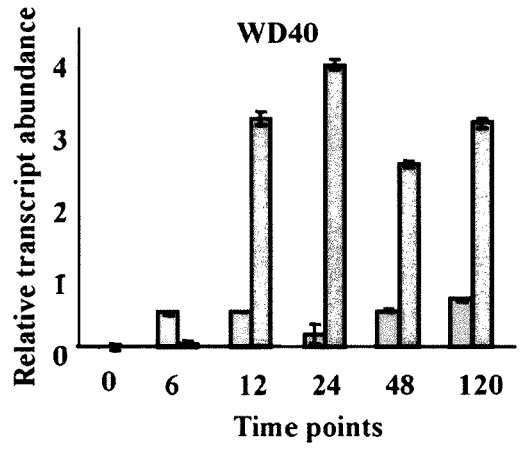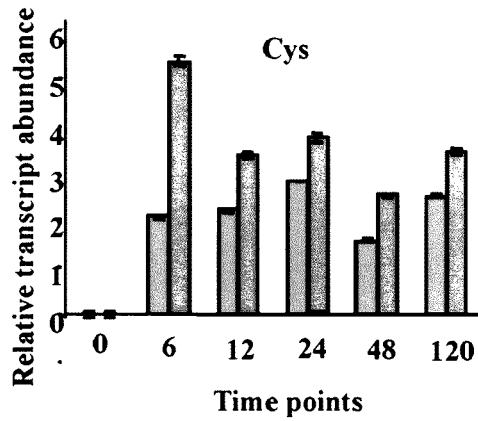
Cys – Cystatin; HP – Hypothetical protein; PE – Pectin esterase; Ext - Extensin
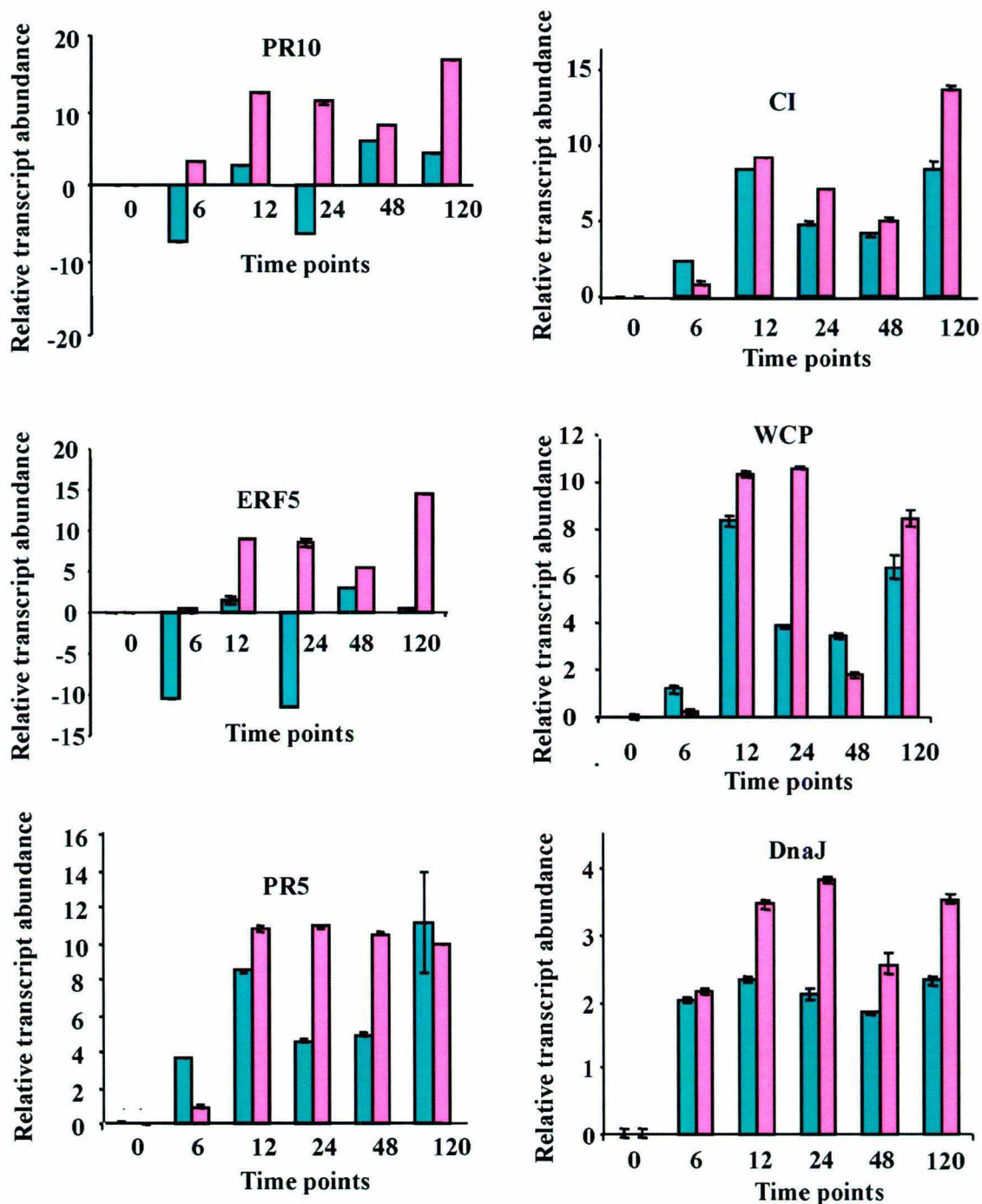
PR10 - Pathogenesis related protein 10; CI - chalcone isomerase; ERF5 - Ethylene responsive factor 5; WCP - water channel protein; PR5 - Pathogenesis related protein 5

Figure 3.11: Bar chart showing the relative transcript level of genes plotted as $\log_2$ ratio against various time points in hours after *Fusarium* infection. Transcript levels were normalized by 18S transcript level. Error bars indicate SD of three real time PCR experiments. Blue color indicates expression ratios of genes in susceptible genotype and pink those in resistant genotype.

between the susceptible and resistant genotypes, a good number is specific to each of the genotypes. Microarray analyses also led to the identification of many genes which showed differential expression either in the susceptible or in the resistant genotype besides those which were common in both. This suggests that although a large part of the differences between susceptibility and resistance might be due to the quantitative differences of common regulatory pathway, there might be some separate branches in this network that make a plant susceptible or resistant. Here we have hypothesized a pathway composed of genes which are differentially expressed in both the genotypes (Figure 3.12) and the pathways composed of genes which are exclusively present in either of the genotypes (Figure 3.13).

As a general mechanism, the recognition of pathogen elicitors initiated by various receptors leads to the signaling events like phosphorylation/dephosphorylation, ion fluxes and oxidative burst. Various protein kinases and receptor kinases were present in the network which might be involved in perceiving the signals and/or relaying them to the downstream proteins. These included leucine-rich repeat (LRR) trans-membrane protein kinase, putative protein kinase, etc. LRR kinases have been implicated in plant defense responses (Dievart and Clark, 2004), however, there is no report of the involvement of APK1 in plant immune responsive pathways. The phosphorylation events might lead to the activation of downstream proteins like 14-3-3 and WD40. 14-3-3 proteins are known to relay phosphorylation events to downstream proteins which in turn perform various cellular functions (Dougherty and Morrison, 2004). Stress often leads to reprogramming of many cellular processes including translational and post translational processes. Many translation associated proteins like ribosomal proteins (L42 and S13) and translation factors like SUI1 showed altered expression in response to stress suggesting their role in immunity. Genes involved in protein modification and turnover like PDI, ubiquitin extension protein and polyubiquitin were also differentially expressed suggesting that these processes play a crucial role in enabling plants to alter their proteome to maximize their chances of survival under stress. Many GTP binding proteins were also present in the network strongly suggesting their role in immune responsive pathways. GTP binding proteins might work by exerting their influence on several processes like translation, transport etc. Transcriptional regulation of stress responses is a well documented phenomenon. In this context we observed many chromatin modifiers and
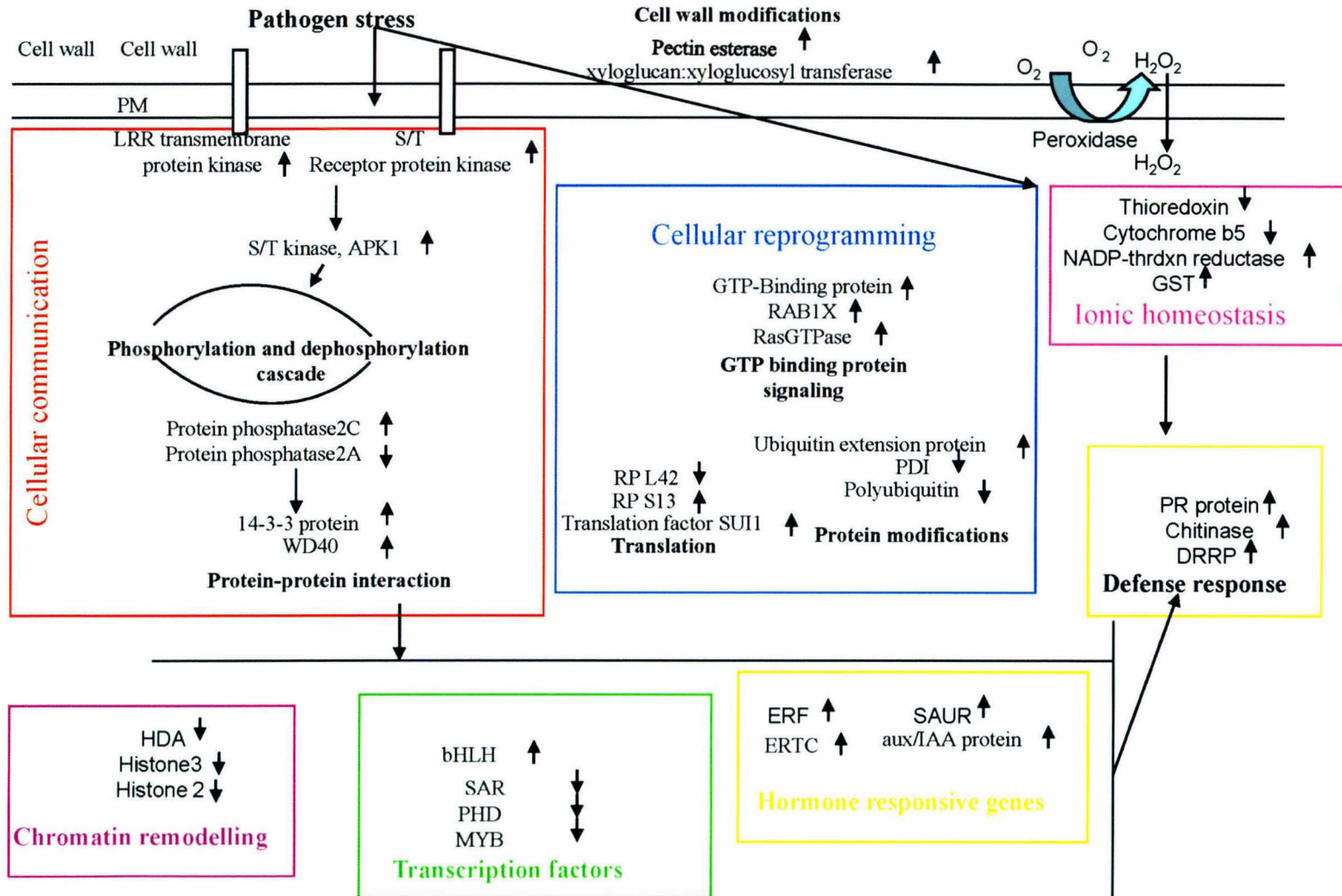
Figure 3.12: Hypothetical immune responsive network constituting genes which showed quantitative differences in expression in the resistant or susceptible genotypes.
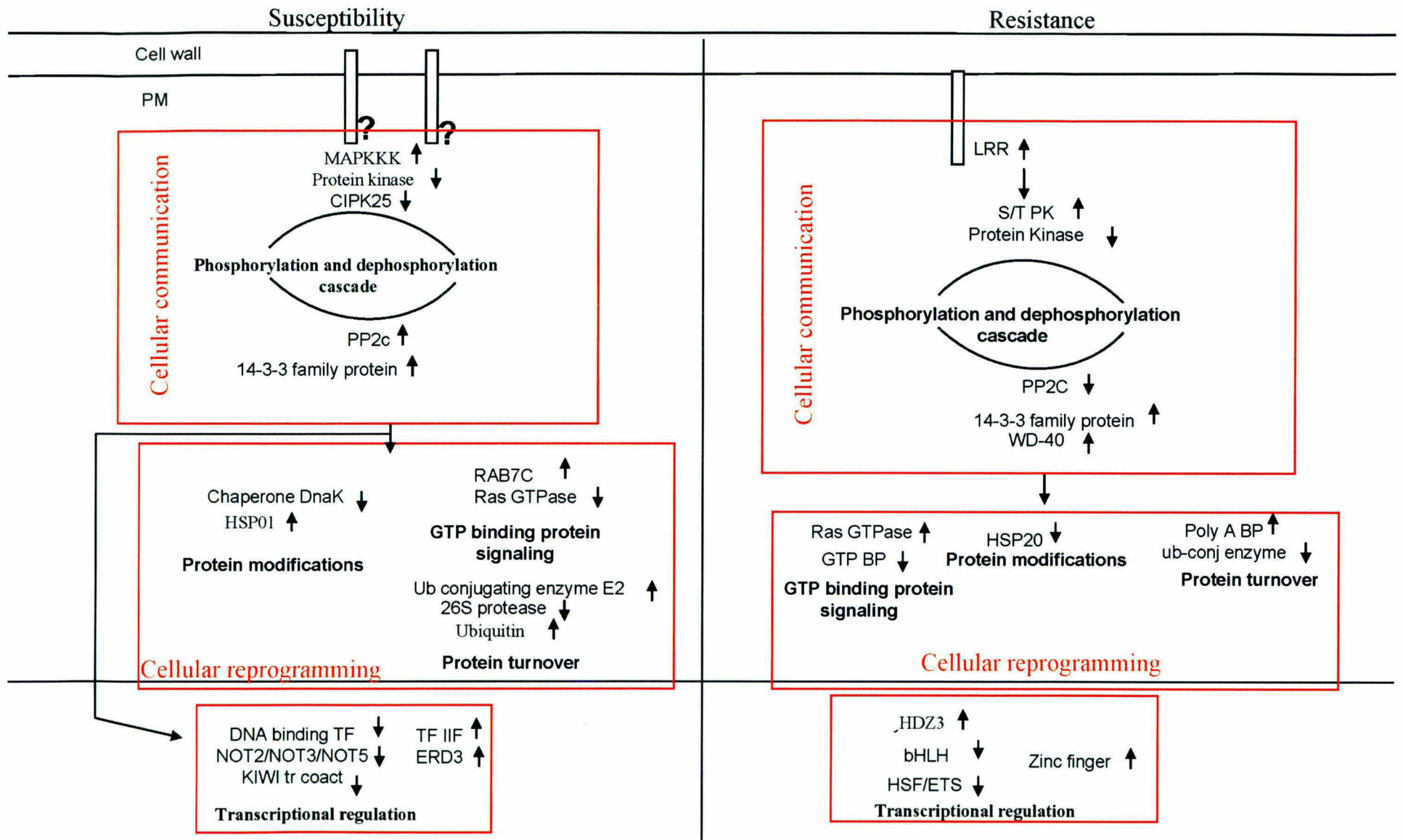
Figure 3.13: Hypothetical immune responsive network constituting genes which were differentaily expressed in either resistant or susceptible genotypes.

transcription factors showing altered expression thus suggesting their role in stress. These include histone 2 and 3, bHLH, Myb, SAR transcription factors. Their expression might be controlled by some kinases present upstream in the network and these in turn might control expression of defense related genes. As a general trend all these genes were differentially expressed in both susceptible and resistant genotypes, however, the extent to which they are expressed or repressed varied between the two genotypes. This suggested that quantitative differences in the expression of such genes might result in susceptibility or resistance.

Apart from the pathway shared by the susceptible and the resistant genotypes, on the basis of genes expressed exclusively in either of the genotypes, we hypothesised a model that might have a role to play in making a plant susceptible or resistant to pathogen stress. We observed that MAPKKK, CIPK and a protein kinase were present only in the susceptible pathway while as genes encoding LRR and some different serine threonine kinases were present in the resistant genotype. Thus, the pathogen elicitors might be perceived by different components in the two genotypes. Downstream to these genes, many cell signalling components like 14-3-3 and RAB7C, were present in the susceptible genotype while as the resistant genotype showed expression of 14-3-3 which is a different member of this gene family. Also a different GTP binding protein was present in the resistant genotype. On the basis of this observation, it can be speculated that different members of gene families might have a role in mediating genotype specific response to pathogen stress. The activity might be controlled by the phosphorylation or dephosphorylation events and these in turn might perform different cellular processes during stress. Several proteins involved in protein modification and turn over were also present and included chaperon DnaK, 26S proteosome in susceptible genotype and HSP20 and ubiquitin conjugating enzyme in the resistant genotype. These proteins might be associated with the degradation of some repressors or in the modification of proteins into the active or the inactive states thereby altering their cellular function. Further, a very important observation was the expression of different classes of chromatin modifiers and transcription factors in the two genotypes. For example, NOT2/NOT3/NOT5 and KIWI transcriptional coactivator showed downregulation and TFIIF and ERD showed upregulation in susceptible genotype. In case of resistant genotype, bHLH gene was repressed and a zinc finger gene induced. This suggests that different transcription factors might

regulate expression of different set of genes which in turn may bring about differences in the host response towards a pathogen thus making it susceptible or resistant.

Conclusion

As the advent of microarray technology made it possible to study expression of thousands of genes, in the present study this technique was applied to study the transcript profile of chickpea in response to *Fusarium* wilt. The comparison between susceptible and resistant genotypes was also carried out so as to gain an insight about the genotype specific host response to the wilt disease. The results showed that although, there was a large overlap between genes expressed in susceptible and resistant genotypes, a significant set of genes was expressed exclusively in either of the genotypes indicating that there might be separate signaling branches leading to susceptibility and resistance. The gene expression dynamics also differed between the genotypes suggesting that temporal and spatial changes might lead to genotype specific response. Further, the genes differentially expressed belonged to varied functional classes suggesting that there is cellular reprogramming during plant stress response. Using K-means clustering, co-expression pattern of genes was studied which revealed that many genes showed a similar pattern of expression and might be responsible for exerting joint influence on stress responsive action. Real time PCR results suggested a high correlation with that of microarray, thus confirming and validating the microarray results. The construction of hypothetical pathway provides a layout of possible wilt responsive or in a broader perspective plant immune responsive regulatory network.

Small Auxin Up RNA (SAUR) – CaSAUR1 a novel defense responsive gene

4.1 Introduction

Auxin is an indispensable phytohormone with a well-documented ability to regulate many aspects of plant development. Auxin modulates diverse processes such as tropic responses to light and gravity, general root and shoot architecture, organ patterning, vascular development and growth in tissue culture (Davies, 1995). Auxin-induced transcripts, auxin biosynthesis, metabolism, and transport together ensure that appropriate auxin levels are in place to orchestrate plant development (Woodward and Bartel, 2005). The effect of auxin on a growing plant depends on the type of auxin applied and its concentration. Auxin influences aspects of cell division, cell elongation and cell differentiation, although exactly how it is involved in each process and to what extent they are intertwined is not completely understood (Teale *et al.*, 2006). Its diverse effects in plants might also extend to animals, as photoactivated auxin seems to have potential as a cytotoxin in cancer therapy (Folkes and Wardman, 2001).

Recently auxin signaling has been shown to play a role in plant-pathogen interactions. Auxin homeostasis acts as an essential constituent of the complex hormone network that modulates plant responses to different pathogens and pests. Auxin is known to have contrasting effects on the progression of disease caused by biotrophic and necrotrophic pathogens. Auxin signalling is shown to have a negative effect on resistance to biotrophic pathogens and repression of auxin signaling is a part of bacterial-induced plant immune response (Navarro *et al.*, 2006 and Wang *et al.*, 2007). It was also demonstrated that a flagellin-induced micro-RNA represses auxin signaling by targeting auxin receptor genes and makes host plants less susceptible to bacterial infection. On the other hand, repression of auxin signalling either through mutations in the auxin pathway or by pharmacological interference with the auxin response has been shown to impair resistance to the necrotrophic fungi (Llorente *et al.*, 2008). Further work is necessary to elucidate the genetic network of auxin responses and signaling cascades.

Auxin functions by regulating a group of primary responsive genes: *Aux/IAA* genes, *GH3* genes, and small auxin-up RNAs (*SAUR*) (Hagen and Guilfoyle, 2002). Members of the *Aux/IAA* gene family have been studied in light regulation of auxin responses (Woodward and Bartel, 2005). *GH3* genes have also been shown to play

role in light-auxin interactions (Nakazawa *et al.*, 2001). GH3 proteins are also implicated in stress responses (Park *et al.*, 2007). However, none of the *SAUR* genes are as yet functionally characterized.

SAUR genes are transcriptionally induced by exogenous auxins within a few minutes after hormone application (Gil *et al.*, 1994). *SAURs* can also be induced by cycloheximide, a translational inhibitor, indicating that their transcription is regulated by a short-lived repressor. The *SAUR* genes were originally characterized in soybean (McClure and Guilfoyle, 1987). Members of this class have also been isolated from mung bean, pea, *Arabidopsis*, tobacco and more recently, maize (Jain *et al.*, 2006). These transcripts have been localized to tissues that are targets for auxin-induced cell elongation (Gee *et al.*, 1991). The appearance of the *SAURs* before auxin-induced cell elongation suggests that they may contribute to the process. The *SAUR* transcripts accumulate in the cells that are destined to elongate, presumably due to a rapid redistribution of endogenous auxin. Their rapid disappearance from cells that are not targeted for enhanced elongation indicates that the *SAUR* transcripts are highly unstable (McClure and Guilfoyle, 1989). The short half-lives of SAUR mRNAs appear to be conferred by downstream elements (DSTs) in the untranslated region of the messages (Sullivan and Green, 1996). There is no evidence that auxin affects the function of DST sequences or the stability of the *SAUR* transcripts (Newman *et al.*, 1993). Therefore, if DST sequences act to destabilize the SAUR transcripts, they may do so constitutively so as to allow the SAUR mRNA level to adjust rapidly in response to decreases and increases in the auxin concentration. The DST elements are represented by GGA---TAGAT-----T--GTA. In all cases, however, the sequences TAGAT in the central region and GTA in the 3' region of the DST element are invariant.

SAUR is represented by a large multigene family in plants. For example, *Arabidopsis* genome comprises more than 70 members (Hagen and Guilfoyle, 2002). In rice, around 58 *SAUR* genes have been identified which are distributed on 10 out of 12 rice chromosomes (Jain *et al.*, 2006). The coding sequences of *SAURs* do not possess any intron. Most of the predicted SAUR protein sequences harbour a putative nuclear localization signal at their N-terminus. Localized gene duplications appear to be the primary genetic event responsible for SAUR gene family expansion as studied in rice

(Jain *et al.*, 2006). Interestingly, the duplication of Os*SAUR*s was found to be associated with the chromosomal block duplication as well.

The auxin-inducible elements have been characterized in the promoters of soybean *SAURs* (Li *et al.*, 1994) and *Arabidopsis SAUR-AC1* (Gil *et al.*, 1994). Within the promoter regions of the soybean *SAUR* genes, the most prominent conserved elements are the NDE and DUE elements, both of which are also found in the *SAURAC1* promoter, albeit in the opposite order. Several other sequence motifs that have been implicated in auxin-responsive expression, either experimentally or on the basis of sequence conservation, are also present upstream of the *SAURAC1* transcription start site. The NDE present in the SAUR promoter is necessary and sufficient for auxin induction. The NDE contains two adjacent sequences, TCTCTC and CCTCCCAT, which have been previously identified as putative auxin-responsive elements. A *SAUR* gene from *Zea mays* (*ZmSAUR1*) was identified as a Ca/CaM-binding protein suggesting that Ca/CaM might regulate the function of this early responsive gene at the post-translational level (Yang and Poovaiah, 2000). The finding that the *SAUR* gene family encodes Ca/CaM-binding proteins has important implications for understanding the "cross-talk" between the calcium/CaM messenger system and auxin signal transduction.

Till date, there has been no report regarding the involvement of SAUR in plant immune responses. Moreover, no functional role has been assigned to these genes. From our transcript profiling experiments, we observed induction of SAUR in response to *Fusarium* wilt in chickpea. This prompted us to investigate the role of *CaSAUR1* in plant immunity. In the present study, we demonstrated that its expression is modulated by the pathogen and many other hormones apart from auxin thereby adding another dimension to the complex and dynamic role of auxin signalling during plant-pathogen interaction. The present study provides basic genomic information for the chickpea *SAUR* gene and will pave the way for deciphering the precise role of SAURs in plant immunity.

4.2 Material and Methods
4.2.1 Cloning of full length *CaSAUR1*

Full length cDNA clone of *CaSAUR1* was obtained by performing 3'RACE using gene specific primer (SAUR-3'F) and UAP primer provided with the 3'RACE kit

(Invitrogen) as detailed in section 4 of appendix III. The gene specific primer was designed from EST sequence (CaF1_JIE_24_F1) corresponding to SAUR. The amplified product was run on 1% agarose gel, purified with gel extraction kit (Qiagen) and the purified product was then cloned in the pGEM-T Easy vector. The clone was named as p3'*CaSAUR1* and was subsequently sequenced using standard procedure of sequencing. For the amplification of full length clone, gene specific primers were designed from the full length nucleotide sequence as obtained from alignment of EST clone and the 3'*CaSAUR1* sequence. The full length cDNA clone was amplified by PCR using cDNA as template and the gene specific primer pair (SAUR-F1 and SAUR-R1). The PCR product was run on 1% agarose gel, purified by gel extraction kit (Qiagen) and subsequently cloned into the pGEM-T Easy vector. The cloned product was named as p*CaSAUR1*. For the amplification of complete ORF of *CaSAUR1*, the primer pair used was SAUR-F2 and SAUR-R2. The sequence of the primers is given in table 2 of appendix II. The PCR was carried out as described in section 6 of appendix III, and the annealing temperature for these reactions was kept at 50°C. For long term storage, the bacterial cultures were grown overnight and 80% sterile glycerol was added so as to obtain a final concentration of 15% and stored in -80°C.

4.2.2 Amplification of *CaSAUR1* from genomic DNA

The genomic clone of *CaSAUR1* was isolated using genomic DNA as template and SAUR-F2 and SAUR-R2 gene specific primer pair (see table 2 of appendix II) so as to amplify genomic ORF. Genomic DNA was isolated as described in section 2 of appendix III. The PCR was carried out as described in section 6 of appendix III. For this primer pair, annealing temperature was kept at 50°C.

4.2.3 Genomic southern

Genomic southern was performed in order to find out the copy number of *CaSAUR1* in chickpea. 10µg of genomic DNA samples were digested with restriction enzymes *Hind*III, *Not*I and *Eco*RV. The digested samples were separated on 0.8% agarose gel, denatured and blotted on to Genescreenplus membrane (Amersham). For hybridization, ORF region of *CaSAUR1* was amplified using SAUR-F2 and SAUR-R2 primer pair and p*SAUR1* as template. The PCR product was run on gel and purified by gel extraction. This purified product was further used for preparing the

^{32}P-dCTP labelled probe by random labelling using random labelling kit (NEB). The detailed description of southern blot preparation and hybridization is given in section 14 of appendix III.

4.2.4 Northern blotting

Northern-blot analysis was performed to determine the expression pattern of *CaSAUR1* in response to *Fusarium* wilt and its tissue specific expression. For this, 20μg of RNA samples were separated on a 1.5% formaldehyde-agarose gel and then blotted onto Genescreenplus membrane (Amersham). The EST clone of *CaSAUR1* was amplified from the plasmid containing this clone using M13 forward and reverse primers. The primer sequence is given in table 2 of appendix II and the PCR detail in section 6 of appendix III. The amplicon was run on 1% gel, purified by gel extraction and the purified product was used for preparing the ^{32}P-dCTP labelled probe by random labelling using random labelling kit (NEB). The detailed description of northern blot preparation is given in section 15 and that of hybridization in section 14 of appendix III.

4.2.5 Quantitative real time PCR

For quantitative RT-PCR, two week old chickpea seedlings were treated with various hormones and the tissue collected at various time points after treatment. Total RNA was extracted using TRIzol reagent and used for cDNA synthesis by High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). The cDNA was diluted 10 times and qRT-PCR was performed in triplicates in ABI 7500 sequence detection system using SYBR Green Master Mix (Applied Biosystems) and gene specific primers SAURRT-F and SAURRT-R (see table 1 of appendix II). The relative quantification method (ΔΔ-CT) was used to evaluate quantitative variation between the replicates examined. The amplification of 18S RNA was used as an endogenous control to normalize all data. The detailed description of real time PCR is given in section 3.2.8 (chapter 3).

4.2.6 Subcellular localization

The subcellular localization of *CaSAUR1* was studied by performing transient expression assay in onion epidermal cells. For this *CaSAUR1* was fused in frame with 5' terminus of GFP reporter gene in pCAMBIA-1302. The fusion construct was

named as pCAMBIA-*CaSAUR1* and was prepared by amplifying *CaSAUR1* from p*CaSAUR1* using pCamSAUR-F and pCamSAUR-R primer pair (see table 2 of appendix II). The amplicon was cloned in *Nco*I and *Spe*I sites of vector as described in section 17 of appendix III. The fusion construct of CaSAUR1-GFP was bombarded on to the onion peels which were then incubated for 24 hours before visualizing in confocal microscope.

4.3 Results and Discussion

4.3.1 Cloning of full length *CaSAUR1* gene

Full length *CaSAUR1* was amplified by 3' RACE using gene specific primer. For this cDNA was first synthesized from total RNA using 3'RACE kit (Invitrogen) following manufacturer's instructions. With this cDNA as template, 3' end of SAUR was amplified using UAP primer of 3'RACE kit and the gene specific primer (SAUR-3'F) designed from the sequence of partial clone (CaF1_JIE_24_E_05). The amplified product of 0.33 kb was run on the gel, eluted and subsequently cloned in pGEM-T vector (Figure 4.1). The clone was sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the sequence of partial clone. The 3' RACE resulted in 624 bp long cDNA clone consisting of 76 bp 5' and 275 bp 3'-untranslated region (UTR) and 273 bp open reading frame (ORF) (Figure 4.2).

4.3.2 *In silico* analysis of CaSAUR1 encoded protein sequence

The predicted protein product of CaSAUR1 comprises 90 amino acid residues (Figure 4.3A, 4.3C) with a calculated molecular mass of about 10.172 kDa and isoelectric point of 7.78. Many other SAURs, for example those from rice were also found to be basic in nature with pI > 7.0 (Jain *et al.*, 2006). *In silico* analysis with ExPASy ProtParam tool revealed the presence of approximately 11% basic, 10 % acidic and 79% neutral amino acids. The homology search against GenBank database showed that *CaSAUR1* is similar to many other *SAUR* genes with maximum homology with that of *Glycine max*. The deduced protein sequence has an auxin inducible domain represented by amino acid 1 to 87 (Figure 4.3B). The sequence of *CaSAUR1* gene was also analyzed for the presence of DST element. DST element represents a conserved sequence present around 30 to 100 bp downstream from the coding region of *SAUR* genes and consists of three highly conserved sequences separated by two variable sequences. DST element was identified in 3' UTR of *CaSAUR1* and is represented as:
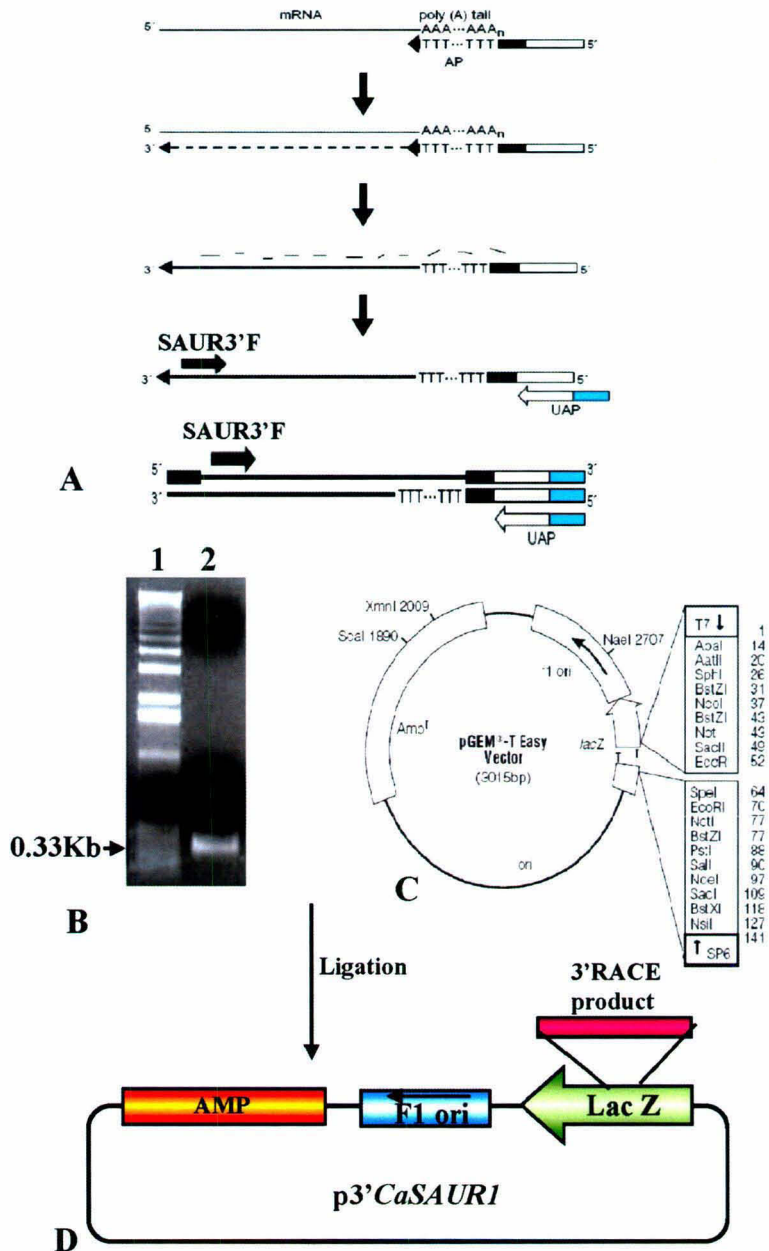
Figure 4.1: 3' RACE of *CaSAUR1*. (A) Flow diagram depicting 3′ RACE procedure. Oligo dT adapter primer (AP) was used to bind the polyA tail of the mRNA. Reverse transcriptase generates single stranded cDNA using mRNA as template. Gene specific primer (SAUR-3′F) and UAP primer specific to 3′ adapter were used to amplify the target 3′cDNA end sequence. (B) 1% agarose/EtBr gel showing 3′ RACE product of *CaSAUR1*. Lane 1 represents 1Kb ladder and lane 2 the PCR product. (C) Map of pGEM-T Easy vector used to clone the amplified product. (D) Diagrammatic representation of 3' RACE product of *CaSAUR1* cloned in pGEM-T Easy vector and the construct was designated as p3'*CaSAUR1*.
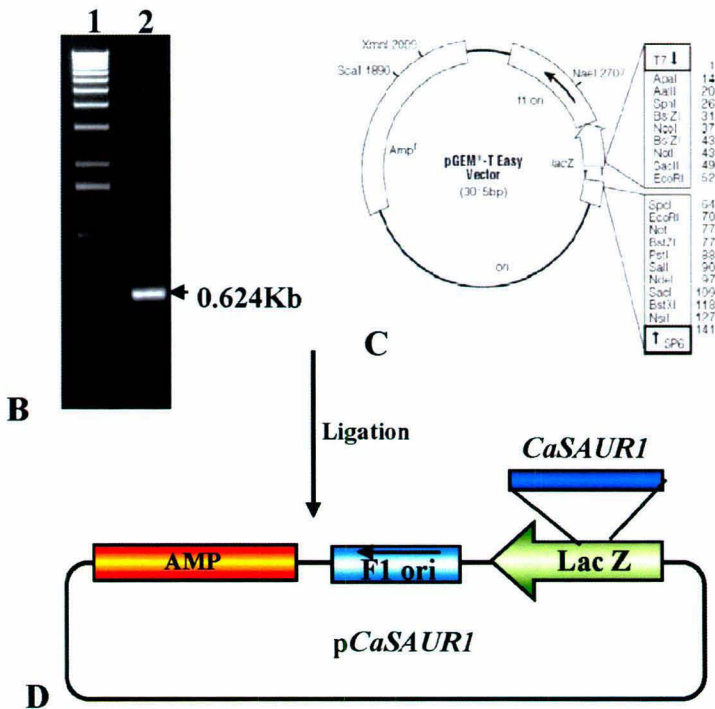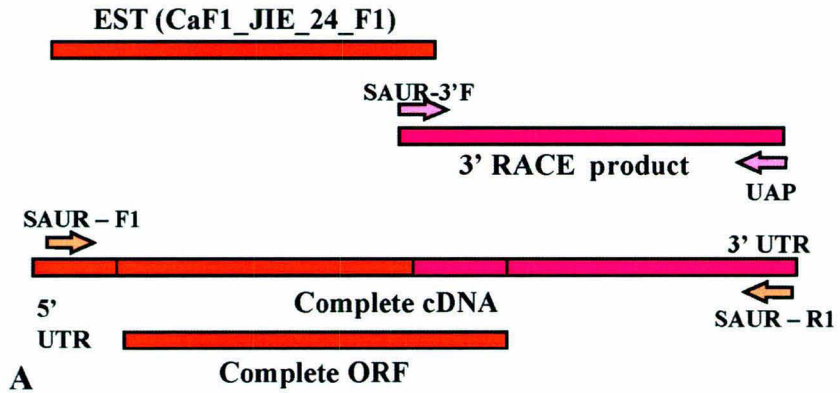
Figure 4.2: Isolation of full length cDNA clone of *CaSAUR1*. (A) Schematic representation of the alignment of sequence of partial clone (EST) and that of 3'RACE product and the primer position for full length cloning. (B) 1% agarose/EtBr gel showing full length cDNA clone. Lane 1 indicates 1Kb ladder and lane 2 PCR product. (C) Map of pGEM-T Easy vector used for cloning the PCR product. (D) Diagramatic representation of full length clone of *CaSAUR1* cloned in pGEMT vector and the construct is designated as p*CaSAUR1*

Figure 4.3: cDNA and deduced amino acid sequence of *CaSAUR1*. (A) Diagrammatic representation of the gene structure. (B) Domain analysis. (C) Nucleotide and the amino acid sequence. Nucleotides in the lower case indicate 5' and 3' UTR and those in the upper case indicate ORF and the deduced amino acid sequence. Region highlighted in blue color indicates the auxin inducible domain. Region in 3'UTR highlighted in pink color indicates sequence of the DST element.

Stop codon – 14 – **GGA** – 6 – **ATAGATT** – 10 – **AAT** – 2 – **GTAT**

The DST element started 14 bp downstream of the stop codon and the sequence of the element resembled that of soyabean *SAUR* (Figure 4.3C).

4.3.3 Multiple sequence alignment and phylogenetic analysis reveal that *SAUR* genes are highly conserved

In order to compare the amino acid sequence of CaSAUR1 with SAUR protein sequence from other plants, we did multiple sequence alignment using MAFFT program LINSI algorithm. The results showed that SAUR proteins are least similar at the extreme N and C termini, however, they share a significant stretch of homology throughout the rest of the length of the protein. Earlier studies have also demonstrated that the core region of *SAUR* genes is highly conserved while the homology at the N and C termini is rather low (Jain *et al.*, 2006). *CaSAUR1* showed maximum homology with SAUR from *Glycine max* followed by that of *Phaseolus* (Figure 4.4).

In order to evaluate the evolutionary relationship of *CaSAUR1*, the deduced amino acid sequence was aligned with twelve orthologs of SAUR from other plants and a phylogram was generated using MAFFT program neighbour joining conserved method. The analysis showed that CaSAUR1 and SAUR from *Glycine max* shared same lineage falling closer to each other in the tree. SAUR from *Phaseolus and Antirrhinum* were also placed closely in the tree showing similarity between them (Figure 4.5). Surprisingly, the *Medicago* ortholog was distantly placed in the tree.

4.3.4 *CaSAUR1* is present as a multiple copy number and an intron less gene

We conducted some preliminary studies to understand the genome organization of *CaSAUR1*. Earlier reports had revealed that *SAUR* is an intron less gene (Gil *et al.*, 1994). In order to detect the presence or absence of introns in *CaSAUR1*, we perfomed PCR using genomic DNA from chickpea as template. Primers were designed from 5' and 3' ends of *CaSAUR1* (appendix II). For the primer combination used, the product obtained with genomic DNA as template was same in size as compared to the corresponding cDNA amplicon (Figure 4.6A). The same size of the genomic and cDNA clones suggests that *CaSAUR1* is an intron-less gene. This was further confirmed by sequencing of the amplified product.

91
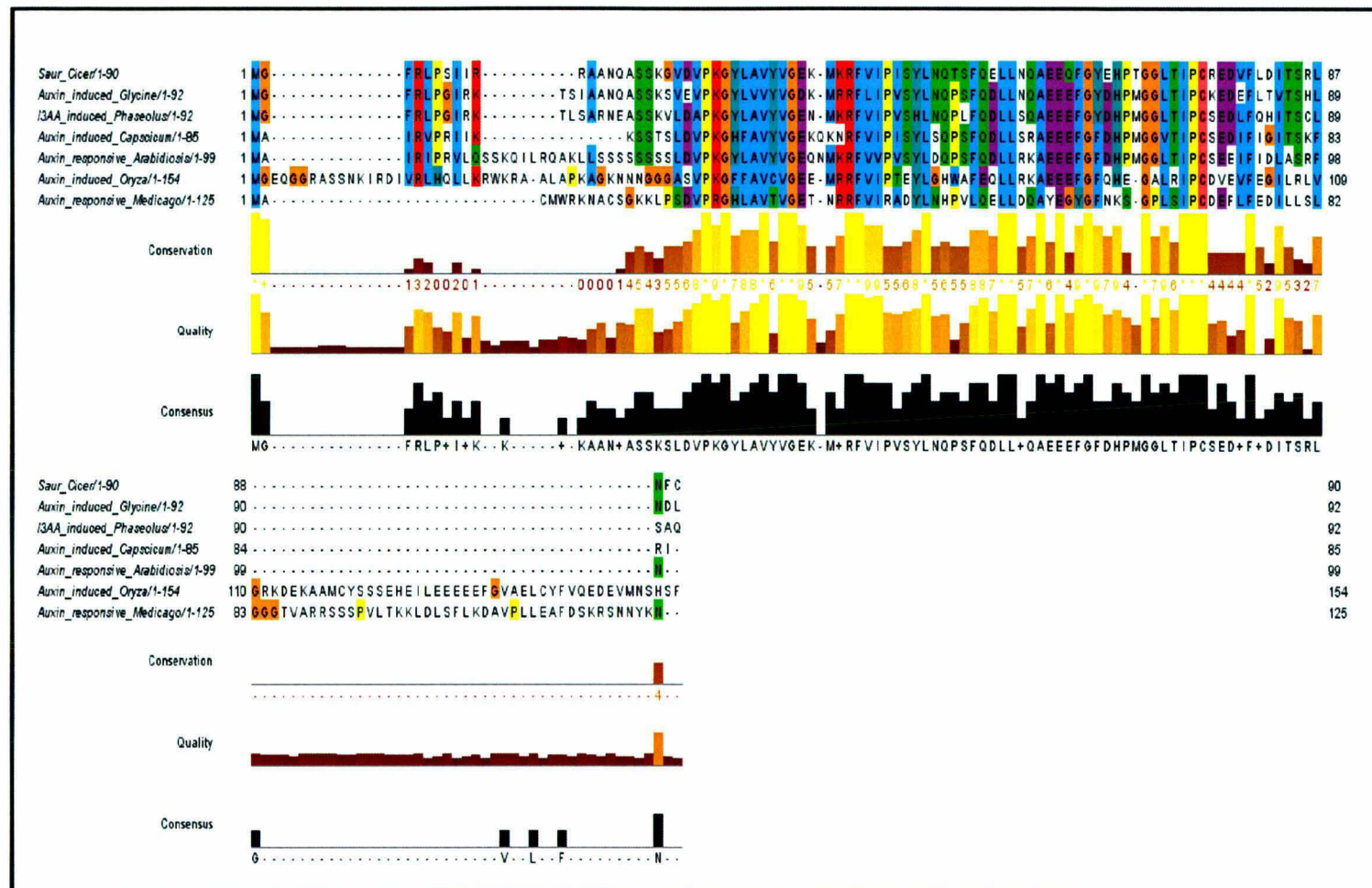
Figure 4.4: Sequence alignment of *CaSAUR1* with the sequences of related proteins from other plant species. The multiple sequence alignment was done using MAFFT program LINSI algorithm. The coloured boxes indicate conserved residues.
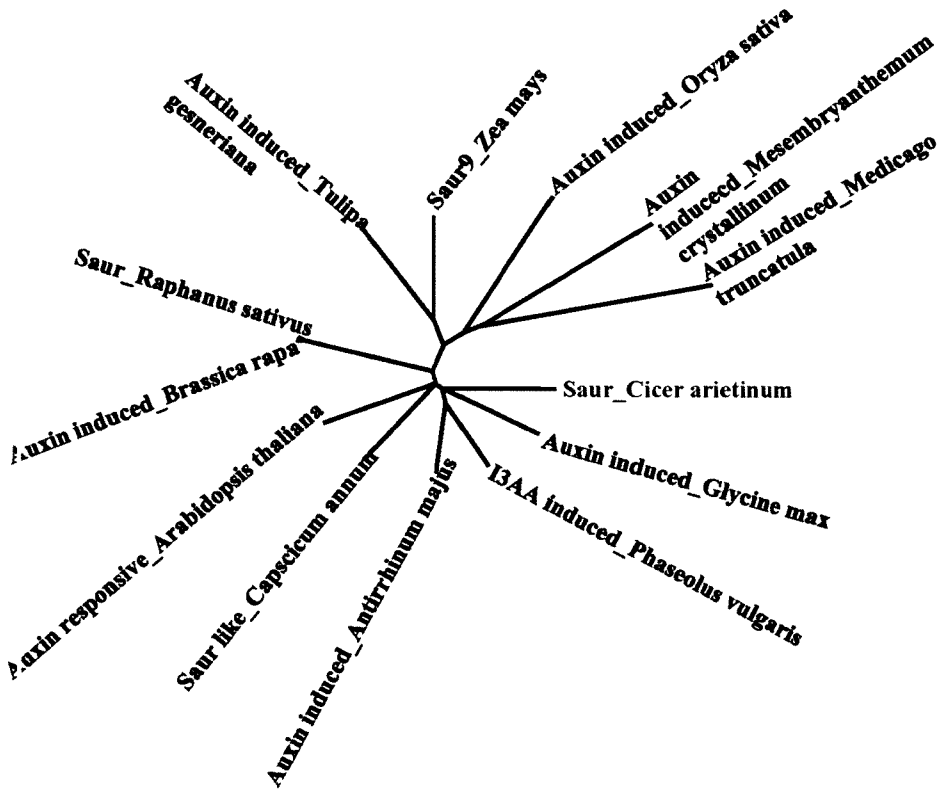
Figure 4.5: Phylogenetic tree showing evolutionary relationship between *CaSAUR1* **and other well-studied auxin inducble family proteins.** The tree was generated using MAFFT program with neighbour joining conserved method .
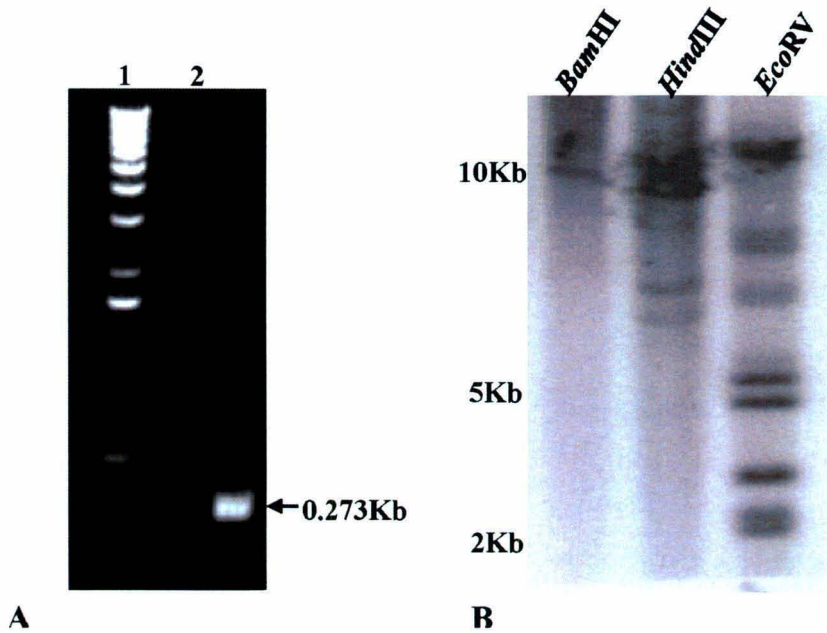
Figure 4.6: Genome organization of *CaSAUR1*. (A) PCR amplified product of *CaSAUR1* using genomic DNA as template and ORF end primers. Lane 1 indicates 1Kb ladder and lane 2 the PCR product. **(B)** Southern blot hybridization analysis for predicting the copy number of *CaSAUR1*. Each lane was loaded with 10μg of chickpea genomic DNA digested with indicated restriction endonucleases and the blot was hybridized with ^{32}P-labeled probe prepared from cDNA clone representing ORF. Size markers in Kb are indicated on left.

To determine the copy number of *CaSAUR1* in chickpea, we performed a genomic DNA southern blot analysis using a cDNA representing ORF of *CaSAUR1* (273 bp) as probe. For this, 10μg of the genomic DNA was digested completely with restriction enzymes: *Hind*III, *Bam*HI and *Eco*RV separately. The digested genomic DNA products were resolved on 0.8% agarose gel and subjected to southern transfer via capillary method. Digestion of genomic DNA with restriction enzyme which does not have any site in the probe should produce single band if single copy of the gene is present while it should give more bands if the gene is present in multiple copies. *Hind*III and *Bam*HI are non-cutters of *CaSAUR1*. Our results showed presence of single band in case of *Bam*HI and four bands in case of *Hind*III digestion. The single site cutter *Eco*RV gave more than one bands after hybridization (Figure 4.6B). Taken together, these results suggest that *CaSAUR1* is present in more than one copy in chickpea. Presence of multiple bands can also be attributed to the fact that SAUR belongs to a multi gene family and the sequence of the family members is highly conserved. It is also possible that in one locus, many SAUR genes are arranged in tandem. Such an arrangement was demonstrated in *SAUR* gene in *Zea mays* (Yang and Poovaiah, 2000).

4.3.5 *CaSAUR1* is a pathogen responsive gene – a novel finding

Earlier observations have shown that SAUR is induced in response to auxin treatment in various plants like *Arabidopsis*, soyabean, rice, maize etc. (Gil *et al.*, 1994; Jain *et al.*, 2006; Knauss *et al.*, 2003). Our gene expression profiling data showed upregulation of *CaSAUR1* in chickpea in response to *Fusarium* infection. In order to further confirm the results, transcript accumulation of the *CaSAUR1* was studied by northern hybridization. For this, total RNA was extracted from root tissue of susceptible and resistant cultivars of chickpea harvested at various time points after *Fusarium* inoculation. The RNA was run on a gel under denaturating conditions, transferred onto the nitrocellulose membrane and the blot was hybridized with a radiolabelled *CaSAUR1* cDNA fragment. The *CaSAUR1* transcript had some basal expression in both the genotypes, however, the expression was induced 12h after *Fusarium* infection and the accumulation of the transcript showed increase in progressive time points (Figure 4.7A and 4.7B). This suggests that although the gene has some basal level expression indicating its role in plant developmental processes, its expression is modulated in response to biotic stress. It was noteworthy to see that

the induction was higher in resistant genotype as compared to the susceptible one indicating the involvement of this gene in plant defense. It would be very interesting to note that earlier also auxin signalling is shown to enhance resistance to necrotrophic pathogens. Thus our results are in confirmation with earlier reports. This gives a new dimension to the involvement of auxin signalling in plant defense responses. In future, it would be of great interest to work out the role of CaSAUR1 in plant immunity.

4.3.6 *CaSAUR1* is preferentially expressed in root tissue

Since *CaSAUR1* gene was found to be differentially regulated, it was of interest to know the tissue specificity of its expression. Total RNA was extracted from root, stem and leaf tissue of 21-day old chickpea seedlings infected with *Fusarium* and the northern blot was prepared and probed with the *CaSAUR1* cDNA fragment. Varying levels of *CaSAUR1* transcripts were detected in stem, leaf and root with expression in root being significantly higher than in other tissues (Figure 4.7C). This may be due to the fact that *Fusarium* infects through roots and the infection is later on spread to other parts of the plant system. Thus root being the first tissue to come in contact with the pathogen leads to increased expression of *CaSAUR1* in this tissue.

4.3.7 Expression of *CaSAUR1* is modulated in response to various hormonal treatments

During pathogen stress plants undergo reprogramming in order to resist the pathogen attack. Defense responses lead to the activation of many hormone signaling pathways which help a plant to prioritize defense over normal cellular functions. Different pathogens may induce different or multiple hormonal pathways to modulate the defense response. In this study we aimed at knowing which hormonal pathway is followed by *CaSAUR1* while mediating defense against *Fusarium* wilt. For this, quantitative real time PCR was performed. Two week old chickpea seedlings were sprayed with various hormones like SA, methyl jasmonate, ACC (precursor of ethylene), 2,4 D, brassinolide, sodium nitrosopruside (SNP) (precursor of NO) and ABA and the tissue was harvested after 1, 3, 6 and 12 hours post treatment. Total RNA was isolated from all these samples and cDNA synthesized which was subsequently used as template for quantitative real time PCR. The results showed that *CaSAUR1* was induced in response to 2,4D as was expected and also in response to
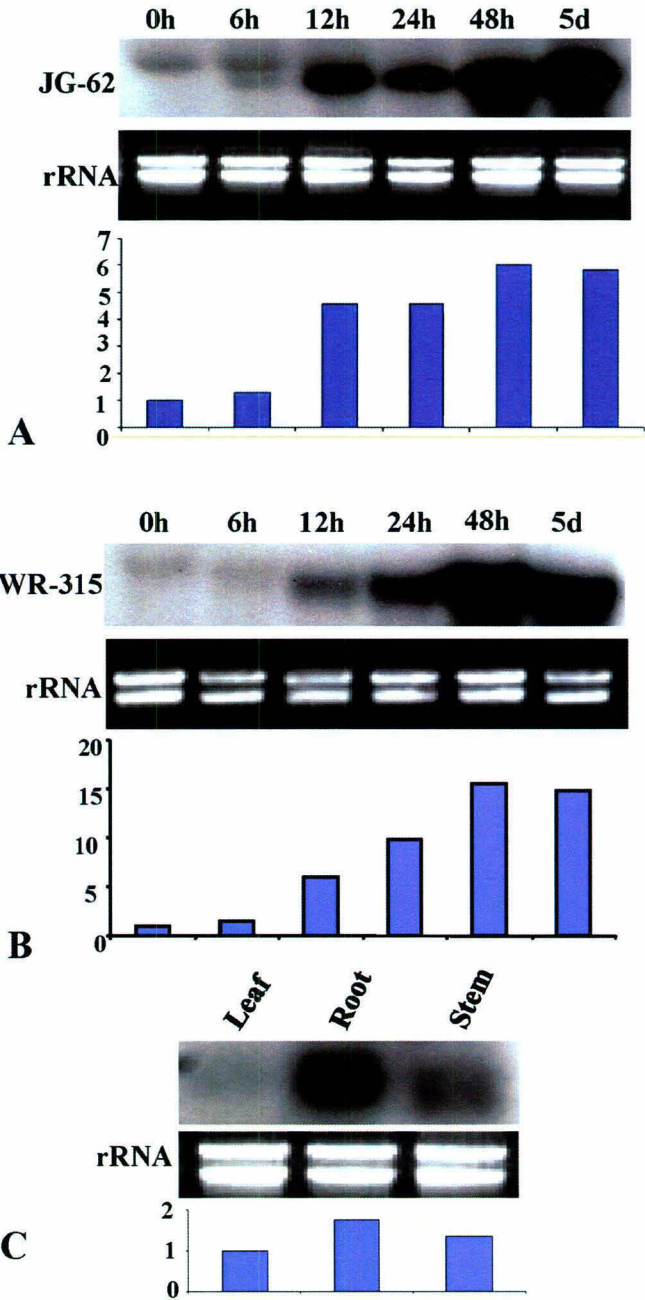
Figure 4.7: RNA blot analysis indicating expression pattern of *CaSAUR1* gene. (A) Expression in response to *Fusarium* infection in susceptible (JG-62) and (B) resistant (WR-315) genotypes of chickpea. 20μg of total root RNA isolated from 25-day old chickpea seedlings harvested at various time points after *Fusarium* infection were separated on 1.5% agarose gel. (C) Expression in different tissues. For tissue specific expression RNA was isolated from different tissues. The northern blot was hybridized with a cDNA fragment of *CaSAUR1* as mentioned in materials and methods. Ribosomal RNAs on lower panels are presented as loading controls. The graphs cited below are indicative of the fold expression in terms of band density.

SA, JA and ET which are the key players involved in defense responsive hormone signalling. Induction was much higher in response to ethylene as compared to SA and JA. Auxin and the gaseous hormone ethylene are known to be intimately linked. In response to brassinosteroid also, *CaSAUR1* was induced in all progressive time points. Auxin response is connected to brassinosteroids (BRs), which act in concert with auxin to promote root gravitropic curvature (Kim *et al.*, 2000). BR and auxin treatments also induce accumulation of many of the same transcripts (Goda *et al.*, 2004; Nemhauser *et al.*, 2004). *CaSAUR1* was also expressed in NO treated plants, however, the level of expression was lesser and in the later stage it showed downregulation. *CaSAUR1* was found to be downregulated in response to ABA. Auxin and ABA are known to have antagonistic effects on plant processes (De Smet *et al.*, 2003) (Figure 4.8).

4.3.8 *CaSAUR1* is localized in nucleus

In order to investigate the subcellular localization of the *CaSAUR1* in the cell, pCAMBIA 35S-Ca*SAUR1*-GFP was constructed (Figure 4.9). In this construct the *CaSAUR1* gene was fused in frame to the 5'terminus of GFP reporter gene. The expression of the fusion gene construct *CaSAUR1*-GFP was driven by the 35S promoter of cauliflower mosaic virus (CaMV-35S). The plasmid construct was named as pCAMBIA-*CaSAUR1*. The plasmids containing vector control DNA and fusion gene were introduced into onion (*Allium cepa*) epidermal cells by particle bombardment. Upon observation under confocal microscope, we found predominant nuclear localization of the fusion protein of CaSAUR1-GFP, whereas the GFP protein alone was distributed in almost all cellular organelles (Figure 4.10). Nuclear localization of SAUR protein was also shown in many earlier reports (Knauss *et al.*, 2003).

4.4 Conclusion

During recent times, emerging evidence have come up which indicate that auxin, an essential plant hormone, involved in regulating growth and development also plays role in plant defense responses. Auxin is known to induce numerous early auxin responsive genes which fall under three categories: *Aux/IAAs*, *SAURs*, and *GH3s*. Though there are reports showing involvement of *Aux/IAAs* and *GH3s* in plant pathogen interaction, no such report has come up as far as *SAURs* are concerned. This
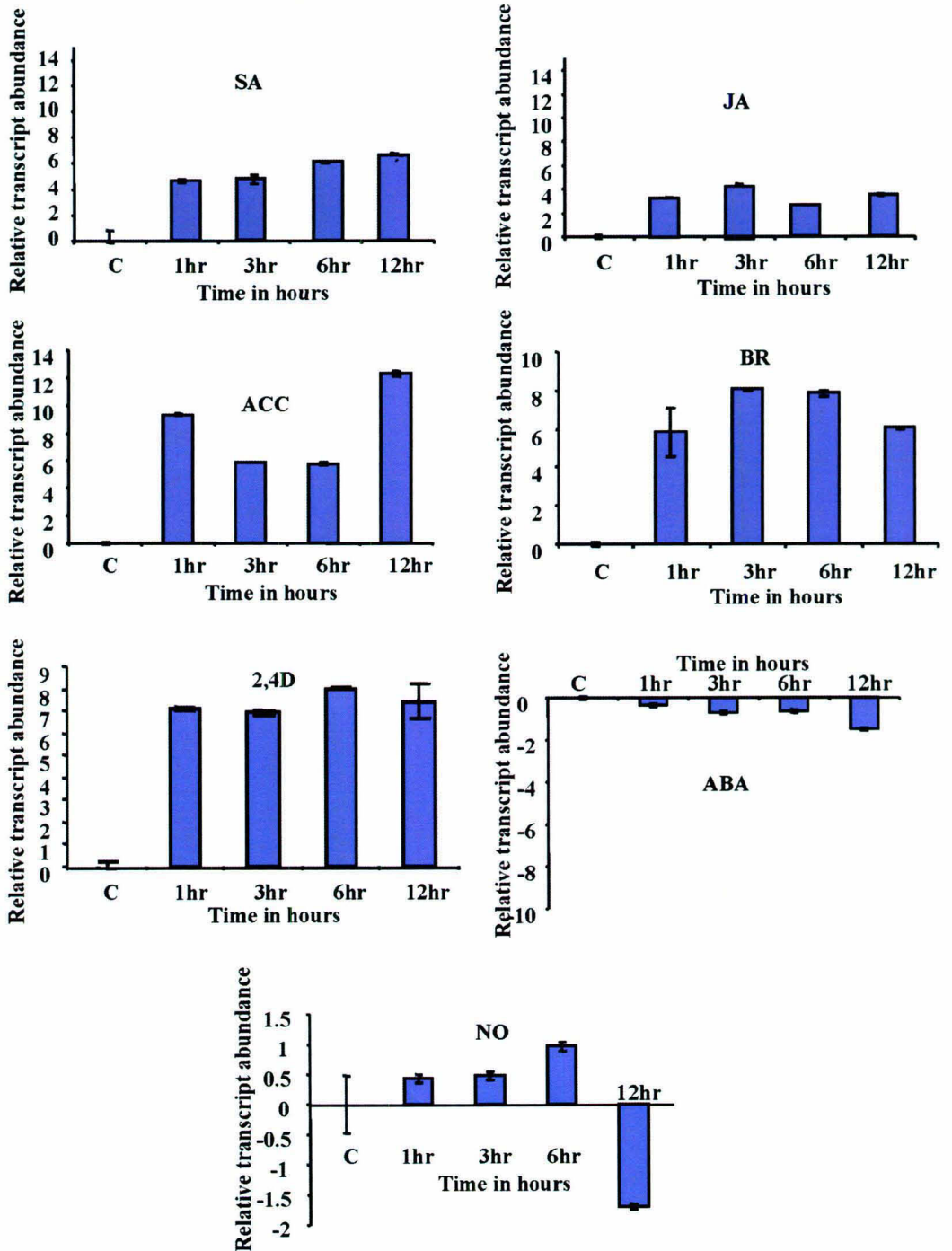
Figure 4.8: Real time PCR showing relative transcript level of *CaSAUR1* in respsonse to various hormones. Transcript levels were normalized by 18S transcript level. Error bars indicate SD of three real time PCR experiments.
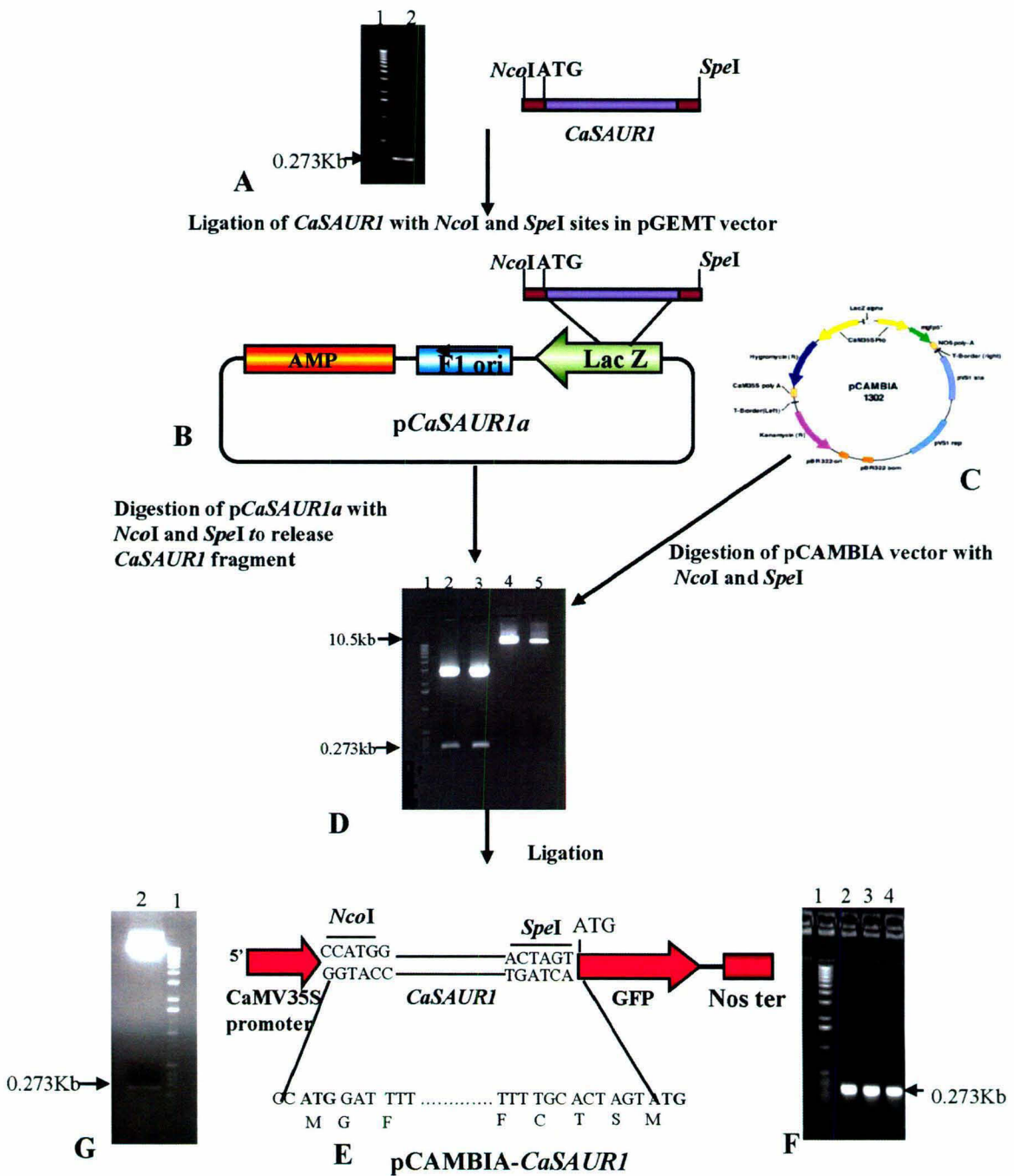
Figure 4.9: Construction of pCAMBIA-*CaSAUR1* construct. (A) The 0.273 kb *CaSAUR1* gene with *NcoI* and *SpeI* sites was PCR amplified. Lane 2 represents the amplified product **(B)** Ligation of *CaSAUR1* into pGEM-T vector and construct named as p*CaSAUR1a*. **(C)** Map of pCAMBIA vector used for cloning. **(D)** Digestion of plasmids p*CaSAUR1a* and pCAMBIA with *SpeI* and *NcoI*. Lanes 2 and 3 represent digested p*CaSAUR1* and lane 4 and 5 represent cut and uncut pCAMBIA vector respectively. **(E)** Ligation of *CaSAUR1* in pCAMBIA. **(F)** Confirmation of positive clones by colony PCR. Lane 2, 3 and 4 represent PCR products obtained from three different colonies. **(G)** Confirmation of positive colonies by digestion. Lane 2 represents the digested products. Lane 1 in all panels represents 1Kb ladder. The construct pCAMBIA-*CaSAUR1* was bombarded in onion cells.
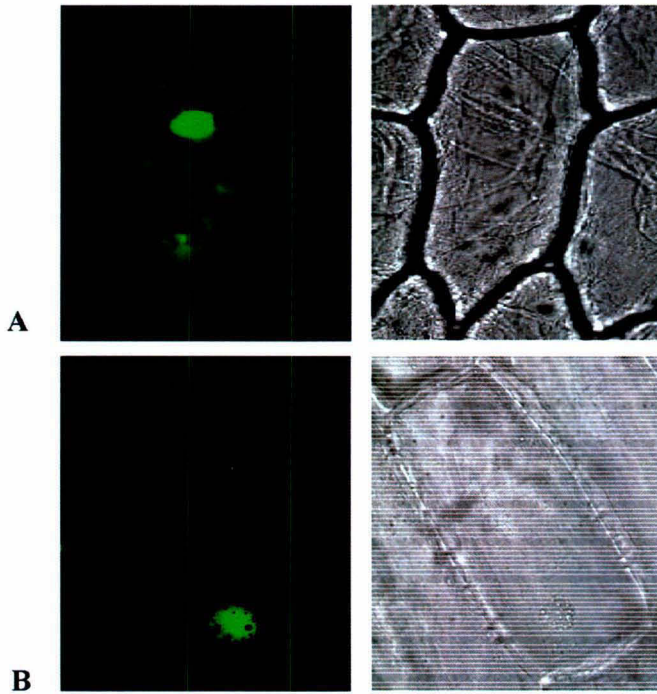
Figure 4.10: Subcellular localization of *CaSAUR1*. Onion epidermal cells bombarded with the **(A)** empty pCAMBIA vector as control and **(B)** pCaMBIA-CaSAUR1-GFP construct. The GFP fluorescence was detected by using Confocal microscope. The right panel shows the corresponding phase contrast image.

study presents first evidence about role of SAURs in plant immune responses. In this study a SAUR encoding gene (*CaSAUR1*) was cloned and its expression pattern in response to wilt pathogen and phytohormones studied. Expression pattern showed that it is expressed to a higher level in resistant genotype, suggesting its role in defense against necrotrophs. Earlier reports have shown that auxin signalling leads to susceptibility to biotrophs and resistance to necrotrophic pathogens. Further, its expression was also modulated in response to auxin and various other hormones. Like most of the *SAUR* genes from other plants, *CaSAUR1* was found to be an intronless gene and localized in nucleus. It is present in multiple copies in chickpea. These results suggest that this gene might have a role to play in transcriptional or posttranscriptional regulation of plant defense.

Cloning and characterization of CabHLH1 transcription factor

5.1 Introduction

Plant defense responses are controlled by dynamic and variable gene expression changes which lead to reprogramming of many cellular functions. The outcome of the induced defense response seems to be finely tuned by cross-talk between various signaling pathways (Koornneef and Pieterse, 2008) and regulatory elements in the promoters of the defense-related genes, resulting in quantitative and/or kinetic effects on the resistance response (Katagiri, 2004). Transcription factors (TFs) are essential components of these signaling pathways and act by regulating the expression of genes encoding PR proteins and enzymes involved in the synthesis of defense related compounds. A number of transcription factors and effectors have been characterized that are critical in the circuitry controlling signal sensitivity and transduction in defense responses, and several may provide a point for crosstalk and signal integration (Chen and Zhu, 2004). These transcription factors belong to different families, for example, Zinc finger, AP2 domain, homeodomain leucine zipper, WRKY, Myb, bHLH etc. and have been shown to play role in plant defense (Eulgem *et al.*, 2005).

Basic helix-loop-helix (bHLH) transcription factors represent a family of proteins that contains a bHLH domain, a motif involved in DNA binding and dimerization (Murre *et al.*, 1989). Members of bHLH protein superfamily have been found to have an ever increasing number of functions in essential physiological and developmental processes in animals as well in plants (Quail, 2000; Ledent and Vervoort, 2001; Toledo-Ortiz *et al.*, 2003; Sonnenfeld *et al.*, 2005). The bHLH domain contains approximately 60 amino acids, with two functionally distinct regions, the basic region and the HLH region. The basic region is located at the N terminus of the bHLH domain and functions as a DNA-binding motif. It consists of approximately 15 amino acids, which typically include six basic residues (Atchley *et al.*, 1999). The HLH region contains two amphipathic α helices with a linking loop of variable lengths. The amphipathic α helices of two bHLH proteins can interact, allowing the formation of homodimers or heterodimers (Nesi *et al.*, 2000). Some bHLH proteins have been shown to bind to sequences containing a consensus core element called the E box (5'-CANNTG-3'), with the G box (5'-CACGTG-3') being the most common form. In addition, the nucleotides flanking the core element may also have a role in binding specificity (Massari and Murre, 2000; Robinson *et al.*, 2000). According to their

phylogenetic relationships, DNA-binding motifs, functional properties and presence or absence of additional domains, known bHLH proteins from eukaryotes have been divided into six main groups named as group A to F (Atchley and Fitch, 1997). It has been suggested that an ancestral HLH sequence most probably came from group B and proteins of this group are the most prevalent type of bHLH proteins in animals. In *Arabidopsis* also, the G-box-binding bHLH proteins (part of group B) are the most abundant group (Toledo-Ortiz *et al.*, 2003). Many additional domains have also been identified in various bHLH proteins most common of which are PAS, orange and leucine zipper domains. These additional domains are mostly present at the carboxy-terminal of the bHLH region. However, the position of the bHLH and additional domains within the sequence of the protein varies widely between different families. Compared to animals, only a small number of plant bHLH proteins have been characterized functionally. In plants, the *R* gene product Lc, which is involved in the control of anthocyanin synthesis in maize, was the first plant protein reported to possess a bHLH motif (Ludwig *et al.*, 1989). In *Arabidopsis* 162 bHLH-encoding genes have been identified from the analysis of genome sequences (Bailey *et al.*, 2003; Heim *et al.*, 2003; Toledo- Ortiz *et al.*, 2003). Also, around 167 bHLH encoding genes were identified in rice genome (Li *et al.*, 2006). Sequence analysis suggests that most of the plant bHLH proteins belonged to Group B which binds to the G-box sequence CACGTG (Atchley and Fitch, 1997).

Family of bHLH proteins is known to perform diverse regulatory functions. The known molecular properties of bHLH proteins suggest a general mechanism by which such regulation may be accomplished. This mechanism involves (1) generation of a high degree of complexity and diversity in transcriptional regulatory activity through variation in the DNA sequence motif recognized by individual bHLH proteins (2) the capacity to combinatorially amplify the spectrum of possible specific protein–DNA interactions through selective heterodimerization between bHLH proteins with different DNA sequence recognition specificity and (3) the capacity to interact with a network of transcriptional coactivators, corepressors, and signaling molecules through selective protein–protein interactions (Toledo- Ortiz *et al.*, 2003).

In animals they act as regulatory factors in different processes such as neurogenesis, cardiogenesis, myogenesis, and hematopoiesis (Jones, 2004). In plants also, bHLH proteins participate in regulating a broad range of growth and developmental

processes at all phases of the life cycle including floral organogenesis, hormone responses, and light signaling. In organogenesis, a bHLH protein was found to be involved in trichome development (Payne *et al.*, 2000) while another gene, *SPATULA*, which encodes a bHLH transcription factor, was found to be involved in carpel and fruit development in *Arabidopsis* (Groszmann *et al.*, 2008). In light signaling, bHLH transcription factors are shown to take a center stage through phytochromes (Duek and Fankhauser, 2005). In addition, bHLH proteins have been shown to positively mediate ABA signaling in *Arabidopsis* (Li *et al.*, 2007). Some bHLH genes were found to be expressed in response to brassinosteroids and were required for normal growth of the plant (Friedrichsen *et al.*, 2002). Beyond their role in normal growth and developmental processes, bHLH proteins are shown to play role in stress responses also. In this direction, a bHLH transcription factor in rice was shown to be involved in signal transduction during cold stress (Wang *et al.*, 2003). Further, a bHLH transcription factor namely *JIN1* was found to mediate jasmonate regulated defense response in *Arabidopsis* (Lorenzo *et al*, 2004).

In the present study, the transcript profiling during *Fusarium* wilt in chickpea led to the identification of a bHLH transcription factor having a unique caspase domain. The gene was found to be differentially expressed and it was of great interest to clone the gene and study its role in defense response. Here we present the expression study of the *CabHLH1* transcription factor in response to *Fusarium* and many plant hormones. We also demonstrate its tissue specific expression, copy number and subcellular localization.

5.2 Material and methods

5.2.1 Cloning of full length *CabHLH1*

Full length cDNA clone of *CabHLH1* was obtained by performing 3' and 5' RACE using gene specific primers and RACE kits (Invitrogen). The gene specific primers were designed from the sequence of EST clone (CaF1_WIE_34_F11). 3'RACE was performed using gene specific primer (bHLH-3'F) and UAP primer provided with the 3'RACE kit as detailed in section 4 of appendix III. 5'RACE was performed by using two gene specific primers (bHLH-5'R1 and bHLH-5'R2) and AP from the kit as detailed in section 5 of appendix III. The amplified products were run on 1% agarose gel, purified with gel extraction kit (Qiagen) and the purified products were then

cloned in the pGEM-T Easy vector. The 3' and 5' RACE products cloned in pGEM-T were named as p3'*CabHLH1* and p5'*CabHLH1* respectively. The clones were sequenced using standard procedure of sequencing. For the amplification of full length clone, gene specific primers were designed from the full length nucleotide sequence as obtained from alignment of the 3'*CabHLH1*, 5'*CabHLH1* and the EST sequences. The full length cDNA clone was amplified by PCR using cDNA as template and the gene specific primer pair (bHLH-F1 and bHLH-R1). The PCR product was run on 1% agarose gel, purified by gel extraction kit (Qiagen) and subsequently cloned into the pGEM-T Easy vector. The cloned product was named as p*CabHLH1*. The sequence of the primers is given in table 2 of appendix II. The PCR was carried out as described in section 6 of appendix III, and the annealing temperature for this primer pair was kept at 60°C. For long term storage, the bacterial cultures were grown overnight and 80% sterile glycerol was added so as to obtain a final concentration of 15% and stored in -80°C.

5.2.2 Amplification of *CabHLH1* from genomic DNA

The genomic clone of *CabHLH1* was isolated using genomic DNA as template and bHLH-F1 and bHLH-R1 gene specific primer pair (see table 2 of appendix II) so as to amplify genomic ORF. For this primer pair, annealing temperature was kept at 60°C. Genomic DNA was isolated as described in section 2 of appendix III. The PCR was carried out as described in section 6 of appendix III.

5.2.3 Genomic southern

Genomic southern was performed in order to find out the copy number of *CabHLH1* in chickpea. 10μg of genomic DNA samples were digested with restriction enzymes *Hind*III, *Not*I, *Nco*I and *Eco*RV. The digested samples were separated on 0.8% agarose gel, denatured and blotted on to Genescreenplus membrane (Amersham). For hybridization, ORF region of *CabHLH1* was amplified using bHLH-F1 and bHLH-R1 primer pair and p*CabHLH1* as template. The amplified product was run on 1% agarose gel and purified by gel extraction and this purified product was further used for preparing the ^{32}P-dCTP labelled probe using random labelling kit (NEB). The detailed description of southern blot preparation and hybridization is given in section 14 of appendix III.

5.2.4 Northern blotting

Northern-blot analysis was performed to determine the expression pattern of *CabHLH1* in response to *Fusarium* wilt and its tissue specific expression. For this, 20μg of RNA samples were separated on a 1.5% formaldehyde-agarose gel and then blotted onto Genescreenplus membrane (Amersham). The EST clone of *CabHLH1* was amplified from the plasmid containing this clone using M13 forward and reverse primers. The primer sequence is given in table 2 of appendix II and the PCR detail in section 6 of appendix III. The amplicon was run on 1% gel, purified by gel extraction and the purified product was used for preparing ^{32}P-dCTP labelled probe by random labelling using random labelling kit (NEB). The detailed description of northern blot preparation is given in section 15 and that of hybridization in section 14 of appendix III.

5.2.5 Quantitative real time PCR

For quantitative RT-PCR, two week old chickpea seedlings were treated with various hormones and the tissue collected at various time points after treatment. Total RNA was extracted using TRIzol reagent and used for cDNA synthesis by High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). The cDNA was diluted 10 times and qRT-PCR was performed in triplicates in ABI 7500 sequence detection system using SYBR Green Master Mix (Applied Biosystems) and gene specific primers bHLHRT-F and bHLHRT-R (see table 1 of appendix II). The relative quantification method (ΔΔ-CT) was used to evaluate quantitative variation between the replicates examined. The amplification of 18S RNA was used as an endogenous control to normalize all data.

5.2.6 Subcellular localization

The subcellular localization of *CabHLH1* was studied by performing transient expression assay in onion epidermal cells. For this *CabHLH1* was fused in frame with 5' terminus of GFP reporter gene in pCAMBIA-1302. The fusion construct was named as pCAMBIA-*CabHLH1* and was prepared by amplifying *CabHLH1* from p*CabHLH1* using pCambHLH-F and pCambHLH-R primer pair (see table 2 of appendix II) and cloning in *Bgl*II and *Spe*I sites of vector as described in section 17 of appendix III. The fusion construct of CabHLH1-GFP was bombarded on to the onion

peels which were then incubated for 24 hours before visualizing in confocal microscope.

5.3 Results and discussion

5.3.1 Cloning of full length *CabHLH1* gene

Full length *CabHLH1* was amplified by 5' and 3' RACE using RACE kits (Invitrogen) and following manufacturer's instructions. For 3' RACE, cDNA was synthesized using oligo dT primer from the kit (AP primer). Using this cDNA as template, 3' end of *CabHLH1* was amplified using UAP primer of the kit and the gene specific primer bHLH-3'F designed from the sequence of the EST (CaF1_WIE_34_F11). The amplified product of approximately 0.65 kb (Figure 5.1) was run on the gel, eluted and subsequently cloned in pGEM-T Easy vector. The clone was sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the sequence of partial clone.

For 5' RACE two gene specific primers were designed namely bHLH-5'R1, and bHLH-5'R2. bHLH-5'R1 was used for making cDNA for the 5'RACE. cDNA was purified with the S.N.A.P. column provided with the kit and the purified cDNA was proceeded for TdT tailing. Tailed cDNA was proceeded for the direct PCR amplification using bHLH-5'R2 primer. The PCR product was diluted (1:100) and 2µl was taken for the secondary PCR using the same primer. The amplified product of ~0.3kb (Figure 5.2) was eluted from gel and cloned in pGEM-T Easy vector. The clone was sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the existing EST sequence. 5' and 3' RACE resulted in a 750 bp long cDNA clone consisting of an open reading frame (ORF) of 687 bp (Figure 5.3).

5.3.2 *In silico* analysis of *CabHLH1* encoded protein sequence

The predicted protein product of *CabHLH1* comprises 228 amino acid residues with a calculated molecular mass of about 25.193 kDa and isoelectric point of 6.63. The schematic representation of the gene structure is given is figure 5.4A. *In silico* analysis with ExPASy ProtParam tool revealed the presence of approximately 14% basic, 14 % acidic and 76 % neutral amino acids. The homology search against GenBank database showed that *CabHLH1* had maximum homology with an unnamed protein of *Vitis vinifera* and bHLH transcription factor of *Nicotiana*. It also showed
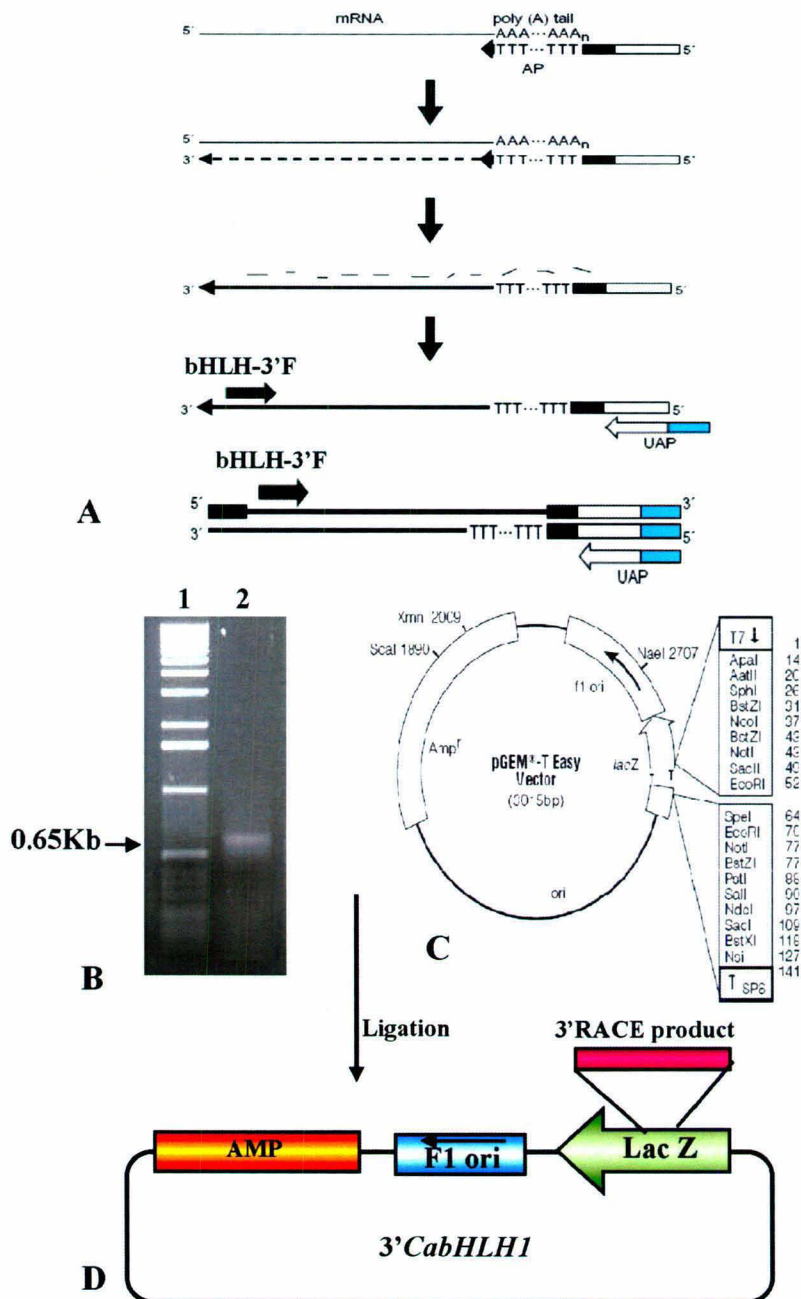
Figure 5.1: 3' RACE of *CabHLH1*. (A) Flow diagram depicting 3' RACE procedure. Oligo dT adapter primer (AP) was used to bind the polyA tail of the mRNA. Reverse transcriptase generates single stranded cDNA using mRNA as template. Gene specific primer (bHLH-3'F) and UAP primer specific to 3' adapter were used to amplify the target 3'cDNA end sequence. **(B)** 1% agarose/EtBr gel showing 3' RACE product of *CabHLH1*. Lane 1 represents 1Kb ladder and lane 2 the PCR product. **(C)** Map of pGEM-T Easy vector used to clone the amplified product. **(D)** Diagrammatic representation of 3' RACE product of *CabHLH1* cloned in pGEM-T easy vector and the construct named as p3'Ca*bHLH1*.
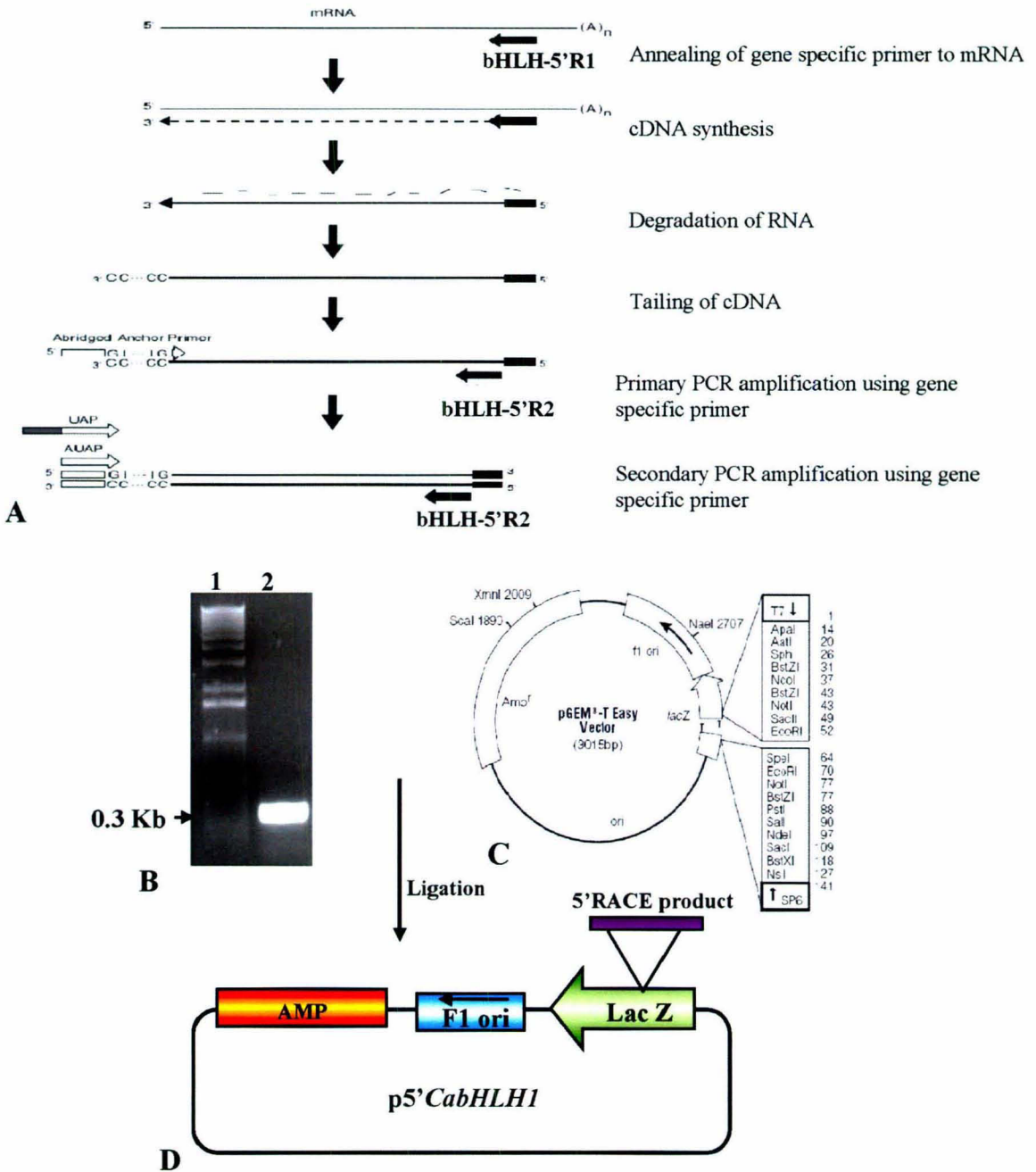
Figure 5.2: 5' RACE of *CabHLH1*. (A) Flow diagram depicting 5' RACE procedure. (B) 1% agarose/EtBr gel showing 5' RACE product of *CabHLH1*. Lane 1 represents 1Kb ladder and lane 2 the PCR product. (C) Map of pGEM-T Easy vector used to clone the amplified product. (D) Diagrammatic representation of 5' RACE product of *CabHLH1* cloned in pGEM-T vector and the construct named as p5'*CabHLH1*.
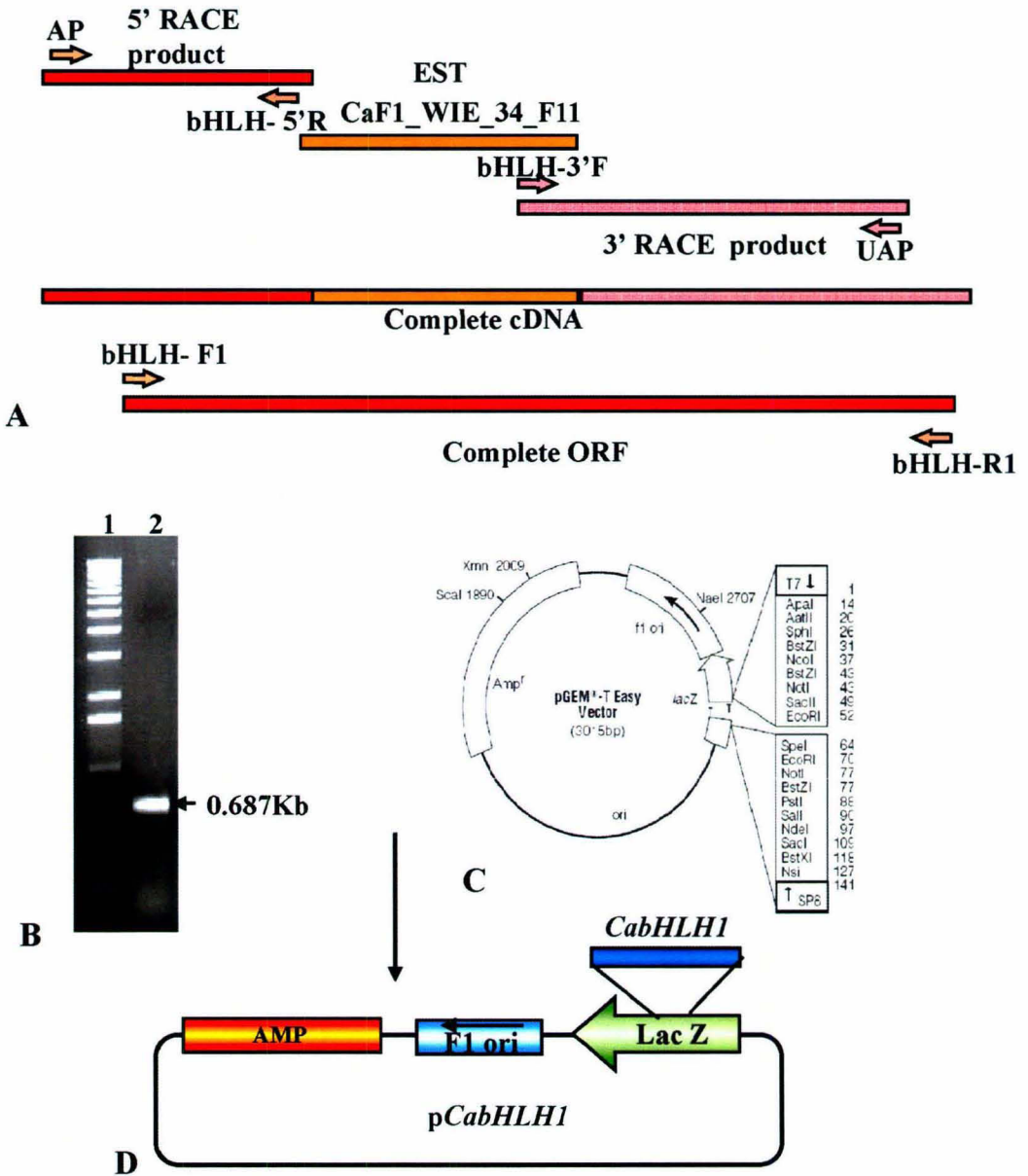
Figure 5.3: Isolation of full length cDNA clone of *CabHLH1*. (A) Schematic representation of the alignment of sequence of partial clone (EST) and that of 3'and 5'RACE products and the primer position for full length cloning. **(B)** 1% agarose/EtBr gel showing full length cDNA clone. Lane 1 indicates 1Kb ladder and lane 2 PCR product. **(C)** Map of pGEM-T Easy vector used for cloning the PCR product. **(D)** Diagrammatic representation of full length clone of *CabHLH1* cloned in pGEM-T Easy vector and the construct is named as p*CabHLH1*.

homology to an unknown protein of *Medicago* and IAA-Leu resistant DNA binding protein of *Arabidopsis*. The deduced protein sequence encodes a transcription factor with bHLH domain represented by amino acid 70 to 119 and a caspase domain represented by amino acids 94 to 169 (Figure 5.4B, 5.4C). The presence of caspase domain in *CabHLH1* raises intriguing possibilities of mechanism and additional aspects of function of this gene. Caspases represent a family of cysteine proteases that have key roles in apoptosis. Further, caspases are known to form dimers and hence it would be interesting to speculate that *CabHLH1* might form homo or hetero dimers through the caspase domain and perform downstream function. Recent progress in caspase research indicated that these proteins are important not only in apoptotic but also in non apoptotic processes. It was shown that caspases independent of their enzymatic action also signal through protein-protein interaction and perform role in innate immunity in animals (Kuranaga and Miura, 2007). Based on these evidences, it could be speculated that *CabHLH1* might play a regulatory role in plant immunity by executing apoptotic or non apoptotic function.

5.3.3 Multiple sequence alignment and phylogenetic analysis of *CabHLH1*

In order to compare the amino acid sequence of CabHLH1 with related protein sequence from other plants, we did multiple sequence alignment using MAFFT program LINSI algorithm. The results showed that there was considerable conservation along the amino acid sequence of orthologs from various organisms particularly in the stretch of amino acids which represented the bHLH domain (Figure 5.5). Earlier studies in *Arabidopsis* and rice have also revealed that the signature bHLH region across the members of bHLH family is highly conserved (Toledo-Ortiz *et al.*, 2003; Li *et al.*, 2006).

In order to evaluate the evolutionary relationship of CabHLH1, its amino acid sequence was aligned with sixteen orthologs of bHLH proteins from other organisms and a phylogram was generated using MAFFT program neighbour joining conserved method. The analysis showed that CabHLH1 shared same lineage as that of unnamed protein of *Vitis vinifera*, bHLH protein from *Nicotiana,* unknown protein from *Medicago* and IAA-Leu resistant DNA binding protein of *Arabidopsis* (Figure 5.6). All these proteins were found to be falling closer to each other in the tree. CabHLH1 was also related, although, distantly to bHLH transcription factors from various other

A

5'
24bp
ATG
ORF
687bp
TAA
3' UTR
39bp

234bp   306   381bp   531

1   25

bHLH domain   Caspase domain

711 750

Full length cDNA of *CabHLH1*

B

Query seq.
Superfamilies

1   25   51   75   100   125   151   175   200   229

HLH superfamily

CASP_C superfamily

C

```
tttccgggaaggttccacccaactATGGTTTCCCCGGAAAACACCAATTGGCTTTTTGAT
                        M  V  S  P  E  N  T  N  W  L  F  D
TACGCTTTGATGGACGATATTCCCGTCCCTGATGGCTCCTTCAACTTCACGTGGCCCCCA
Y  A  L  M  D  D  I  P  V  P  D  G  S  F  N  F  T  W  P  P
CCGGACCCTTCCAATGTTGGTGTTGAAATTGATGGCTCTCTGGGGGATTCTGATGGTATC
P  D  P  S  N  V  G  V  E  I  D  G  S  L  G  D  S  D  G  I
AAAGATCCTGGTTCGAAGAAGAGGGGTAGATCGGATTCATGTGCTCCTTCCAGCTCTAAG
K  D  P  G  S  K  K  R  G  R  S  D  S  C  A  P  S  S  S  K
GCGTGTCGAGAGAAGTTGCGGAGGGATAGGCTGAATGACAAGTTTGTTGAATTAGGCTCA
A  C  R  E  K  L  R  R  D  R  L  N  D  K  F  V  E  L  G  S
ATCTTGGAGCCTGGAAGGCCTCCCAAAACAGACAAAGCTGCAATTCTGATTGACGCTGTC
I  L  E  P  G  R  P  P  K  T  D  K  A  A  I  L  I  D  A  V
CGAATGGTGACACAATTGCGAGGTGAAGCTCAGAAGTTGAAAGCCTCAAATTCAAGTCTT
R  M  V  T  Q  L  R  G  E  A  Q  K  L  K  A  S  N  S  S  L
CAAGAGAAGATTAAAGAGTTAAAGGTCGAGAAGAATGAACTCCGTGATGAGAAACAGAGG
Q  E  K  I  K  E  L  K  V  E  K  N  E  L  R  D  E  K  Q  R
CTTAAGGCCGATAAAGAGAAGTTGGAGCAGCAGGTGAAATCAATGAATACCCAACCTGGT
L  K  A  D  K  E  K  L  E  Q  Q  V  K  S  M  N  T  Q  P  G
TTCTTGACACCCCCTCCTGCAATCCCTGCTGCATTTGCTCCACAAGGCCAAGCCCCTAGC
F  L  T  P  P  P  A  I  P  A  A  F  A  P  Q  G  Q  A  P  S
AACAAGTTAATGCCTTTCATTAGTTACCCTGGAGTTGCAATGTGGCAATTTGTGCCACCG
N  K  L  M  P  F  I  S  Y  P  G  V  A  M  W  Q  F  V  P  P
GCAACCATGGATATCTCACAGGATCATGAACTCCGCCCACCAGTTGCCTAAgttggcact
A  T  M  D  I  S  Q  D  H  E  L  R  P  P  V  A  -
gactgtacctcggccgcgaccacgccacgc
```

Figure 5.4: cDNA and deduced amino acid sequence of *CabHLH1*. (A)
Diagrammatic representation of gene structure of *CabHLH1*. **(B)** Domain analysis.
(C) Nucleotide and the amino acid sequence. Nucleotides in lower case indicate the
5'UTR and 3'UTR. Those in upper case indicate complete ORF with the start and
stop codons and the deduced amino acid sequence. Nucleotides highlighted in sea
green colour indicate the helix loop helix domain and those with blue color indicate
caspase domain. The green region signifies the region shared by the two domains.
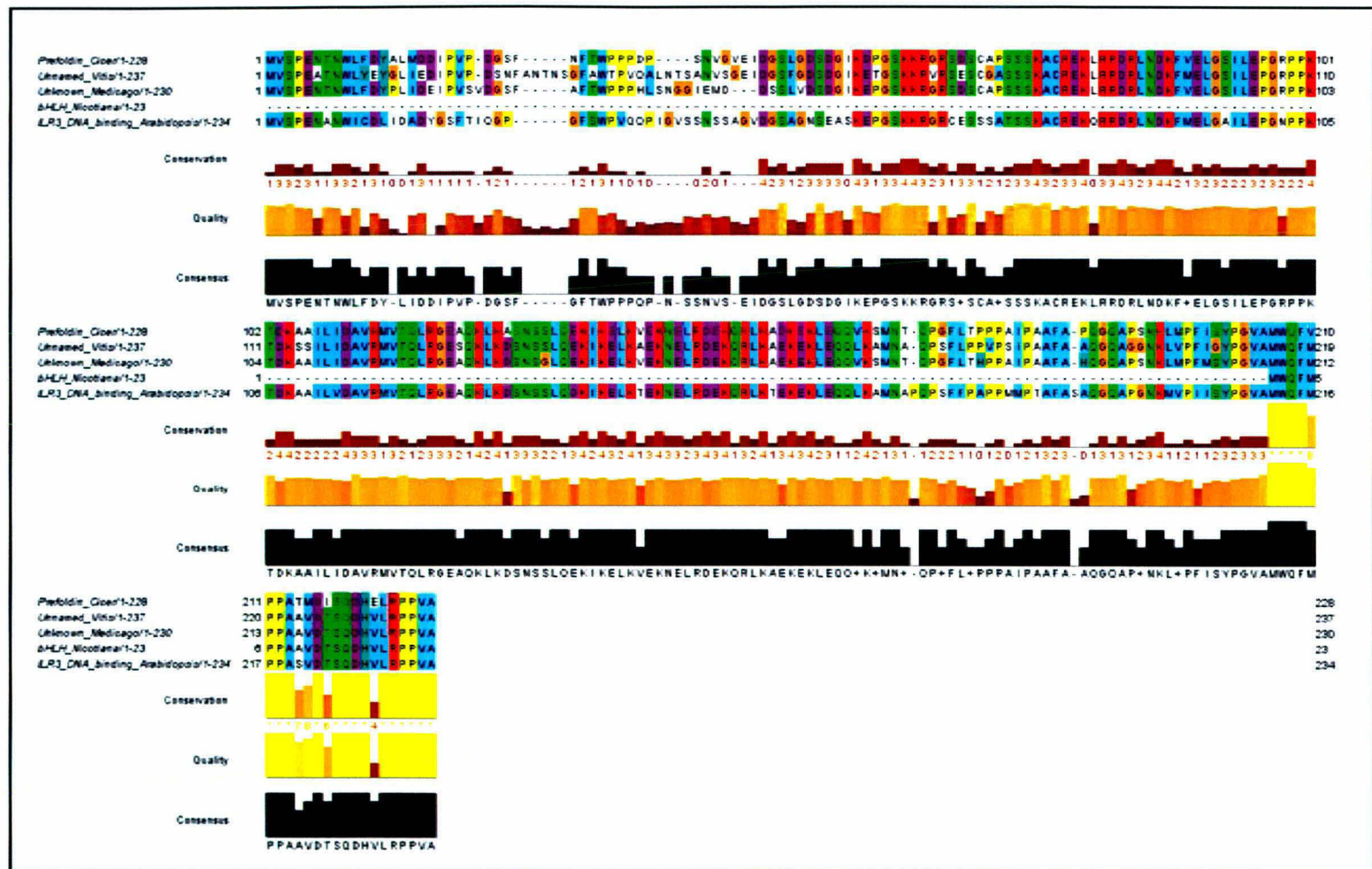
Figure 5.5: Sequence alignment of *CabHLH1* with the sequences of orthologs from other plants. The multiple sequence alignment was done using MAFFT program LINSI algorithm.
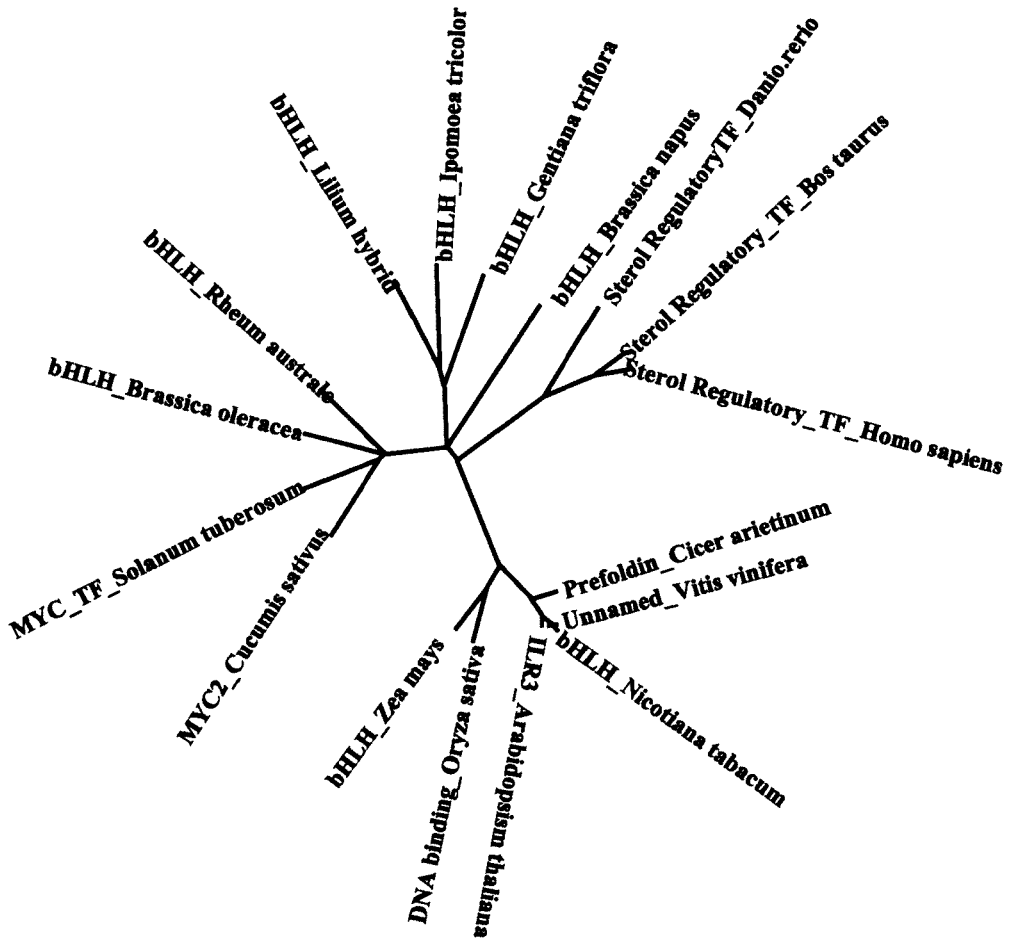
Figure 5.6: Phylogenetic tree showing evolutionary relationship between *CabHLH1* **and other similar proteins.** The tree was generated using MAFFT program neighbour joining conserved method .

plants like oryza, maize, Brassica etc. Further, CabHLH1 was placed quite far away from its human ortholog. Interestingly, CabHLH1 was more closely related to its ortholog from *Vitis* and *Nicotiana* rather than its legume counterpart.

5.3.4 Genome organization of *CabHLH1*

We aimed at obtaining initial information about the genome organization (copy number and intron number and size) of *CabHLH1* gene. In this direction, we performed PCR using genomic DNA from chickpea as template. Primers were designed from 5' and 3' ends of *CabHLH1* cDNA in order to amplify the ORF. The PCR gave a band of same size as that of the cDNA clone suggesting that *CabHLH1* is an intronless gene (Figure 5.7A). Earlier reports on presence and distribution of introns in bHLH family proteins has shown that 80% of the members of this family in *Arabidopsis* and rice contain introns and the intron position is conserved even though the number varies (Toledo-Ortiz *et al.*, 2003; Li *et al.*, 2006). However, in this study the absence of intron suggests that this gene family might differ in genome organization in legumes. Interestingly, a *myc* gene in mouse was found to be intronless and acted as a strong inducer of apoptosis (Sugiyama *et al.*, 1998).

Southern transfer of chickpea genomic DNA digested separately with various restriction enzymes was also performed and hybridized with ORF region of *CabHLH1* cDNA as probe under high stringent conditions. *Hind*III, and *Not*I are non-cutters of *CabHLH1* and gave double and single bands respectively after hybridization. The single site cutters *Eco*RV and *Nco*I gave three bands each (Figure 5.7B). Taken together, these results suggest that *CabHLH1* is present in more than one copies in chickpea.

5.3.5 Expression of *CabHLH1* in response to *Fusarium* wilt

Expression pattern of *CabHLH1* in response to *Fusarium* wilt in chickpea was studied by RNA gel blot. The experiment was performed with susceptible (JG-62) and resistant (WR-315) genotypes of chickpea. The results showed that *CabHLH1* gene had some basal expression in both the genotypes indicating its requirement in normal developmental processes. However, the transcript level increased in response to *Fusarium* infection suggesting that the gene is regulated by biotic stress (Figure 5.8A and 5.8B). The expression reached maximum level within 6 hours of *Fusarium* infection indicating its role in early phase of defence. However, the expression did not
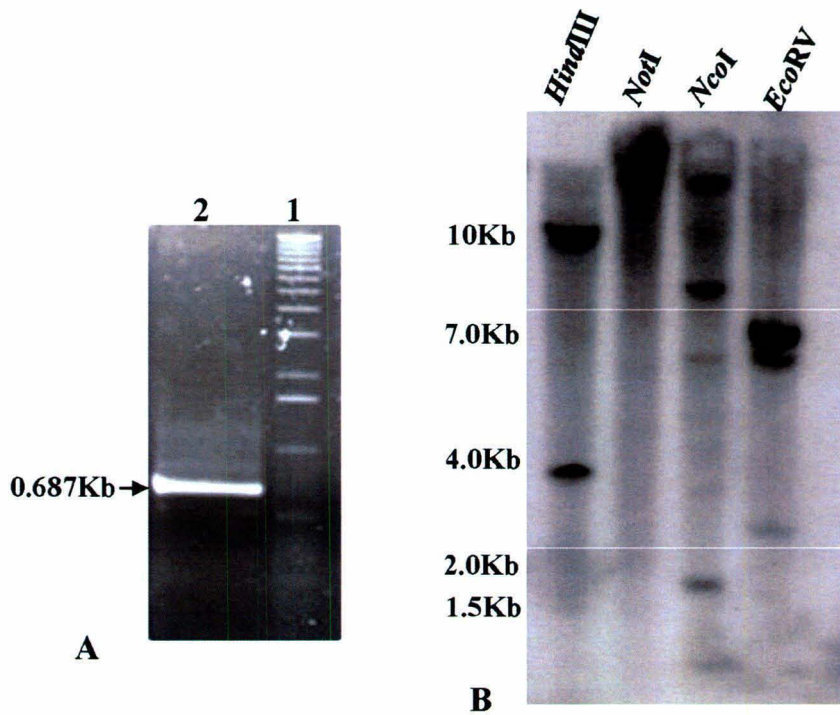
Figure 5.7: Genome organization of *CabHLH1*. (A) PCR amplified product of *CabHLH1* using genomic DNA as template and ORF end primers. Lane 1 and 2 indicate 1Kb ladder and PCR product respectively **(B)** Southern blot hybridization analysis for predicting the copy number. Each lane was loaded with 10μg of chickpea genomic DNA digested with indicated restriction endonucleases and the blot was hybridized with ^{32}P-labeled probe prepared from full length cDNA clone of *CabHLH1*. Size markers in Kb are indicated on left.

show any major difference in susceptible and resistant genotypes. This suggested that although this gene might be involved in defense response against *Fusarium* wilt, its expression kinetics does not bear any direct effect on susceptibility or resistance of a plant. Moreover, as suggested by the expression pattern, *CabHLH1* might be involved in very early events and might be followed by a long series of other important events that would in turn have a direct effect on making a plant susceptible or resistant to pathogen attack. Also this gene might be involved in basal defense.

5.3.6 Analysis of tissue-specific expression of *CabHLH1*

In order to know about the tissue specific expression of *CabHLH1*, RNA gel blot was performed with RNA isolated from root, stem and leaf tissue of *Fusarium* infected chickpea plant. The blot was hybridized with probe prepared from cDNA clone of *CabHLH1*. The results showed that the gene was expressed in all the three tissue types, however, the expression was more in root followed by stem and leaf (Figure 5.8C). This suggested that the gene plays regulatory role against wilt in multiple organs, however, root being the first organ to come in contact with the fungus expresses this gene to a higher level.

5.3.7 Expression of *CabHLH1* in response to various hormonal treatments

We investigated the expression of *CabHLH1* gene in response to various hormonal treatments by quantitative real time PCR. We observed increased *CabHLH1* transcript accumulation in response to SA, transcript level reaching its maximum at 6 hours after treatment. The expression also increased during progressive time points after JA and ACC treatment with maximum level at 12 hours after treatment. Brassinosteroid and NO also induced its expression, however, in case of NO, there was an initial downregulation of the *CabHLH1* followed by upregulation at 3 and 6 hours after the treatment. The expression again showed downregulation at 12 hours post treatment. Further, the gene was found to be downregulated in response to ABA except a slight increase at later time points (Figure 5.9). These results suggest that *CabHLH1* is induced in response to all the hormones which are known to play role in plant defense pathways, thereby further confirming its role in plant immunity. Earlier studies have also shown the modulation of expression of bHLH genes in response to various hormones. For example, Myc transcription factors belonging to bHLH class are known to play key role in jasmonate signaling (Boter *et al.*, 2004). bHLH genes are
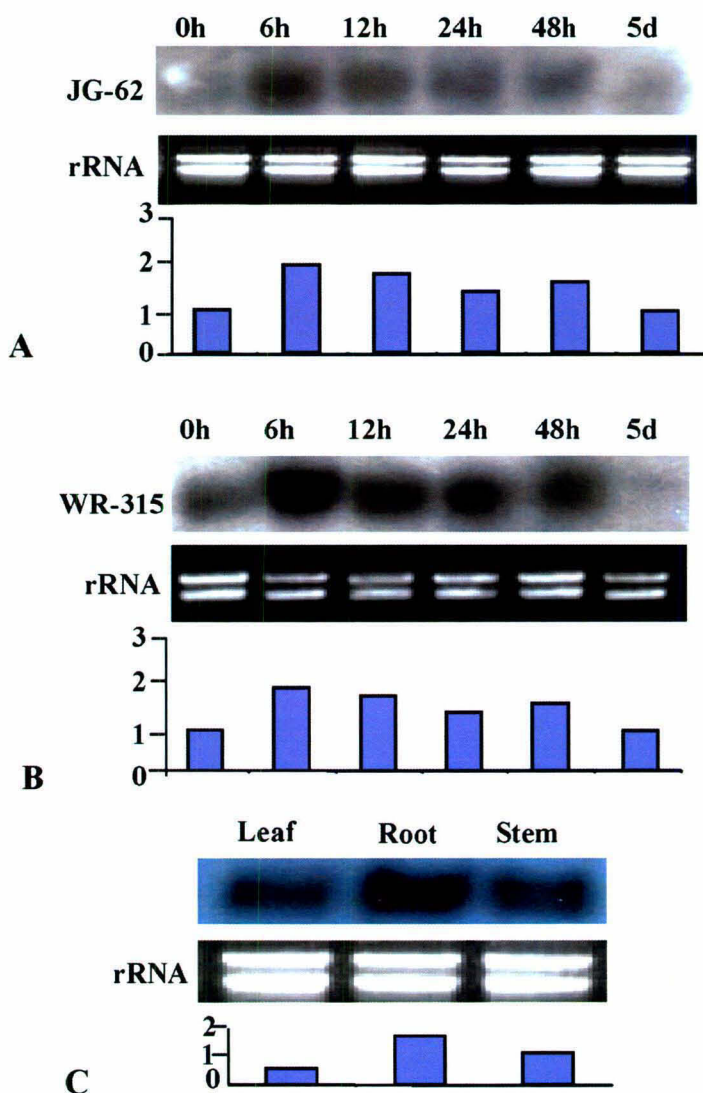
Figure 5.8: RNA blot analysis indicating expression pattern of *CabHLH1* gene. (A) expression in response to *Fusarium* infection in susceptible (JG-62) and (B) resistant (WR-315) genotypes of chickpea. 20µg of total root RNA isolated from 25-day old chickpea seedlings harvested at various time points after *Fusarium* infection were separated by 1.5% agarose gel. (C) expression in different tissues. For tissue specific expression RNA was isolated from different tissues. The northern blot was hybridized with a cDNA fragment of *CabHLH1* as mentioned in materials and methods. Ribosomal RNAs at lower panels represent loading controls. The graphs indicate fold expression in terms of band intensity.
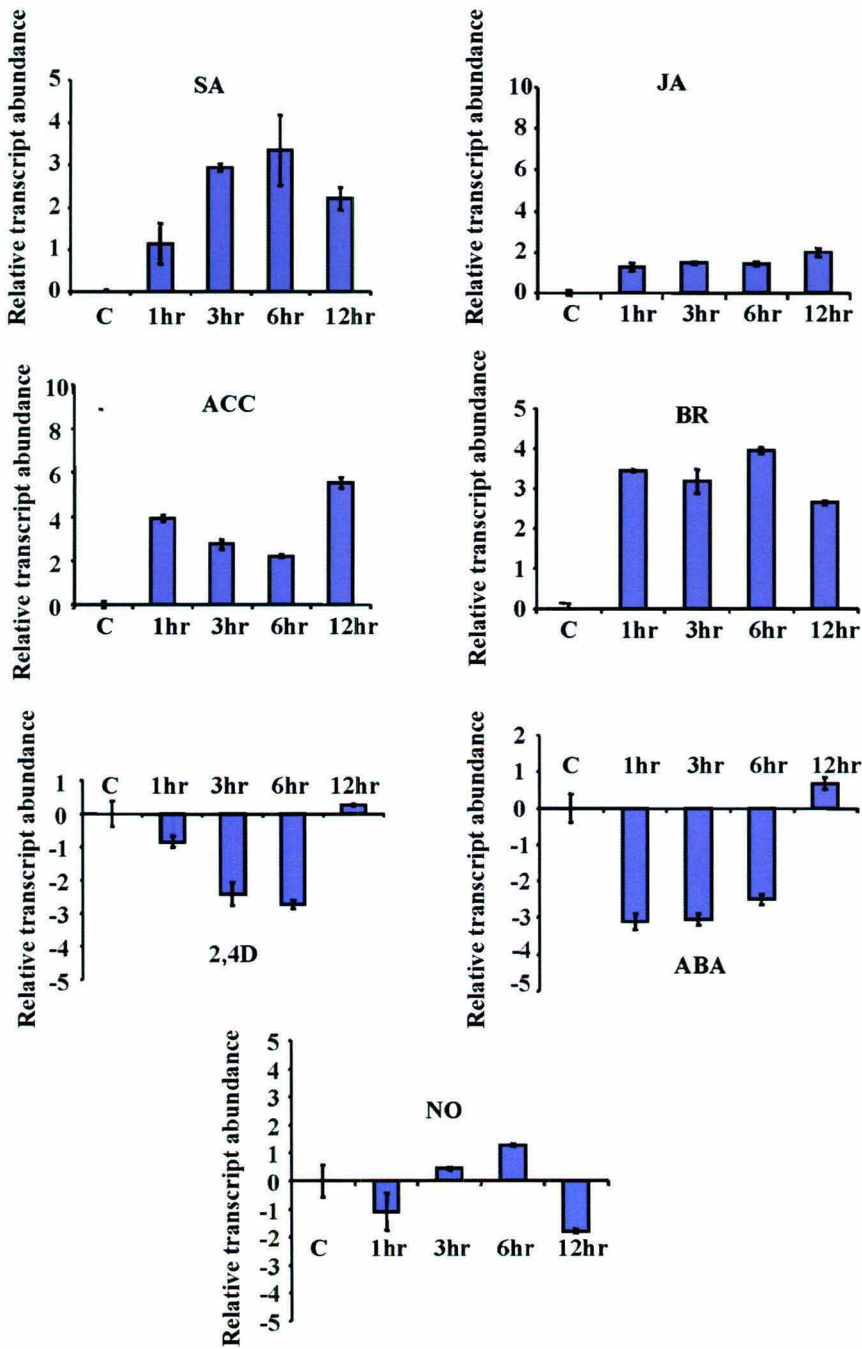
Figure 5.9: Real time PCR showing relative transcript level of *CabHLH1* in response to various hormones. Transcript levels were normalized by 18S transcript level. Error bars indicate SD of three real time PCR experiments

also shown to be involved in brassinosteroid mediated responses (Friedrichsen *et al.*, 2002).

5.3.8 Subcellular localization of *CabHLH1*

In order to investigate the subcellular localization of the *CabHLH1* in the cell, pCAMBIA 35S-*CabHLH1*-GFP was constructed (Figure 5.10). The expression of the fusion gene *CabHLH1*-GFP was driven by the 35S promoter of cauliflower mosaic virus (CaMV-35S). The plasmid construct was named as pCAMBIA-*CabHLH1*. The plasmids containing vector control DNA and fusion gene were introduced into onion (*Allium cepa*) epidermal cells by particle bombardment. Upon observation under confocal microscope, we found predominant nuclear localization of the fusion protein of *CabHLH1*-GFP, whereas the GFP protein alone from the vector control was distributed in almost all cellular organelles (Figure 5.11). The nuclear localization of *CabHLH1* was in confirmation with earlier results which showed that the group of bHLH proteins are present in nucleus where they bind to DNA and perform their regulatory functions (Massari and Murre, 2000).

5.4 Conclusion

Transcription control of plant defense responses is very well documented and members of different families of transcription factors are known to play role in such regulation. Several members of bHLH family of TFs are also implicated in defense. Here we cloned a member of bHLH family (*CabHLH1*). The gene harboured a bHLH domain like other members of this family, however, it also had a caspase domain which made it different from rest of its counterparts. It was also quite surprising to find the gene intronless. Preliminary analysis regarding the genome organization of the gene as conducted by southern hybridization showed that the gene is present in more than one copies in chickpea. The gene was found to be induced in response to *Fusarium* wilt in chickpea and its expression was also altered in response to hormones which are usually known to play role in defense pathways. The gene was localized in nucleus. Taken together, these results suggest a possible role of this gene in plant defense against vascular wilt.
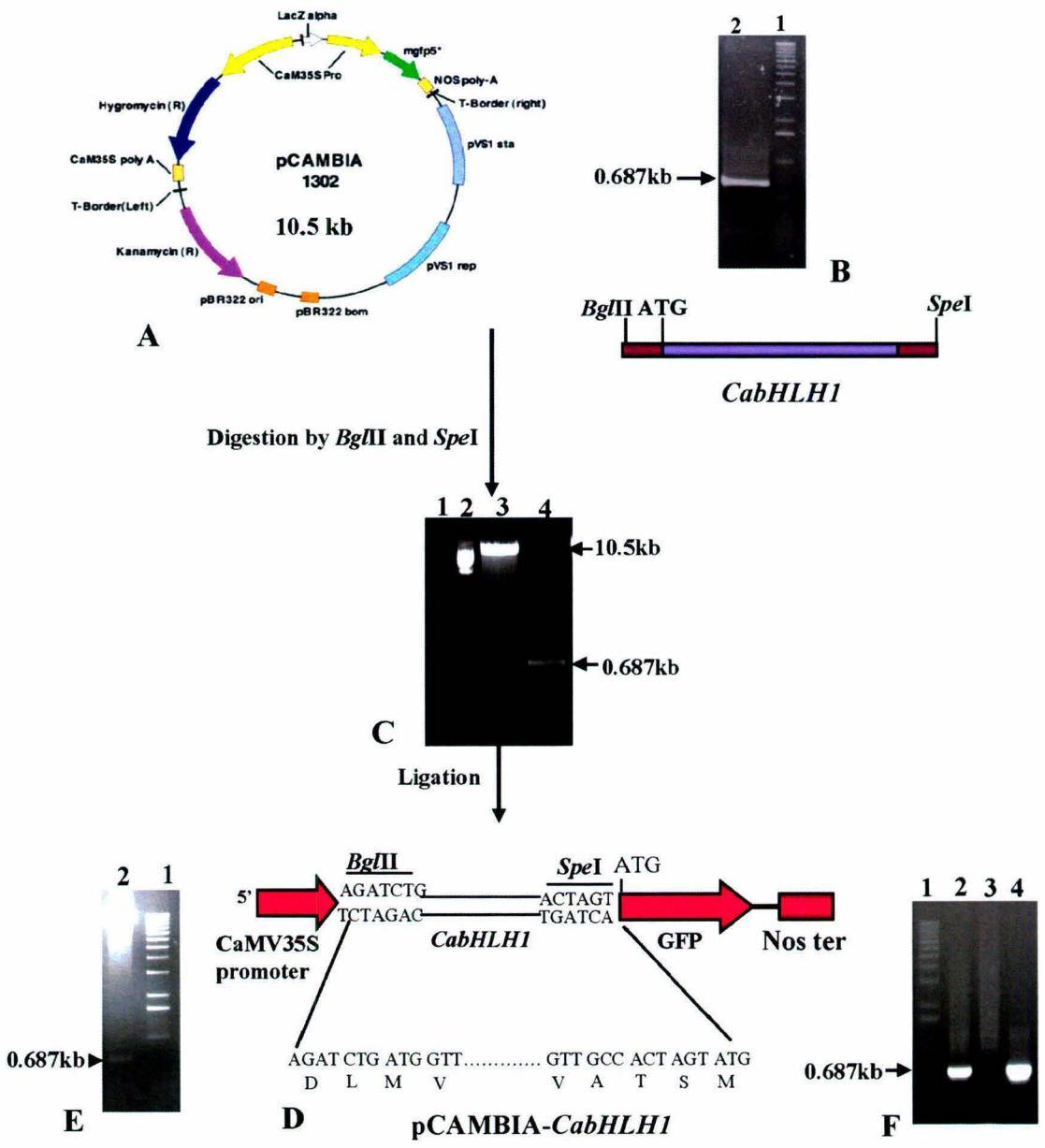
Figure 5.10: Construction of pCAMBIA-*CabHLH1*. (A) Map of pCAMBIA vector used for cloning. **(B)** PCR amplification of *CabHLH1* gene with cloning sites containing sequence for restriction enzymes *Bgl*II and *Spe*I. Lane 2 represents the PCR product **(C)** Digestion of pCAMBIA vector and *CabHLH1* amplicon with *Bgl*II and *Spe*I. Lanes 2 and 3 represent uncut and digested pCAMBIA vector respectively. Lane 4 represents digested *CabHLH1* amplicon. **(D)** Ligation of *CabHLH1* in pCAMBIA. **(E)** confirmation of positive clones by digestion. Lane 2 represents digested product of pCAMBIA-*CabHLH1* **(F)** confirmation of positive colonies by colony PCR. Lanes 2, 3, and 4 represent PCR products obtained from three different colonies. The construct pCAMBIA-*CabHLH1* was bombarded in onion cells.
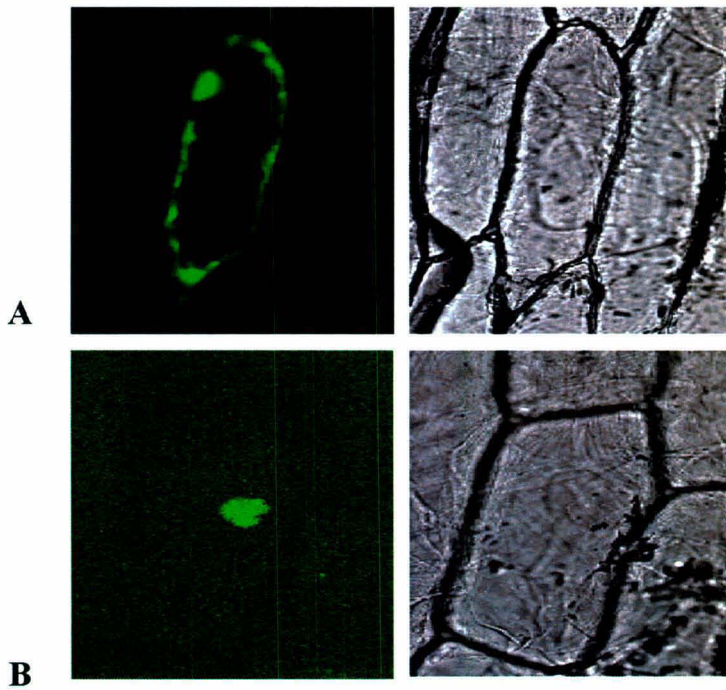
Figure 5.11: Subcellular localization of *CabHLH1*. Onion epidermal cells bombarded with **(A)** empty vector for control **(B)** pCaMBIA-*CabHLH1*-GFP. The GFP fluorescence was detected by using Confocal microscope. The right panel shows the corresponding phase contrast image.

Isolation and study of expression pattern of Ca14-3-3-1 gene

6.1 Introduction

All biological processes are controlled by signal transduction and metabolic regulation. One hallmark of signal transduction events is the phosphorylation-induced transition of a protein from one activity state to another. Kinases, phosphatases, transcription factors, and enzymes all can be influenced by phosphorylation. In many cases, to complete their regulatory actions, these phosphorylated proteins must physically associate with the specialized adapter proteins, which are known as 14-3-3 proteins (Dougherty and Morrison, 2004). 14-3-3 proteins generally act as adapters, chaperones, activators or repressors and interact physically with target proteins phosphorylated to execute an important step in signal transduction and metabolism (Ferl *et al.*, 2002). Moreover, the discovery that kinases and phosphatases themselves can be bound and regulated by 14-3-3s (Aitken *et al.*, 1995; Camoni *et al.*, 1998) firmly establishes 14-3-3s as integral components of signal transduction pathways. It should be noted, however, that 14-3-3s can also bind certain non phosphorylated targets (Henriksson *et al.*, 2002), indicating that they also have roles outside the context of phosphorylation mediated signal transduction.

6.1.1 Characteristic structural features of 14-3-3s

14-3-3 is a highly conserved and ubiquitously expressed protein family. In mammals, there are at least seven isoforms each encoded by a distinct gene (Aitken *et al.*, 1995). Up to 15 isoforms are present in plants and two isoforms have been identified in yeast, *Drosophila melanogaster* and *Caenorhabditis elegans* (Dougherty and Morrison, 2004). 14-3-3 molecules form homo and heterodimers that can interact with a wide variety of cellular proteins. The 14-3-3 protein sequence can be roughly divided into three regions: divergent amino terminus, conserved core region and divergent carboxyl terminus. The presence of divergent termini and few amino acid changes that occur in the core region result in multiple isoforms in most organisms and present the potential for target specific interaction occurring in distinct cellular locations. The conserved middle core region of the 14-3-3s encodes amphipathic groove that forms the main functional domain forming a cradle for interacting with target proteins. The divergent C terminus harbours a divalent metal binding site near the point at which the C terminus hinges to the main structure. The binding of divalent cations often is a requirement for 14-3-3 interaction with targets. Analyses of known 14-3-3 binding

sites, together with the use of peptide libraries, have defined two high-affinity phosphorylation-dependent binding motifs that are recognized by all 14-3-3 isotypes: RSXpSXP (mode 1) and RXXXpSXP (mode 2), where pS represents phosphor seirne. However, phosphorylation-dependent sites that diverge significantly from these motifs have also been described. Whether the interactions are dependent on phosphorylation or not, all targets appear to interact with the same binding domain on 14-3-3 (Sehnke *et al.*, 2002)

6.1.2 14-3-3s resolve signal transduction

The underlying paradigm for 14-3-3 participation in signal transduction events is the phosphorylation-dependent association of 14-3-3s with their targets. 14-3-3 acts as a 'molecular anvil' that causes conformational changes in the binding partner that can alter its enzymatic activity; mask or reveal functional motifs that regulate its localization, activity, phosphorylation state and/or stability. This entire regulatory process would be contingent on the cellular levels of 14-3-3s, the kinase and phosphatase that act on the target enzyme, and divalent cations. This complex contingency, as well as the large number of possible 14-3-3 isoform combinations allows for multiple regulatory controls on the target activity (Sehnke *et al.*, 2002). The phosphorylation of some 14-3-3-binding motifs appears to be constitutive, whereas others are highly regulated, phosphorylation being mediated by kinases that are activated under specific conditions. In addition, some sites are substrates of multiple kinases, the enzyme involved depending on the cell type and conditions. An emerging theme from the study of kinase regulators is that the effect of 14-3-3 on a particular process can, in some cases, be attributed primarily to the activity of a particular kinase or group of kinases. Of equal importance are the phosphatases that dephosphorylate the 14-3-3 docking sites. Not surprisingly, the ubiquitously expressed serine/threonine protein phosphatases PP2A and PP1 found in all eukaryotic organisms are important players in this process (Dougherty and Morrison, 2004).

6.1.3 Functions of 14-3-3 proteins

A wide array of biological functions involving kinase-mediated signal transduction, growth and developmental regulation and response to environmental stress have been attributed to members of the 14-3-3 family. The notable recurrent themes in these diverse systems are the involvement of protein–protein interactions, divalent cations,

kinases, and phosphatases, and the role of 14-3-3s continues to center on direct participation in signal transduction events (Sehnke *et al.*, 2002).

6.1.3.1 Regulation of intracellular protein localization

One important means by which 14-3-3 regulates cellular processes is by modulating protein localization. In most cases, 14-3-3 binding sequesters the target protein in a particular subcellular compartment, and the release of 14-3-3 then allows the target to relocate. This relocation is often due to the exposure of an intrinsic subcellular targeting sequence that was masked by the 14-3-3 dimer. This mechanism of regulation contributes to the nuclear retention of proteins such as human telomerase reverse transcriptase (TERT), Txl-2 and Chk1. This regulatory mechanism also applies to proteins that shuttle from the cytoplasm to the plasma membrane, particularly proteins that are involved in Ras and heterotrimeric G-protein signaling. 14-3-3 has also been implicated in the endoplasmic reticulum (ER)-to-plasma-membrane trafficking of certain multimeric complexes, including the KCNK and KATP potassium channels (Dougherty and Morrison, 2004).

6.1.3.2 Growth and developmental signaling

In several organisms, it has become apparent that 14-3-3s play a pivotal role in the growth and development of the cell. A clear example of this comes from the study of the yeast *Saccharomyces cerevisiae* 14-3-3s, known as BMH1 and BMH2 (van Heusden *et al.*, 1995). Strains with disrupted BMH genes grow more slowly on minimal medium and double mutants are lethal.

6.1.3.3 Kinase-mediated signal transduction

Signal transduction via protein phosphorylation is a common pathway for many organisms. Perhaps the best studied of these is the GTP-dependent Ras pathway. The stimulation of cell division by extracellular growth factors involves the receptor-based production of active Ras, which then turns on a series of protein kinases, including Raf-1, which activate enzymes in the nucleus that are critical for message transduction, including the transcription factors. In 1994, several investigators independently identified 14-3-3s within the Ras pathway as activators of Raf-1 and 14-3-3 proteins were found to associate with Raf-1 *in vivo* while completing its activation and recruitment to the membrane. Since these initial reports, other kinases

have been shown to either bind to or be activated by 14-3-3s including the notable calcium dependent protein kinase (Sehnke *et al.*, 2002).

6.1.3.4 Structure and movement

There are examples of 14-3-3 proteins interacting with proteins that would not normally be considered enzymes or be subject to signal-induced transitions in activity. Keratin intermediate filaments are expressed in simple-type epithelia and are responsible for cell structural integrity and 14-3-3s associate with keratins where they act as solubility factors. Other structural roles for 14-3-3s are indicated by the localization of 14-3-3s to the mitotic spindle apparatus, by associations with centrosomes, and by interactions with other cell scaffold–type. In certain contexts, 14-3-3s serve to alter the subcellular localization of their targets. The 14-3-3s contain a nuclear export signal, such that interaction between 14-3-3s and targets within the nucleus serves to assist in the nuclear export of the target (Sehnke *et al.*, 2002).

6.1.3.5 14-3-3 and human diseases

14-3-3 proteins have been shown to be associated with many human diseases. A role for 14-3-3 in human cancer is not unexpected given the interaction of these proteins with components of both signal transduction and cell-cycle regulatory pathways. The binding of 14-3-3 protects p53 from Mdm2-mediated ubiquitinylation, thereby stabilizing its levels (Yang *et al.*, 2003). 14-3-3 binding also promotes p53 tetramerization, resulting in increased transcriptional activity and thus acts as a tumor suppressor, and loss of its function may be a crucial event in the progression of certain human cancers. Evidence is accumulating that 14-3-3 might contribute to both Alzheimer's disease and Parkinson's disease. 14-3-3 proteins have been identified in the characteristic pathological lesions associated with these diseases (Dougherty and Morrison, 2004).

6.1.3.6 14-3-3s and plant stress

14-3-3 proteins function as regulators of a wide range of target proteins that are involved in responses to abiotic and biotic stress in plants, by regulating target proteins with functions of either signaling or transcription activation or defense. The study of stress induction in different organisms has led to the identification of 14-3-3s as integral components of response pathways. Environmental conditions affect 14-3-3

proteins directly, because external stimuli such as cold and increased salt are shown to have led to altered regulation of plant 14-3-3s. Biological interactions between organisms also engage 14-3-3s, as in the *Pseudomonas aeruginosa* exotoxin S, which requires a cellular 14-3-3 for activity. A direct relationship exists between 14-3-3s and the receptor for the wilt-inducing phytotoxin fusicoccin (FC). Fusicoccin (FC) binds to the complex of 14-3-3s and H^+ATPase, resulting in continuous high-activity state of H^+ATPase, guard cell solute uptake and excessive transpiration that leads to leaf wilt (Sehnke at al, 2002).

14-3-3s function in a plethora of cellular processes in plant and animal systems, their major role being in signal transduction pathways. We observed induction of this protein in *Fusarium* infected chickpea plants. Earlier report had shown involvement of a 14-3-3 protein in modulating the activity of H^+ATPase which ultimately led to excessive transpiration and wilting in *Fusarium* infected plants. We were very inquisitive to know whether this gene had a similar role to play or was associated with pathogen stress in a different way since 14-3-3 proteins are known to perform diverse roles in cellular processes. In this direction, we have cloned a 14-3-3 gene from chickpea (*Ca14-3-3-1*) and performed its expression studies. We have also studied its subcellular localization and genome organization in chickpea.

6.2 Material and methods
6.2.1 Cloning of full length *Ca14-3-3-1*

Full length clone of *Ca14-3-3-1* was obtained by performing 3' and 5' RACE using gene specific primers and RACE kits (Invitrogen). The gene specific primers were designed from the sequence of the EST clone (CaF1_JIE_9_F4). 3'RACE was performed using gene specific primer (14-3-3-3'F) and UAP primer provided with the 3'RACE kit as detailed in section 4 of appendix III. 5'RACE was performed by using two gene specific primers (14-3-3-5'R1 and 14-3-3-5'R2) and AP from the kit as detailed in section 5 of appendix III. The amplified products were run on 1% agarose gel, purified with gel extraction kit (Qiagen) and the purified products were then cloned in the pGEM-T Easy vector. The 3' and 5' RACE products cloned in pGEM-T were named as p3'*Ca14-3-3-1* and p5'*Ca14-3-3-1* respectively. The clones were sequenced using standard procedure of sequencing. For the amplification of full length clone, gene specific primers were designed from the full length nucleotide

sequence as obtained from alignment of the 3'*Ca14-3-3-1*, 5'*Ca14-3-3-1* and the EST sequences. The full length cDNA clone was amplified by PCR using cDNA as template and the gene specific primer pair (14-3-3-F1 and 14-3-3-R1). The PCR product was run on 1% agarose gel, purified by gel extraction kit (Qiagen) and subsequently cloned into the pGEM-T Easy vector. The cloned product was named as p*Ca14-3-3-1*. The sequence of the primers is given in table 2 of appendix II. The PCR was carried out as described in section 6 of appendix III, and the annealing temperature for these reactions was kept at 53°C. For long term storage, the bacterial cultures were grown overnight and 80% sterile glycerol was added so as to obtain a final concentration of 15% and stored in -80°C.

6.2.2 Amplification of *Ca14-3-3-1* from genomic DNA

The genomic clone of *Ca14-3-3-1* was isolated using genomic DNA as template and 14-3-3-F1 and 14-3-3-R1 gene specific primer pair (see table 2 of appendix II) so as to amplify genomic ORF. For this primer pair, annealing temperature was kept at 53°C. Genomic DNA was isolated as described in section 2 of appendix III. The PCR was carried out as described in section 6 of appendix III.

6.2.3 Genomic southern

Genomic southern was performed in order to find out the copy number of *Ca14-3-3-1* in chickpea. 10μg of genomic DNA samples were digested with restriction enzymes *Hind*III, *Bam*HI and *Eco*RV; separated on 0.8% agarose gel; denatured and blotted on to Genescreenplus membrane (Amersham). For hybridization, ORF region of *Ca14-3-3-1* was amplified using 14-3-3-F and 14-3-3-R primer pair and p*14-3-3-1* as template which was further used for preparing the ^{32}P-dCTP labelled probe for hybridization. The detailed description of southern blot preparation and hybridization is given in section 14 of appendix III.

6.2.4 Northern blotting

Northern-blot analysis was performed to determine the expression pattern of *Ca14-3-3-1* in response to *Fusarium* wilt and its tissue specific expression. For this, 20μg of RNA samples were separated on a 1.5% formaldehyde-agarose gel and then blotted onto Genescreenplus membrane (Amersham). The EST clone of *Ca14-3-3*-1 was amplified from the plasmid containing this clone using M13 forward and reverse

primers. The primer sequence is given in table 2 of appendix II and the PCR detail in section 6 of appendix III. The amplicon was run on 1% gel, purified by gel extraction and the purified product was used for preparing the ^{32}P-dCTP labelled probe by random labelling using random labelling kit (NEB). The detailed description of northern blot preparation is given in section 15 and that of hybridization in section 14 of appendix III.

6.2.5 Quantitaive real time PCR

For quantitative RT-PCR, two week old chickpea seedlings were treated with various hormones and the tissue collected at various time points after treatment. Total RNA was extracted using TRIzol reagent and used for cDNA synthesis by High-Capacity cDNA Reverse Transcription kit (Applied Biosystems). The cDNA was diluted 10 times and qRT-PCR was performed in triplicates in ABI 7500 sequence detection system using SYBR Green Master Mix (Applied Biosystems) and gene specific primers 14-3-3RT-F and 14-3-3RT-R (see table 1 of appendix II). The relative quantification method (ΔΔ-CT) was used to evaluate quantitative variation between the replicates examined. The amplification of 18S RNA was used as an endogenous control to normalize all data. The procedure of real time PCR is detailed in section 3.2.8 (chapter 3).

6.2.6 Subcellular localization

The subcellular localization of *Ca14-3-3-1* was studied by performing transient expression assay in onion epidermal cells. For this *Ca14-3-3-1* was fused in frame with 5' terminus of GFP reporter gene in pCAMBIA-1302. The fusion construct was named as pCAMBIA-*Ca14-3-3-1* and was prepared by amplifying *Ca14-3-3-1* from p*Ca14-3-3-1* using pCam14-3-3-F and pCam14-3-3-R primer pair (see table 2 of appendix II) and cloning in *Nco*I and *Spe*I sites of vector as described in section 17 of appendix III. The fusion construct was bombarded on to the onion peels which were then incubated for 24 hours before visualizing in confocal microscope.

6.3 Results and discussion
6.3.1 Cloning of full length *Ca14-3-3-1* Gene

Full length *Ca14-3-3-1* was amplified by 5' and 3' RACE using gene specific primers. For amplification of the 3' end, a gene specific forward primer was designed towards

the 3' end of the EST sequence (CaF1_JIE_9_F4). cDNA was made using AP primer from the RACE kit (Invitogen) followed by the PCR amplification with the help of GSP primer (14-3-3-3'F). Around 500bp amplified product was obtained as a result of final PCR amplification (Figure 6.1). PCR product was ligated into pGEM-T Easy vector, sequenced and analyzed.

For 5' RACE two gene specific primers were designed which were 184bp and 149bp interior of the 5'end of the partial clone respectively. GSP1 which was named as 14-3-3-5'R1 was used for making cDNA for the 5'RACE and GSP2 (14-3-3-5'R2) was used for the direct PCR amplification. The amplified product of ~0.6 kb (Figure 6.2) was eluted from gel and cloned in pGEM-T Easy vector. The clone was sequenced by standard procedures of sequencing and confirmed by the presence of overlap with the existing clone. The 5' and 3' RACE resulted in 1.019 kb long full length cDNA clone of *Ca14-3-3-1* with 792bp ORF (Figure 6.3).

6.3.2 *In silico* analysis of the *Ca14-3-3-1* encoded protein sequence

The predicted protein product of *Ca14-3-3-1* consists of 263 amino acid residues with a calculated molecular mass of about 30.148 kDa and isoelectric point of 4.83. The schematic representation of the gene structure is given in figure 6.4A. *In silico* analysis with ExPASy ProtParam tool revealed the presence of approximately 12.5% basic, 18.25 % acidic and 69.25 % neutral amino acids. The homology search using BLASTX against GenBank database showed that *Ca14-3-3-1* is similar to many other 14-3-3 genes with maximum homology with that of *Glycine max* and *Medicago*. The deduced protein sequence has a 14-3-3 domain which is represented by amino acid 8 to 243 (Figure 6.4B, 6.4C).

6.3.3 Multiple sequence alignment and phylogenetic analysis of *Ca14-3-3-1*

We performed a comparison of predicted amino acid sequence of Ca14-3-3-1 with other 14-3-3 protein sequences in the database using MAFFT program LINSI algorithm. We observed that all 14-3-3 proteins shared a significant stretch of homology in the core region of the protein whereas, the N and C termini showed some variability. Ca14-3-3-1 showed maximum homology with its ortholog from *Medicago* followed by *Glycine max* (Figure 6.5).
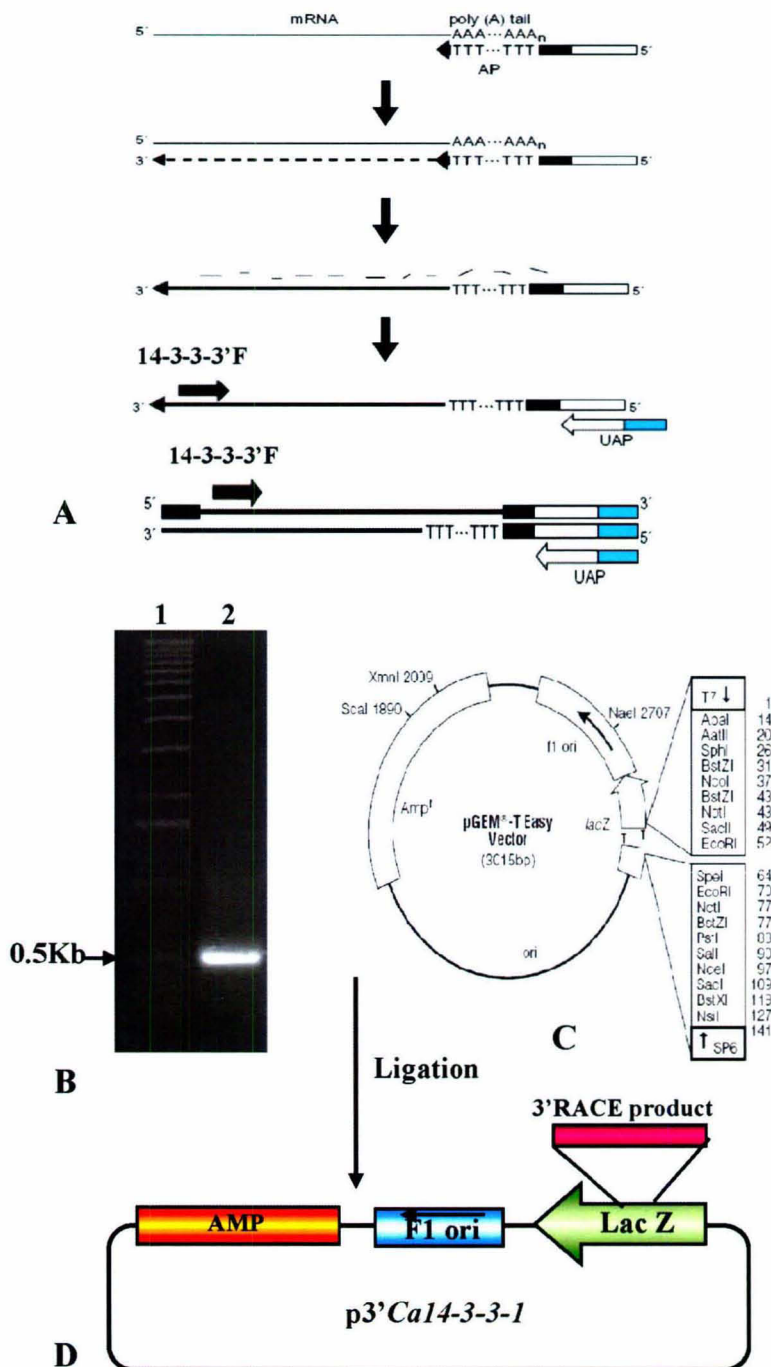
Figure 6.1: 3' RACE of *Ca14-3-3-1*. (A) Flow diagram depicting 3' RACE procedure. Oligo dT adapter primer (AP) was used to bind the polyA tail of the mRNA. Reverse transcriptase generates single stranded cDNA using mRNA as template. Gene specific primer (14-3-3-3'F) and UAP primer specific to 3' adapter were used to amplify the target 3'cDNA end sequence. **(B)** 1% agarose/EtBr gel showing 3' RACE product of *Ca14-3-3-1*. Lane 1 represents 1Kb ladder and lane 2 the PCR product. **(C)** Map of pGEM-T Easy vector used to clone the amplified product. **(D)** Diagrammatic representation of 3' RACE product of *Ca14-3-3-1* cloned in pGEM-T Easy vector and the construct named as p3'*Ca14-3-3-1*.
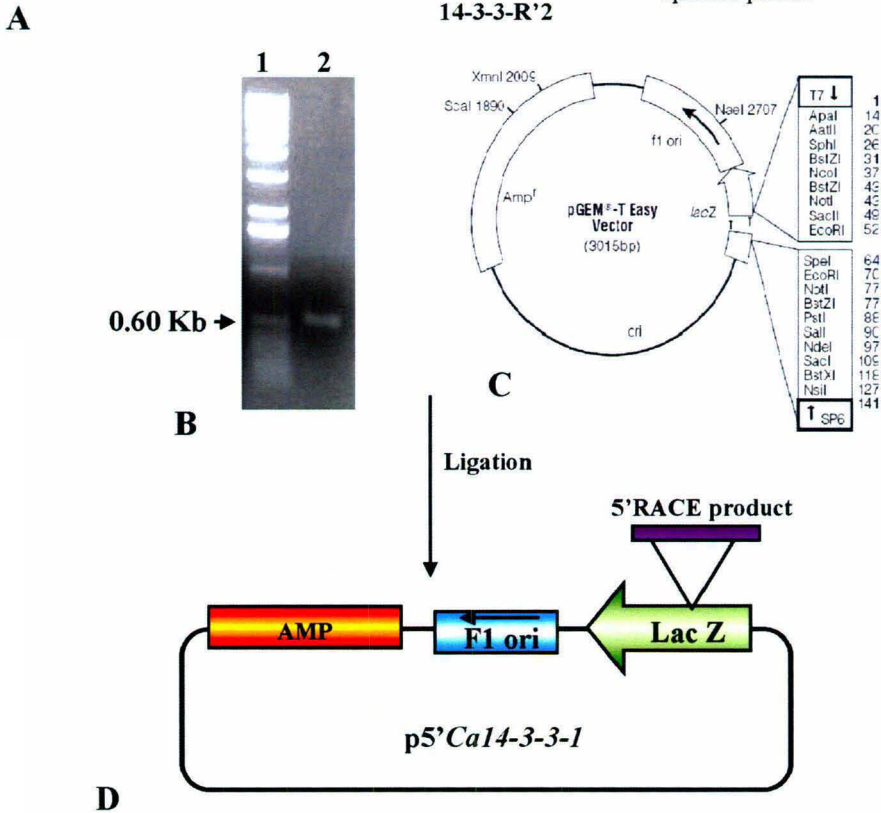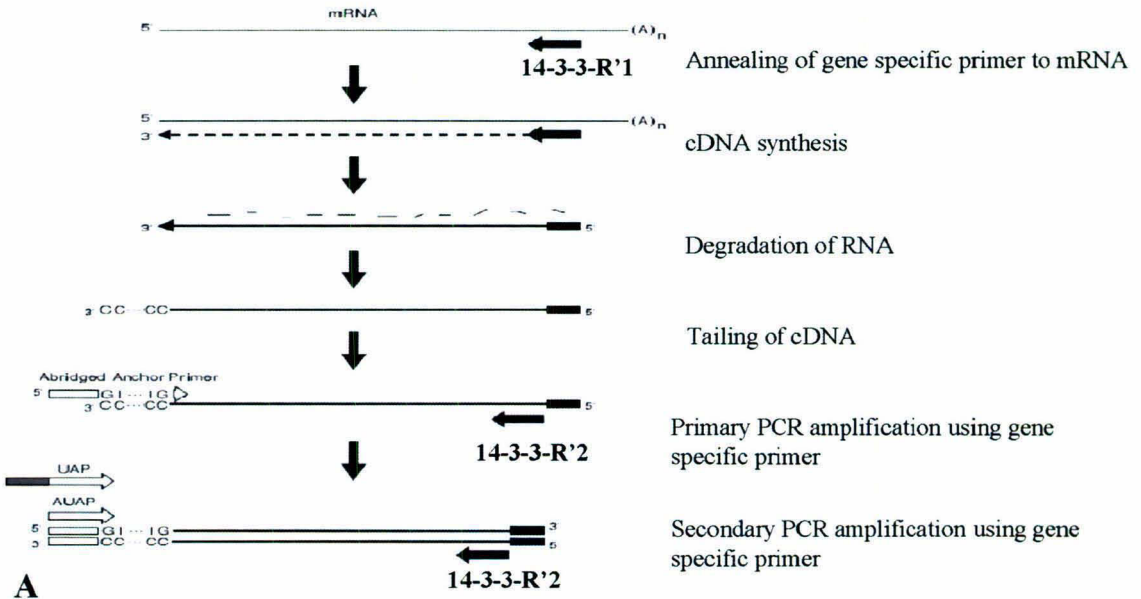
Figure 6.2: 5' RACE of *Ca14-3-3-1*. (A) Flow diagram depicting 5' RACE procedure. (B) 1% agarose/EtBr gel showing 5' RACE product of *Ca14-3-3-1*. Lane 1 represents 1Kb ladder and lane 2 the PCR product. (C) Map of pGEM-T Easy vector used to clone the amplified product. (D) Diagrammatic representation of 5' RACE product of *Ca14-3-3-1* cloned in pGEM-T vector and the construct named as p5'*Ca14-3-3-1*.
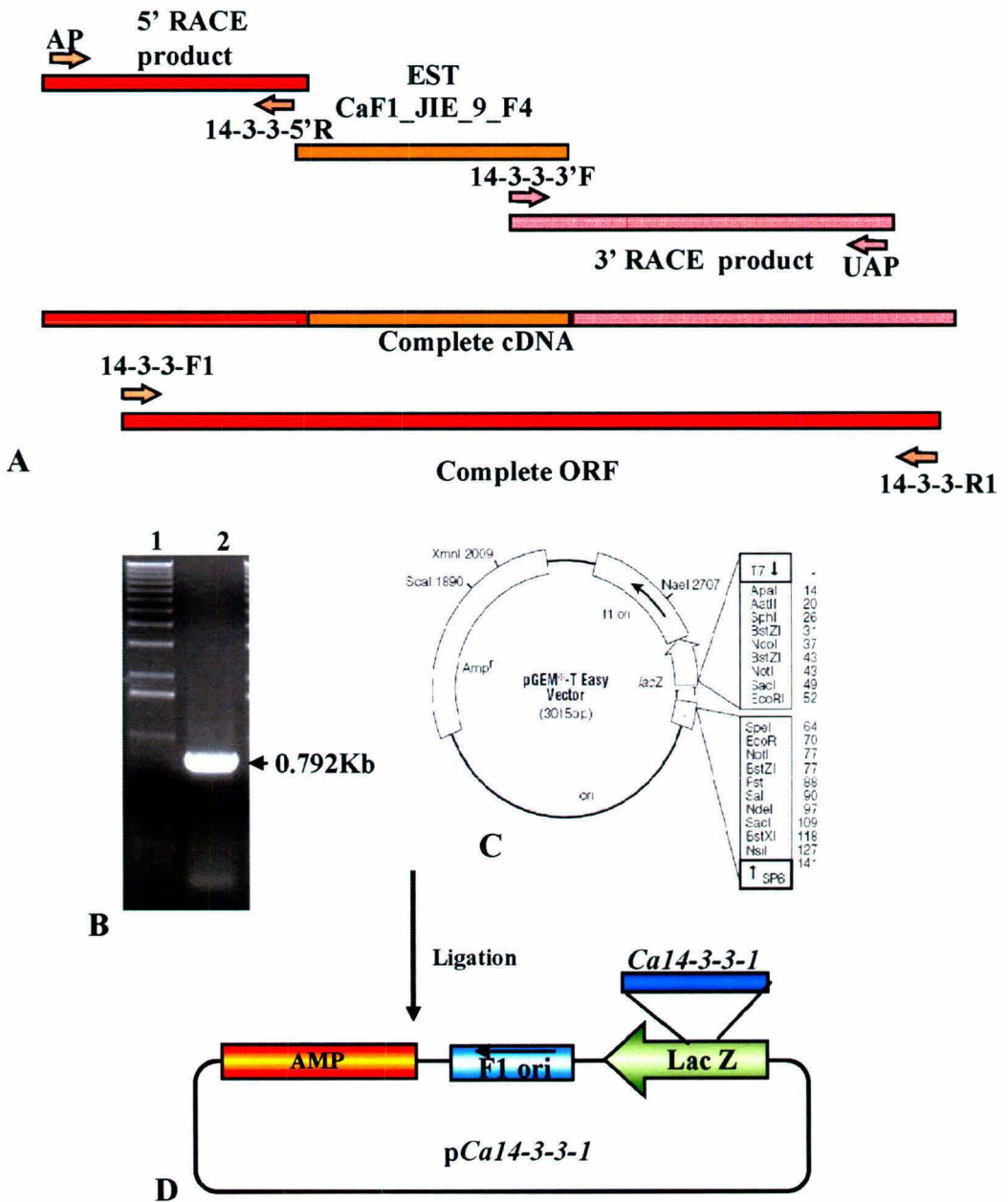
Figure 6.3: Isolation of full length cDNA clone of *Ca14-3-3-1*. (A) Schematic representation of the alignment of sequence of partial clone (EST) and that of 3'and 5'RACE products and the primer position for full length cloning. **(B)** 1% agarose/EtBr gel showing full length cDNA clone of *Ca14-3-3-1*. Lane 1 indicates 1Kb ladder and lane 2 PCR product. **(C)** Map of pGEM-T Easy vector used for cloning the PCR product. **(D)** Diagrammatic representation of full length clone of *CaSAUR1* cloned in pGEM-T Easy vector and named as p*Ca14-3-3-1*.

A

Full length cDNA of *Ca14-3-3-1*

B
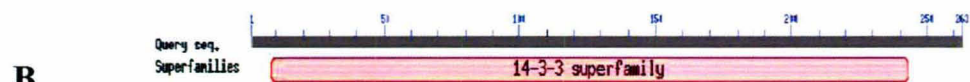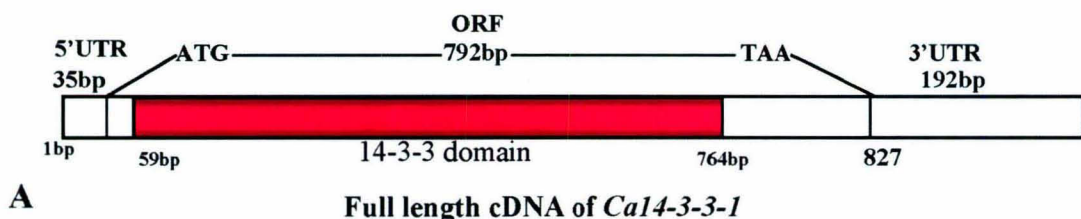
aacactgcaactttgagcagagagaaatttcagcc

```
ATGGCTTCTTCCACCAACGTCCGTGTGAACTTTGTCTATGTTGCAAAGCTAGCCGAGCAA
 M  A  S  S  T  N  V  R  V  N  F  V  Y  V  A  K  L  A  E  Q
GCTGAACGCTATGATGAAATGGTGGAAGCAATGAAGAAACTAGCAAAGATGGATGTTGAA
 A  E  R  Y  D  E  M  V  E  A  M  K  K  L  A  K  M  D  V  E
TTGAGTGTGGAAGAGAGAAACTTGTTCTCTGTTGGGTACAAGAATGTGGTGGGATCCAGA
 L  S  V  E  E  R  N  L  F  S  V  G  Y  K  N  V  V  G  S  R
AGAGCTTCATGGAGGATCCTATCATCAATAGAACAGAAAGAGGAATCAAAAGGGAATGAG
 R  A  S  W  R  I  L  S  S  I  E  Q  K  E  E  S  K  G  N  E
TTGAATGCGAAGCGTATTAAGGATTACAAGCACAAGGTGGAGTTGGAGCTTTCCAACATT
 L  N  A  K  R  I  K  D  Y  K  H  K  V  E  L  E  L  S  N  I
TGCAATGACATTATGATTCTTTTAGATGAACATCTTATTCCATCCACTAATGTTGCTGAA
 C  N  D  I  M  I  L  L  D  E  H  L  I  P  S  T  N  V  A  E
TCCACAGTGTTTTATTATAAGATGAAAGGAGATTATTATCGATACTTAGCTGAATTCAAA
 S  T  V  F  Y  Y  K  M  K  G  D  Y  Y  R  Y  L  A  E  F  K
GATGGTAATGAAAAGAAAGAGGTAGCAGATCAGTCACTTAAAGCGTATCAGACAGCTTCT
 D  G  N  E  K  K  E  V  A  D  Q  S  L  K  A  Y  Q  T  A  S
ACCACTGCTGAGAGTGAATTACAACCTACGCATCCTATTCGATTGGGTCTGGCTCTAAAC
 T  T  A  E  S  E  L  Q  P  T  H  P  I  R  L  G  L  A  L  N
TTCTCGGTGTTTTATTATGAGATAATGAATTCACCTGAAAGGGCCTGCCACCTTGCAAAG
 F  S  V  F  Y  Y  E  I  M  N  S  P  E  R  A  C  H  L  A  K
CAAGCCTTTGATGATGCTGTCTCAGAGCTGGATACCCTAAACGAAGATTCTTACAAGGAC
 Q  A  F  D  D  A  V  S  E  L  D  T  L  N  E  D  S  Y  K  D
AGTACCTTGATCATGCAGCTGTTGAGAGACAACCTTACTTTATGGACTTCTGATATCCCT
 S  T  L  I  M  Q  L  L  R  D  N  L  T  L  W  T  S  D  I  P
GAAGATGGTGAGGACCAAAAAATGGAAAGCACAGGTAGGAGTGGTCAAGATGAAGATGAA
 E  D  G  E  D  Q  K  M  E  S  T  G  R  S  G  Q  D  E  D  E
TTGGGTCGCTAA
 L  G  R  -
```

gttgtctgtcataactataagattgatgtgttaattttttttcagaatatttttttccaatgt
aatgaatatttgaatcatgtacttcatttgtatgctattgttgtttgtgtaacaataaccc
tattggtttgtaatgaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagtactctgcgttgata
ccactgctt

C

Figure 6.4: cDNA and deduced amino acid sequence of *Ca14-3-3-1*. (A)
Diagrammatic representation of gene structure of *Ca14-3-3-1*. **(B)** Domain analysis.
(C) Nucleotide and the amino acid sequence. Nucleotides in lower case indicate the
5'UTR and 3'UTR. Those in upper case indicate complete ORF with the start and
stop codons and the deduced amino acid sequence. Nucleotides highlighted in color
indicate 14-3-3 domain.

In order to understand the evolutionary relationship of Ca14-3-3-1, the amino acid sequence was aligned with eight orthologs of 14-3-3 from other organisms and a phylogram was generated using MAFFT program neighbour joining conserved method program. The analysis showed that Ca14-3-3-1 and 14-3-3 from *Medicago* and *Glycine max* shared same lineage falling closer to each other in the tree (Figure 6.6). 14-3-3 from *Nicotiana* and *Arabidopsis* were also placed closely in the tree showing similarity between them. Surprisingly, 14-3-3 from *Vicia faba*, a legume was placed distantly, suggesting variation within leguminaceae family.

6.3.4 Genome organization of *Ca14-3-3-1*

In an effort to gain an understanding of genome organization of *Ca14-3-31* gene, we conducted experiment to look for the presence of introns and detect the copy number of the gene. We did a PCR using genomic DNA of chickpea as template and primers from 5' and 3' end of the ORF. For the primer combination used (14-3-3-F1 and 14-3-3-R1), the product obtained with genomic DNA as template was same in size as the corresponding cDNA amplicon (Figure 6.7A). The same size of the genomic and cDNA clones of *Ca14-3-3-1* suggests that it is an intron-less gene.

To determine the copy number of *Ca14-3-3-1*, we performed a genomic DNA southern blot analysis using a full-length *Ca14-3-3-1* cDNA (792bp) as probe. For this, 10μg of the genomic DNA was digested completely with restriction enzymes: *Hind*III, *Bam*HI, *and Eco*RV separately. The digested genomic DNA products were resolved on 0.8% agarose gel and subjected to southern transfer via capillary method. A non cutter enzyme should produce single band if single copy of the gene is present while a single cutter enzyme should produce two bands. In this study, *Hind*III and *Bam*HI are non cutters and produced single bands. *Eco*RV, a single cutter produced two bands (Figure 6.7B). Taken together, these results suggest that *Ca14-3-3-1* is present as a single copy in chickpea.

6.3.5 Expression of *Ca14-3-3-1* in response to *Fusarium* wilt

In order to study the transcript accumulation of the *Ca14-3-3-1* in response to *Fusarium* wilt in chickpea, 25 day old chickpea plants were infected with *Fusarium* spore culture and root tissue was collected from the infected plants at various time points. RNA was isolated and run on a gel under denaturing conditions and transferred onto the nitrocellulose membrane. The blot was hybridized with a
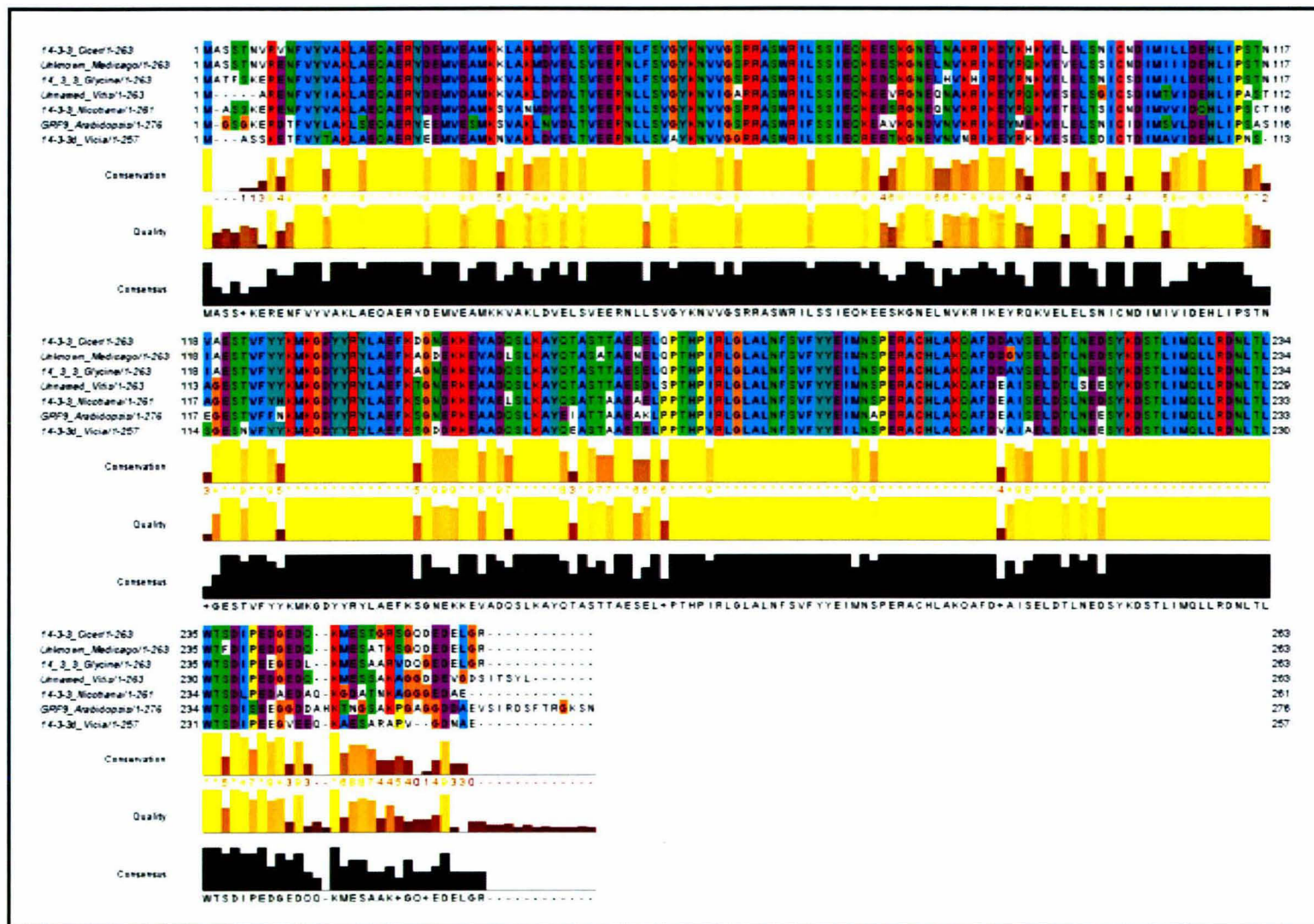
Figure 6.5: Sequence alignment of *Ca14-3-3-1* with the sequences of orthologs from other plant species. The multiple sequence alignment was done using MAFFT program LINSI algorithm.
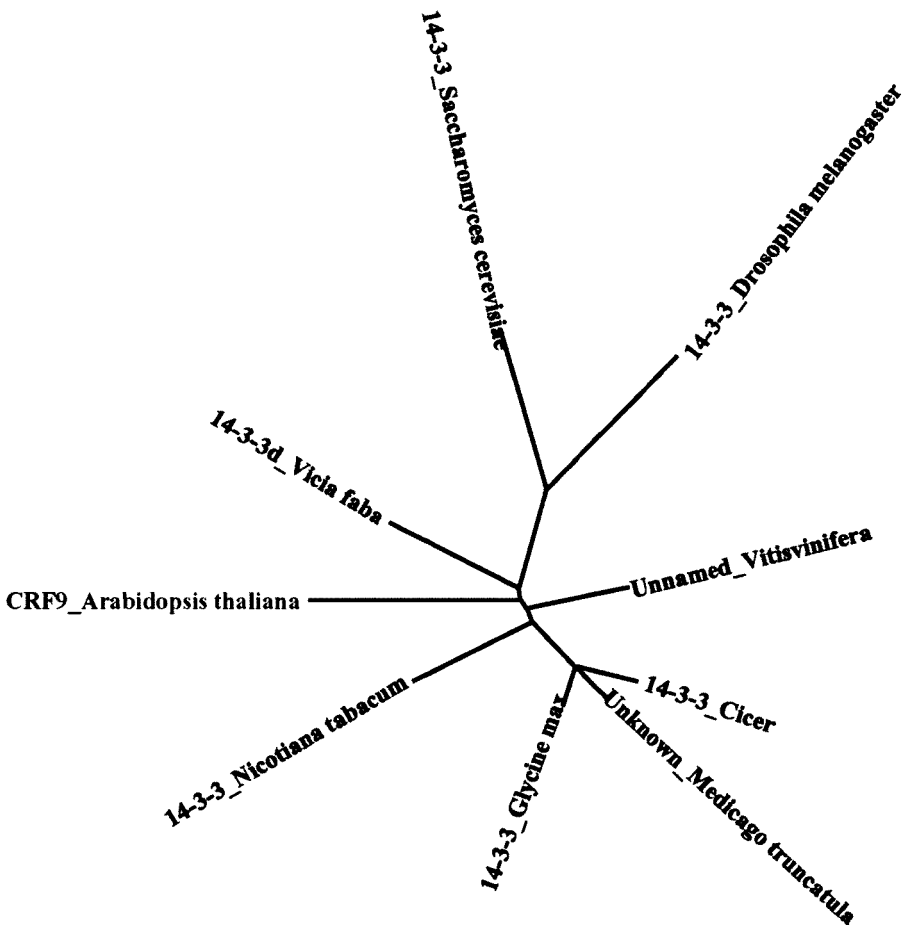
Figure 6.6: Phylogenetic tree showing evolutionary relationship between *Ca14-3-3-1* **and other 14-3-3 family proteins.** The tree was generated using MAFFT program neighbour joining conserved method.
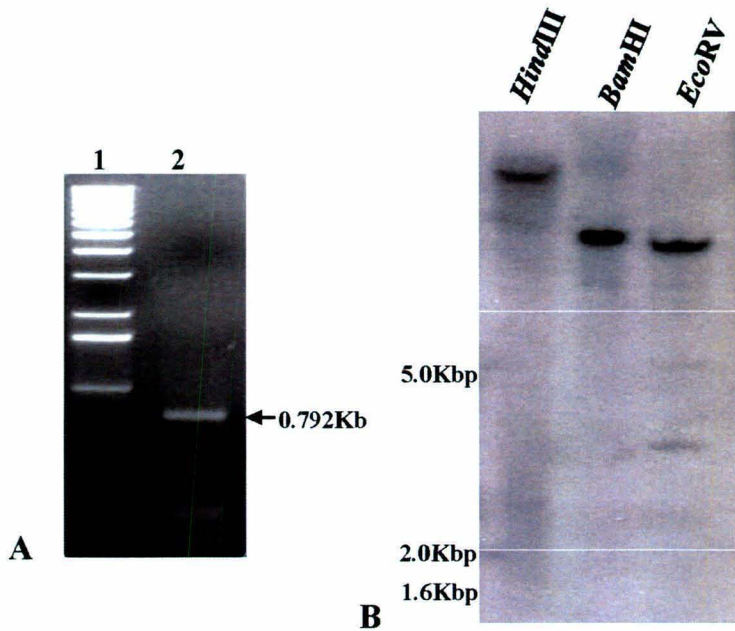
Figure 6.7: Genome organisation of *Ca14-3-3-1*. (A) PCR amplified product of *Ca14-3-31* using genomic DNA as template and ORF end primers. Lane 1 and 2 represent 1Kb ladder and PCR product respectively **(B)** Southern blot analysis. Chickpea genomic DNA (approximately 10μg) was digested with indicated restriction endonucleases, and hybridized with ^{32}P-labeled probe prepared from full length cDNA clone of *Ca14-3-3-1*.

radiolabeled *Ca14-3-3-1* cDNA fragment. It was quite interesting to observe that the expression of *Ca14-3-3-1* varied between the susceptible and resistant genotypes of chickpea. While it showed induction at 24 hours after *Fusarium* infection in the resistant genotype, the gene was expressed only after 5 days in the susceptible genotype (Figure 6.8A, 6.8B). The results suggest that this gene is involved in plant defense pathways and its delayed expression might be a cause to succumb to the disease in case of the susceptible genotype. Earlier also it has been suggested that delayed expression of defense related genes causes susceptibility (Tao *et al.*, 2003).

6.3.6 Analysis of tissue-specific expression of *Ca14-3-3-1*

In order to study the tissue specific expression of *Ca14-3-3-1*, total RNA was extracted from 21-day old chickpea seedlings infected with *Fusarium* and the northern blot was prepared and probed with the *Ca14-3-3-1* cDNA fragment. Varying levels of *Ca14-3-3-1* transcripts were detected in stem, leaf and roots with expression in root being higher than in other tissues (Figure 6.8C). This might be because of the fact that in case of *Fusarium* wilt, root is the first site of infection and hence may induce expression of defense related genes to a higher level.

6.3.7 Expression of *Ca14-3-3-1* in response to various hormonal treatments

In order or study the expression of *Ca14-3-3-1* in response to various hormonal treatments, we performed quantitative real time PCR. The results indicated that *Ca14-3-3*-1 showed upregulation in response to SA, JA and ACC. Induction was higher in response to SA and ACC compared to JA. More interestingly, *Ca14-3-3-1* was also induced in response to brassinosteroid and NO (Figure 6.9). Recently both these hormones were shown to mediate plant defense responses. Also, 14-3-3 proteins have been shown to play essential role in mediating brassinosteroid signal transduction in *Arabidopsis* (Gampala *et al.*, 2007). Expression of *Ca14-3-3-1* was much higher in NO treated plants indicating its strong role in NO mediated signalling during plant defense.

6.3.8 Subcellular localization of *Ca14-3-3-1*

14-3-3 proteins are widely localized in diverse subcellular compartments, implicating their multifunctional roles in cellular processes. We were interested in investigating the subcellular localization of the *Ca14-3-3-1*. For this, pCAMBIA-Ca*14-3-31* was
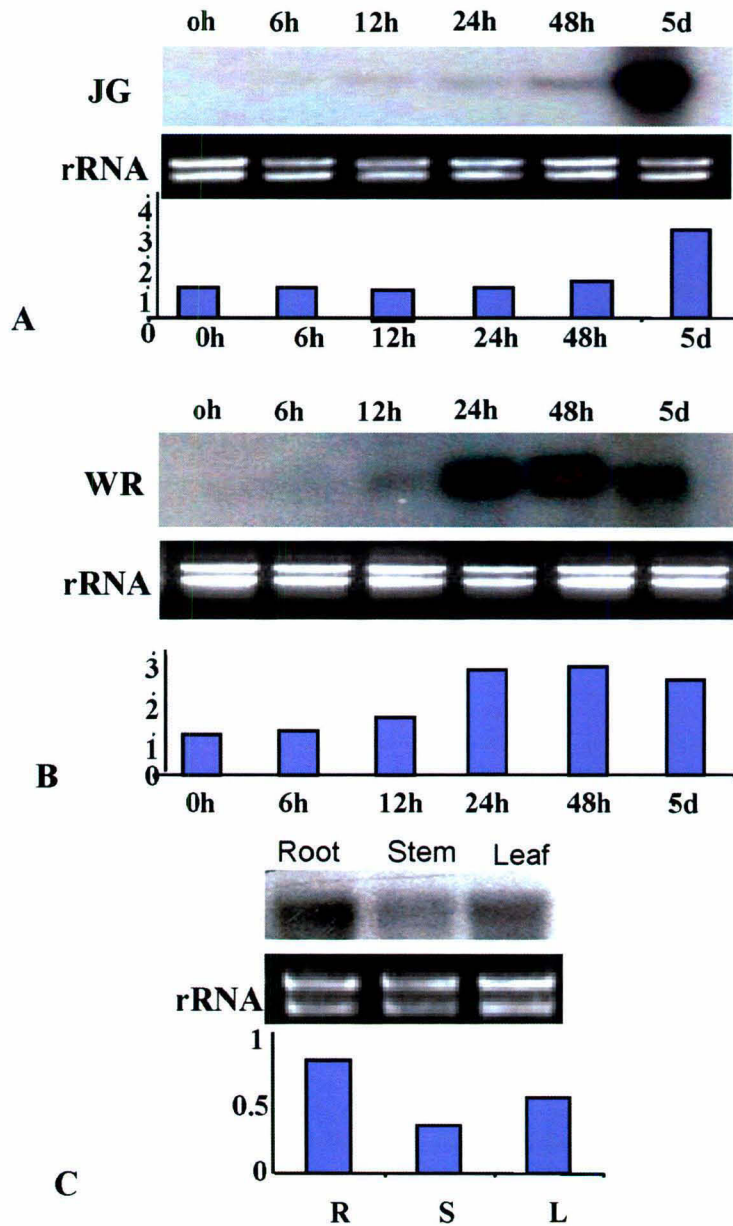
Figure 6.8: RNA blot analysis indicating expression pattern of *Ca14-3-3-1 gene.*
(A) Expression in response to *Fusarium* infection in (A) susceptible (JG-62) and **(B)** resistant (WR-315) genotypes of chickpea. 20 µg of total root RNA isolated from 25-day old chickpea seedlings harvested at various time points after *Fusarium* infection were separated by 1.5% agarose gel. **(C)** Expression in different tissues. For tissue specific expression RNA was isolated from different tissues. The northern blot was hybridized with a cDNA fragment of *Ca14-3-3-1* as mentioned in materials and methods. Ribosomal RNAs at lower panels are presented as loading controls. The graphs cited below are indicative of the fold expression in terms of band density.
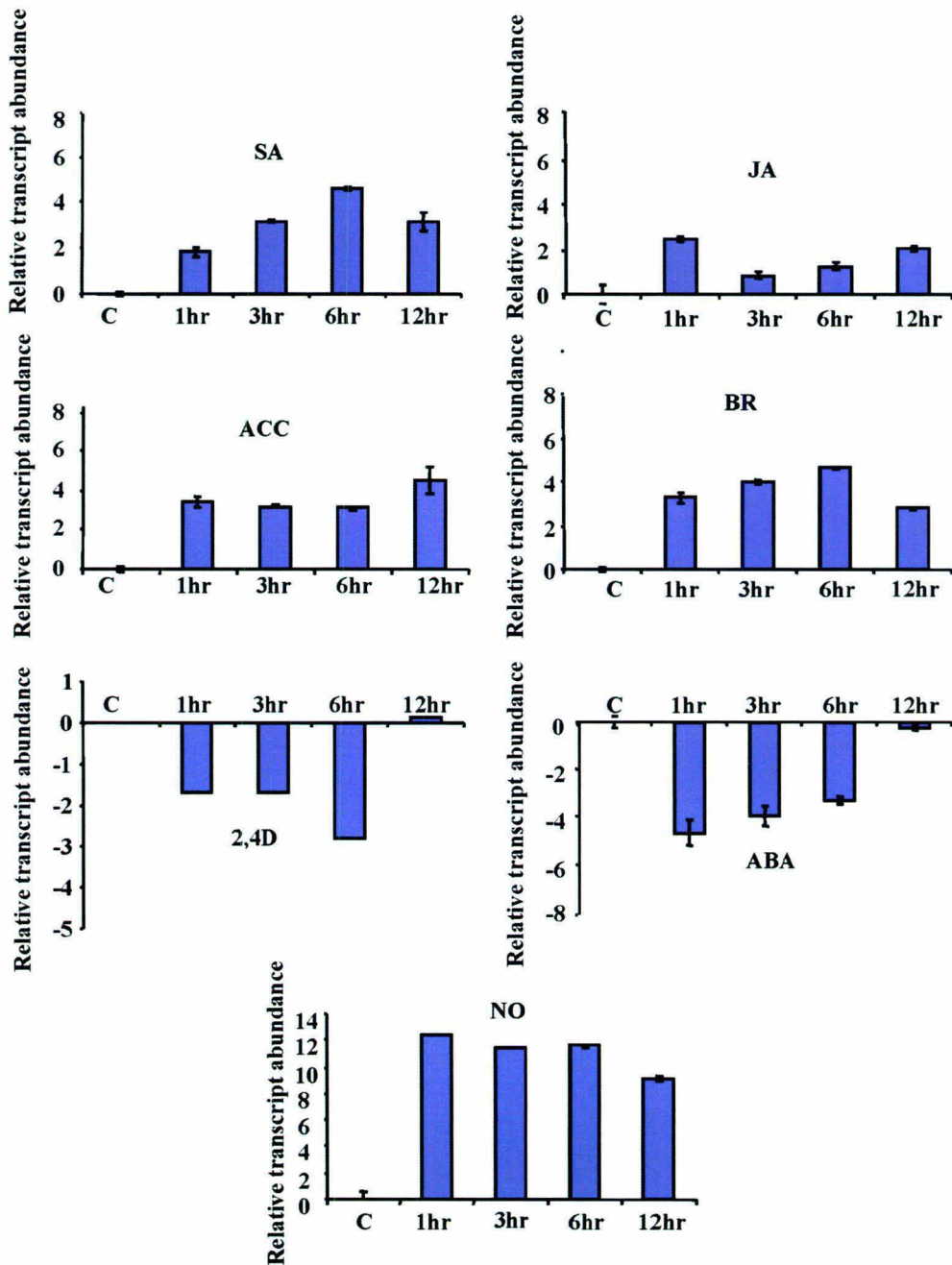
Figure 6.9: Real time PCR showing relative transcript level of *Ca14-3-3-1* in repsonse to various hormones. Transcript levels were normalized by 18S transcript level. Error bars indicate SD of three real time PCR experiments.

constructed (Figure 6.10) which has *Ca14-3-3-1* gene fused in frame to the 5'terminus of the green fluorescent protein (GFP) reporter gene. The expression of the fusion gene *Ca14-3-3-1-GFP* was driven by the 35S promoter of cauliflower mosaic virus (CaMV-35S). The plasmids containing vector control DNA and fusion gene were introduced into onion (*Allium cepa*) epidermal cells by particle bombardment. Upon observation under confocal microscope, we observed that Ca14-3-3-1-GFP was present in nucleus whereas the GFP protein alone was distributed in almost all cellular organelles (Figure 6.11).

6.4 Conclusion

The family of 14-3-3 proteins is known to be in involved in a range of signal transduction events and many functional roles have been attributed to these proteins. In this study a 14-3-3 gene from chickpea (*Ca14-3-3-1*) was cloned and its expression studied in response to *Fusarium* infection and various hormonal treatments. The results demonstrated that the gene is expressed in resistant cultivar of chickpea in early time points as compared to the susceptible genotype suggesting that timely expression of this gene is required for resisting the pathogen attack. The gene was also induced to very high levels in response to NO treatment indicating the involvement of NO mediated signaling pathway in defense. Further, the nuclear localization of *Ca14-3-3-1* suggests its transcriptional control in defense response.
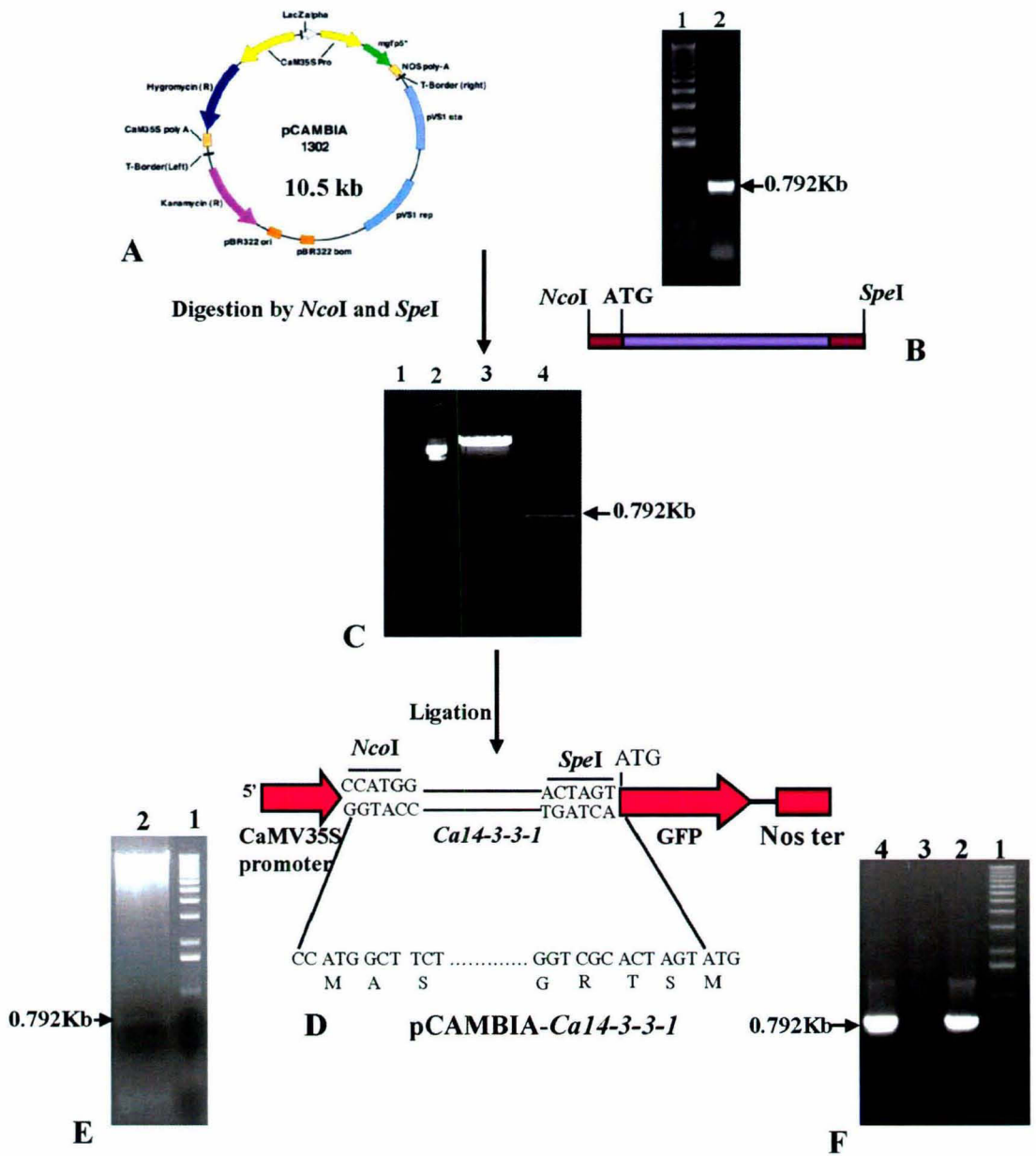
Figure 6.10: Construction of pCAMBIA-*Ca14-3-3-1*. (A) Map of pCAMBIA vector used for cloning. (B) PCR amplification of *Ca14-3-3-1* gene with cloning sites containing sequence for restriction enzymes *Nco*I and *Spe*I. Lane 2 represents the PCR product (C) Digestion of pCAMBIA vector and *Ca14-3-3-1* amplicon with *Nco*I and *Spe*I. Lanes 2 and 3 represent uncut and digested pCAMBIA vector respectively. Lane 4 represents digested *Ca14-3-3-1* amplicon. (D) Ligation of *Ca14-3-3-1* in pCAMBIA. (E) Confirmation of positive clones by digestion. Lane 2 represents digested product of pCAMBIA-*14-3-3-1* (F) Confirmation of positive colonies by colony PCR. Lanes 2, 3, and 4 represent PCR products obtained from three different colonies. The construct pCAMBIA-*Ca14-3-3-1* was bombarded in onion cells.
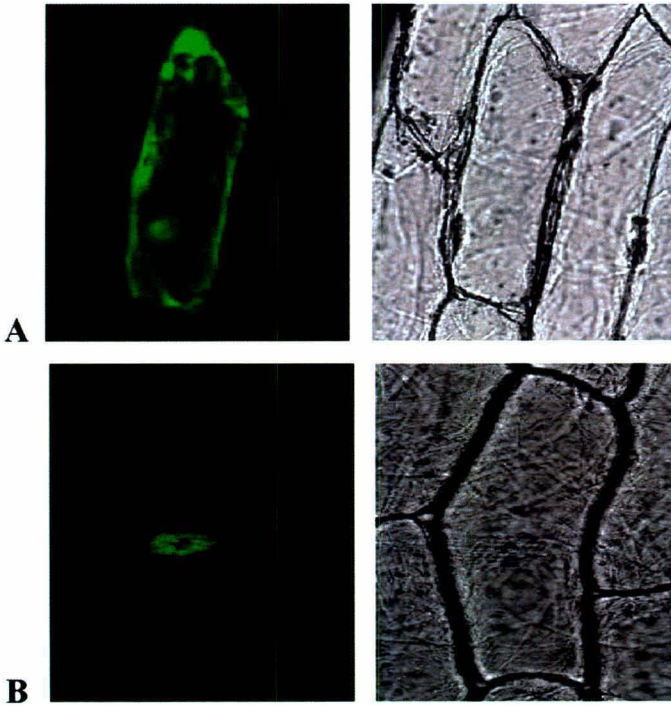
Figure 6.11: Subcellular localization of *Ca14-3-31*. Onion epidermal cells bombarded with **(A)** empty pCAMBIA vector as control and **(B)** vector pCaMBIA-*Ca14-3-31*-GFP. The GFP fluorescence was detected by using Confocal microscope. The right panel shows the corresponding phase contrast image.

Chapter 7

Summary

Summary

Environmental stresses bear serious threat to crop production. Among various stresses, biotic stress is an important factor and leads to 31-42% yield loss. Therefore, it is important to address research hypotheses in context of various forms of biotic stress, host-pathogen relationship and gene repertoire expressed during defense or disease and regulatory networks underlying these responses. The genomics technologies made it possible to study in detail the area of plant-pathogen interactions, and has provided unparalleled perception into the mechanisms underlying gene-for-gene resistance and basal defense, host vs nonhost resistance, biotrophy vs necrotrophy, and pathogenicity of vascular vs nonvascular pathogens, among many others. In this way, genomics technologies have facilitated a system-wide approach to unifying themes and unique features in the interactions of hosts and pathogens.

Much of the study conducted so far on immune responsive pathways is based on the biotrophic pathogens and very little is known on how necrotrophs are perceived and which immune pathways come into play to resist the attack by them. Therefore it is important to direct the research to the study of necrotrophic pathogens. *Fusarium* is an important necrotrophic pathogen infecting a wide range of crop plants which makes its study very important.

Chickpea is the world's third most important legume and is important protein rich food and an increasingly valuable traded commodity. This makes it an important crop for genomic prospecting. Chickpea is affected by various biotic and abiotic factors which limit the yield of this crop. One of the biotic constraints to increased and stable yields of chickpea is *Fusarium* wilt caused by (*Fusarium oxysporum* f.sp. *ciceri*). Various breeding efforts were directed towards the study of chickpea wilt but chickpea being recalcitrant to breeding methods did not yield much success in this regard. Therefore, it became necessary to *ride the genomics wave* for dissecting the molecular mechanism of chickpea wilt and frame the strategies to develop resistant cultivars.

Towards this, the present study provides an initial platform for the functional genomics of chickpea. We developed a chickpea EST database enriched with wilt responsive genes. These ESTs were functionally annotated and classified into various functional classes. Comparative genomics involving comparison of chickpea ESTs

with other legume ESTs led to the identification of many new chickpea specific genes. Comparison of ESTs from this study with other known stress related genes led to the discovery of a catalogue of non canonical genes which gave a new dimension to stress biology. SNPs were also identified between the susceptible and resistant genotypes of chickpea which may assist in studying genotype specific response to *Fusarium* wilt. Various chickpea gene families were also identified and may help in understanding the functional novelty of members of these families, their functional convergence or divergence.

The transcript profiling during chickpea-*Fusarium* interaction, conducted in the present study with the susceptible and resistant genotypes led to the identification of a set of differentially expressed genes among which some were common to both the genotypes while a subset of genes were specific to either of the genotype. This finding suggested that as has been mentioned in the previous reports, a major part of the signal transduction network might be shared during compatible and incompatible interaction; however, the genes specific to either of the genotype might form a separate branch and bear a direct effect on susceptibility or resistance. Based on the microarray expression data, we hypothesised a comprehensive immune responsive regulatory network.

Moving towards the functional and application genomics, rest part of the research was directed towards cloning and characterization of few of the stress responsive genes so as to understand the role played by these genes and plan strategy to develop the resistant cultivars. Three genes belonging to different functional classes were cloned. Identification of *CaSAUR1* as pathogen stress inducible gene was a novel discovery, since this group of genes has not been yet characterized and they have never been implicated in defense. *CabHLH1* was unique from other known bHLH genes in having a caspase domain. This gene was induced in both susceptible and resistant genotypes, however, there was no significant difference in the expression of this gene in two genotypes suggesting that this gene might have a role in basal defense during early signalling. *Ca14-3-3-1* was induced during early time points in resistant genotype, indicating its role in plant defense response. Thus it will be quite interesting to study the functional role of these genes and how they work in association with other genes in the regulatory network to mount immune response against pathogen attack.

Future plan

In future it would be interesting to know how the above mentioned genes work in the immune responsive networks. Towards this, the identification and characterization of interacting partners of the important genes would elucidate the protein interaction network that comes into play during vascular wilt. Also functional characterization of such genes using overexpression and antisense lines would be of great interest to know how these genes are involved in plant immunity. Furthermore, the genes belonging to no significant homology, hypothetical proteins and the non canonical groups would help in elucidating immune pathways which have not been identified so far.

Appendices

Appendix I: Materials and their sources

Plant and fungal strain
Cicer arietinum L.
Fusarium oxysporum ciceri

Bacterial Strain
Escherichia coli (DH5a)

Vectors
pGEM-T Easy	Promega
pCAMBIA-1302	Cambia labs

Kits
Perfectprep Plasmid 96 Vac Direct bind kit	Eppendorf
Wizard plus SV minipreps DNA purification system	Promega
Perfectprep PCR clean up kit	Eppendorf
Micromax TSA labeling and detection kit	PerkinElmer
3′ RACE System for Rapid Amplification of cDNA Ends	Invitrogen
5′ RACE System for Rapid Amplification of cDNA Ends	Invitrogen
Advantage 2 polymerase Mix	Clontech
DNeasy plant maxi kit	Qiagen

Marker
1kb Ladder DNA	Invitrogen

Microarray consumables
poly-L-lysine coated slides	Sigma
Hybridization chambers	Corning
Deep well culture plates	Nunc
96-well and 384-well plates	Axygen

Membrane and filter papers
Nitrocellulose, 3MM Whatman sheet	Amersham

Radioisotope
$[a^{32}P]dCTP$	Perkin Elmer

Enzymes
Restriction enzymes	NEB
Klenow	NEB

T4 DNA ligase	NEB
T4 DNA Polymerase	Perkin-Elmer, Life Sciences
General chemicals	Sigma, USB, Pharmacia, Qualigens
Antibiotic	**Concentration**
Ampicillin (Sigma)	50µg/ml
Kanamycin (Sigma)	50µg/ml
Media and solutions	
Germination medium	MS medium, pH 5.6
LB medium (1 liter)	24g/l
LB agar	32g/l
PDB	24g/l
80% Glycerol	80 ml ultra pure glycerol and 20 ml water mixed and filtered
50X TAE buffer (1 liter)	242 g Tris 28.25 g Na_2-EDTA 47 ml Acetic acid 752 ml MilliQ H_2O
6X Endo-R	30% Ficoll 400 60 mM EDTA (pH 8.0) 0.6% SDS
1X TE (pH 8.0)	10 mM Tris-HCl (pH 8.0) 1 mM EDTA (pH 8.0)
TNT buffer	0.1M Tris-HCl 0.15M NaCl 0.05% Tween -20
1M Tris-HCl	12.1 g/l of tris dissolved in water pH 7.5 adjusted by adding HCl and volume made upto 1liter
1M NaOH	4g/100ml
4M NaCl	23.3g/ml
1M CaCl$_2$	14.7g/ml
DEPC-treated water (1 litre)	1 ml of DEPC was added to 1 litre of MQ water, mixed, incubated overnight and autoclaved

121

10X MOPS buffer	200 mM 3-[*N*-morpholino] propane- sulfonic acid (MOPS) 50 mM sodium acetate 10 mM EDTA pH adjusted to 6.5–7.0 with NaOH and filter sterilized
20X SSC	175.3g NaCl 88.2g Trisodium citrate pH 7.0
20%SDS	20g/100ml
50%Dextran sulphate	25g dissolved in 30ml of sterile MQ by heating at 65°C and volume made upto 50ml
X-gal	20 mg/ml in Dimethyl formamide (DMF)
IPTG	200 mg/ml in H_2O
Salicyclic acid (1M)	0.138g dissolved in ethanol and volume made upto 1 ml by adding water.
Methyl jasmonate (1M)	0.224ml dissolved in acetone and volume made upto 1ml by adding water
2,4D (0.1M)	0.02g dissolved in ethanol and volume made upto 1ml by adding water
ACC (1mM)	0.1g dissolved in NaOH and volume made upto 1ml by adding water
Brassinolid (1M)	0.48g dissolved in ethanol and volume made upto 1ml with water
Sodium nitrosopruside (1M)	0.297g dissolved in 1ml of water
ABA (10mM)	0.0132g dissolved in ethanol and volume made upto 5ml with water

Controls used in microarray

Positive controls

Disease resistance response
protein
Foxy transposable element
pfkB-type carbohydrate kinase
copper amine oxidase
Fructose-bisphosphate aldolase

Negative controls

CaNAG2
Human actin
SSD
3X SSC
50% DMSO
1.5X SSC+50% DMSO
Spike controls

Appendix II: Sequences of the oligonucleotide primers

Table 1. Sequence of oligonucleotide primers used for quantitative real time PCR.

S.No.	Name of primer	Sequence of primer
3	PR10 - F	5'-TCCTTCACCATCCAGCATACC-3'
4	PR10 - R	5'-TGACGGGCTCGAGAAAATACA-3'
5	CI - F	5'-AGCCAAGGGTGATGGTCTTTT-3'
6	CI - R	5'-AATTGTAATTAGGATTGGCCAAACA-3'
7	ERF - F	5'-GGGCAGGTGTCGACTCTCAT-3'
8	ERF - R	5'-CGAGTCATGGTTGTTGATGGA-3'
9	WCP - F	5'-TCATCGGCAGTGTTGGAGACT-3'
10	WCP - R	5'-AACAGCGCTTTAAATCAGGAGAA-3'
11	PR5 - F	5'-CCGTGTTAGGAGCGGGATTT-3'
12	PR5 - R	5'-CATCCCCGCAGTTTTGTGT-3'
13	DnaJ - F	5'-GCTTCCAGTTTACTGTCACACATCTT-3'
14	DnaJ - R	5'-TTCCCCAGGGTTTGATTTGA-3'
15	Cys - F	5'-CTTGGTGGAGTTCGCGATGT-3'
16	Cys - R	5'-TAGCGGGCGAGACTATCGAT-3'
17	WD40 - F	5'-CGGATGCATTTGAGAAAAACC-3'
18	WD40 - R	5'-GGAAGAGCCTTATCCAGTGAAATC-3'
19	FOT5 - F	5'-GGTTTTCTCCTGCAAGGGTTT-3'
20	FOT5 - R	5'-GCCAGTAGCTGCAACCCTAGA-3'
21	HP - F	5'-TTGGTGGGTGAACAATCCAA-3'
22	HP - R	5'-TCGGCTCTTCCTATCATTGTGA-3'
23	PE - F	5'-GCAGATTGCTCATCCCACAA-3'
24	PE - R	5'-TGTCTAGCATGATGGAGCAGTTTT-3'
25	EXT - F	5'-CCAGTTTCACCACCTTACCACTACT-3'
26	EXT - R	5'-GAGGTAAGGCTTCTTGGTTGGA-3'
27	SAURRT - F	5'-CATCAGAAGGGCTGCAAACC-3'
28	SAURRT - R	5'-GCAAGATATCCCTTTGGCACAT-3'
29	bHLHRT - F	5'-GGAAGGCCTCCCAAAACA-3'
30	bHLHRT - R	5'-CCATTCGGACAGCGTCAA-3'
31	14-3-3RT - F	5'-GGAATCAAAAGGGAATGAGTTGA-3'
32	14-3-3RT - R	5'-CCAACTCCACCTTGTGCTTGT-3'
33	18S - F	5'-CCCCGTGTTAGGATTGGGTAA-3'
34	18S - R	5'-CGGCTACCACATCCAAGGA-3'

Table 2. Sequence of oligonucleotide primers used for cloning work

S.No.	Name of primer	Sequence of primer
1	SAUR-3'F	5'-TCTCACAATTCCTTGCAGAGAAGAC-3'
2	SAUR-F1	5'-AGGTCAGAAACTTGAGTCTTCTATA-3'
3	SAUR-R1	5'-AAGCAGTGGTATCAACGCAGAGTAC-3'
4	SAUR-F2	5'-ATGGGTTTTCGTTTACCTAG-3'
5	SAUR-R2	5'-TCAGCAAAAATTCAAACGAG-3'
6	pCamSAUR-F	5'-CATGCCATGGGTTTTCGTTTACCTAGT-3'
7	pCamSAUR-R	5'-GGACTAGTGCAAAAATTCAAACGAGA-3'
8	bHLH-3'F	5'-AAGAAGAGGGGTAGATCGGATTCAT-3'
9	bHLH-5'R1	5'-ATTAGGCTCAATCTTGGAGCCTGGA-3'
10	bHLH-5'R2	5'-GTTGGTGTTGAAATTGATGGCTCTCTGGGG-3'
11	bHLH-F1	5'-ATGGTTTCCCCGGAAAACACCAATTGGCTTTT-3'
12	bHLH-R1	5'-TTAGGCAACTGGTGGGCGGA-3'
13	pCAMbHLH-F	5'-GAAGATCTGATGGTTTCCCCGGAAAACACCAATTGG3'
14	pCAMbHLH -R	5'-GGACTAGTGGCAACTGGTGGGCGGAGTTCAT-3'
15	14-3-3-3'F	5'-GCGTATCAGACAGCTTCTACCACTG-3'
16	14-3-3-5'R1	5'-GCGTATCAGACAGCTTCTACCACTGCTGAG-3'
17	14-3-3-5'R2	5'-ATGAAAAGAAAGAGGTAGCAGATCAGTCAC-3'
18	14-3-3-F1	5'-ATGGCTTCTTCCACCAACGTCCGTG-3'
19	14-3-3-R1	5'-TTAGCGACCCAATTCATCTTCATCT-3'
20	pCAM14-3-3-F	5'-CATGCCATGGGCTTCTTCCACCAACGT-3'
21	pCAM14-3-3-R	5'-GGACTAGTGCGACCCAATTCATCTTCAT-3'
22	M13-F	5'-GTTTTCCCAGTCACGACGTTG-3'
23	M13-R	5'-TGAGCGGATAACAATTTCACACAG-3'

The underlined bases represent the sites of various restriction enzymes.

Appendix III: Molecular biology techniques

1. RNA isolation

Total RNA was isolated using TRIzol reagent (Invitrogen) as per the following procedure:

1. 0.2g of frozen tissue was crushed in liquid nitrogen and suspended in TRIzol reagent (1 ml/0.1g tissue; v/w) and vortexed for 30 seconds.

2. 200μl chloroform/ml Trizol reagent was added to the suspension, mixed gently and incubated for 10 min at room temperature (RT).

3. The contents were centrifuged at 13,000 rpm for 15 min at 4°C in a microcentrifuge.

4. The upper aqueous phase was carefully transferred to a fresh microfuge.

5. 500μl iso-propanol/ml Tripure reagent was added and incubated for 10 min at RT to precipitate the total RNA.

6. The pellet was collected by centrifugation at 10,000 rpm for 10 min at 4°C, washed with 70% ethanol and dried.

7. RNA was dissolved in 20μl of DEPC treated MQ water to yield a total of around 40μg of RNA with a concentration of 2μg/μl.

2. Isolation of Genomic DNA

Genomic DNA was isolated from two week old chickpea seedlings using DNeasy plant maxi kit (Qiagen)

1. 1g of frozen tissue sample was crushed using mortar and pestle.

2. 5ml of preheated buffer AP1 and 10μl RNase A were added to the crushed tissue sample and incubated for 10 min and during incubation the contents were mixed by inverting the tubes.

3. 1.8ml of AP2 buffer was added and incubated for 10 min on ice.

4. The contents were centrifuged at 5000g for 5 minutes

5. The supernatant was decant into a QIAshredder Maxi spin column in a 50ml tube and centrifuged at 5000g for 5 minutes

6. The flowthrough was transferred into a new 50ml tube and 1.5 volumes of AP3/E buffer added followed by votexing.

7. The sample was then transfered into DNeasy Maxi spin column in a 50ml collection tube and centrifuged at 5000g for 5 minutes.

8. 12ml of AW buffer was added and again centrifuged for 10 min at 5000 x g and the flowthrough was discarded.

9. The spin column was transferred to a new 50ml tube and 1ml of AE buffer added to elute the DNA by first incubating for 5 min at room temperature and then centrifuging for 5 minutes at 5000 x g.

3. Isolation of plasmid DNA

Plasmid DNA was isolated using Wizard plus SV minipreps DNA purification system (Promega) following manufacturer's instruction as mentioned below:

1. 5ml of overnight grown culture of bacterial cells was harvested by centrifugation at maximum speed for 1 minute.

2. The pellet of bacterial cells was resuspended in 150µl of resuspension buffer provided with the kit.

3. 150µl of lysis buffer was then added and the contents mixed by inverting the tubes containing the contents.

4. 10µl of protease was added and the contents incubated for 5 minutes.

5. 350µl of neutralization buffer was added and tubes again inverted for mixing.

6. The contents were centrifuged at 10,000 rpm for 10 minutes.

7. The supernatant was transferred to spin column and centrifuged for 1 minute so as to bind DNA to the membrane.

8. The flowthrough was discarded and 750µl of washing solution was added to the column followed by centrifugation for one minute.

9. The flowthrough was again discarded and 250µl of washing solution added to the column and a spin given for one minute.

10. After discarding the flowthrough, a spin was again given for one minute in order to remove any residual ethanol present in the washing solution.

11. The spin column was finally transferred to a fresh 1.5ml eppendorf and 20µl of elution buffer added to the column followed by centrifugation at 10,000 rpm for one minute.

12. The quality of the plasmid DNA was checked by agarose gel electrophoresis.

4. 3' RACE

Approximately, 2µg of total RNA isolated from WR-315 cultivar of chickpea was used to set up the 3'RACE reaction using 3'RACE kit (Invitrogen). For cDNA

synthesis, RNA was primed with oligo (dT) primer (AP) by adding 1μl of this primer to 2μg of RNA and volume made upto 11μl followed by the steps given below:

1. RNA-primer mixture was heated to 70°C for 10 min and chilled on ice for 1 minute. The following components were then added:

   10X PCR buffer   2.0μl

   25mM MgCl₂   2.0μl

   10mM dNTP mix   1.0μl

   0.1M DTT   2.0μl

2. The reaction mixture was tapped gently and equilibrated to 42°C for 5 min.

3. 1μl SuperScript™ II Reverse Transcriptase was added and incubation done at 42°C for 50 min.

4. The reaction was terminated by incubating at 70°C for 15 min.

5. Reaction mixture was chilled on ice, 1μl of RNase H was added, and incubated for 20 min at 37°C.

The cDNA prepared was used for PCR amplification of the 3' ends of all the three genes (*CaSAUR1*, *CabHLH1* and *Ca14-3-3-1*) by using gene specific primers which were designed from sequence of respective EST clones and the UAP primer provided with the kit. The sequence of the primers is given in table 2 of appendix II. PCR products were run on 1% agarose gel, purified by gel extraction using gel extraction kit (Qiagen) and ligated to pGEMT-Easy cloning vector (Promega). The positive clones were further sequenced and analysed.

5. 5' RACE

5' RACE was done in following steps:

First strand cDNA synthesis

5' RACE was done using total RNA isolated from WR-315 cultivar of chickpea seedlings. 2μg of total RNA was mixed with 1μl of 10μM stock of gene specific primer (GSP1) and the volume was made up to 15.5μl. The mixture was incubated at 70°C for 10 minutes to denature RNA followed by quick chill for 2 min on ice. The contents were collected in the tube by brief centrifugation and the following components were added to the reaction mix:

10X PCR buffer	2.5µl
25mM MgCl2	2.5µl
10mM dNTP mix	1.0µl
0.1M DTT	2.5µl

The contents were mixed gently and collected by brief centrifugation. Reaction mixture was incubated for 1 min at 42°C. 1µl of SuperScript™ II RT was added, mixed gently and the reaction mixture was incubated for 50 min at 42°C followed by incubation at 70°C for 15 min to terminate the reaction. A brief spin was given and reaction was again incubated at 37°C for 30 minutes after adding 1µl of RNase.

S.N.A.P. column purification of cDNA

cDNA was purified using S.N.A.P column provided with the kit as follows:

1. The binding solution was equilibrated to room temperature.
2. 120µl of binding solution (6M NaI) was added to the first strand reaction, and the mixture was transferred to S.N.A.P. column and centrifuged at 13,000 x g for 20 s.
3. 0.4ml of cold (4°C) 1X wash buffer was added to the spin cartridge and centrifuged at 13,000 x g for 20 s. The flowthrough was discarded and this step was repeated twice more for washing the column.
4. The spin cartridge was washed with cold 70% ethanol twice.
5. After removing the final ethanol from the tube, the spin catridge was transfered to a fresh tube and 50µl of pre-warmed MQ was added to elute the cDNA by centrifuging at 13,000 x g for 20 s.

TdT tailing of cDNA

1. The purified cDNA was used for TdT-tailing reaction. For the reaction, following components were mixed:

DEPC-treated water	6.5µl
5X tailing buffer	5.0µl
2mM dCTP	2.5µl
S.N.A.P-purified cDNA sample	10.0µl
final volume	24.0µl

2. The mixture was incubated for 3 min at 94°C, chilled for 1 minute on ice, and the contents of the tube were collected by brief centrifugation

3. 1µl of TdT was added and after gentle mixing incubated for 10 min at 37°C.

4. The enzyme was heat inactivated at 65°C and tailed cDNA was used for the direct PCR amplification.

PCR of dC-tailed cDNA

For the PCR amplification, following components were mixed to prepare the reaction mixture:

Sterile distilled water	31.5µl
10X PCR buffer	5.0µl
25mM MgCl$_2$	3.0µl
10mM dNTP mix	1.0µl
Nested GSP2 (10µM)	2.0µl
Abridged Anchor Primer (10µM)	2.0 µl
dC-tailed cDNA	5.0µl
DNA polymerase (5 units/µl)	0.5µl

PCR was performed using following cycling conditions:

1. Initial denaturation	94°C for 2 minutes
2. Denaturation	94°C for 0.5 min
3. Annealing of primers	60°C for 1 min
4. Primer extension Steps 2 to 4 repeated 30 times	72°C for 1 min
5. Final extension	72°C for 10 min
6. Indefinite hold	4°C, until samples are

removed.

Nested amplification

10µl of 5' RACE primary product was analysed on ethidium bromide agarose gel. The PCR product was diluted (1:100) and 2µl taken for the secondary PCR in order to increase the specificity of the PCR product. For this, the reaction mixture consisted of following components:

Sterilized distilled water	33.5µl
10X PCR buffer	5.0µl
25mM MgCl$_2$	3.0µl
10mM dNTP mix	1.0µl

Nested GSP2 (10μM)	1.0μl
AUAP or UAP (10μM)	1.0μl
Diluted primary PCR product	2.0μl
DNA polymerase (5 units/ μl)	0.5μl

PCR was performed using the following conditions:

1. Initial denaturation	94°C for 2 min
2. Denaturation	94°C for 0.5 min
3. Annealing of primers	60°C for 1 min
4. Primer extension Steps 2 to 4 repeated 30 times	72°C for 1 min
5. Final extension	72°C for 10 min
6. Indefinite hold	4°C

The PCR product was run on gel, purified by gel extraction using gel extraction kit (Qiagen) and ligated to pGEMT-Easy (Promega) cloning vector. The ligation mixture was transformed, positive colonies were screened and plasmids were isolated for positive colonies which were further sequenced and analysed. For all the genes, the gene specific primers were designed from the sequence of the respective EST clones. The sequence of the primers is given in table 2 of appendix II.

6. Polymerase chain reaction

All the PCR reactions for gene cloning from cDNA and genomic DNA were carried out using gene specific primers. The sequence of the primers is given in table 2 of appendix II. The reaction mixture of 50μl was prepared using various components as per the concentration given below:

10X PCR buffer	5.0μl
25mM $MgCl_2$	3.0μl
10mM dNTP mix	1.0μl
Gene specific forward primer (10μM)	1.0μl
Gene specific reverse primer (10μM)	1.0μl
Template (cDNA/genomic DNA) (50ng)	1.0μl
DNA polymerase (5 units/μl)	0.5μl
Sterile MQ water	37.5μl

Typically the amplification reactions were done for 30 cycles with specific alterations arrived at, empirically to optimize the yield. The PCR programme employed is given below:

1. Initial denaturation	94°C for 2 min
2. Denaturation	94°C for 30 seconds
3. Annealing of primers	X°C for 30 seconds
4. Primer extension Step 2 to 4 were cycled 30 times	72°C for 1 min
5. Final extension	72°C for 10 min
6. Indefinite hold:	4°C

"X" represents the annealing temperature which varied as per the primer combination used as mentioned in chapters 4, 5 and 6.

7. Colony PCR

Putative recombinant clones were identified by colony PCR. A single colony to be tested was suspended in 25μl of water and incubated at 100°C for 5 min in a thermocycler. The mixture was centrifuged and 3μl of supernatant was used as template for PCR using a combination of specific forward and reverse primers as mentioned in section 6.

8. Agarose gel electrophoresis

The quality of the DNA/RNA was checked by agarose gel electrophoresis as described in Maniatis *et al.*, (1982). Agarose gel electrophoresis was routinely carried out in 1X TAE both in gel as well as running buffer. For visualizing genomic and plasmid DNA, 0.8% gel was used while for the PCR products and for checking the quality of RNA 1% gel was used. Ethidium bromide was added to a final concentration of 0.5μg/ml in the gel and the gel was run at 100 V in case of DNA samples and 70 V in case of RNA samples until the dye reached 2/3rd of the gel. The gel was visualized under UV light and the image was captured using alpha imager-EC (Alpha Innovator Corporation).

9. Spectrophotometric Estimation of Nucleic acids

The quality and quantity of nucleic acids was determined by measuring the absorbance at 230, 260 nm and 280 nm. The amount was calculated as per the

formula 1.0 A_{260} = 50μg/ml for DNA and 1.0 A_{260} = 40μg/ml for RNA. The purity of nucleic acid was determined by calculating the ratio A_{260}/A_{280} ratio for each sample. For high quality, the ratios should be equal to or greater than 1.8.

10. Gel extraction

The PCR products were sliced from the gel and eluted using gel extraction kit (Qiagen) as follows:

1. The DNA band was excised from the agarose gel with a clean, sharp scalpel and 3 volumes of buffer QG were added to the gel in a tube.
2. The gel was incubated at 50°C for 10 min in a water bath for solubilisation of agarose and binding of the DNA.
3. After the agarose was solubilized, one volume of isopropanol was added, mixed properly and then the mixture was transferred to the column provided with the kit and centrifuged for 1 minute at 13,000 rpm.
4. The flowthrough was discarded and 500μl of QG buffer was again added to the column followed by centrifugation at 13,000 rpm for 1 minute.
5. Then a volume of 750μl of washing buffer PE was added to the column and centrifuged for 1 minute at maximum speed.
6. The flowthrough was discarded and the column again given a spin to remove the residual wash solution.
7. 10μl of elution buffer was added to the column and centrifuged at maximum speed for 1 minute in order to elute the purified PCR product and was stored at - 20.

11. Restriction digestion

Enzymatic manipulation of DNA like restriction digestion was carried out essentially as described in Sambrook *et al.*, 1989. All the preparative digestions for the preparation of inserts and vectors were generally set up in 50μl volume with following components:

1μg of DNA	Xμl
10X buffer	5μl
10X BSA	5μl
Enzyme	1μl
Water	39-Xμl

Typically the digestion reactions were carried out at 37°C overnight. Digests were resolved on 1.0% agarose gel (see section 8) and appropriate DNA fragments were cut out from the gel and eluted as mentioned in section 10.

12. Ligation

Ligation reactions were set up at vector: insert molar ratio of 1: 3 (in general) in 10μl reaction volume containing 1X ligase buffer and 1μl (10 units) T4 DNA ligase. The reactions were carried out at 16°C for 16 hours.

13. Preparation of competent cells and bacterial transformation

13.1 Competent cell preparation

DH5α competent cells were prepared by the protocol of Hanahan (1985). A single colony of DH5α was inoculated into 5ml LB medium and grown overnight at 37°C. 1ml of this culture was inoculated freshly into 100 ml of LB and grown at 37°C till an OD_{600} of 0.4-0.5 was obtained (2-3 hours). The bacterial culture was chilled on ice for two hours. The cells were harvested by centrifugation at 3,000 g for 10 min at 4°C. The pellet was re-suspended in 50ml of sterile ice-cold 100mM $CaCl_2$ solution and incubated on ice for 45 min. The cells were again centrifuged and the pellet re-suspended in 4 ml of sterile ice-cold 100mM $CaCl_2$. All these steps were carried out in a laminar flow hood under sterile conditions. The competent cells were then mixed with 80% glycerol solution to achieve a final concentration of 15%. The cell suspension (0.1 ml) was aliquoted into eppendorf tubes, frozen immediately in liquid nitrogen and stored at -80^0C.

13.2 Bacterial transformation

For transformation, the ligation mix was added to the competent cells and mixed by tapping and then incubated for 45 min on ice. This mixture was subjected to heat shock by incubating at 42°C for 90 sec, immediately chilled on ice and allowed to stand for 10 min. The cells were allowed to recover by adding 1ml of LB medium to the tube and incubated at 37°C for 1 hour with slow shaking. The transformed competent cells were then plated onto LB plate containing the required antibiotic.

13.3 Plating of transformation mix

1ml of transformation mix was centrifuged at 6000 rpm for 3 minutes and the pellet resuspended in around 100µl volume and subsequently spread on 82mm plate containing 25ml LB agar media and 50µg/ml ampicillin. In case of blue white selection, 40µl of 20mg/ml of X-gal and 4µl of 200mg/ml IPTG was spread on the plate prior to plating the transformation mix. The plates were incubated in 37°C incubator until colonies appeared (16 hr). In case of cloning into pCAMBIA vector, kanamycin (50µg/ml) was used as selectable marker.

14. Southern transfer and hybridization

The process of southern hybridization was carried out in various steps as mentioned below:

14.1 Restriction digestion of genomic DNA

1. 10µg of genomic DNA was digested in a 50µl reaction volume at standard reaction condition of each enzyme. All the reactions were incubated at 37°C for 16 hours.
2. 2µl reaction mixture was checked for complete digestion on 0.8% agarose gel.
3. After confirming the complete digestion, rest of the digestion reaction mixture was loaded on 15cm x 15cm 0.8% agarose gel containing 0.5 µg/ml ethidium bromide along with uncut chickpea DNA and DNA molecular weight marker.
4. The gel was run at 80 V constant voltage, until the dye reached 2/3rd of the gel.

14.2 Transfer of target DNA onto membrane

The gel was rinsed with MQ water after the run.
1. Depurination of the DNA within the gel was done by immersing the gel in 200ml of 0.25N HCl for 15-30 min in a baking dish with mild shaking at 50 rpm on a gyratory platform shaker till bromophenol blue changed to yellow colour. This was followed by rinsing with water.
2. Denaturation of DNA was performed by immersing the gel in 200ml of denaturing solution (1.5M NaCl and 0.5M NaOH) and incubated for 30 min under similar conditions followed by washing with sterile water.

3. The gel was neutralized by immersing in 200ml of 1.5M NaCl and 0.5M Tris-HCl (pH 8.0) for 30 min.
4. It was briefly equilibrated in transfer buffer (10X SSC).
5. Genescreenplus (Perkin Elmer) hybridization membrane and 3mm whatman paper was also pre-wet by floating on transfer buffer.
6. DNA was transferred onto the Genescreenplus membrane by the capillary blotting method for 16 h.

14.3 Fixation of target DNA onto the membrane

1. After the transfer was complete, the gel wells were marked on the membrane with a pencil.
2. The membrane was put into a fresh 10X SSC transfer buffer, gently rubbed off with damp cotton to remove any agarose gel residue.
3. The wet membrane was placed on a 3 mm Whatman filter paper and DNA was cross-linked at 1200 j/cm for 30 sec using UV cross linker (Stratagene) and air dried.

14.4 Pre-hybridization of membrane

The membrane was prehybridized in 50% deionized formamide, 10% dextran sulphate, 1% SDS, 1X SSC at 42°C for 4 h in hybridization oven with 0.035 ml of pre-hybridization solution /cm^2 of the membrane in hybridization bottle

14.5 Probe preparation

The probe was prepared using random primer labelling kit (NEB) in 50µl volume as described below:

50ng of DNA was mixed with sterile MQ water, volume made up to 36.5µl and denatured at 100°C for 5 min in a water bath and then chilled on ice for 5 minutes. Following components were then added:

Labelling buffer (10X)	5µl
dNTPs(-dCTP) (0.5mM)	2µl (each)
α^{32} dCTP (10µCi)	0.5µl
Klenow DNA polymerase enzyme (5U/µl)	2 µl

The reaction mix and incubated at 37°C for 2 h.

14.6 Probe purification

The radiolabelled probe was purified using Sephadex G-50 column. The base of a 1ml syringe was plugged with autoclaved glass wool using a plunger. The syringe was placed in a 15ml polypropylene falcon tube containing a de-capped eppendorf tube at its bottom. The syringe was filled with Sephadex G-50 beads till the column volume was around 0.9ml and then centrifuged at 3,000 rpm for 2 min. The packed column was washed with 50μl of TE buffer (pH 8.0) and centrifuged at 3,000 rpm for 2 min. This equilibration procedure was repeated until the volume of TE eluted was equal to the volume of TE loaded (50μl). Radiolabelled DNA was loaded onto the spin column and centrifuged at 3,000 rpm for 2 min. The radiolabelled DNA probe was eluted out from spun column whereas the free radioactive label was retained in the column.

14.7 Hybridization procedure

1. A total count of the purified probe was determined using Geiger muller counter.
2. Sheared salmon sperm DNA (200μg/ml) was added to the purified probe and the mixture was denatured by incubating at 100°C for 5 min in a water bath followed by chilling on ice for another 5 min.
3. This denatured probe was then added to the pre-hybridization solution.
4. Hybridization was carried out for 18 h at 42°C in a hybridization oven.

14.8 Post hybridization washing and autoradiography

1. The blot was washed with 2X SSC for 10 min at RT.
2. High stringency washing was done twice using 0.2X SSC and 0.1% SDS at 42°C for 10 min each.
3. After washing, the membrane was briefly blotted on dry filter paper and sealed in a plastic bag.
4. The membrane was exposed to Kodak X-ray films (GE Biosciences) within intensifying screens at -80°C and autoradiographed.

15. Northern blotting

Northern-blot analysis was performed to determine the expression pattern of genes in response to *Fusarium* wilt and their tissue specific expression. For this, total RNA was extracted from root tissue of susceptible and resistant cultivars of chickpea harvested at 1h, 6h, 12h, 24h, 48h and 5d after *Fusarium* inoculation. For tissue

specific expression RNA was isolated from root, stem and leaf tissue of resistant cultivar of chickpea infected with *Fusarium*. The tissue samples were collected after 48h of infection. RNA was isolated using TRIzol reagent as mentioned in section1.

Formaldehyde denatured RNA gel electrophoresis was performed as described by Sambrook *et al.*, 1989. 50ml of 1.5% denatured agarose gel was prepared by boiling 0.75g of agarose in 36ml of DEPC treated water and 5ml of 10X MOPS. After cooling, 9ml of formaldehyde was added (final concentration 2.2M) and the gel was allowed to solidify. For sample preparation (30μl volume), 20μg of total RNA (volume made upto 8μl in DEPC treated water) was mixed with 3μl of 10X MOPS, 10μl of formamide and 6μl of formaldehyde; incubated at 65°C for 10 min and then chilled on ice for 5 min. 2μl of bromophenol blue dye and 1μl of ethidium bromide (final concentration 0.5 μg/ml) was added to the sample mixture and proceeded as follows:

1. RNA was loaded onto the 1.5% (w/v) agarose gel prepared as mentioned above.
2. The gel was run at 50V until the dye migrated to three-fourth the length of gel.
3. After electrophoresis, the gel was rinsed thrice in DEPC treated water.
4. After rinsing, the gel was neutralized in 2X SSC.
5. RNA was transferred to GeneScreen plus membrane (Amersham) using 20X SSC by capillary transfer for 16 h.
6. After the transfer, the gel wells were marked on the membrane with a pencil. The membrane was put into a fresh 2X SSC transfer buffer, gently rubbed off with damp cotton to remove any agarose gel residue.
7. The wet membrane was placed on a 3mm Whatman filter paper and RNA was fixed by UV cross linking at 1200 j/cm for 30 sec using UV cross linker (Stratagene) and air dried.

Probe preparation, nucleic acid hybridization, post hybridization washing and autoradiography were done as mentioned in section 14.

16. Quantitative real time PCR

The expression of genes in response to various hormonal treatments was studied by performing quantitative real time PCR. For this, two week old chickpea seedlings were sprayed with various hormones. The hormones used and their respective concentration is given below:

Salicyclic acid	1mM
Methyl jasmonate	100µM
ACC (precursor of ethylene)	0.5mM
Sodium nitrosoprusside SNP (releases NO)	1mM
Brassinolid	50µM
ABA	100µM
2,4D	50µM

The tissue samples were collected after 1h, 3h, 6h and 12h post treatment. For control, seedlings were sprayed with water. Total RNA was isolated from all the tissue samples and quality of RNA was checked by agarose gel electrophoresis as mentioned in section 1 and 8 of appendix III respectively. The RNA samples were quantified using a NanoDrop Spectrophotometer (Nanodrop Technologies) and equal concentration of RNA (2µg) isolated from various tissue samples was taken for cDNA preparation using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) as mentioned in section 3.2.8 (chapter 3). The cDNA was diluted 10 times and qRT-PCR was performed in triplicates in ABI 7500 sequence detection system as mentioned in section 3.2.8. The relative transcript level of each gene was determined by comparing its level with that of 18S RNA. The sequences of the primers used for the real time are given in table 1 of appendix II. The relative quantification method (ΔΔ-CT) was used to evaluate quantitative variation between the replicates examined.

17. Subcellular localization
17.1 Plant material

The inner epidermal peels (2 × 2 cm) of *Allium cepa* cv. bulbs were placed on agar plates containing Murashige and Skoog (MS) salts, 0.3% sucrose and 2% agar-Type A (Sigma), pH 5.7. The peels were bombarded within 1h of transfer to agar plates.

17.2 GFP construct

GFP construct optimized for bright fluorescence in plant cells (mGFP5 in pCAMBIA 1302, with the 35S cauliflower mosaic virus promoter and nos terminator), was used to localize the GFP signal to intracellular compartment. In order to prepare the fusion construct of gene of interest (GOI) with GFP, the coding region of GOI was amplified

by PCR using gene specific primers where, in the reverse primer, the natural termination codon of the gene was deleted. For directional cloning, *Spe*I site was introduced into the reverse primer for all the genes whereas; *Nco*I and *Bgl*II sites were introduced into the forward primer before the start codon in case of *CaSAUR1*, *Ca14-3-3-1* and *CabHLH1* respectively. In case of *CaSAUR1*, the PCR amplified product was ligated in pGEM-T to yield p*CaSAUR1a* before proceeding for the preparation of fusion construct. In order to make the GOI-GFP fusion product, either the PCR product (in case of *CabHLH1* and *Ca14-3-3-1*) or p*CaSAUR1a* were digested by the respective restriction enzymes in a 50µl volume as follows:

p*CaSAUR1a*/PCR amplicon of the gene	Xµl (1µg)
10X NEB buffer	5.0µl
RE [*Nco*I (10U/µl), *Bgl*II (10U/µl)]	0.5µl
RE *Spe*I (10U/µl)	0.5µl
10X BSA	5.0µl
MQ	39.0-X µl

The pCAMBIA 1302 was also double digested with the respective restriction enzymes as per the above reactions. All the reactions were incubated at 37°C for 16 h. The digested products were run on a 0.8% agarose gel and the gene fragment as well as the digested and linearized vector was purified from the gel using gel extraction kit (Qiagen). The purified vector and the insert were subsequently ligated in 1: 3 molar ratio and transformed in DH5α competent cells. The positive colonies were confirmed by colony PCR followed by restriction digestion and finally by sequencing.

Bibliography

References

Adams M D, Kerlavage A R, Fleischmann R D, Fuldner R A, Bult C J (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature* **377**: 3-174.

Adie B A T, Perez-Perez J (2007) ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in Arabidopsis. *Plant Cell* **19**: 1665-1681.

Agizzio A P, Carvalho A O, Ribeiro Sde F, Machado O L, Alves E W, Okorokov L A, Samarão S S, Bloch C Jr, Prates M V, Gomes V M (2003) A 2S albumin-homologous protein from passion fruit seeds inhibits the fungal growth and acidification of the medium by *Fusarium oxysporum. Arch. Biochem. Biophys.* **416**: 188-195.

Agrios G N (1988) Plant Pathology, 3rd. ed. *Academic Press, Inc.: New York.* 803.

Aharoni A, Vorst O (2001) DNA microarrays for functional plant genomics *Plant Molecular Biology* **48**: 99-118.

Aitken A, Howell S, Jones D, Madrazo J, Patel Y (1995) 14-3-3 alpha and delta are the phosphorylated forms of raf-activating 14-3-3 beta and zeta. In vivo stoichiometric phosphorylation in brain at a Ser-Pro-Glu- Lys MOTIF. *J. Biol. Chem.* **270**: 5706-5709.

Albert M, Werner M, Proksch P, Fry S C, Kaldenhoff R (2004) The cell wall-modifying xyloglucan endotransglycosylase/hydrolase LeXTH1 is expressed during the defence reaction of tomato against the plant parasite *Cuscuta reflexa. Plant Biol. (Stuttg).* **6**: 402-407.

Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P (2001) Molecular biology of the cell, 4th edition.

Andrews J, Bouffard G G, Cheadle C, Lü J, Becker K G, Oliver B (2000) Gene discovery using computational and microarray analysis of transcription in the Drosophila melanogaster testis. *Genome Res,* **10**: 2030-2043.

Arabidopsis Genome Initiative data. (2000) *Nature* **408**: 79685.

Arnholdt-Schmitt B, Costa J H, de Melo D F (2006) AOX-a functional marker for efficient cell reprogramming under stress? *Trends Plant Sci.* **11**: 281-287.

Asamizu E, Nakamura Y, Sato S, Tabata S (2000) Generation of 7137 non-redundant expressed sequence tags from a legume, Lotus japonicus. *DNA Res.* **7**: 127-130.

Asselbergh B, De Vleesschauwer D, Höfte M (2008) Global switches and fine-tuning - ABA modulates plant pathogen defense. *Mol. Plant Microbe Interact.* **21**: 709-719.

Ashraf N, Ghai D, Barman P, Nagaraju G, Mandal M K, Chakraborty N, Datta A, Chakraborty S Comparative analyses of genotype dependent expressed sequence tags and stress-responsive transcriptome of chickpea illustrates predicted and unexpected genes and novel regulators of plant immunity. *BMC Genomics*. In press

Atchley W R, Fitch W M (1997) A natural classification of the basic helixloop-helix class of transcription factors. *Proc. Natl. Acad. Sci.* USA **94**: 5172-5176.

Atchley W R, Terhalle W, Dress A (1999) Positional dependence, cliques, and predictive motifs in the bHLH protein domain. *J. Mol. Evol.* **48**: 501-516.

Ausubel F M (2005) Are innate immune signaling pathways in plants and animals conserved? *Nat. Immunol.* **6**: 973–979.

Bailey P C, Martin C, Toledo-Ortiz G, Quail P H, Huq E, Heim M A, Jakoby M, Werber M, Weisshaar B (2003) Update on the basic helix-loop-helix transcription factor gene family in Arabidopsis thaliana. *Plant Cell* **15**: 2497-2502.

Barbazuk W B, Emrich S J, Chen H D, Li L, Schnable P S. (2007) SNP discovery via 454 transcriptome sequencing. *Plant J.* 51: 910–918.

Bartel D P (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**: 281–297.

Barth C, Moeder W, Klessig D F, Conklin P L (2004) The timing of senescence and response to pathogens is altered in the ascorbate-deficient Arabidopsis mutant vitamin c-1. *Plant Physiol.* **134**: 1784-1792.

Belkhadir Y, Chory J (2006) Brassinosteroid signaling: a paradigm for steroid hormone signaling from the cell surface. *Science* **314**: 1410-1411.

Bent A F (1996) Plant disease resistance genes: function meets structure. *Plant Cell* **8**: 1757-1771.

Bent AF, Mackey D (2007) Elicitors, Effectors, and *R* Genes: The New Paradigm and a Lifetime Supply of Questions. *Annu. Rev. Phytopathol.* **45**: 399-436.

Berrocal-Lobo M, Molina A (2007) Arabidopsis defense response against *Fusarium oxysporum*. *Trends Plant Sci.* **13**: 145-150.

Beyer A, Bandyopadhyay S, Ideke*r* T (2007) Integrating physical and genetic maps: from genomes to interaction networks. *Nat. Rev. Gen.* **8**: 699-710.

Bhushan D, Pandey A, Choudhary M K, Datta A, Chakraborty S, Chakraborty N (2007) Comparative proteomic analysis of differentially expressed proteins in chickpea extracellular matrix during dehydration stress. *Mol. Cell. Proteomics* **6**: 1868-1884.

Birnbaum K, Shasha D E, Wang J Y, Jung J W, Lambert G M. (2003) A gene expression map of the Arabidopsis root. *Science* **302**: 1956-60.

Boller T, Felix G (2009) A Renaissance of Elicitors: Perception of Microbe-Associated Molecular Patterns and Danger Signals by Pattern-Recognition Receptors Annu. *Rev. Plant Biol.* **60**: 379-406.

Bostock R M (2005) Signal crosstalk and induced resistance: Straddling the Line between cost and benefit. *Annu. Rev. Phytopathol.* **43**: 545-580.

Boter M, Ruı′z-Rivero O, Abdeen A, Pratl S (2004) Conserved MYC transcription factorsplay a key role in jasmonate signalling both in tomato and *Arabidopsis. Genes Dev.* **18**: 1577-1591.

Brinda S, Ravikumar R L (2005) Inheritance of wilt resistance in chickpea - A molecular marker analysis. *Curr. Sci.* **88**: 701-702.

Brodersen P, Petersen M, Nielsen H B, Zhu S, Newman M, Shokat K M, Rietz S, Parker J, Mundy J (2006) *Arabidopsis* MAP kinase 4 regulates salicylic acid- and jasmonic acid/ethylene-dependent responses via EDS1 and PAD4. *Plant J.* **47**: 532-546.

Brown D W, Butchko R A E, Proctor R H (2006) Fusarium genomic resources: Tools to limit crop diseases and mycotoxin contamination. *Mycopathologia* **162**: 191-199.

Brown P O, Botstein D (1999) Exploring the new world of the genome with DNA microarrays. *Nat. Genet.* **21:** 33-37.

Buchanan B B, Gruissem W, Jones R (2000) Biochemistry and molecular biology of plants. *American Society of Plant Physiologists.*

Bucher P (1999) Regulatory elements and expression profiles. *Curr.Opin. Struct. Biol.* **9**: 400-407.

Camoni L, Harper J F, Palmgren M G (1998) 14-3-3 proteins activate a plant calcium-dependent protein kinase (CDPK). *FEBS Lett.* **430**: 381-384.

Cantero A, Barthakur S, Bushart T J, Chou S, Morgan R O, Fernandez M P, Clark G B, Roux S J (2006) Expression profiling of the Arabidopsis annexin gene family during germination, de-etiolation and abiotic stress: *Plant Physiol. Biochem.* **44**: 13-24.

Caplan J, Padmanabhan M, Dinesh-Kumar S P (2008) Plant NB-LRR Immune Receptors: From Recognition to Transcriptional Reprogramming. *Cell Host Microbe* **3**: 126-135

Carr J P, Marsh L E, Lomonossoff G P, Sekiya M E, Zaitlin M (1992) Resistance to tobacco mosaic virus induced by the 54-kDa gene sequence requires expression of the 54-kDa protein. *Mol. Plant Microbe Interact.* **5**: 397-404.

Chen B S, Yang S K, Lan C Y, Chuang Y J (2008) A systems biology approach to construct the gene regulatory network of systemic inflammation via microarray and databases mining. *BMC Med. Genomics* **1**: 46.

Chen F, Li Q, Sun L, He Z (2006) The Rice 14-3-3 Gene Family and its Involvement in Responses to Biotic and Abiotic Stress. *DNA Res.* **13**: 53-63.

Chen W J, Zhu T (2004) Networks of transcription factors with roles in environmental stress response. *Trends Plant Sci.* **9**: 591-596.

Cheung F, Haas B J, Goldberg S M, May G D, Xiao Y, To'wn C D (2006) Sequencing Medicago truncatula expressed sequenced tags using 454 Life Sciences technology. *BMC Genomics* **7**: 272.

Cheung V G, Spielman R S (2002) The genetics of variation in gene expression. *Nat. Genet.* **32**: 522-525.

Chinchilla D, Zipfel C, Robatzek S, Kemmerling B, Nurnberger T, Jones J D G, Felix G, Boller (2007) A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defence. *Nature* **448**: 497-500.

Chinchilla D, Bauer Z, Regenass M, Boller T, Felix G (2006) The Arabidopsis receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *Plant Cell* **18**: 465-476.

Choi H K, Mun J H, Kim D J, Zhu H, Baek J M, Mudge J, Roe B, Ellis N, Doyle J, Kiss G B, Young N D, Cook D R (2004) Estimating genome conservation between crop and model legume species. *Proc. Natl. Acad. Sci. U S A* **43**: 15289-15294.

Chu L H, Chen B S (2008) Construction of a cancer-perturbed protein-protein interaction network for discovery of apoptosis drug targets. *BMC Syst. Biol.* **2**: 56.

Chung K M, Igari K, Uchida N, Tasaka M (2008) New Perspectives on Plant Defense Responses through Modulation of Developmental Pathways. *Mol. Cells* **26**: 107-112.

Collins N C, Thordal-Christensen H, Lipka V, Bau S, Kombrink E, Qiu J L, Hückelhoven R, Stein M, Freialdenhoven A, Somerville S C, Schulze-Lefert P (2003) SNARE-protein-mediated disease resistance at the plant cell wall. *Nature* **425**: 973-977.

Coram T E, Pang E C K (2005) Isolation and analysis of candidate ascochyta blight defence genes in chickpea. Part II. Microarray expression analysis of putative defense related ESTs. *Plysiol. Mol. Plant Pathol.* **66**: 201-210.

Croser J S, Ahmad F, Clarke H J, Siddique K H M (2003) Utilizatea of wild cicer in chickpea improvement-progress, constraints and prospects. *Acot. J. Agric. Res.* **54**: 429-444.

Cushman J C, Meyer G, Michalowski C B, Schmitt J M, Bohnert H J (1989) Salt stress leads to differential expression of two isogenes of phosphoenolpyruvate carboxylase during CAM induction in the common ice plant. *Plant Cell.* **1**: 715-725.

Dahan J, Etienne P, Petitot A S, Houot V, Blein J P, Suty L (2001) Cryptogein affects expression of alpha3, alpha6 and beta1 20S proteasome subunits encoding genes in tobacco. *J. Exp. Bot.* **52**: 1947-1948.

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

Davies P J (1995) Plant hormones. *Dordrecht: Kluwer Academic Publishers.*

Davletova S, Rizhsky L, Liang H, Shenggiang Z, Oliver D J, Coutu J, Shulaev V, Schlauch K, Mittler R (2005) Cytolosic ascorbate peroxidase 1 is a central component of thereactive oxygen gene network of Arabidopsis. *Plant Cell* **17**: 268-281.

De Lorenzo G, Ferrari S (2002) Polygalacturonase-inhibiting proteins in defense against phytopathogenic fungi. *Curr. Opin. Plant Biol.* **5**: 295-299.

De Lucca A J (2007) Harmful fungi in both Agriculture and Medicine. *Rev. Micol.* **24**: 3-13.

De Smet I, Signora L, BeeckmanT, Inze' D, Foyer C H, Zhang H (2003). An abscisic acid-sensitive checkpoint in lateral root development of Arabidopsis. *Plant J.* **33**: 543-555.

Deavours B E, Dixon R A (2005) Metabolic engineering of isoflavonoid biosynthesis in alfalfa (*Medicago sativa* L.). *Plant Physiol.* **138**: 2245-2259.

Delahaye N F, Colte N, Puthier D, Barbier M, Benech P, Joly F, Iraqi F A, Grau G E, Nguyen C, Rihet P (2007) Gene expression analysis reveals early changes in several molecular pathways in cerebral malaria-susceptible mice versus cerebral malaria-resistant mice. *BMC Genomics* **8**: 452.

Delessert C, Kazan K, Wilson I W, Van Der S D, Manners J, Dennis E S, Dolferus R (2005) The transcription factor ATAF2 represses the expression of pathogenesis-related genes in Arabidopsis. *Plant J.* **43**: 745-757.

DeRisi J L, Iyer V R, Brown P O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**:680-686.

Desikan R, A-H-Mackerness S, Hancock, J T (2001) Regulation of the Arabidopsis transcriptome by oxidative stress. *Plant Physiol.***127**: 159-172.

Deslandes L, Olivier J, Peeters N, Feng DX, Khounlotham M, Boucher, Somssich I, Genin S, Marco Y (2003) Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc. Natl. Acad. Sci.* USA **100**: 8024-8029.

Devoto A, Muskettz P R, Shirasu K (2003) Role of ubiquitination in the regulation of plant defence against pathogens. *Curr. Opin.Plant Biol.* **6**: 307-311.

Di Matteo A, Federici L, Mattei B, Salvi G, Johnson K A, Sauino C, De Lorenzo G, Tsernoglou D, Cervone F (2003) The crystal structure of polygalacturonose-inhibiting protein (PGIP), a leucine-rich repeat protein involved in plant defense. *Proc. Natl. Acad. Sci. USA* **100**: 10124-10128.

Diévart A, Clark S E (2004) LRR-containing receptors regulating plant development and defense. *Development* **131**: 251-261.

Dittrich M T, Klau G Y, Rosenwald A, Dandekar T, Müller T (2008) Identifying functional modules in protein–protein interaction networks: an integrated exact approach. *Bioinformatics* **24**: 223-231.

Dixon R A, Achnine L, Kota P, Liu C J, Reddy M S S, Wang L (2002) The phenylpropanoid pathway and plant defense – a genomics perspective. *Mol. Plant Pathol.* **3**: 371-390.

Dong X (2004) NPR1, all things considered. *Curr. Opin. Plant Biol.* 7: 547-52

Dos Remedios C G, Chhabra D, Kekic M, Dedova I V, Tsubakihara M, Berry D A, Nosworthy N J (2003) Actin Binding Proteins: Regulation of Cytoskeletal Microfilaments. *Physiol. Rev.* **83**: 433-473.

Dougherty M K, Morrison D K (2004) Unlocking the code of 14-3-3. *J. Cell Sci.* **117**: 1875-1884.

Dowd C, Wilson L W, McFadden H (2004) Gene expression profile changes in cotton root and hypocotyl tissues in response to infection with *Fusarium oxysporum f. sp. vasinfectum. Mol. Plant Microbe Interact.* **17**: 654-667.

Duek P D, Fankhauser C (2005) bHLH class transcription factors take centre stage in phytochrome signalling. *Trends Plant Sci.* **10**: 2.

Duke J A (1981) Handbook of legumes of world economic importance. pp. 52-57, Plenum Press, New York.

Ellis J G, Dodds P N, Lawrence G J (2007) Flax rust resistance gene specificity is based on direct resistance-avirulence protein interactions. *Annu. Rev. Phytopathol.* **45**: 289-306.

Ellis J R, Burke J M (2007) EST-SSRs as a resource for population genetic analyses. *Heredity* **99**: 125-132.

Ellis J, Dodds P, Pryor T (2000) Structure, function and evolution of plant disease resistance genes. *Curr. Opin. Plant Biol.* **3**: 278-284.

Eulgem T (2005) Regulation of the Arabidopsis defense transcriptome. *Trends Plant Sci.* **10**: 2.

Eulgem T (2006) Dissecting the WRKY web of plant defense regulators. *PLoS Pathog.* **2**: 126.

Eulgem T, Rushton P J, Robatzek S, Somssich I E (2000) The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* **5**: 199-206.

Ewing B, Hillier L, Wendl M C, Green P (1998) Base-calling of automatedsequencer traces using Phred. I. Accuracy assessment. *Genome Res.* **8**: 175-185.

Ewing R M, Kahla A B, Poirot O, Lopez F, Audic S, Claverie J M (1999) Large-scale statistical analyses of rice ESTs reveal correlated patterns of gene expression. *Genome Res.* **9**: 950-959.

Fait A, Angelovici R, Less H, Ohad I, Urbanczyk-Wochniak E, Fernie A R, Galili G (2006) Arabidopsis seed development and germination is associated with temporally distinct metabolic switches. *Plant Physiol.* **142**: 839-854.

Fan J B, Chen X, Halushka M K, Berno A, Huang X, Ryder T, Lipshutz R J, Lockhart D J, Chakravarti A (2000) Parallel genotyping of human SNPs using generic high-density oligonucleotide tag arrays. *Genome Res.* **10**: 853-860.

FAOSTAT data (2004) http://faostat.fao.org .

Fehrmann R S N, Li X, van der Zee A G J, de Jong S, Meerman G J, de Vries E G E, Crijns A P G (2007) Profiling Studies in Ovarian Cancer: A Review. *The Oncologist* **12**: 960-966.

Felix G, Boller T (2003) Molecular sensing of bacteria in plants. The highly conserved RNA-binding motif RNP-1 of bacterial cold shock proteins is recognized as an elicitor signal in tobacco. *J. Biol. Chem.* **278**: 6201-6208.

Ferl R J, Manak M S, Reyes M F (2002) The14-3-3s *Genome Biol.* **3**: 3010.1-3010.7.

Folkes L K, Wardman P (2001) Oxidative activation of indole-3-acetic acids to cytotoxic species - a potential new role for plant auxins in cancer therapy. *Biochem. Pharmacol.* **61**:129-136.

Friedrichsen D M, Nemhauser J, Muramitsu T, Maloof J N, Alonso J, Ecker J R, Furuya M, Chory J (2002) Three Redundant Brassinosteroid Early Response Genes Encode Putative bHLHT ranscription Factors Required for Normal Growth. *Genetics* **162**: 1445-1456.

Fujiwara M, Umemura K, Kawasaki T, Shimamoto K (2006) Proteomics of Rac GTPase Signaling Reveals Its Predominant Role in Elicitor-Induced Defense Response of Cultured Rice Cells. *Plant Physiol.* **140**: 734-745

Fukamatsu Y, Naoto Y, Hasunuma K (2003) Arabidopsis NDK1 is a component of ROS signaling by interacting with three catalses. *Plant Cell Physiol.* **44**: 982-989.

Futamura N, Ishiiminami N, Hayashida N, Shinohara K (1999) Expression of DnaJ homologs and Hsp70 in the Japanese willow (Salix gilgiana Seemen). *Plant Cell Physiol.* **40**: 524-31.

Gampala S S, Kim T W, He J X, Tang W, Deng Z, Bai M, Guan S, Lalonde S, Sun Y, Gendron J M, Chen H, Shibagaki N, Ferl RJ, Ehrhardt D, Chong K, Burlingame A L, and Wang Z Y (2007) An Essential Role for 14-3-3 Proteins in Brassinosteroid Signal Transduction in Arabidopsis. *Dev. Cell* **13**: 177-189.

Gee M A, Hagen G, Guilfoyle T J (1991) Tissue-specific and organ-specific expression of soybean auxin-responsive transcripts GH3 and SAURs. *Plant Cell* **3**: 419-430.

Ghai D (2003) In Ph.D thesis: Search for pathogenesis related genes in chickpea. Jawahar Lal Nehru university, New Delhi-110067, India.

Gil P, Liu Y, Orbovic V, Verkamp E, Poff K L, Green P J (1994) Characterization of the auxin-inducible SAUR-AC1 gene for use as a molecular genetic tool in *Arabidopsis. Plant Physiol.* **104**: 777-784.

Glazebrook J (2005) Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. *Annu. Rev. Phytopathol.* **43**: 205-227.

Glazebrook J, Chen W, Estes B, Chang H S, Nawrath C (2003) Topology of the network integrating salicylate and jasmonate signal transduction derived from global expression phenotyping. *Plant J.* **34**: 217-228.

Goda H, Sawa S, Asami T, Fujioka S, Shimada Y, Yoshida S (2004) Comprehensive comparison of auxin-regulated and brassinosteroid- regulated genes in Arabidopsis. *Plant Physiol.* **134**: 1555-1573.

Gomez-Gomez L, Boller T (2002) Flagellin perception: a paradigm for innate immunity. *Trends Plant Sci.* **7:** 251-256.

Gong Z, Dong C H, Lee H, Zhu J, Xiong L, Gong D, Stevenson B, Zhu J K (2005) A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in Arabidopsis. *Plant Cell* **17**: 256-67.

Gordon T R, Martyn R D (1997) The evolutionary biology of *Fusarium oxysporum. Annu. Rev. Phytopathol.* 35:111–128.

Gowri G, Bugos R C, Campbell W H, Maxwell C A, Dixon R A (1991) Stress responses in Alfalfa (Medicago sativa L): X-molecular cloning and expression of S-adenosyl-I-methionine; Caffeic acid 3-0-methyltransferase, a key enzyme of lignin biosynthesis. *Plant Physiol.* **97**: 7-14.

Graham P H, Vance C P (2003) Legumes: importance and constraints to greater use. *Plant Physiol.* **131**: 872-877.

Grant J J, Loake G J (2000) Role of reactive oxygen intermediates and cognate redox signaling in disease resistance. *Plant Physiol.* **124**: 21-29.

Grant M R, Jones J D G (2009) Hormone (Dis) harmony Moulds Plant Health and Disease. *Science* **324**: 750-752.

Gray W M, Muskett P R, Chuang H W, Parker J E (2003) The Arabidopsis SGT1b protein is required for SCFTIR1-mediated auxin response. *Plant Cell* **15**:1310-1319.

Groszmann M, Paicu T, Smyth D R (2008) Functional domains of SPATULA, a bHLH transcription factor involved in carpel and fruit development in Arabidopsis *Plant J.* **55**: 40-52.

Grusak M A (2002) Enhancing mineral content in plant food products. *J. Am. Coll. Nutr.* **21**: 178-183.

Gu Y_Q, Wildermuth M_C, Chakravarthy S, Loh Y_T, Yang C_M, He X_H, Han Y, Martin G_B (2002) Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in Arabidopsis. *Plant Cell* **14**: 817-831.

Gupta, P K, Rustgi, S (2004) Molecular markers derived from expressed/transcribed portion of the genome in higher plants. *Funct. Integr. Genomics* **4**: 139-162.

Hagen G, Guilfoyle T J (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Mol. Biol.* **49**: 373-385.

Hammond-Kosack K E, Jones J D G (1997) Plant disease resistance genes. *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.* **48**: 575-607.

Harlizius B, van Wijk R, Merks J W M (2004) Genomics for food safety and sustainable animal production. *J. Biotechnol.* **113**: 33-42.

Harmer S L, Hogenesch L B, Straume M, Chang H S, Han B, Zhu T, Wang X, Kreps J A, Kay S A (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science* **290**: 2110–2113.

Haware, M P (1990) Fusarium wilt and other important diseases of chickpea in the Mediterranean area. *Options Méditerr. Sér. Sémin.* **9**: 61-64.

Heim M A, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey P C (2003) The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Mol. Biol. Evol.* **20**: 735-747.

Henriksson M L, Francis M S, Peden A, Aili M, Stefansson K, Palmer R, Aitken A Hallberg B (2002) A nonphosphorylated 14-3-3 binding motif on exoenzyme S that is functional in vivo. *Eur. J. Biochem.* **269**: 4921-4929.

Hocquette J F (2005) Where are we in genomics? *J. Physiol. Pharmacol.* **56**: 37-70.

Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biol.* **5**: 85.

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. *Genome Res.* **9**: 868-877.

Hughes T R, Marton M J, Jones A R, Roberts C J, Stoughton R, Armour C D, Bennett H A, Coffey E, Dai H, He Y D (2000) Functional discovery via a compendium of expression profiles. *Cell* **102**: 109-126.

Iyer V R, Horak C E, Scafe C S, Botstein D, Synder M, Brown O P (2001) Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**: 533-538.

Jain M, Tyagi A K, Khurana J P (2006) Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*). *Genomics* **88**: 360-371.

Jakoby M, Weisshaar B, Droge-Laser W, Vicente-Carbajosa J, Tiedemann J, Kroj T, Parcy F (2002) bZIP transcription factors in *Arabidopsis*. *Trends Plant Sci.* **7**: 106-111.

Jang J Y, Rhee J Y, Kim D G, Chung G C, Lee J H, Kang H. (2007) Ectopic expression of a foreign aquaporin disrupts the natural expression patterns of endogenous aquaporin genes and alters plant responses to different stress conditions. *Plant Cell Physiol.* **48(9)**:1331-9.

Jantasuriyarat C, Gowda M, Haller K, Hatfield J, Lu G, Stahlberg E, Zhou B, Li H, Kim H, Yu Y, Dean R A, Wing R A, Soderlund C, Wang G L (2005) Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. *Plant Physiol.* **138**: 105-115.

Jiang Y, Deyholos M K (2006) Comprehensive transcriptional profiling of NaCl-stressed Arabidopsis roots reveals novel classes of responsive genes. *BMC Plant Biol.* **6**: 25.

Jiménez-Díaz R M, Alcalá-Jiménez A R, Hervás A, Trapero-Casas J L (1993) Pathogenic variability and host resistance in the *Fusarium oxysporum* f. sp. *ciceris/Cicer arietinum* pathosystem. In: *Fusarium* Mycotoxins, Taxonomy, Pathogenicity, Host Resistance. *Proc. Eur. Seminar, 3rd ed. E. Arseniuk and T. Goral, eds. Plant Breed. Acclim. Inst., Radzikóv, Poland.* 87-94.

Jin S, Cheng Y, Guan Q, Liu D, Takano T, Liu S (2006) A metallothionein-like protein of rice (rgMT) functions in E. coli and its gene expression is induced by abiotic stresses. *Biotechnol. Lett.* **28**: 1749-53.

Johansen C, Baldev B, Brouwer JB, Erskine W, Jermyn WA, Li-Juan L, Malik BA, Ahad Miah A, Silim S N (1994) Biotic and abiotic stresses constraining productivity of cool season food legumes in Asia, Africa and Oceania. *Dordrecht, The Netherlands: Kluwer Academic Publishers* 75-194.

Jones J D, Dangl J L. (2006) The plant immune system. *Nature* **444**: 323-329.

Jones S (2004) An overview of the basic helix-loop-helix proteins. *Genome Biol.* **5**: 226.

Kader J C (1996) Lipid-transfer proteins in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **47**: 627-654.

Kagan-Zur V, Tieman D M, Marlow S J, Handa A K (1995) Differential regulation of polygalacturonase and pectin methylesterase gene expression during and after heat stress in ripening tomato (*Lycopersicon esculentum* Mill.) fruits. *Plant Mol Biol.* **29**: 1101-1110.

Kaku H, Nishizawa Y, Ishii-Minami N, Akimoto-Tomiyama C, Dohmae N, Takio K, Minami E, Shibuya N (2006) Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci.* **103**:11086-11091.

Kamoun S (2006). A catalogue of the effector secretome of plant pathogenic oomycetes.*Annu. Rev. Phytopathol.* **44**: 41-60.

Karlova R, Boeren S, Russinova E, Aker J, Vervoort J, De Vries S (2006) The Arabidopsis somatic embryogenesis receptor-like kinase 1 protein complex includes brassinosteroid-insensitive 1. *Plant Cell* **3**: 626-638.

Katagiri F (2004) A global view of defense gene expression regulation - a highly interconnected signaling network *Curr. Opin. Plant Biol.* **7**: 506-511.

Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A Jr, Zhu J K, Staskawicz B J, Jin H L (2006) A pathogen-inducible endogenous siRNA in plant immunity. *Proc. Natl. Acad. Sci.* USA **103**: 18002-18007.

Kazan K, Schenk P M, Wilson I, Manners J M (2001) DNA microarrays: new tools in the analysis of plant defence responses. *Mol. Plant Pathol.* **2**: 177-185.

Khandelwal A, Elvitigala T, Ghosh B, Quatrano R (2008) Arabidopsis transcriptome reveals control circuits regulating redox homeostasis and the role of an AP2 transcription factor. *Plant Physiology* **148**: 2050-2058.

Kim S H, Hong J K, Lee S C, Sohn K H, Jung H W, Hwang B K (2004) CAZFP1, Cys2/His2-type zinc-finger transcription factor gene functions as a pathogen-induced early-defense gene in Capsicum annuum. *Plant Mol. Biol.* **55**: 883-904.

Kim S K, Chang S C, Lee E J,Chung W S, Kim Y S, Hwang S, Lee J S. (2000). Involvement of brassinosteroids in the gravitropic response of primary root of maize. *Plant Physiol.* **123**:997–1004.

Kitagawa K, Skowyra D, Elledge S J, Harper J W, Hieter P (1999) SGT1 encodes an essential component of the yeast kinetochore assembly pathway and a novel subunit of the SCF ubiquitin ligase complex. *Mol Cell* **4**: 21-33.

Knauss S, Rohrmeier T, Lehle L (2003) The auxin-induced maize gene ZmSAUR2 encodes a short-lived nuclear protein expressed in elongating tissues. *J. Biol. Chem.* **278**: 23936-23943.

Knudsen J, Jensen M V, Hansen J K, Faergeman N J, Neergaard T B, Gaigg B (1999) Role of acylCoA binding protein in acylCoA transport, metabolism and cell signalling. *Mol. Cell Biochem.* **192**: 95-103.

Kobae Y, Sekino T, Yoshioka H, Nakagawa T, Martinoia E, Maeshima M (2006) Loss of AtPDR8, a plasma membrane ABC transporter of Arabidopsis thaliana, causes hypersensitive cell death upon pathogen infection. *Plant Cell Physiol.* **47**: 309-18.

Koornneef A, Pieterse C M J (2008) Cross Talk in Defense Signaling. *Plant Physiol.* **146**: 839-844.

Kotchoni S O, Gachomo E W (2006) The reactive oxygen species network pathways: an essential prerequisite for perception of pathogen attack and the acquired disease resistance in plants. *J. Biosci.* **31**: 389-404.

Kuranaga E, Miura N (2007) Nonapoptotic functions of caspases: caspases as regulatory molecules for immunity and cell-fate determination. *Trends Cell Biol.* **17**: 135-144.

Kwon Y R, Lee H J, Kim K H, Hong S W, Lee S J, Lee H (2008) Ectopic expression of Expansin3 or Expansinbeta1 causes enhanced hormone and salt stress sensitivity in Arabidopsis. *Biotechnol. Lett.* **30**: 1281-1288.

Laitinen R A, Immanen J, Auvinen P, Rudd S, Alatalo E, Paulin L, Ainasoja M, Kotilainen M, Koskela S, Teeri T H, Elomaa P (2005) Analysis of the floral transcriptome uncovers new regulators of organ determination and gene families related to flower organ differentiation in Gerbera hybrida (Asteraceae). *Genome Res.* **15**: 475-86.

Laxalt A M, Cassia R O, Sanllorenti P M, Madrid E A, Andreu A B, Daleo G R, Conde R D, Lamattina L (1996) Accumulation of cytosolic glyceraldehyde-3-phosphate dehydrogenase RNA under biological stress conditions and elicitor treatments in potato. *Plant Mol. Biol.* **30**: 961-972.

Lazo G R, Chao S, Hummel D D, Edwards H, Crossman C C, Lui N, Matthews D E, Carollo V L, Hane D L, You F M (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delineated map. *Genetics* **168**: 585-93.

Ledent V, Vervoort M (2001) The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis. *Genome Res.* **11**: 754-770.

Lee M H, Lee S H, Kim H, Jin J B, Kim D H, Hwang I (2006) A WD40 repeat protein, Arabidopsis Sec13 homolog 1, may play a role in vacuolar trafficking by controlling the membrane association of AtDRP2A. *Mol. Cells* **22**: 210-219.

Lee S J, Suh M C, Kim S, Kwon J K, Kim M, Paek K H, Choi D, Kim B D (2001) Molecular cloning of a novel pathogen-inducible cDNA encoding a putative acyl CoA synthetase from capsicum annum. *Plant Mol. Biol.* **46**: 661-671.

Li H Y, Chye M L (2004) Arabidopsis Acyl-CoA-binding protein ACBP2 interacts with an ethylene-responsive element-binding protein, AtEBP, via its ankyrin repeats. *Plant Mol. Biol.* **54**: 233-43.

Li H, Sun J, Xu Y, Jiang H, Wu X, Li C (2007) The bHLH-type transcription factor AtAIB positively regulates ABA response in Arabidopsism. *Plant Mol Biol.* **65**: 655-665.

Li J, Brader G, Palva E T (2004) The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. *Plant Cell* **16**: 319-331.

Li X, Duan X, Jiang H, Sun Y, Tang Y, Yuan Z, Guo J, Liang W, Chen L, Yin J, Ma H, Wang Zhang D (2006) Genome-Wide Analysis of Basic/Helix-Loop-Helix Transcription Factor Family in Rice and Arabidopsis1. *Plant Physiology* **141**: 1167-1184.

Li Y, Strabala T J, Hagen G, Guilfoyle T J (1994) The soybean SAUR open reading frame contains a cis element responsible for cycloheximide-induced mRNA accumulation. *Plant Mol. Biol.* **24**: 715-723.

Lim G T T, Wang G P, Hemming M N, Basuki S, McGrath D J, Carroll B J Jones D A (2006) Mapping the I-3 gene for resistance to Fusarium wilt in tomato: application of an I-3 marker in tomato improvement and progress towards the cloning of I-3. *Australas. Plant Pathol.* **35**: 671-680.

Lipka V, Panstruga R (2005) Dynamic cellular responses in plant-microbe interactions. *Curr. Opin. Plant Biol.* **8**: 625-631.

Liu J H, Nada K, Honda C, Kitashiba H, Wen X P, Pang X M, Moriguchi T (2006) Polyamine biosynthesis of apple callus under salt stress: importance of the arginine decarboxylase pathway in stress response. *J. Exp. Bot.* **57**: 2589-2599.

Liu J, Wang X J. (2006) An integrative analysis of the effects of auxin on jasmonic acid biosynthesis in *Arabidopsis thaliana. J. Integr. Plant Biol.* **48**:99–103.

Ljung K, Hull A K, Celenza J, Yamada M, Estelle M (2005) Sites and regulation of auxin biosynthesis in Arabidopsis roots. *Plant Cell* **17**: 1090-1104.

Llorentea F, Muskettb P, Sa´ nchez-Valleta A, Lo´ peza G, Ramosa B, Sa´ nchez C, Gueza R, Jorda´ L, Parkerb J, Molinaa A (2008) Repression of the Auxin Response Pathway Increases Arabidopsis Susceptibility to Necrotrophic Fungi. *Mo. Plant* **1**: 496-509.

Lohar D L, Sharopova N, Endre G, Penuela S, Samac D, Town C, Silverstein K A T, Vandenbosch K A (2006) Transcript Analysis of Early Nodulation Events in Medicago truncatula. *Plant Physiol* **140**: 221-234.

Lorenzo O, Chico J M, Sanchez-Serrano J J, Solano R (2004) *JASMONATEINSENSITIVE1* encodes a MYC transcription factor essential to discriminate between different jasmonate-regulated defense responses in Arabidopsis. *Plant Cell* **16**: 1938-1950.

Lorenzo O, Piqueras R, Sánchez-Serrano J J, Solano R (2003) Ethylene Response Factor1 integrates signals from ethylene and jasmonate pathways in plant defense. *Plant Cell* **15**: 165-178.

Lu D P, Christopher D A (2008) Endoplasmic reticulum stress activates the expression of a sub-group of protein disulfide isomerase genes and AtbZIP60 modulates the response in Arabidopsis thaliana. *Mol. Genet .Genomics* (In press).

Lucas J A (1998) *Plant Pathology and Plant Pathogens*. Oxford, UK: Blackwell Sci. pp.274.

Ludwig S R, Habera L F, Dellaporta S L, Wessler S R (1989) Lc, a member of the maize R gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the mychomology region. *Proc. Natl. Acad. Sci.* USA **86**: 7092-7096.

Mackey D, Holt B F, Wiig A, Dangl J L (2002) RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for *RPM1* mediated resistance in Arabidopsis. *Cell* **108**: 743-754.

Maleck K, Levine A, Eulgem T, Morgan A, Schmid J, Lawton K A, Dangl J L, Dietrich R A (2000) The transcriptome of *Arabidopsis thaliana* during systemic acquired resistance. *Nat. Genet.* **26**: 403-410.

Mantri N L, Ford R, Coram T E, Pang E C K (2007) Transcriptional profiling of chickpea genes differentially regulated in response to high-salinity, cold and drought. *BMC Genomics* **8**: 303.

Martin G B, Bogdanove A J, Sessa G (2003) Understanding the functions of plant disease resistance proteins *Annu. Rev. Plant Biol.* **54**: 23-61.

Massari M E, Murre C (2000) Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol. Cell Biol.* **20**: 429-440.

Mathers J C (2004) What can we expect to learn from genomics? *Proc. Nutr. Soc.* **63**: 1-4.

Mauch-Mani B, Slusarenko A J (1994) Systemic acquired resistance in *Arabidopsis thaliana* induced by a predisposing infection with a pathogenic isolate of *Fusarium oxysporum. Mol. Plant Microbe Interact.* **7**: 378-383.

McClure B A, Guilfoyle T J (1987) Characterization of a class of small auxininducible soybean polyadenylated RNAs. *Plant Mol. Biol.* **9:** 611-623.

McDowell J M, Dangl J L (2000) Signal transduction in the plant immune response. *Trends Biochem. Sci.* **25**: 79-82.

Merkouropoulos G, Shirsat A H (2003) The unusual Arabidopsis extensin gene atExt1 is expressed throughout plant development and is induced by a variety of biotic and abiotic stresses: *Planta* **217**: 356-66.

Meyer A J (2007) The integration of glutathione homeostasis and redox signaling. *J Plant Physiol.* **165**: 1390-1403.

Miklis M, Consonni C, Bhat R A, Lipka V, Schulze-Lefert P, Panstruga R (2007) Barley MLO Modulates Actin-Dependent and Actin-Independent Antifungal Defense Pathways at the Cell Periphery. *Plant Physiol.* **144**: 1132-1143.

Mittler R, Vanderauwera S, Gollery M, Van Breusegem F (2004) The reactive oxygen gene network of plants. *Trends Plant Sci.* **9**: 490-498.

Molina C, Rotter B, Horres R,Udupa S M, Besser B, Bellarmino L, Baum M, Matsumura H, Terauchi R, Kahl G, Winter P (2008) SuperSAGE: the drought stress-responsive transcriptome of chickpea roots. *BMC Genomics* **9**: 553.

Morgante M, Hanafey M, Powell W (2002) Microsatellites are preferentially associated with non repetitive DNA in plant genomes. *Nat. Genet.* **30**: 94-200.

Muehlbauer, F.J., R.J. Redden, A.M. Nassib, L.D. Robertson and J.B. Smithson. 1988. Population improvement in pulse crops: an assessment of methods and techniques. p.943-966. In: R.J. Summerfield (ed.), World Crops: Cool Season Food Legumes. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Mur L A J, Kenton P, Atzorn R, Miersch O, Wasternack C (2006) The outcomes of concentration-specific interactions between salicylate and *jasmonate* signalling include synergy, antagonism, and oxidative stress leading to cell death. *Plant Physiol.* **140**: 249-262

Murre C, McCaw P S, Baltimore D (1989) A new DNA binding and dimerization motif in immunoglobulin enhancer binding, *daughterless*, *MyoD*, and *myc* proteins. *Cell* **56**: 777-783.

Nakazawa M, Yabe N, Ichikawa T, Yamamoto Y Y, Yoshizumi T, Hasunuma K, Matsui M (2001) DFL1, an auxinresponsive GH3 gene homologue, negatively regulates shoot cell elongation and lateral root formation, and positively regulates the light response of hypocotyl length. *Plant J.* **25**: 213-221.

Navarro L Bari R, Achard C, Lison P, Nemri A, Harberd N T, Jones JDG (2008) DELLAs control plant immune responses by modulating the balance of jasmonic acid and salicylic acid signaling. *Curr. Biol.* **18**:650-655.

Navarro L, Dunoyer P, Jay F, Arnold B, Dharmasiri N, Estelle M, Voinnet O, Jones J D (2006) A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. *Science* **312**: 436-439.

Navarro L, Zipfel C, Rowland O, Keller I, Robatzek S, Boller T, Jones J D (2004) The transcriptional innate immune response to flg22. Interplay and overlap with Avr gene-dependent defense responses and bacterial pathogenesis. *Plant Physiol.* **135**: 1113-1128.

Nawy T, Lee J Y, Colinas J, Wang J Y, Thongrod S C (2005) Transcriptional profile of the Arabidopsis root quiescent center. *Plant Cell* **17**: 1908-1925.

Ndamukong I, Abdallat A A, Thurow C, Fode B, Zander M, Weigel R, Gatz C(2007) SA-inducible *Arabidopsis* glutaredoxin interacts with TGA factors and suppresses JA-responsive *PDF1.2* transcription. *Plant J.* **50**: 128-139.

Nemhauser J L, Mockler T C, Chory J (2004). Interdependency of brassinosteroid and auxin signaling in Arabidopsis. *PLoS Biol.* **2**: 1460-1471.

Nene Y L, Reddy M V (1987) Chickpea diseases and their control. *CAB International, Oxon, United Kingdom.* 233-270.

Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L. (2000) The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *Plant Cell* **12**: 1863–1878

Newman T C, Ohme-Takagi M, Taylor C B, Green P J (1993) DST sequences, highly conserved among plant SAUR genes, target reporter transcripts for rapid decay in tobacco. *Plant Cell* **5**: 701-714.

Nimbalkar, S B, Harsulkar, A M, Giri, A P, Sainani, M N, Franc, V, Gupta S (2006) Differentially expressed gene transcripts in roots of resistant chickpea plant (Cicer arietinum L.) upon *Fusarium oxysporum* infection. *Physiol. Mol. Plant Pathol.* **68**: 176-188.

Nissan G, Manulis-Sasson S, Weinthal D, Mor H, Sessa G, Barash I (2006) The type III effectors HsvG and HsvB of gall-forming Pantoea agglomerans determine host specificity and function as transcriptional activators. *Mol. Microbiol.* **61**: 1118-1131.

Norman-Setterblad C, Vidal S, Palva E T (2000) Interacting signal pathways control defense gene expression in Arabidopsis in response to cell wall-degrading enzymes from *Erwinia carotovora. Mol. Plant Microbe Interact.* **13**: 430-438.

Oakey R J, Beechey C V (2002) Imprinted genes: identification by chromosome rearrangements and post-genomic strategies. *Trends Genet.* **18**: 359-366.

Onate-Sanchez L, Singh K B (2002) Identification of Arabidopsis ethylene-responsive element binding factors with distinct induction kinetics after pathogen infection. *Plant Physiol.* **128**: 1313-1322.

Osborn T C, Pires J C, Birchler J A, Auger D L, Chen Z J, Lee H S, Comai L, Madlung A, Doerge R W, Colot V (2003) Understanding mechanisms of novel gene expression in polyploids. *Trends Genet.* **19**: 141-147.

Pande S, Rao J N, Sharma M (2007) Establishment of the Chickpea Wilt Pathogen *Fusarium oxysporum* f. sp. *Ciceris* in the Soil through Seed Transmission. *Plant Pathol. J.* **23**: 3-6.

Pandey A, Chakraborty S, Datta A, Chakraborty N (2008) Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* **7**: 88-107.

Pavet V, Quintero C, Cecchini N M, Rosa A L, Alvarez M E (2006) *Arabidopsis* Displays Centromeric DNA Hypomethylation and Cytological Alterations of Heterochromatin Upon Attack by *Pseudomonas syringae. Mol. Plant Microb Interect.* **19**: 577-587.

Park J E, Park J Y, Kim Y S, Staswick P E, Jeon J, Yun J, Kim S Y, Kim J, Lee H W, Park C M. (2007) GH3-mediated Auxin Homeostasis Links Growth Regulation with Stress Adaptation Response in *Arabidopsis. The Journal of Biological Chemistry* **282(13)**: 10036–10046.

Payne C T, Zhang F, and Lloyd, A M (2000).*GL3*encodes a bHLH protein that regulatestrichome development in Arabidopsis throughinteraction with GL1 and TTG1. *Genetics* **156**: 1349-1362.

Peart J R, Lu R, Sadanandom A (2002) Ubiquitin ligase-associated protein SGT1 is required for host and nonhost disease resistance in plants. *Proc. Natl. Acad. Sci.* USA **99**: 10865-10869.

Penninckx I A M A, Thomma B P H J, Buchala A, Métraux J P, Broekaert W F (1998) Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in *Arabidopsis*. *Plant Cell* **10**: 2103-2113.

Persson S, Wei H, Milne J, Page G P, Somerville C R (2005) Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. USA* **102**: 8633-8638.

Petersen M, Brodersen P, Naested H, Andreasson E, Lindhart U (2000) Arabidopsis MAPkinase 4 negatively regulates systemic acquired resistance. *Cell* **103**: 1111-1120.

Pieterse C M J, Leon-Reyes A, Van der Ent S, Van Wees S C M (2009) Networking by small-molecule hormones in plant immunity. *Nat. Chem. Biol.* **5**: 308-316.

Potters G, Pasternak T P, Guisez Y, Palme K J, Jansen M A (2007) Stress-induced morphogenic responses: growing out of trouble? *Trends Plant Sci.* **12**: 98-105.

Qin H, Chan M W, Liyanarachchi S, Balch C, Potter D, Souriraj I J, Cheng A S, Agosto-Perez F J, Nikonova E V, Yan P S, Lin H J, Nephew K P, Saltz J H, Showe L C, Huang T H, Davuluri R V (2009) An Integrative ChIP-chip and Gene Expression Profiling to Model SMAD Regulatory Modules. *BMC Syst. Biol.* **3**: 73-78.

Quail P H (2000) Phytochrome-interacting factors. *Semin. Cell Dev. Biol.* **11**: 457-466.

Raabe R D, Conners I L, Martinez A P (1981) Checklist of plant diseases in Hawaii: including records of microorganisms, principally fungi, found in the state. *Hawaii Institute of Tropical Agriculture and Human Resources (CTAHR), Information Text Series*. **022**: 313.

Rafalski A (2002) Applications of single nucleotide polymorphisms in crop genetics. *Curr. Opin. Plant Biol.* **5**: 94-100.

Ralph S G, Jancsik S, Bohlmann J (2007) Dirigent proteins in conifer defense II: Extended gene discovery, phylogeny, and constitutive and stress-induced gene expression in spruce (*Picea* spp.). *Phytochemistry* **68**: 1975-91.

Ramírez M, Graham M A, Blanco-López L, Silvente S, Medrano-Soto A, Blair M W, Hernández G, Vance C P, Lara M (2005) Sequencing and analysis of common bean ESTs. Building a foundation for functional genomics. *Plant Physiol.* **137**: 1211-27.

Reimers P J, Guo A, Leach J E (1992) Increased Activity of a Cationic Peroxidase Associated with an Incompatible Interaction Between *Xanthomonas oryzae* pv. *oryzae* and Rice (*Oryza sativa*). *Plant Physiol.* **99**: 1044-1050.

Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannet N, Kanin E, Volkart T L, Wolson C J, Bell S P, Young R A (2000) Genome-wide location and function of DNA binding proteins. *Science* **290**: 2306-2309.

Rensink W A, Buell C R (2005) Microarray expression profiling resources for plant genomics. *Trends Plant Sci.* **10**: 603-609.

Revenkova E, Masson J, Koncz C, Afsar K, Jakovleva L, Paszkowski J (1999) Involvement of Arabidopsis thaliana ribosomal protein S27 in mRNA degradation triggered by genotoxic stress. *EMBO J.* **18**: 490-499.

Richmond T, Somerville S (2000) Chasing the dream: plant EST microarrays. *Curr. Opin. Plant Biol.* **3**: 108-116.

Riechmann J L, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe O J, Samaha R R (2000) Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* **290**: 2105-2110.

Roberts M R, Bowles D J (1999) Fusicoccin, 14-3-3 proteins, and defense responses in tomato plants. *Plant Physiol.* **119**: 1243-1250.

Robinson K A, Koepke J I, Kharodawala M, Lopes J M (2000) A network of yeast basic helix-loop-helix interactions. *Nucleic Acids Res.* **28**: 4460-4466.

Romano P G, Edvardsson A, Ruban A V, Andersson B, Vener A V, Gray J E, Horton P (2004) Arabidopsis AtCYP20-2 is a light-regulated cyclophilin-type peptidyl-prolyl cis-trans isomerase associated with the photosynthetic membranes. *Plant Physiol.* **134**: 1241-1243.

Rowland O, Ludwig A A, Merrick C J, Baillieul F, Tracy F E, Durrant W E, Fritz-Laylin L, Nekrasov V, Sjölander K, Yoshioka H, Jones J D G (2005) Functional Analysis of *Avr9/Cf-9 Rapidly Elicited* Genes Identifies a Protein Kinase, ACIK1, That Is Essential for Full Cf-9–Dependent Disease Resistance in Tomato. *Plant Cell* **17**: 295-310.

Roy A, Das K, Rao K B V, Dilaji (2001) A new chickpea variety for Hills zone of Assam, India. International Chickpea and Pigeonpea. *News Letter.* **8**: 6-7.

Rudd S, Schoof H, Mayer K (2005) Plant Markers - a database of predicted molecular markers from plants. *Nucleic Acids Res.* **33**: 628-632.

Sambrook J, Fritsch E F, Maniatis T (1989) Molecular Cloning: A Laboratory Manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY*

Sasaki Y, Asamizu E, Shibata D, Nakamura Y, Kaneko T, Awai K, Amagai M, Kuwata C, Tsugane T, Masuda T, Shimada H, Takamiya K-I, Ohta H, Tabata S (2001) Monitoring of methyl jasmonate-responsive genes in *Arabidopsis* by cDNA macroarray: self-activation of jasmonic acid biosynthesis and crosstalk with other phytohormone signaling pathways. *DNA Res.* **8**: 153-161.

Scheideler M, Schlaich N L, Fellenberg K, Beissbarth T, Hauser N C, Vingron M, Slusarenko A J, Hoheisel J D (2002) Monitoring the switch from housekeeping to pathogen defense metabolism in *Arabidopsis thaliana* using cDNA arrays. *J. Biol. Chem.* **277**: 10555-10561

Schenk P M, Kazan K, Wilson I, Anderson J P, Richmond T, Somerville S C, Manners J M (2000) Coordinated plant defense responses in Arabidopsis revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA* **97**: 1655-1660.

Schrage Y M, Hameetman L, Szuhai K, Cleton-Jansen A M, Taminiau A H, Hogendoorn P C, Bovée J V (2009) Aberrant heparan sulfate proteoglycan localization, despite normal exostosin, in central chondrosarcoma. *Am. J. Pathol.* **174**: 979-988.

Sehnke P C, DeLille J M, Ferl R J (2002) Consummating Signal Transduction: The Role of 14-3-3 Proteins in the Completion of Signal-Induced Transitions in Protein Activity. *Plant Cell* S339–S354.

Shendure J, Ji H (2000) Next generation DNA sequencing. *Nat. Biotech.* **26**: 1135-1145.

Shin S, Mackintosh C A, Lewis J, Heinen S J, Radmer L, Dill-Macky R, Baldridge G D, Zeyen R J, Muehlbauer G J (2008) Transgenic wheat expressing a barley class II chitinase gene has enhanced resistance against Fusarium graminearum. *J. Exp. Bot.* **59**: 2371-2378.

Siemens J, Keller I, Sarx J, Kunz S, Schuller A, Nagel W, Schmulling T, Parniske M, Ludwig-Muller J (2006) Transcriptome analysis of *Arabidopsis* clubroots indicate a key role for cytokinins in disease development. *Mol. Plant Microbe Interact.* **19**: 480-494.

Simmons C R, Fridlender M, Navarro P A, Yalpani N (2003) A maize defense-inducible gene is a major facilitator superfamily member related to bacterial multidrug resistance efflux antiporters. *Plant Mol. Biol.* **52**: 433-46.

Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Lee T V, Bleeker M, Onstenk J, de Both M, Haring M, Mes J, Cornelissen B, Zabeau M, Vos P (1998) Dissection of the Fusarium I2 gene cluster in tomato reveals six homologs and one active gene copy. *Plant Cell* **10**: 1055-1068.

Singh K B, Foley R C, Oñate-Sánchez L (2002)Transcription factors in plant defense and stress responses. *Curr. Opin. Plant Biol.* **5**: 430-436.

Singh K B, Ocampo B, Robertson L D (1998) Diversity for abiotic and biotic stress resistance in the wild annual cicer species. *Genet. Res. Crop. Eval.* **45**: 9-17.

Singh, K B (1987) Chickpea breeding. *The Chickpea. CAB International, UK.* pp.127-162.

Singla B, Chugh A, Khurana J P. Khuranay P (2006) An early auxin-responsive Aux/IAA gene from wheat (*Triticum aestivum*) is induced by epibrassinolide and differentially regulated by light and calcium. *J. Expt. Bot.* **57**: 4059-4070.

Song W Y, Wang G L, Chen L L, Kim H S, Pi L Y (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* **270**: 1804-1806.

Sonnenfeld M J, Delvecchio C, Sun X (2005) Analysis of the transcriptional activation domain of the Drosophila tango bHLH-PAS transcription factor. *Dev. Genes Evol.* **215**: 221-229.

Spencer M W B, Casson S A, Lindsey K (2007) Transcriptional Profiling of the Arabidopsis Embryo. *Plant Physiol.* **143**: 924-940.

Stracke R, Werber M, Weisshaar B (2001) The R2R3-MYB gene family in *Arabidopsis thaliana. Curr. Opin. Plant Biol.* **4**: 447-456.

Stros M, Launholt D, Grasser K D (2007) The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell. Mol. Life Sci.* **64**: 2590-2606.

Sugiyama A, Noguchi K, Kitanaka C, Katou N (1999) Molecular cloning and chromosomal mapping of mouse intronless *myc* gene acting as a potent apoptotic inducer. *Genes* **226**: 273-283.

Sullivan J A, Shirasu K, Deng X W (2003) The diverse roles of ubiquitin and the 26S proteasome in the life of plants. *Nat. Rev. Genet.* **4**: 948-958.

Sullivan M L, Green P J (1996) Mutational analysis of the DST element in tobacco cells and transgenic plants: identification of residues critical for mRNA instability. *RNA 2:* 308-315.

Sun W, Dunning F M, Pfund C,Weingarten R, Bent A F (2006) Within-species polymorphism among *Xanthomonas campestris* pv. *campestris* flagellins alters elicitation of *Arabidopsis FLS2*-dependent defenses. *Plant Cell* **18**: 764-779.

Suzuki K, Suzuki N, Ohme-Takagi M (1998) Immediate early induction of mRNAs for ethylene-responsive transcription factors in tobacco leaf strips after cutting. *Plant J.* **15**: 657-665.

Takahashi H, Hayashi M, Goto F, Sato S, Soga T, Nishioka T, Tomita M, Kawai-Yamada M, Uchimiya H (2006) Evaluation of metabolic interaction in transgenic rice over expressing dihydroflruone-4-reductase. *Ann. Bot.* **98**: 819-825.

Tao Y, Xie Z, Chen W, Glazebrook J, Chang H-S, Han B, Zhu T, Zou G, Katafiri F (2003) Quantitative Nature of Arabidopsis Responses during Compatible and Incompatible Interactions with the Bacterial Pathogen *Pseudomonas syringae Plant Cell* **15** : 317-330.

Teale W D, Paponov I A, Palme K (2006) Auxin in action: signalling, transport and the control of plant growth and development. *Nat. Rev. Mol. Cell. Biol.* **7**: 847-859.

Thomma B P, Penninckx I A, Broekaert W F, Cammue B P (2001) The complexity of disease signaling in Arabidopsis. *Curr. Opin. Immunol.* **13**: 63-68.

Thurow C, Schiermeyer A, Krawczyk S, Butterbrodt T, Nickolov K,Gatz C (2005) Tobacco bZIP transcription factor TGA22 and related factor TGA21 have distinct roles in plant defense responses and plant development. *Plant J.* **44**: 100-113.

Tili E, Michaille J J, Calin G A (2008) Expression and function of micro RNAs in immune cells during normal or disease state. *Int. J. Med. Sci.* **5**: 73-79.

Tjamos E C, Beckman C H (1989) Vascular Wilt Diseases of Plants: Basic Studies and Control. *Berlin: Springer-Verlag.*

Toledo-Ortiz G, Huq E, Quail P H (2003) The Arabidopsis basic/helix-loop helix transcription factor family. *Plant Cell* **15**: 1749-1770.

Trusov Y, Rookes J E, Chakravorty D, Armour D, Schenk P M, Botella J R (2006) Heterotrimeric G Proteins Facilitate Arabidopsis Resistance to Necrotrophic Pathogens and Are Involved in Jasmonate Signaling. *Plant Physiol.* **140**: 210-220.

Udall J A, Swanson J M, Haller K, Rapp R A, Sparks M E, Hatfield J, Yu Y, Wu Y, Dowd C, Arpat A B, Sickler B A, Wilkins T A (2006) A global assembly of cotton ESTs. *Genome Res.* **16**: 441-450.

Vailleau F, Daniel X, Tronchet M, Montillet J L, Triantaphylides C, Roby D (2002) A R2R3-MYB gene, AtMYB30, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. *Proc. Natl. Acad. Sci. USA* **99**: 10179-10184.

Van Heusden G P, Griffiths D J, Ford J C, Chin A W T F, Schrader P A, Carr A M, Steensma H Y (1995) The 14-3-3 proteins encoded by the BMH1 and BMH2 genes are essential in the yeast *Saccharomyces cerevisiae* and can be replaced by a plant homologue. *Eur. J. Biochem.* **229**: 45-53.

Van Loon L C, Geraats B P J, Linthorst H J M (2006) Ethylene as a modulator of disease resistance in plants. *Trends Plant Sci.* **11**: 184-191.

van Wees S C, Chang H S, Zhu T, Glazebrook J (2003) Characterization of the early response of Arabidopsis to *Alternaria brassicicola* infection using expression profiling. *Plant Physiol.* **132**: 606-617.

Varshney R K, Horres R, Molina C, Nayak S, Jungmann R, Swamy P, Winter P, Jayashree B, Kahl G, Hoisington D A (2007) Extending the repertoire of microsatellite markers for genetic linkage mapping and germplasm screening in chickpea. *SATeJournal* **5**: 1-3.

Varshney R K, Graner A, Sorrells M E (2005) Genic microsatellite markers in plants: features and applications. *Trends Biotechnol.* **23**: 48-55.

Vercauteren I, de Almeida Engler J, De Groodt R, Gheysen G. (2002) An *Arabidopsis thaliana* pectin acetylesterase gene is upregulated in nematode feeding sites induced by root-knot and cyst nematodes. *Mol. Plant Microbe Interact.* **15**: 404-407.

Wan J F, Mark D, Andrew F B. (2002) Probing plant-pathogen interactions and downstream defense signaling using DNA microarrays. *Funct. Integr. Genomics* 2: 259–273.

Wang D, Amornsiripanitch N, Dong X (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLos Pathogens* **2**: 1042-1050.

Wang D, Pajerowska-Mukhtar K, Hendrickson Culler A, Dong X (2007) Salicylic acid inhibits pathogen growth in plants through repression of the auxin signaling pathway.*Curr. Biol.* **17**:1784-1790.

Wang Y J, Zhang Z G, He X J, Zhou H L, Wen Y X, Dai J X, Zhang J S, Chen S Y (2003) A rice transcription factor OsbHLH1 is involved in cold stress response. *Theor. Appl. Genet.* **107**: 1402-1409.

Wei S M, Xie C G, Abe Y, Cai J T (2009) ADP-ribosylation factor like 7 (ARL7) interacts with alpha-tubulin and modulates intracellular vesicular transport. *Biochem. Biophys. Res. Commun.***384**: 352-356.

Wellmer F, Alves-Ferreira M, Dubois A, Riechmann J L, Meyerowitz E M (2006) Genome-wide analysis of gene expression during early Arabidopsis flower development. *PLoS Genetics* **2**: 1012-1024.

Weltmeier F, Ehlert A, Mayer C S, Dietrich K, Wang X, Schütze K, Alonso R, Harter K, Vicente-Carbajosa J, Dröge-Laser W (2006) Combinatorial control of Arabidopsis proline dehydrogenase transcription by specific heterodimerisation of bZIP transcription factors. *EMBO J.* **25**: 3133-43.

White J A, Todd J, Newman T, Focks N, Girke T, de Ilarduya M, Jaworski J G, Ohlrogge J B, Benning C (2000) A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil. *Plant Physiol.* **124**: 1582-1594.

Woodward A W, Bartel B (2005) Auxin: Regulation, Action, and Interaction. *Annals of Botany* **95**: 707-735.

Wu C T, Bradford K J (2003) Class I chitinase and beta-1,3-glucanase are differentially regulated by wounding, methyl jasmonate, ethylene, and gibberellin in tomato seeds and leaves. *Plant Physiol.* **133**: 263-73.

Xu L, Liu F, Lechner E, Genschik P, Crosby W L (2002). The SCF (COI1) biquitin-ligase complexes are required for jasmonate response in Arabidopsis. *Plant Cell* **14**: 1919-1935.

Yang H Y, Wen Y Y, Chen C H, Lozano G, Lee M H (2003) 14-3-3 sigma positively regulates p53 and suppresses tumor growth. *Mol. Cell. Biol.* **23**: 7096-7107.

Yang T, Poovaiah B W (2000) Molecular and biochemical evidence for the involvement of calcium/calmodulin in auxin action. *J. Biol. Chem.* **275:** 3137-3143.

Yasuda M. *et al.* (2008) Antagonistic interaction between systemic acquired resistance and the abscisic acid-mediated abiotic stress response in *Arabidopsis*. *Plant Cell* **20**: 1678–1692.

Yang Y, Shah J, Klessig D F (1997). Signal perception and transduction in plant defense responses. *Genes Dev.* **11**: 1621-1639.

Zabala G, Zou1 J, Tuteja J, Gonzalez1 D O, Clough S J, Vodkin L O (2006) Transcriptome changes in the phenylpropanoid pathway of Glycine max in response to Pseudomonas syringae infection. *BMC Plant Biol.* **6**: 26.

Zeng L R, Vega-Sánchez M E, Zhu T, Wang G L (2006) Ubiquitination-mediated protein degradation and modification: an emerging theme in plant-microbe interactions *Cell Res.* **16**: 413-426.

Zhang S, Klessig D F (2001) MAPK cascades in plant defense signaling. *Trend Plant Sci.* **6**: 520-527.

Zhang Y, Tessaro M J, Lassner M, Li X (2003) Knockout analysis of Arabidopsis transcription factors TGA2, TGA5, and TGA6 reveals their redundant and essential roles in systemic acquired resistance. *Plant Cell* **15**: 2647-2653.

Zhang Y, Wang Y, Jiang L, Xu Y, Wang Y, Lu D, Chen F (2007). Aquaporin JcPIP2 is involved in drought responses in *Jatropha curcas. Acta. Biochim. Biophys. Sin. (Shanghai)* **39**: 787-794.

Zhou C, Zhang L, Duan J, Miki B, Wu K (2005) *HISTONE DEACETYLASE19* is Involved in Jasmonic Acid and Ethylene Signaling of Pathogen Response in Arabidopsis. *Plant Cell* **17**: 1196-1204.

Zhu H, Choi H-K, Cook D R, Shoemaker R C (2005) Bridging model and crop legumes through comparative genomics. *Plant Physiol.* **137**: 1189-1196.

Zipfel C, Felix G (2005) Plants and animals: a different taste for microbes? *Curr. Opin. Plant Biol.* **8**: 353-360.

Zipfel C, Robatzek S, Navarro L, Oakeley E J, Jones J D, Felix G, Boller T (2004) Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature* **428**: 764-767.

Zipfel C, Kunze G, Chinchilla D, Caniard A, Jones J D G, Boller T, Felix G (2006) Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell* **125**: 749-760.