The Pennsylvania State University

The Graduate School

Graduate Degree Program in Integrative Biosciences

FUNCTIONAL ANALYSIS OF NON EXPRESSOR OF PR1 (NPR1) AND ITS PARALOG NPR3 IN *THEOBROMA CACAO* AND *ARABIDOPSIS THALIANA*

A Dissertation in

Integrative Biosciences

by

Zi Shi

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2010

The dissertation of Zi Shi was reviewed and approved* by the following:

Mark J. Guiltinan
Professor of Plant Molecular Biology
Dissertation Advisor
Chair of Committee

Sarah M. Assmann
Waller Professor of Biology

Hong Ma
Distinguished Professor of Biology

Yinong Yang
Associate Professor of Plant Pathology

Siela N. Maximova
Associate Professor of Horticulture

Peter Hudson
Willaman Professor of Biology
Director of Huck Institutes of the Life Sciences

*Signatures are on file in the Graduate School

ABSTRACT

Arabidopsis NON EXPRESSOR OF PR1 (NPR1) is a key transcription regulator of the salicylic acid (SA) mediated defense signaling pathway. The *NPR* gene family consists of *NPR1* and five other *NPR1-like* genes in *Arabidopsis*. This research focuses on the functional analysis of an *NPR1* ortholog from *Theobroma cacao* L. and characterization of one of the *NPR1* paralogs, *NPR3*, in both *Arabidopsis* and cacao. To identify the function of NPR3 in *Arabidopsis*, I first examined the gene expression pattern of *NPR3* and found it to be strongly expressed in developing flower tissues. Interestingly, an *npr3* knockout mutant displayed enhanced resistance to *Pseudomonas syringae* tomato pv. DC3000 (*P.s.t.*) infection of immature flowers. Gene expression analysis also revealed increased basal and induced levels of *PR1* transcripts in *npr3* developing flowers. To investigate the possible mechanism of NPR3-dependent negative regulation of defense response, I tested the physical interactions of NPR3 with both TGA2 and NPR1 *in vivo*, which suggests that NPR3 represses NPR1-dependent transcription by inhibiting the nuclear localization of NPR1 through direct binding to TGA2 and NPR1. To characterize the *NPR1* ortholog from cacao, I isolated *TcNPR1* gene from genotype of Scavina6, and demonstrated that it expresses constitutively in all the tested tissues. To functionally analyze this gene, a bacterial growth assay was carried out with *npr1-2* transgenic lines overexpressing *TcNPR1*, and a reduced level of bacterial growth demonstrated that *TcNPR1* can partially complement *Arabidopsis* the *npr1-2* mutation. In addition, TcNPR1 was shown to translocate into nuclei upon SA treatment in a manner identical to *Arabidopsis* native NPR1. To further explore the *NPR* gene family in cacao, I identified a total of four *NPR-like* genes from the cacao genome, and phylogenetic analysis indicated that the duplications of three clades in this gene family occurred before the divergence of *Arabidopsis* and cacao. To identify the functional ortholog of *Arabidopsis* NPR3, I isolated a putative *TcNPR3* gene and demonstrated that its expression level was higher in un-open flowers and older leaves, a pattern similar to *Arabidopsis* NPR3. A complementation test of TcNPR3 expressed in the *Arabidopsis npr3-3* null mutant showed that *TcNPR3* can functionally substitute for the *Arabidopsis NPR3* gene, demonstrating that TcNPR3 is the functional ortholog of AtNPR3. To obtain the genome-

wide transcriptional responses of SA treatment in cacao, I used microarray analysis to measure gene expression in two cacao genotypes (ICS1 and Scavina6), three leaf developmental stages (A, C and E) and two treatments (water and SA). After validating the microarray results with RT-PCR, I identified differentially expressed genes from all twenty-four pair-wise comparisons. Interestingly, chloroplast and mitochondrial genes are enriched in SA-induced Scavina6 but those genes are underrepresented in ICS1, suggesting that the oxidative burst and hypersensitive response during defense response may vary between the two genotypes. In all, this research will not only offer us the knowledge of defense response mechanism and signal transduction regulation in *Arabidopsis* and cacao, but also provide molecular tools for selecting cultivars with enhanced disease resistance for cacao breeders and farmers.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank all people who has helped me and encouraged me during my doctoral study.

I owe my deepest gratitude to my advisor Mark Guiltinan for his great support and guidance for the past years. His patience, enthusiasm, and immense knowledge had motivated me all the time of research and writing of this thesis. I could not have come to this far without his inspiration and great efforts. Besides my advisor, I would like to thank my thesis committee members, Dr. Hong Ma, Dr. Sarah Assamnn, Dr. Yinong Yang, for their insightful comments and suggestions. Special thanks go to Dr. Siela Maximova for her continual encouragement, support and advice, and this work would not have been possible without her. I would like to thank Joseph Verica, Ann Yuong, Sharon Pishak, Marna Yandeau-Nelson, Luis Mejia and Yufan Zhang from Guiltinan's lab for their numerous technical support, valuable suggestions and generous help. I feel thankful to Nicole Zembower at Penn State Cytometry Facility for providing help with microscopy throughout my thesis study.

I want to thank my husband Yuannian Jiao for his support, consideration and encouragement not only in my daily life, but also academically. I thank my friends in the USA: Yi Liu, Hui Zhang, Xiaoying Meng, Xuan Ma, Fang Cong and Rong Xiao, for all the joy we have had in the past years.

Finally, and most importantly, I wish to thank my parents for their boundless love and tremendous assistance throughout my life. To them I dedicate this thesis.

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1 Plant Immunity and Systemic Signaling

In contrast to animals, plants are unable to move during their lifetime, thus, plants have evolved a number of specialized defense mechanisms to combat microbial pathogens and herbivorous insect invasions. Plant innate immunity is based on an amazingly complicated defense response that is highly flexible allowing the recognition of and resistance against different pathogens (Pieterse and Van Loon, 2004). Besides the pre-existing physical and chemical barriers, plants have developed inducible defense mechanisms that are triggered by invader's attack. Plants have also evolved the capacity to establish systemic signals, which can confer resistance to secondary infections of pathogens in the uninfected or distal parts of plants. Several systemic signaling pathways have been characterized in plants. These include systemic acquired resistance (SAR), which is activated by necrotizing pathogens (Sticher et al., 1997; Durrant and Dong, 2004); induced systemic resistance (ISR), which is triggered by colonization of roots by nonpathogenic bacteria (Van der Ent et al., 2008); and wound induced defense which is triggered by tissue damage causing by feeding insects (Kessler and Baldwin, 2002). Hormonal signal molecules, such as salicylic acid (SA), jasmonic acid (JA) and ethylene (ET) play major roles in regulating those signal transduction pathways (Durrant and Dong, 2004; Grant and Lamb, 2006; Vlot et al., 2008; Leon-Reyes et al., 2009).

In *Arabidopsis thaliana*, it has been shown that SA, JA and ET regulate separate defense pathways effective against different invaders (Pieterse and Van Loon, 2004; Glazebrook, 2005; Leon-Reyes et al., 2009). SA is required for the onset of SAR (Durrant and Dong, 2004) and SA-dependent responses typically deter the growth of biotrophic pathogens, such as *Pseudomonas syringae*. In contrast, necrotrophic pathogens are commonly more sensitive to ISR, which is a SA-independent process but needs JA and/or ET as signal molecules (Feys and Parker, 2000; Spoel et al., 2003). In nature, plants often face the invasions of different aggressors simultaneously, which may affect the primary induced defense response of the host plants (Stout et al., 2006; Poelman et al., 2008). Since the activation of defense response can be associated with fitness costs (van Hulten et al.,

2006), plants have to regulate the defense mechanism to effectively acclimate to the changes in their environments. Cross-talk between different defense pathways provides a means to manipulate the signaling interactions and fine tune the defense outcomes depending on different pathogens (Kunkel and Brooks, 2002; Spoel et al., 2003; Koornneef and Pieterse, 2008; Spoel and Dong, 2008), and hence, to minimize the physiological costs during defense responses (Bostock, 2005).

1.2 The Function of NPR1 in SA-dependent Defense Signaling Pathway

1.2.1 Activation of NPR1 by SA

Systemic acquired resistance (SAR) protects plants against a broad spectrum of pathogens (Dong, 2004; Loake and Grant, 2007). The onset of SAR requires the accumulation of SA and it involves the integrated action of numbers of pathogenesis-related (PR) proteins at the molecular level. Many of the PR proteins are believed to have antimicrobial activities *in vitro*, and SAR is considered as the harmonious effect of many PR proteins (Loon and Strien, 1999). Although the exact role of each PR-protein in defense response is not clearly known, *PR* genes still serve as vital molecular markers of SAR (Durrant and Dong, 2004; Grant and Lamb, 2006). Exogenous application of SA or its chemical analog 2,6-dichloroiosnicotinic acid (INA) and benzothiadiazole S-methyl ester (BTH) can initiate SAR in the absence of pathogen challenge. It has been well demonstrated that the increase of endogenous SA induced after infection is involved in the signal transduction pathways of the defense response, as transgenic *Arabidopsis NahG* plants that express the bacterial salicylate hydroxylase cannot accumulate SA, resulting in lower *PR* expression and impaired disease resistance (Pieterse et al., 1996; van Wees and Glazebrook, 2003).

Non Expressor of PR1 (NPR1) was identified in a number of genetic screens for plants that are unable to express PR genes after SAR induction (Cao et al., 1994; Ryals et al., 1997; Shah et al., 1997). NPR1 has been characterized as a transcription coactivator that plays a vital role in the plant defense response (Delaney et al., 1995; Glazebrook et al., 1996; Grant and Lamb, 2006; Loake and Grant, 2007). The *Arabidopsis npr1* mutant was severely

impaired in resistance to bacterial and fungal pathogens (Glazebrook et al., 1996; Cao et al., 1997; Ryals et al., 1997). The disease symptoms after infection by *Pseudomonas syringae pv. maculicola* (*P.s.m.*) ES4326 in wild type *Arabidopsis* was confined to the bacterial infiltration site, however, the lesions spread dramatically on the infected leaf of *npr1* mutants. In addition, the growth of bacteria was significantly increased in *npr1* mutant plants compared to wild type. Upon the treatment of SA or its chemical analog, the *PR1* gene was highly induced in wild type but was only slightly induced in the *npr1* mutant. Microarray analysis showed that NPR1 was required for the activation or down-regulation of 2,248 target genes in *Arabidopsis* in response to 8h treatment with SA analog, BTH (Wang et al., 2006). In another study, 193 genes were rapidly activated in a NPR1-dependent manner by 2.5h SA incubation, identified by the high throughput Complete Arabidopsis Transcriptome MicroArray (CATMAv2) chip (Blanco et al., 2009). Taken together, these data confirm that NPR1, as a positive regulator of disease resistance, regulates numerous downstream genes in defense signaling pathways.

DNA sequence analysis revealed four exons and three introns in the *NPR1* gene. The protein encoded by the *NPR1* gene contains a broad complex, tramtrack and bric a brac/pox virus and zinc finger (BTB/POZ) domain at the N-terminus and ankyrin-repeats in the central region, both of which mediate protein-protein interactions in other proteins (Cao et al., 1997; Stogios et al., 2005). Interestingly, NPR1 shares homology with the human IκBα transcription inhibitors (Ryals et al., 1997) that functions as a negative regulator of the NF-κB signaling pathway that controls the innate immunity response (Ghosh et al., 1998). The transcription factor NF-κB normally resides in the cytoplasm in an inactive form bound to IκB. Upon reception of a signal initiating the immune response (e.g. inflammatory cytokines), IκB is degraded releasing NF-κB. NF-κB then enters the nucleus where it affects transcription of genes that function in innate immunity (Baldwin, 1996).

To identify NPR1 interacting factors, a yeast-two-hybrid screen was performed and two proteins from the *Arabidopsis* TGA family of basic leucine zipper (bZip) transcription factors, TGA2 and TGA3, have been shown to physically interact with NPR1 (Despres et al., 2000; Fan and Dong, 2002). It was also demonstrated that TGA2, TGA5 and TGA6 are functional redundant, as single mutants do not display any phenotype and only *tga2tga5tga6* triple mutant showed *npr1*-like phenotype (Zhang et al., 2003). The

interaction between NPR1 and TGA2 is indispensable for the onset of SAR, as there is no *PR1* induction and increased susceptibility to pathogens results if the interaction is abolished (Despres et al., 2000; Fan and Dong, 2002; Johnson et al., 2003; Ndamukong et al., 2007). Moreover, TGA transcription factors are strongly recruited to the cis-element LS7 in *PR1* promoter by a SA-and NPR1-dependent manner to induce *PR* expression in defense response.

1.2.2 NPR1-Dependent Signaling Pathway

Tremendous efforts have been made to elucidate the mechanism of NPR1-dependent signaling pathway in *Arabidopsis* (Mou et al., 2003; Dong, 2004; Rochon et al., 2006; Loake and Grant, 2007; Vlot et al., 2008; Mukhtar et al., 2009). In the absence of pathogens, NPR1 is present as inactive oligomers in the cytoplasm of the cell through the cysteine-based disulfide bonds within or between molecules. Upon pathogen induction, endogenous SA is accumulated which alters the redox state of the cell resulting in the reduction of NPR1 to its active monomeric form. In leaves, NPR1 monomer translocates into the nucleus, where it interacts with TGA transcription factors and the complex binds to the promoter region and induces the *PR* gene expression. It was shown that the BTB/POZ domain facilitates dimerization of NPR1, and the ankyrin-repeat region is involved in the interaction with the TGA subfamily of bZIP transcription factors (Zhang et al., 1999). Although NPR1 does not function as a transcription factor *per se,* as it has no DNA-binding domain, it acts via specific associations with the TGA subfamily of transcription factors (Johnson et al., 2003; Ndamukong et al., 2007; Johnson et al., 2008). Nuclear localization of NPR1 is required for the activation of *PR* gene expression upon pathogen infection or elevated SA level (Kinkema et al., 2000). Computer-based analysis identified three potential nuclear localization signal (NLSs) at the C-terminal region of NPR1 protein, and five basic amino acids in this region facilitate this localization.

Recent studies shed light on an additional level of complexity in the mechanism governing NPR1 function. It has been shown that the turnover of NPR1 protein plays a crucial role in regulating the transcription of target genes (Mukhtar et al., 2009; Spoel et al.,

2009). In the absence of inducers, nuclear NPR1 is constantly cleared in a proteasome-based degradation pathway, which may prevent spurious activation of downstream targets. Pathogen challenge promotes the phosphorylation of NPR1, which in turn assists the recruitment of NPR1 to cullin3-based ubiquitin ligase, as NPR1 protein is more stable in *cul3a cul3b* double mutant. Degradation of NPR1 following the activation of target genes allows fresh NPR1 to be enrolled for a new initiation of transcription, resulting in the fulfillment of SAR. Thus, the proteasome-dependent degradation of NPR1 is important for both preventing and activating SAR, which provides insights of how NPR1 protein is regulated.

1.2.3 The Role of NPR1 in Pathway Cross-talk

There is a great deal of "Cross-talk" among different defense signaling pathways that involves both positive and negative feedback mechanisms (Kunkel and Brooks, 2002; Bostock, 2005; Koornneef and Pieterse, 2008). The signal transduction pathways dependent on SA, JA and ET are highly regulated and concerted to ensure the most efficient pathway will be activated upon the infection of different pathogens (Feys and Parker, 2000). Defense against biotrophic leaf pathogens commonly involves SA-dependent signaling, which is mediated by NPR1, whereas induced systemic resistance against chewing insects and necrotrophic pathogens generally depends on JA-mediated signaling (Beckers and Spoel, 2006; Loake and Grant, 2007; Balbi and Devoto, 2008). However, more complex physiological scenarios emerge because of the environmental conditions and the combination of different lifestyles for various pathogens (Spoel and Dong, 2008). SA and JA signaling pathways are shown to interact either synergistically or antagonistically depending on their respective concentrations (Koornneef and Pieterse, 2008), indicating that the negative crosstalk between JA and SA can be utilized by pathogens to enhance their pathogenesis. One well-known example is hemibiotrophic bacterium *Pseudomonas syringae*, which exploits phytotoxin coronatine, a chemical similar to jasmonate-isoleucine to suppress SA-dependent pathway, thus increasing plant susceptibility (Laurie-Berry et al., 2006; Uppalapati et al., 2007). This complication makes the cross-talk between different

defense pathways fundamental to combat a series of invaders and enhances the rational conservation of energy.

It has been suggested that NPR1 protein also plays a role in JA/ET-dependent ISR signaling (Feys and Parker, 2000). Recruitment of NPR1 to these pathways appears to rely on the nature of pathogen input signal and data suggest that NPR1 plays a role as signal module and controls particular outputs. Studies on the molecular mechanism of the antagonistic effect of SA and JA signaling reveal that *Arabidopsis* plants unable to accumulate SA have 25-fold higher endogenous JA level and exhibit increased expression of JA-responsive genes *lipoxygenase 2* (*LOX2*), *plant defensin 1.2* (*PDF1.2*) and *vegetative storage protein 2* (*VSP2*) in response to *Pseudomonas syringae* pv tomato DC3000 (Spoel et al., 2003). This SA-mediated suppression of JA-responsive genes is dependent on NPR1, as the repression is abolished in *npr1* mutant (Spoel et al., 2003; Spoel et al., 2007). Though nuclear localization is essential for the onset of SAR, it is not required for the suppression of JA response, indicating a novel function of NPR1 in cytosol. A similar function has been reported for the rice *NPR1* homolog (*OsNPR1*) (Yuan et al., 2007). Overexpression of *OsNPR1* in cytosol results in the suppression of JA-inducible genes and enhanced susceptibility to insects. In addition, increased SA level and high susceptibility to herbivore attack have been reported in *NPR1*-silenced tobacco plants (Rayapuram and Baldwin, 2007). It is proposed that cytosolic NPR1 is required for the repression of SA accumulation upon insect herbivory infection, which in turn negatively regulates SA/JA cross-talk to allow the release of JA signaling. These results indicate the regulatory role of NPR1 in induced resistance and signal crosstalk. Collectively, the complex regulation and interplay among these pathways may allow plants to fine-tune defense responses appropriately to the specific invading pathogen and to conserve metabolic energy for non-defense needs.

1.3 Costs and Benefits of Defense Response

Plants evolved constitutive and inducible resistance to defend against offending harmful microbes and insects. It is commonly believed that the evolution of induced resistance is a strategy to conserve energy under enemy-free conditions, avoiding the significant energenic costs of activation of defense responses unnecessarily (van Hulten et

al., 2006). The costs of defense against biotic factors can result from the allocation of resources to defense compounds and gene activation and/or the toxicity of defense products (Heil, 2002). Moreover, external factors can also contribute to the costs during defense, if the defensive traits affect the beneficial interaction between plant and its biotic/abiotic environment. Thus, it is generally presumed that the plants will activate induced resistance only if its benefits outweigh the costs of the response.

Various studies have demonstrated costs associated with JA-induced resistance against herbivores, which affect the growth and reproductive traits of the plants in the absence of invaders, including flowering time and numbers of pollen grains per flower in wild radish (Agrawal et al., 1999) and less seed production in coyote tobacco (Baldwin, 1998). Many studies also have pointed out the ecological benefits and enhanced plant performance of JA-inducible defense upon infections. It has been shown that induction of defense in wild radish positively correlated with seed production (Agrawal, 1998) and increased seed production in wild populations of tobacco upon herbivory attacks (Baldwin, 1998). In addition, the cost-benefit balance of SA-induced defense has also been proposed and studied. SA-induced defense in wheat plants results in lower biomass and less seed set without pathogen infection (Heil et al., 2000) and exogenous application of SA in the absence of enemy attack reduces 15% of seed production in *Arabidopsis* (Cipollini, 2002). Heidel et al. (Heidel et al., 2004) performed field assays on *Arabidopsis* mutants with impaired SA-induced defense(*npr1*) and mutants that constitutively expresses SA-induced defense (*cpr1*, *cpr5* and *cpr6*). It was demonstrated that although *npr1* mutation has no effect on fitness in a growth chamber, it does result in decreased fitness in the field. The constitutive activation of SAR by constitutive defense mutant results in reduced fitness in both field and growth chamber conditions, indicating the pathway is costly and fitness consequence is highly contingent on the environmental cues. Taken together, these results suggest that optimal fitness will be reached at intermediate defense levels that balance the costs and benefits.

Defense response in the absence of pathogen infection is an additional energetic cost to plants and directly induced resistance is demonstrated to reduce relative growth rate and lower seed production without infection (van Hulten et al., 2006), which might be the reason that plants have developed negative regulators to control the activation of

defense responses, and thus have evolved mechanisms to monitor the resources allocations and maintain low costs in the absence of hostile conditions.

1.4 Other *NPR1-like* Genes in *Arabidopsis*

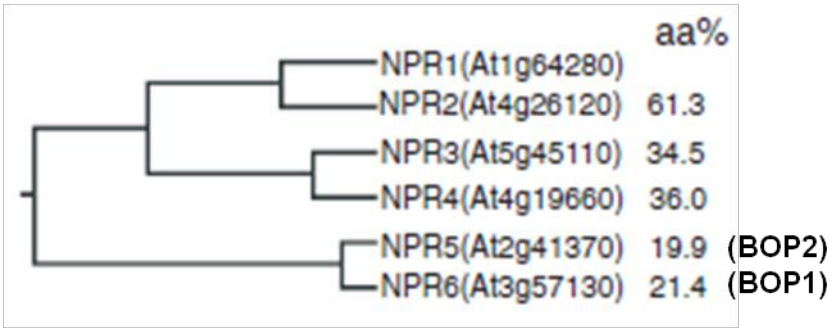


Figure 1-1 Phylogenic tree of six *NPR1-like* gene in *Arabidopsis* (adapted from Liu et al., 2005). The amino acid identity to NPR1 protein is shown as percentage (aa%). BOP: BLADE-ON-PETIOLE.

There are another five *NPR1-like* gene identified in *Arabidopsis* genome, named *NPR1-like 2* (*NPR2*), *NPR3*, *NPR4*, *BLADE-ON-PETIOLE 2* (*BOP2*; also named *NPR5*) and *BOP2* (also named *NPR6*) (Hepworth et al., 2005; Liu et al., 2005). Each *NPR1-like* gene was given a name based on the similarity to NPR1, for instance, NPR2 shares the highest similarity (61.3% identity), followed by NPR3, NPR4 and so on. Phylogenetic studies showed that they form three distinct clades and there is high homology within each pair (Figure 1-1). Similar to the NPR1 protein, each of the five predicted proteins contains BTB/POZ domain at N-terminus followed by the putative ankyrin repeats and C-terminal NLS, and they also possess a set of highly conserved cysteine residues that are thought to be involved in redox regulation(Hepworth et al., 2005; Norberg et al., 2005).

1.4.1 NPR2, NPR3 and NPR4

Although NPR2 exhibits the highest similarity with NPR1, very little is known about it and no peer reviewed manuscript describing its function has been published to date. A poster abstract was published describing NPR2 as a regulator for NPR1-dependent and

NPR1-independent PR expression (Racki et al., 2003). According to this report, *Arabidopsis npr2* mutant shows no SA-dependent 3-1,3- glucanase (*BGL2*) expression and it displays constitutive *PR1* expression, indicating NPR2 is involved in SA/pathogen-dependent *BGL* expression. Although *npr2* mutant itself does not respond to JA treatment, *npr1 npr2* double mutant exhibits enhanced *PDF1.2* gene expression, indicating the role of NPR2 in SA/JA signaling cross-talk and NPR1-dependent pathway.

Two independent studies suggest NPR3 and NPR4 play roles in plant defense response, though their results are somewhat contradictory (Liu et al., 2005; Zhang et al., 2006). An *Arabidopsis npr4* T-DNA mutant exhibits more susceptibility than wild type against the *Pseudomonas syringe pv. tomato (P.s.t.)* DC3000 and the fungal pathogen *Erysiphe cichoracearum* (Liu et al., 2005), and the mutant is dramatically affected in the SA-induced priming, resulting in slower *PR1* accumulation compared to wild type if treating plants with SA 24 hours prior to pathogen infection. In addition, the authors suggested that NPR4 can interact with the same spectrum of TGA transcription factors as NPR1 and it predominantly localizes in the nucleus of transiently transformed onion cells. It was also demonstrated that NPR4, as a positive regulator of SA signaling, might be involved in mediating the cross-talk between JA and SA pathways. However, the findings that the basal level of *PR1* is dramatically elevated in *npr3 npr4* double mutant and that *npr3npr4* double mutant shows enhanced resistance against virulent *Pseudomonas syringae* pv. *maculicola* ES4326(*P.s.m.*) and oomycete pathogen *Hyaloperonospora parasitica* Noco2 (Zhang et al., 2006) indicate that NPR3 and NPR4 play negative roles in defense response, which is inconsistent with previous results. It is suggested that the enhanced disease resistance observed in *npr3 npr4* double mutants is partially dependent on NPR1, as PR1 up-regulation in the double mutant is partly compromised in *npr1 npr3 npr4* triple mutant. Moreover, NPR3 was shown to interact with a set of TGA transcription factors in yeast two hybrid assays and bimolecular fluorescence complementation test (BiFC) suggested that NPR3 and TGA2 interact mainly in the nucleus of transiently transformed onion epidermal cells.

1.4.2 BLADE-ON-PETIOLE 1(BOP1) and BOP2

Consistent with their position as the farthest group from NPR1 on the phylogenetic tree, no evidence implicates that BOP1 and BOP2 are involved in defense response. However, a large amount of data has demonstrated that they function in vegetative development in plants (Ha et al., 2003; Ha et al., 2004; Hepworth et al., 2005; Norberg et al., 2005; Ha et al., 2007; McKim et al., 2008; Jun et al., 2010). BOP1 and BOP2 encode two proteins sharing 81% amino acid identity, which suggests that their function may be highly redundant in terms of leaf development. *BOP* genes are expressed in proximal parts of lateral organs during vegetative and reproductive development. No visible phenotype was found for the *bop2* single mutant and only a weak phenotype could be detected for *bop1* in short day conditions with ectopic leaves on rosette leaves. The double mutant *bop1 bop2* showed delayed growth and ectopic growth of blades, thus forming larger leaves without petioles. Furthermore, the flowers of this double mutant are often surrounded by bract-like organs and they have an open structure with the abaxial sepal missing (Hepworth et al., 2005; Norberg et al., 2005). It is believed that the formation of bracts in *bop1 bop2* was partially due to loss of suppression of class 1 *knox* genes by *BOP* genes, indicating the importance of BOPs in the leaf development signal pathway. BOP2 is localized in both the nucleus and cytoplasm and it can interact with the TGA factor PERIANTHIA (PAN), suggesting that BOP proteins control leaf development via direct interaction with a TGA factor, in a manner similar to the NPR1 signaling mechanism (Hepworth et al., 2005). It also has been demonstrated BOP1 and BOP2 regulate the spatial expression of a set of genes involved in organ polarity determination to specify adaxial-abaxial organ identity, such as direct activation of *ASYMMETRIC LEAVES2* (*AS2*) and suppression of the ectopic expression of abaxial identity gene *KANADI 1* (*KAN1*) (Ha et al., 2007; Jun et al., 2010). RBOP proteins are essential to establish abscission zones in flower tissues and cauline leaves and that they are required for the cell differentiation in the proximal regions of inflorescence lateral organs (McKim et al., 2008). The precise degree to which NPR- and BOP-mediated signal transduction pathways share similar mechanisms remains to be investigated.

1.5 *NPR1-like* Genes in Other Species

A large amount of evidence supports the conclusion that the SA-dependent, NPR1-mediated defense pathway is also conserved in other plant species across wide phylogenetic distances. Transgenic tomato plants expressing *Arabidopsis NPR1* (*AtNPR1*) display enhanced resistance to four important tropical disease, including bacterial wilt (BW) and gray leaf spot (GLS) (Lin et al., 2004). Similarly, overexpression of *AtNPR1* in wheat, results in the increased resistance to Fusarium head blight in susceptible cultivar Bobwhite (Makandar et al., 2006).

Recently, homologs of *Arabidopsis NPR1* have been cloned and characterized in several species, including grape (Le Henanff et al., 2009), apple (Malnoy et al., 2007), banana (Endah et al., 2008), cotton (Zhang et al., 2008), and rice (Yuan et al., 2007). Two *NPR1-like* genes from *Vitis vinifer*, *VvNPR1.1* and *VvNPR1.2* have been isolated, which show 55% and 40% identity to AtNPR1 respectively (Le Henanff et al., 2009). Similar to AtNPR1, the expression of those two genes in grape leaves is induced by treatment with SA. In addition, VvNPR1.1/1.2-GFP fusion proteins were shown to be predominantly localized in nucleus and triggered the accumulation of PR1 and PR2 proteins when transiently expressed in *Nicotiana benthamiana* leaves. An *NPR1* homolog, *MpNPR1-1* was cloned from apple (Malus x domestica) (Malnoy et al., 2007). Overexpression of this gene in two apple cultivars, Galaxy and M26, confers *PR* genes activation and significant reduction in bacterial and fungal pathogen growth. In tobacco and tomato, the down regulation of *NPR1-like* gene results in enhanced susceptibility to tobacco mosaic virus and bacterial wilt, respectively (Liu et al., 2002; Chen et al., 2009). Two *NPR1-like* genes in banana (*MNPR1A* and *MNPR1B*) are differentially expressed in response to SA treatment and pathogen infection, but both of them are shown to be directly related to *PR* gene expression (Endah et al., 2008). In cotton, GhNPR1 shares similar gene structure and 53% protein identity to AtNPR1 (Zhang et al., 2008). *GhNPR1* is constitutively expressed in all tissues and can be induced by signaling molecules (SA, MeJA and ET) as well as by pathogen attack, in a manner similar to *AtNPR1*, indicating GhNPR1 may play important roles in defense response in cotton.

Major efforts have been made to elucidate the NPR1-dependent signaling pathway in rice. The facts that rice plants display induced resistance against rice blast fungus after

SA analog probenazole treatment (Sakamoto et al., 1999) and that rice plants expressing bacterial salicylate hydrolase (*nahG*) exhibit increased susceptibility to rice blast (Yang et al., 2004), along with the evidence that overexpression of the *Arabidopsis NPR1* gene in rice leads to enhanced resistance to the bacterial pathogen *Xanthomonas oryzae* pv. *oryzae* (Fitzgerald et al., 2004) indicate NPR1-mediated defense mechanism is also conserved in monocots. An NPR1 ortholog (*OsNPR1/NH1*) has been cloned from rice, whose overexpression in rice confers enhanced resistance to bacterial and oomycete pathogens (Chern et al., 2005). Furthermore, OsNPR1 is able to restore the phenotype of *Arabidopsis npr1-1* mutant in a complementation assay (Yuan et al., 2007). These authors observed that *OsNPR1* is constitutively expressed and stays in cytoplasm in the absence of pathogen and it can be translocated into nucleus after the application of a reducing agent, suggesting it may share the same activation mechanism as AtNPR1. Protein sequence analysis reveals that two cysteines (C82 and C216 in AtNPR1), which are essential for the redox regulation of AtNPR1, are also conserved in OSNPR1, and the mutation of these two cycteines results in the constitutive nuclear localization of the protein. Taken together, the SA- and NPR1-dependent defense signaling is likely to be conserved through all plant lineages.

1.6 The Tropical Tree *Theobroma cacao* and Its Diseases

Theobroma cacao L is a small tropical evergreen tree in the family of Malvaceae, native to the deep tropical region of Americas. Its seeds are used to make chocolate and cocoa powder, providing an income for small-scale farmers in central and southern America, the Caribbean, Malaysia, Indonesia and West Africa. The study of cacao genetics shows that it originated in Amazon and distributed by humans to Central America and other places (Motamayor et al., 2002; McNeil, 2006). The tree is originally found growing in the Amazon and Orinoco river basins at elevation about 200-400m (Wood and Lass, 1985). It is an understory tree, growing best under some shade and it requires high humidity with regular rainfall. The leaves are alternate with 10-40cm long and 5-20 cm width. The flowers are very small and produced on the main trunk and older branches. Its pods are ovoid and turn yellow to orange during ripening. Three major cultivar groups of cacao are used to produce cocoa and chocolate. The most rare and expensive one is Criollo

group, because of its less bitter and more aromatic taste than other beans. Only 10% of chocolate is made from Criollo. About 80% of chocolate is made from Forastero Group, as its yield and resistance to disease is significantly higher than Criollo trees. Trinitario, a hybrid of Criollo and Forastero, is used in about 10% of chocolate.

Pathogens are a major problem of cacao production, and about 30-40% world production is lost annually (Evans, 2007; Hebbar, 2007; Ploetz, 2007; Argout et al., 2008). In Amazon basin, its centre of diversity, cacao is susceptible to several potentially destructive pathogens, such as fungal pathogen *Moniliophthora perniciosa*, the causal agent of witches' broom disease (Purdy and Schmidt, 1996), fungal pathogen *Moniliophthora roreri*, the causal agent of frosty pod rot (Wood and Lass, 1985; Evans and Priori, 1987; Phillips-Mora and Wilkinson, 2007), and oomycete pathogen *Phytophthora* spp., the causal agent of black pod disease (Bailey et al., 2005; Bailey et al., 2005). Cacao is also susceptible to a number of contagious pathogens, such as fungal pathogen *Ceratocystis cacaofunesta*, the causal agent of Ceratocystis wilt and fungal pathogen *Oncobasidium theobromae*, the causal agent of Vascular streak dieback (Buddenhagen, 1977; Ploetz, 2007; Schnell et al., 2007; Guiltinan et al., 2008).

Evans described the "trilogy of crippling fungal diseases"-black pod, frosty pot rot and witches' broom- that affects cacao production in the tropics (Evans, 2007). Black pod disease accounts for 20-25% world production losses, making it the major constraint to production before the arrival of witches' broom disease in Brazil; it is still a major threat to cacao in West Africa (Watling, 2002). Although frosty pod rot disease is still mainly limited to the Americas, it is considered a potential international threat to cacao production(Phillips-Mora and Wilkinson, 2007). Witches' broom disease, another serious pathogen resulting in huge yield reductions, has been the focus of much research over the past century, but its basic biology and ecology are still not fully understood (Meinhardt et al., 2008; Mondego et al., 2008).

1.7 Defense Response in Cacao tree

The study of cacao-pathogen interaction may give better understanding of how cacao plants respond to pathogens and the results can be used in cacao breeding programs

to select cultivars with higher resistance, and thus enhance cocoa production. Therefore, it is not surprising that efforts have been made to elucidate the defense response in cacao. While witches' broom disease, caused by C-biotype *Moniliophthora perniciosa*, is of major economic importance, the mechanism of infection and plant response are difficult to study because of long life cycle and limited genetic resources of cacao trees (Cronshaw and Evans, 1978; Evans, 1978; Marelli et al., 2009). However, the evolutionarily closely related S-biotype of *M. perniciosa* infects solanaceous hosts, such as tomato, which is more adaptable for studying the host-pathogen interaction compared to cacao (Marelli et al., 2009). Marelli et al. demonstrated that the major symptoms observed in both tomato and cacao are swelling infected shoots and proliferation of axillary meristem after infection with their compatible biotypes. In addition, increased cell size, increased numbers of xylem vessels and phloem parenchyma are also observed in infected stem. The study of *M. perniciosa*-tomato interaction reveals that the fungal colonization is biotrophic during the first phase of infection, with calcium oxalate crystals formation associated with hyphal growth, indicating that the pathogenicity mechanisms and physiological symptoms are conserved between two different biotypes. Moreover, a major quantitative trait loci (QTL) was identified from the F_2 population resulting from the cross of two cacao genotype, 'Scavina-6' (resistant) and 'ICS-1' (susceptible), which explain 35% of the resistance to witches' broom (Queiroz et al., 2003). This QTL has great potential for use in marker-assisted cocoa breeding program, since it is a dominant component of resistance to this pathogen.

Towards the task of reducing the devastating effects of pathogenic diseases on cacao production, a number of genomic and genetic resources have been developed to facilitate researches. These resources include the identification of disease associated QTL on a high-density linkage maps (Bennett, 2003). A study of gene discovery and array analysis identified 1,380 unigenes from over 4,000 expressed sequence tags (ESTs) of untreated cacao tissues, in which 46 sequences have been demonstrated as defense-related genes (Jones et al., 2002). In addition, suppressive subtractive hybridization (SSH) cDNA libraries and microarray analysis have identified 330 unigenes induced in exogenous BTH, MeJA and necrosis-inducing protein NEP1 treatments (Verica et al., 2004). Sequencing and annotation of those genes suggested many may play roles in plant defense, cell wall development, gene regulation and signal transduction. Those defense-related genes are

differentially expressed in different developmental stages (Bailey et al., 2005; Bailey et al., 2005). It is shown that MeJA induces a caffeine synthase (*TcCaf-1*) and a class VII chitinase (*TcChiB*) in young leaves but a type III peroxidase (*TcPer-1*) and *TcChiB* in mature green leaves, suggesting that cacao response to signaling molecules was greatly affected by developmental stages and it is important to take developmental stages into consideration when studying cacao defense response. To further confirm the function of genes isolated from large scale screening, cacao transformation system has been applied to overexpress class I chitinase gene (*TcChi1*) (Maximova et al., 2006). The resulting transgenic cacao plants exhibit reduced susceptibility to the fungal pathogen *Colletotrichum gloeosporioides* compared to control plants. Recently, an ornithine decarboxylase (*TcODC*) and an arginine decarboxylase (*TcADC*), associated with drought stress, were shown to be responsive to fungal protein NEP1 treatment and *Phytophthora megakarya* infection (Bae et al., 2008), indicating those proteins may be involved in convergence of multiple stresses signaling. Furthermore, genes sharing sequence homology with known *PR* genes have been identified from cacao and some of them are responsive to SA and BTH induction. In summary, cacao has evolved specific defense response against different offending pathogens and recent evidence indicates that cacao may possess SAR pathway as well; however, the extent of conservation of this pathway is still unknown and more work is required to elucidate the details.

The research presented in this thesis was conducted in an effort to further understand the function of *Arabidopsis* NPR3 in defense response and whole plant fitness, which suggests that NPR3 negatively regulates the transcription of defense genes during early flower development. Our data indicate the neofunctionalization of NPR3 from the duplication of the NPR ancestor and the evolutionary importance of this gene family. Moreover, this thesis begins to dissect the mechanisms of resistance in cacao. Using a translational biology approach, I have identified four NPR1-like genes of cacao (TcNPR1-4) and have functionally tested two of them in transgenic *Arabidopsis* plants. My results show that the NPR1-mediated defense response pathway is highly conserved between *Arabidopsis* and cacao and that both TcNPR1 and TcNPR3 are functional orthologs of *Arabidopsis* native NPR1 and NPR3, acting as regulators of disease resistance in cacao.

These results of this study will be useful for breeding or engineering of cacao varieties with enhanced disease resistance.

1.8 References

Agrawal AA (1998) Induced responses to herbivory and increased plant performance. Science **279:** 1201-1202

Agrawal AA, Strauss SY, Stout MJ (1999) Costs of induced responses and tolerance to herbivory in male and female fitness components of wild radish. Evolution **53:** 1093-1104

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M, Kuhn D, Verica J, Courtois B, Loor G, Babin R, Sounigo O, Ducamp M, Guiltinan MJ, Ruiz M, Alemanno L, Machado R, Phillips W, Schnell R, Gilmour M, Rosenquist E, Butler D, Maximova S, Lanaud C (2008) Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao L.* generated from various tissues and under various conditions. BMC Genomics **9:** 512

Bae H, Kim SH, Kim MS, Sicher RC, Lary D, Strem MD, Natarajan S, Bailey BA (2008) The drought response of Theobroma cacao (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses. Plant Physiol Biochem **46:** 174-188

Bailey BA, Bae H, Strem MD, Antunez de Mayolo G, Guiltinan MJ, Verica JA, Maximova SN, Bowers JH (2005) Developmental expression of stress response genes in *Theobroma cacao* leaves and their response to Nep1 treatment and a compatible infection by *Phytophthora megakarya*. Plant Physiol Biochem **43:** 611-622

Bailey BA, Strem MD, Bae H, de Mayolo GA, Guiltinan MJ (2005) Gene expression in leaves of *Theobroma cacao* in response to mechanical wounding, ethylene, and/or methyl jasmonate. Plant Science **168:** 1247-1258

Balbi V, Devoto A (2008) Jasmonate signalling network in *Arabidopsis thaliana*: crucial regulatory nodes and new physiological scenarios. New Phytol **177:** 301-318

Baldwin AS, Jr. (1996) The NF-kappa B and I kappa B proteins: new discoveries and insights. Annu Rev Immunol **14:** 649-683

Baldwin IT (1998) Jasmonate-induced responses are costly but benefit plants under attack in native populations. Proc Natl Acad Sci U S A **95:** 8113-8118

Beckers GJ, Spoel SH (2006) Fine-tuning plant defence signalling: Salicylate versus jasmonate. Plant Biol **8:** 1-10

Bennett AB (2003) Out of the Amazon: Theobroma cacao enters the genomic era. Trends Plant Sci **8:** 561-563

Blanco F, Salinas P, Cecchini NM, Jordana X, Van Hummelen P, Alvarez ME, Holuigue L (2009) Early genomic responses to salicylic acid in *Arabidopsis*. Plant Mol Biol **70:** 79-102

Bostock RM (2005) Signal crosstalk and induced resistance: straddling the line between cost and benefit. Annu Rev Phytopathol **43:** 545-580

Buddenhagen IW (1977) Resistance and Vulnerability of Tropical Crops in Relation to Their Evolution and Breeding. Annals of the New York Academy of Sciences **287:** 309-326

Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an Arabidopsis mutant that is nonresponsive to inducers of systemic acquired resistance. Plant Cell **6:** 1583-1592

Cao H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The *Arabidopsis NPR1* gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell **88:** 57-63

Chen YY, Lin YM, Chao TC, Wang JF, Liu AC, Ho FI, Cheng CP (2009) Virus-induced gene silencing reveals the involvement of ethylene-, salicylic acid- and mitogen-activated protein kinase-related defense pathways in the resistance of tomato to bacterial wilt. Physiol Plant **136:** 324-335

Chern M, Fitzgerald HA, Canlas PE, Navarre DA, Ronald PC (2005) Overexpression of a rice *NPR1* homolog leads to constitutive activation of defense response and hypersensitivity to light. Mol Plant Microbe Interact **18:** 511-520

Cipollini DF (2002) Does competition magnify the fitness costs of induced responses in Arabidopsis thaliana? A manipulative approach. Oecologia **131:** 514-520

Cronshaw DK, Evans HC (1978) Witches-broom disease of cocoa *Crinipellis-perniciosa* in Ecuador Part 2 methods of infection. Annals of Applied Biology **89:** 193-200

Delaney TP, Friedrich L, Ryals JA (1995) *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. Proc Natl Acad Sci U S A **92:** 6602-6606

Despres C, DeLong C, Glaze S, Liu E, Fobert PR (2000) The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. Plant Cell **12:** 279-290

Dong X (2004) NPR1, all things considered. Curr Opin Plant Biol **7:** 547-552

Durrant WE, Dong X (2004) Systemic acquired resistance. Annu Rev Phytopathol **42:** 185-209

Endah R, Beyene G, Kiggundu A, van den Berg N, Schluter U, Kunert K, Chikwamba R (2008) Elicitor and *Fusarium*-induced expression of *NPR1-like* genes in banana. Plant Physiol Biochem **46:** 1007-1014

Evans H, Priori C (1987) Cocoa pod diseases: causal agents and control. Outlook on Agriculture **16:** 35-41

Evans HC (1978) Witches-broom disease of cocoa *Crinipellis-perniciosa* in Ecuador Part 1 the fungus. Annals of Applied Biology **89:** 185-192

Evans HC (2007) Cacao diseases-the trilogy revisited. Phytopathology **97:** 1640-1643

Fan W, Dong X (2002) In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis. Plant Cell **14:** 1377-1389

Feys BJ, Parker JE (2000) Interplay of signaling pathways in plant disease resistance. Trends Genet **16:** 449-455

Fitzgerald HA, Chern MS, Navarre R, Ronald PC (2004) Overexpression of (At)NPR1 in rice leads to a BTH- and environment-induced lesion-mimic/cell death phenotype. Mol Plant Microbe Interact **17:** 140-151

Ghosh S, May MJ, Kopp EB (1998) NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol **16:** 225-260

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

Glazebrook J, Rogers EE, Ausubel FM (1996) Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening. Genetics **143:** 973-982
Grant M, Lamb C (2006) Systemic immunity. Curr Opin Plant Biol **9:** 414-420
Guiltinan M, Verica J, Zhang D, Figueira A (2008) Genomics of *Theobroma cacao*, "The foods of the Gods". *In* P.M. Moore, R Ming, eds, Genomics of Tropical Crop Plants. New York: Springer
Ha CM, Jun JH, Nam HG, Fletcher JC (2004) *BLADE-ON-PETIOLE1* encodes a BTB/POZ domain protein required for leaf morphogenesis in *Arabidopsis thaliana*. Plant Cell Physiol **45:** 1361-1370
Ha CM, Jun JH, Nam HG, Fletcher JC (2007) BLADE-ON-PETIOLE1 and 2 control Arabidopsis lateral organ fate through regulation of LOB domain and adaxial-abaxial polarity genes. Plant Cell **19:** 1809-1825
Ha CM, Kim GT, Kim BC, Jun JH, Soh MS, Ueno Y, Machida Y, Tsukaya H, Nam HG (2003) The *BLADE-ON-PETIOLE 1* gene controls leaf pattern formation through the modulation of meristematic activity in *Arabidopsis*. Development **130:** 161-172
Hebbar PK (2007) Cacao diseases: a global perspective from an industry point of view. Phytopathology **97:** 1658-1663
Heidel AJ, Clarke JD, Antonovics J, Dong X (2004) Fitness Costs of Mutations Affecting the Systemic Acquired Resistance Pathway in *Arabidopsis thaliana*. Genetics **168:** 2197-2206
Heil M (2002) Ecological costs of induced resistance. Curr Opin Plant Biol **5:** 345-350
Heil M, Hilpert A, Kaiser W, Linsenmair KE (2000) Reduced Growth and Seed Set Following Chemical Induction of Pathogen Defence: Does Systemic Acquired Resistance (SAR) Incur Allocation Costs? Journal of Ecology **88:** 645-654
Hepworth SR, Zhang Y, McKim S, Li X, Haughn GW (2005) BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. Plant Cell **17:** 1434-1448
Johnson C, Boden E, Arias J (2003) Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. Plant Cell **15:** 1846-1858
Johnson C, Mhatre A, Arias J (2008) NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2. Biochim Biophys Acta **1779:** 583-589
Jones PG, Allaway D, Gilmour DM, Harris C, Rankin D, Retzel ER, Jones CA (2002) Gene discovery and microarray analysis of cacao (*Theobroma cacao L.*) varieties. Planta **216:** 255-264
Jun JH, Ha CM, Fletcher JC (2010) BLADE-ON-PETIOLE1 coordinates organ determinacy and axial polarity in Arabidopsis by directly activating ASYMMETRIC LEAVES2. Plant Cell **22:** 62-76
Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. Annu Rev Plant Biol **53:** 299-328
Kinkema M, Fan W, Dong X (2000) Nuclear localization of NPR1 is required for activation of *PR* gene expression. Plant Cell **12:** 2339-2350
Koornneef A, Pieterse CM (2008) Cross talk in defense signaling. Plant Physiol **146:** 839-844
Kunkel BN, Brooks DM (2002) Cross talk between signaling pathways in pathogen defense. Curr Opin Plant Biol **5:** 325-331

Laurie-Berry N, Joardar V, Street IH, Kunkel BN (2006) The *Arabidopsis thaliana JASMONATE INSENSITIVE 1* gene is required for suppression of salicylic acid-dependent defenses during infection by *Pseudomonas syringae*. Mol Plant Microbe Interact **19:** 789-800

Le Henanff G, Heitz T, Mestre P, Mutterer J, Walter B, Chong J (2009) Characterization of *Vitis vinifera NPR1* homologs involved in the regulation of pathogenesis-related gene expression. BMC Plant Biol **9:** 54

Leon-Reyes A, Spoel SH, De Lange ES, Abe H, Kobayashi M, Tsuda S, Millenaar FF, Welschen RA, Ritsema T, Pieterse CM (2009) Ethylene modulates the role of NONEXPRESSOR OF PATHOGENESIS-RELATED GENES1 in cross talk between salicylate and jasmonate signaling. Plant Physiol **149:** 1797-1809

Lin WC, Lu CF, Wu JW, Cheng ML, Lin YM, Yang NS, Black L, Green SK, Wang JF, Cheng CP (2004) Transgenic tomato plants expressing the *Arabidopsis NPR1* gene display enhanced resistance to a spectrum of fungal and bacterial diseases. Transgenic Res **13:** 567-581

Liu G, Holub EB, Alonso JM, Ecker JR, Fobert PR (2005) An *Arabidopsis NPR1-like* gene, *NPR4*, is required for disease resistance. Plant J **41:** 304-318

Liu Y, Schiff M, Marathe R, Dinesh-Kumar SP (2002) Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for N-mediated resistance to tobacco mosaic virus. Plant J **30:** 415-429

Loake G, Grant M (2007) Salicylic acid in plant defence--the players and protagonists. Curr Opin Plant Biol **10:** 466-472

Loon LCv, Strien EAv (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol.Mol.Plant Pathol. **55:** 85-97

Makandar R, Essig JS, Schapaugh MA, Trick HN, Shah J (2006) Genetically engineered resistance to *Fusarium* head blight in wheat by expression of *Arabidopsis NPR1*. Mol Plant Microbe Interact **19:** 123-129

Malnoy M, Jin Q, Borejsza-Wysocka EE, He SY, Aldwinckle HS (2007) Overexpression of the apple *MpNPR1* gene confers increased disease resistance in *Malus* x *domestica*. Mol Plant Microbe Interact **20:** 1568-1580

Marelli J-P, Maximova S, Gramacho K, Kang S, Guiltinan M (2009) Infection Biology of *Moniliophthora perniciosa* on *Theobroma cacao* and Alternate Solanaceous Hosts. Tropical Plant Biology **2:** 149-160

Maximova SN, Marelli JP, Young A, Pishak S, Verica JA, Guiltinan MJ (2006) Over-expression of a cacao class I chitinase gene in *Theobroma cacao L.* enhances resistance against the pathogen, *Colletotrichum gloeosporioides*. Planta **224:** 740-749

McKim SM, Stenvik GE, Butenko MA, Kristiansen W, Cho SK, Hepworth SR, Aalen RB, Haughn GW (2008) The *BLADE-ON-PETIOLE* genes are essential for abscission zone formation in *Arabidopsis*. Development **135:** 1537-1546

McNeil CL (2006) Chocolate in Mesoamerica : a cultural history of cacao. University Press of Florida, Gainesville

Meinhardt LW, Rincones J, Bailey BA, Aime MC, Griffith GW, Zhang D, Pereira GA (2008) *Moniliophthora perniciosa*, the causal agent of witches' broom disease of cacao: what's new from this old foe? Mol Plant Pathol **9:** 577-588

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, Garcia O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araujo MR, de Moraes MH, Castro LA, Gramacho KP, Goncalves MS, Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA (2008) A genome survey of *Moniliophthora perniciosa* gives new insights into Witches' Broom Disease of cacao. BMC Genomics **9:** 548

Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, Lanaud C (2002) Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity **89:** 380-386

Mou Z, Fan W, Dong X (2003) Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. Cell **113:** 935-944

Mukhtar MS, Nishimura MT, Dangl J (2009) NPR1 in plant defense: It's not over 'til It's turned over. Cell **137:** 804-806

Ndamukong I, Abdallat AA, 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

Norberg M, Holmlund M, Nilsson O (2005) The *BLADE ON PETIOLE* genes act redundantly to control the growth and development of lateral organs. Development **132:** 2203-2213

Phillips-Mora W, Wilkinson MJ (2007) Frosty pod of cacao: a disease with a limited geographic range but unlimited potential for damage. Phytopathology **97:** 1644-1647

Pieterse CM, Van Loon LC (2004) NPR1: the spider in the web of induced resistance signaling pathways. Curr Opin Plant Biol **7:** 456-464

Pieterse CM, van Wees SC, Hoffland E, van Pelt JA, van Loon LC (1996) Systemic resistance in Arabidopsis induced by biocontrol bacteria is independent of salicylic acid accumulation and pathogenesis-related gene expression. Plant Cell **8:** 1225-1237

Ploetz RC (2007) Cacao diseases: important threats to chocolate production worldwide. Phytopathology **97:** 1634-1639

Poelman EH, Broekgaarden C, Van Loon JJ, Dicke M (2008) Early season herbivore differentially affects plant defence responses to subsequently colonizing herbivores and their abundance in the field. Mol Ecol **17:** 3352-3365

Purdy LH, Schmidt RA (1996) STATUS OF CACAO WITCHES' BROOM: biology, epidemiology, and management. Annu Rev Phytopathol **34:** 573-594

Queiroz VT, Guimaraes CT, Anhert D, Schuster I, Daher RT, Pereira MG, Miranda VRM, Loguercio LL, Barros EG, Moreira MA (2003) Identification of a major QTL in cocoa (*Theobroma cacao L.*) associated with resistance to witches' broom disease. Plant Breeding **122:** 268-272

Racki LR, Yorgey P, Dewdney J, Ausubel FM (2003) *NPR2*: A Novel *Arabidopsis* Defense Response Gene. Poster, Plant Biology 2003, Honolulu, Hawaii USA

Rayapuram C, Baldwin IT (2007) Increased SA in NPR1-silenced plants antagonizes JA and JA-dependent direct and indirect defenses in herbivore-attacked *Nicotiana attenuata* in nature. Plant J **52:** 700-715

Rochon A, Boyle P, Wignes T, Fobert PR, Despres C (2006) The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. Plant Cell **18:** 3670-3685

Ryals J, Weymann K, Lawton K, Friedrich L, Ellis D, Steiner HY, Johnson J, Delaney TP, Jesse T, Vos P, Uknes S (1997) The Arabidopsis NIM1 protein shows homology to the mammalian transcription factor inhibitor I kappa B. Plant Cell **9:** 425-439

Sakamoto K, Tada Y, Yokozeki Y, Akagi H, Hayashi N, Fujimura T, Ichikawa N (1999) Chemical induction of disease resistance in rice is correlated with the expression of a gene encoding a nucleotide binding site and leucine-rich repeats. Plant Mol Biol **40:** 847-855

Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W, Amores FM, Motamayor JC (2007) Development of a marker assisted selection program for cacao. Phytopathology **97:** 1664-1669

Shah J, Tsui F, Klessig DF (1997) Characterization of a salicylic acid-insensitive mutant (*sai1*) of *Arabidopsis thaliana*, identified in a selective screen utilizing the SA-inducible expression of the *tms2* gene. Mol Plant Microbe Interact **10:** 69-78

Spoel SH, Dong X (2008) Making sense of hormone crosstalk during plant immune responses. Cell Host Microbe **3:** 348-351

Spoel SH, Johnson JS, Dong X (2007) Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. Proc Natl Acad Sci U S A **104:** 18842-18847

Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Metraux JP, Brown R, Kazan K, Van Loon LC, Dong X, Pieterse CM (2003) NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. Plant Cell **15:** 760-770

Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P, Dong X (2009) Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. Cell **137:** 860-872

Sticher L, Mauch-Mani B, Metraux JP (1997) Systemic acquired resistance. Annu Rev Phytopathol **35:** 235-270

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG (2005) Sequence and structural analysis of BTB domain proteins. Genome Biol **6:** R82

Stout MJ, Thaler JS, Thomma BP (2006) Plant-mediated interactions between pathogenic microorganisms and herbivorous arthropods. Annu Rev Entomol **51:** 663-689

Uppalapati SR, Ishiga Y, Wangdi T, Kunkel BN, Anand A, Mysore KS, Bender CL (2007) The phytotoxin coronatine contributes to pathogen fitness and is required for suppression of salicylic acid accumulation in tomato inoculated with *Pseudomonas syringae* pv. tomato DC3000. Mol Plant Microbe Interact **20:** 955-965

Van der Ent S, Verhagen BW, Van Doorn R, Bakker D, Verlaan MG, Pel MJ, Joosten RG, Proveniers MC, Van Loon LC, Ton J, Pieterse CM (2008) MYB72 is required in early signaling steps of rhizobacteria-induced systemic resistance in Arabidopsis. Plant Physiol **146:** 1293-1304

van Hulten M, Pelser M, van Loon LC, Pieterse CM, Ton J (2006) Costs and benefits of priming for defense in *Arabidopsis*. Proc Natl Acad Sci U S A **103:** 5602-5607

van Wees SC, Glazebrook J (2003) Loss of non-host resistance of *Arabidopsis NahG* to *Pseudomonas syringae* pv. phaseolicola is due to degradation products of salicylic acid. Plant J **33:** 733-742

Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ (2004) Isolation of ESTs from cacao (*Theobroma cacao L.*) leaves treated with inducers of the defense response. Plant Cell Rep **23:** 404-413

Vlot AC, Klessig DF, Park SW (2008) Systemic acquired resistance: the elusive signal(s). Curr Opin Plant Biol **11:** 436-442

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 Pathog **2:** e123

Watling R (2002) Tropical mycology. CABI, Wallingford

Wood G, Lass R (1985) Cocoa. Longman Scientific & Technical Press, New York

Yang Y, Qi M, Mei C (2004) Endogenous salicylic acid protects rice plants from oxidative damage caused by aging as well as biotic and abiotic stress. Plant J **40:** 909-919

Yuan Y, Zhong S, Li Q, Zhu Z, Lou Y, Wang L, Wang J, Wang M, Yang D, He Z (2007) Functional analysis of rice *NPR1-like* genes reveals that *OsNPR1/NH1* is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility. Plant Biotechnol J **5:** 313-324

Zhang Y, Cheng YT, Qu N, Zhao Q, Bi D, Li X (2006) Negative regulation of defense responses in *Arabidopsis* by two *NPR1* paralogs. Plant J **48:** 647-656

Zhang Y, Fan W, Kinkema M, Li X, Dong X (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene. Proc Natl Acad Sci U S A **96:** 6523-6528

Zhang Y, Tessaro MJ, 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 X, Cheng C, Gao Q, Liu J, Guo X (2008) Molecular cloning and characterization of *GhNPR1*, a gene implicated in pathogen responses from cotton (*Gossypium hirsutum L.*). Biosci Rep **28:** 7-14

CHAPTER 2: NPR1-LIKE PROTEIN 3 (NPR3) IS A NEGATIVE REGULATOR OF THE TRANSCRIPTIONAL DEFENSE RESPONSE DURING EARLY FLOWER DEVELOPMENT IN *ARABIDOPSIS*

Zi Shi[a], Siela N. Maximova[b], Yi Liu[a], Joseph Verica[b], and Mark J. Guiltinan[a,b,1]

[a]Huck Institute of Life Sciences, The Pennsylvania State University, University Park, PA 16802

[b]The Department of Horticulture, The Pennsylvania State University, University Park, PA 16802

[1]Corresponding author

2.1 Abstract

Arabidopsis NON-EXPRESSOR OF PR1 (NPR1) is a transcription coactivator that plays a central role in regulating the transcriptional response to plant pathogens. The *NPR* family consists of *NPR1* and five *NPR1-like* genes. Here, we analyzed the role of *NPR1* paralog, *NPR3*, the most highly expressed member of the *NPR* family. Although *NPR3* was expressed in most tissues, expression was strongest in the petals and sepals of developing flowers and declined after flower opening. An *npr3* knockout mutant displayed enhanced resistance to *Pseudomonas syringae* infection of immature flowers. Gene expression analysis revealed that developing flowers of *npr3* mutant had increased levels of basal and induced *PR1* transcript accumulation. In *NPR3* over-expression lines, the extent of bacterial colonization was positively correlated with *NPR3* expression level. Bimolecular fluorescence complementation analysis revealed evidence of protein-protein interactions between NPR3 and both NPR1 and TGA2 *in vivo*. While the NPR1-TGA2 complex became localized in the nucleus after *P. syringae* infection, the NPR3-TGA2 complex did not. When NPR3, TGA2, and NPR1 were co-expressed in transgenic plants, the NPR1-TGA2 heterodimer failed to enter the nucleus after pathogen infection. Thus, we propose that NPR3 represses NPR1-mediated transcription of defense response genes by inhibiting the nuclear localization of NPR1 through direct binding to NPR1 and TGA2.

2.2 Introduction

NPR1 (NON-EXPRESSOR OF PR1) is a transcription coactivator that is a central regulator of the plant defense response (Delaney et al., 1995; Glazebrook et al., 1996; Shah et al., 1997; Beckers and Spoel, 2006; Grant and Lamb, 2006; Loake and Grant, 2007). In *Arabidopsis*, NPR1 is required for the activation or down-regulation of 2248 genes in response to treatment with the defense signaling molecule, salicylic acid (SA) analog, BTH (Wang et al., 2006). Blanco et al. used the Complete Arabidopsis Transcriptome MicroArray (CATMAv2) to demonstrate that 193 genes are rapidly induced in a SA- and NPR1-dependent manner (Blanco et al., 2009). Furthermore, *npr1* mutants exhibit reduced expression of defense response genes and increased susceptibility to pathogens (Cao et al., 1994; Shah et al., 1997). The *NPR1* gene encodes a protein that contains a BTB/POZ domain at the N-terminus and ankyrin-repeats in the central region, both of which mediate protein-protein interactions in animal proteins (Stogios et al., 2005). The NPR1 BTB/POZ domain facilitates dimerization of NPR1, and the ankyrin-repeat region is involved in the interaction with the TGA subfamily of bZIP transcription factors (Zhang et al., 1999).

NPR1 shares sequence similarity of functional features with the human IκBα transcription inhibitors (Ryals et al., 1997), negative regulators of the NF-κB signaling pathway that governs the innate immune response (Ghosh et al., 1998). IκB binds to and inactivates the homo- or heterodimers of transcription factor NF-κB in the cytoplasm of non-stimulated cells via the ankyrin-repeat region (Baldwin, 1996). Upon exposure to a signal that initiates the immune response (e.g. inflammatory cytokines), IκB is degraded, releasing NF-κB. NF-κB then enters the nucleus, where it affects the transcription of genes that function in innate immunity. Whereas human IκBα is a negative regulator of transcription that dissociates from the NF-κB transcription factor upon infection, *Arabidopsis* NPR1 is a transcription coactivator that binds to and activates transcription factors upon infection. Under uninduced conditions, NPR1 is present as inactive oligomers in the cytoplasm of the cell. Upon induction of systemic acquired resistance (SAR), the redox state of the cell is altered, resulting in the reduction of NPR1 to its active monomeric form (Mou et al., 2003; Dong, 2004; Loake and Grant, 2007; Mukhtar et al., 2009). In leaves, monomeric NPR1 migrates to the nucleus, where it interacts with TGA transcription factors

and induces the expression of *PATHOGENESIS-RELATED* (*PR*) genes. Because NPR1 itself has no DNA-binding domains, it does not function as a transcription factor *per se,* but acts via specific associations with the TGA subfamily of transcription factors (Despres et al., 2000; Zhou et al., 2000; Fan and Dong, 2002; Johnson et al., 2003; Ndamukong et al., 2007). An additional level of complexity in the mechanism governing NPR1 function involves the regulation of its half-life via the proteasome-dependent degradation pathway (Mukhtar et al., 2009; Spoel et al., 2009). In the absence of inducers, NPR1 is constantly cleared from the nucleus by the proteasome to inhibit the onset of SAR. Upon pathogen challenge, the phosphorylation of NPR1 assists the recruitment of NPR1 to cullin3-based ubiquitin ligase, as more NPR1 accumulates in *cul3a cul3b* double mutant upon induction. The cullin3-mediated degradation of NPR1 is found to be required for the fulfillment of SAR. Thus, the turnover of coactivator NPR1 is indispensable for both repressing and activating the transcription of target gene in plant defense response.

The *NPR* family of *Arabidopsis* consists of *NPR1* and five *NPR1*-like genes encoding proteins with significant similarity to NPR1, named *NPR1-like 2* (*NPR2*)*, NPR3, NPR4, BLADE-ON-PETIOLE2 (BOP2*; also named *NPR5*)*,* and *BOP1* (also named *NPR6*) (Hepworth et al., 2005; Liu et al., 2005; Norberg et al., 2005). Each of the predicted proteins contains conserved BTB/POZ and ankyrin protein-protein interaction domains (Stogios et al., 2005). In addition, each of the *Arabidopsis NPR* family members contains a set of highly conserved cysteine residues that are thought to be involved in the redox control mechanism (Hepworth et al., 2005).

The functional roles of several of the NPR1-like proteins are starting to emerge. BOP1 and BOP2 have functionally redundant roles in the regulation of organ determinacy and symmetry during leaf morphogenesis (Ha et al., 2004; Hepworth et al., 2005; Norberg et al., 2005; McKim et al., 2008; Jun et al., 2010). The *bop1 bop2* double mutant exhibits leafy petioles, loss of floral abscission, and asymmetric flowers subtended by a bract. Like NPR1, BOP proteins localize in both cytoplasm and nucleus, and interact with the TGA class bZIP transcription factor PERIANTHIA.

Liu et al., 2005 and Zhang et al., 2006 presented conflicting evidence concerning the functions of NPR3 and NPR4, but both studies suggested roles in the plant defense response pathway. An *Arabidopsis npr4* mutant was found to be more susceptible to infection by *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*) and the fungal pathogen *Erysiphe cichoracearum* (Liu et al., 2005) than wild-type controls, suggesting that NPR4 is an activator of the basal defense response. However, the finding that *npr3 npr4* double mutants exhibit elevated basal *PR1* expression and display enhanced resistance to *Pseudomonas syringae* pv. *maculicola* ES4326(*P.s.m.*) and oomycete pathogens (Zhang et al., 2006), suggest that NPR3 and/or NPR4 function as negative regulators of the basal defense response. One limitation of these studies is a focus only on leaf tissue.

We report that NPR3 negatively regulates the transcription of defense response genes in developing flowers. *PR1* is up-regulated in developing flowers of the *npr3* knockout mutant, and this is associated with increased pathogen resistance. We demonstrate that NPR3 interacts with TGA2 *in vivo*. *P.s.t.* infection-dependent nuclear translocation of NPR1 was inhibited in plants overexpressing NPR3. Our data suggest that NPR3 acts as a repressor of NPR1 function via competitive interactions with NPR1 and the bZIP transcription factor TGA2.

2.3 Results

Gene structure and phylogenic analysis of *NPR* family members

NPR1 and five *NPR1* paralogs are highly similar to each other (Figure 1A and Figure S1). *NPR1* through *NPR4* each contain four exons and encode predicted proteins that are similar in size (i.e., 574 to 600 amino acids). *BOP2* (*NPR5*) and *BOP1* (*NPR6*), which each contain two exons and encode predicted proteins that are significantly smaller than those encoded by *NPR1-4*, are more closely related to each other than to *NPR1-NPR4*. All six predicted NPR proteins contain a BTB/POZ domain near their N-terminals and an ankyrin repeats in their central region. Furthermore, the presence of five conserved cysteine residues in all six NPR proteins suggests that these proteins adopt a similar structural conformation (Figure S1).

Two of these conserved cysteines (C82 and C216 in NPR1) play a role in the redox-regulated activation and nuclear localization of NPR1 (Mou et al., 2003). The C-terminal region of NPR1 contains a nuclear localization sequence that directs NPR1 monomers into the nucleus, and five basic amino acids in this region facilitate this localization (Kinkema et al., 2000). An alignment of NPR1 through NPR4 revealed that three of these five basic amino acids are identical in these proteins (Figure S1). In addition, the C-terminal regions of these proteins are rich in the basic amino acids argentine and lysine. These similarities suggest that, like NPR1, NPR2 through NPR4 may also contain functional nuclear localization sequences. The C-terminal regions of BOP2 and BOP1 are rich in the basic amino acid histidine, but do not contain the conserved basic amino acids found in the other NPR proteins.

We analyzed the promoter region (~1.5 kb upstream of the ATG start codon) of the *NPR1* family members for potential cis-acting elements (Figure S2). All six promoter regions contained putative TATA and CAAT boxes. In addition, except for *BOP1*, all other promoters contained at least one copy of the putative W-box (TTGAC), which is required for the SA induction of the *Nicotiana tabacum* (tobacco) class I chitinase gene (Yang et al., 1999). These findings suggest that, like *NPR1*, the other *NPR* family members may also be regulated by SA.

A phylogenetic analysis of NPR1-like proteins from five species using the neighbor-joining algorithm is presented in Figure 1B. The analysis supports the hypothesis that an ancient duplication resulted in the divergence of two main clades, NPR and BOP. The BOP clade appears to have undergone subsequent duplication resulting in BOP1 and BOP2 in *Arabidopsis* and perhaps into three orthologs in rice. The ancestral NPR gene is likely to have undergone a second round of duplication resulting in the NPR1/NPR2 and the NPR3/NPR4 clades. This duplication is likely to have occurred prior to the monocot dicot divergence, as ortholog specific features of the two clades have been conserved between *Arabidopsis* and rice. At some point after the monocot-dicot split, a final round of gene duplication events involving each of the NPR1/NPR2, NPR3/NPR4 and BOP clade progenitor genes, resulting in the current state of three sets of related gene pairs: NPR1-NPR2; NPR3-NPR4, and BOP1-BOP2 in *Arabidopsis*. The fact that all five species possess at

least one copy in each of the three clades indicates that the duplication of this gene family is an ancient event, suggesting the importance of this gene family in defense response and development.

NPR Family Expression Profiles

To begin to understand the function of NPR1-4 during *Arabidopsis* development and disease response, we examined the expression profiles of *NPR* family members using the microarray database resource Genevestigator (Zimmermann et al., 2004; Zimmermann et al., 2005). All six members were constitutively expressed in a wide range of tissues examined, although the levels of expression varied substantially (Figure S3). Based on microarray data alone, the relative expression of *NPR1* was moderate and was relatively high in leaf tissues, especially in senescent, cauline, and adult leaves. The expression of *NPR2* was relatively low compared to that of the other family members, and its expression was highest in senescent leaves. *NPR3* appeared to be the most highly expressed gene in this family, showing strong expression in petals, sepals, and roots. The absolute expression level of *NPR3* transcript in almost all other organs was relatively high. The microarray data suggest that *NPR4* is highly expressed in leaf tissues and roots; however, its expression was relatively low in other organs. *BOP2* (*NPR5*) was strongly expressed in the petiole, pollen, shoot apex, and hypocotyl, but weakly expressed in all other organs, with only 1/5 of the expression level observed in the hypocotyl. The *BOP1* (*NPR6*) gene had the weakest expression. It was expressed at a very low level in most organs, but at a relatively high level in the petiole, consistent with a role in leaf development and validating, in part, the accuracy of the microarray data resource.

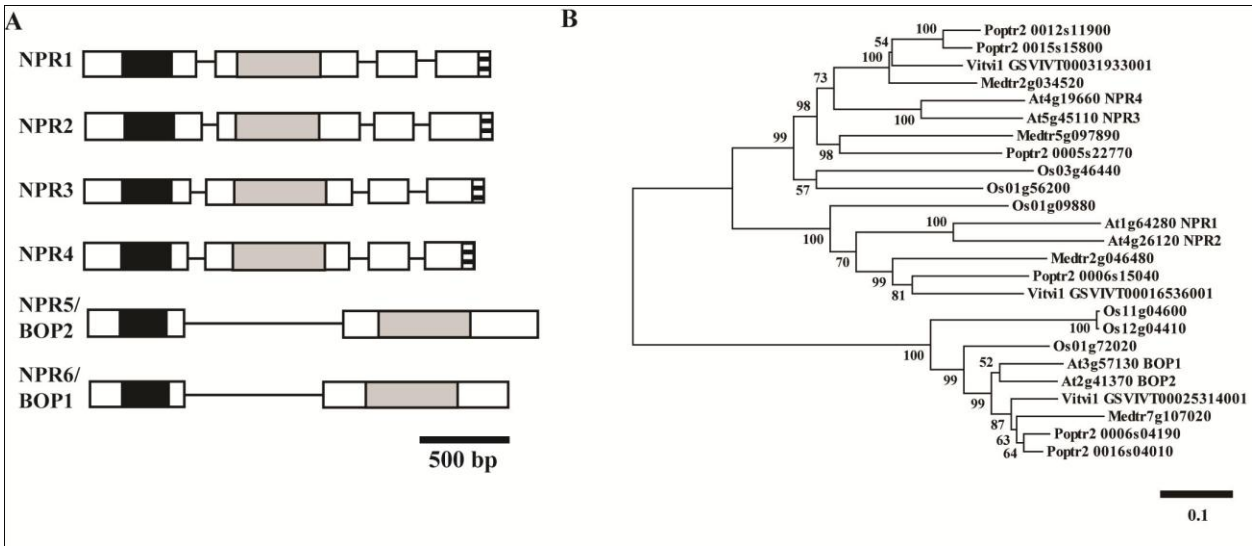


Figure 2-1 Gene structure and phylogenetic analysis of *NPR* family members. A, Graphic representation of the six members of the *Arabidopsis NPR* gene family. Exons are represented by boxes and lines represent introns. Regions with homology to known functional domains are colored (black - BTB/POZ domains; gray - ankyrin repeat regions; horizontally hatched - nuclear localization sequences). B, Phylogenetic analysis of *NPR1* gene family in five species. The tree was constructed based on the full length protein sequences of *NPR1-like* genes in five species A full length protein sequence of *Arabidopsis* NPR1 was used to search the Phytozome database to obtain *NPR1-like* genes from popular, Medicago, grape and rice using the TBLASTN program (E-value cutoff $1e^{-20}$). Multiple protein sequences alignment of total 25 sequences was carried out by MUSCLE software. The phylogenetic tree was constructed with MEGA program using the neighbor-joining method with the Poisson correction distance. The scale bar represents 0.1 substitutions per site and the values next to the nodes are the bootstrap values from 2,000 replicates.

Confirmation of microarray data mining: Promoter-GFP fusions

As a second approach to gain information regarding functions of each of the *NPR* family members we constructed transcriptional fusions with the promoters of each *NPR* family member driving the fluorescent reporter gene EGFP (designated $NPR1_{pro}$:EGFP through $BOP1_{pro}$:EGFP). Transgenic *Arabidopsis* plants harboring these transgenes were observed by confocal microscopy. In leaves, the *NPR1* and *NPR2* promoters drove expression of

EGFP only to relatively low levels. Expression from the *NPR3* promoter was intermediate, and the *NPR4, BOP2* and *BOP1* promoters promoted the highest levels of EGFP accumulation (Figure 2A). GFP expression was strongest in the leaves of plants transformed with $BOP1_{pro}$:EGFP, which had expression levels comparable to those of plants transformed with $CaMV35S_{pro}$:EGFP. The results show that all *NPR* genes expressed to some degree in leaves.

Using low magnification fluorescence stereomicroscopy, we evaluated the expression pattern of the *NPR* promoters in whole organs (Figure 2C). In contrast to the microarray results, we can only observe *NPR1* and *NPR2* promoter-driven expression in leaves under confocal microscope. Consistent with the microarray data, the *NPR4* promoter fusion produced a relatively high level of GFP expression in mature siliques and roots, but low to undetectable expression in other tissues. Also in agreement with the microarray data, GFP was detected at high levels in the hypocotyls and shoot apical meristems of plants transformed with $BOP2_{pro}$:EGFP, and in the petioles, which is consistent with *BOP2* having a function in leaf development (Ha et al., 2004). In addition to the relatively high level of activity in leaves (Figure 2A), the *BOP1* promoter displayed moderate activity in siliques (Figure 2C), which was not reflected in the microarray data.

Modulation of NPR3 Expression at Flower Maturation

Microarray data suggested that *NPR3* had a very high expression level in flower tissues. To verify this, we examined the expression of each promoter fusion in flowers using epi-fluorescence microscopy. Flowers of plants transformed with $NPR1_{pro}$:EGFP and $NPR2_{pro}$:EGFP showed no fluorescence in flowers, and those of plants transformed with $NPR4_{pro}$:EGFP and $BOP2_{pro}$: EGFP exhibited only slight fluorescence (1000-ms exposure, Figure 2B). The $BOP1_{pro}$:EGFP transformants exhibited weak tissue-specific expression in the vasculature of petals. As expected, both the opened and developing flower petals of $CaMV35S_{pro}$:EGFP plants exhibited high levels of fluorescence. However, the petals of plants transformed with $NPR3_{pro}$:EGFP exhibited high levels of fluorescence only in the unopened state, and low levels of fluorescence in the mature, open state, chiefly in the vasculature of

petals. This tissue- and stage-specific expression pattern of $NPR3_{pro}$:EGFP is consistent with the microarray data, and suggested a specific function of NPR3 in early flower development.

Quantitative PCR measurements of *NPR3* transcript levels

Q-PCR confirmed that *NPR3* expression was highest in developing flowers (Figure 2D) and was two to three-fold lower in RNA extracts from opened flowers consistent with the transcriptional fusion experiment (Figure 2B). In general the Q-PCR and microarray data were in agreement, except that *NPR3* expression in roots appeared relatively low in Q-PCR (Figure 2D), but it was shown to be high level in microarray (Figure S3).

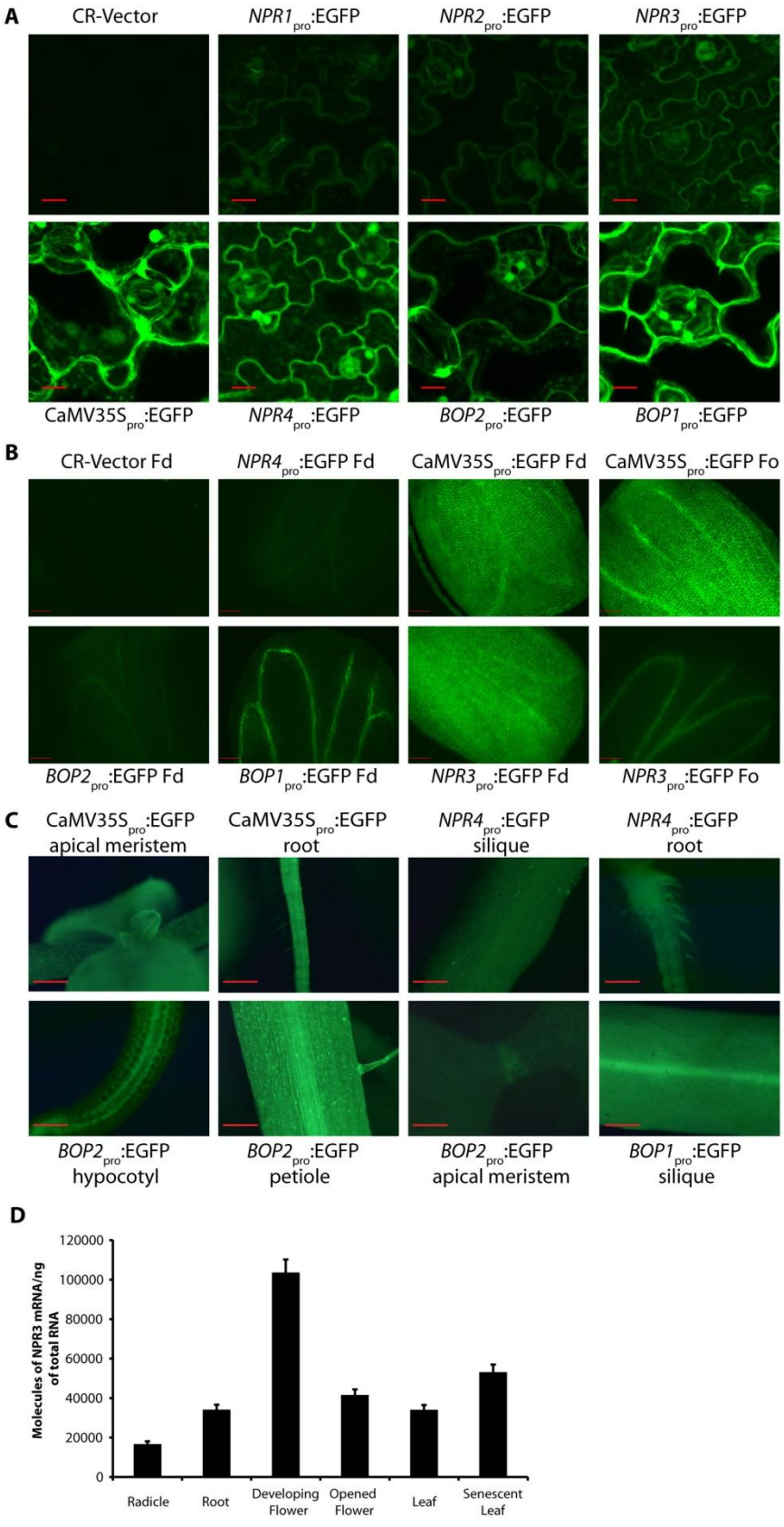

A
CR-Vector
NPR1pro:EGFP
NPR2pro:EGFP
NPR3pro:EGFP
CaMV35Spro:EGFP
NPR4pro:EGFP
BOP2pro:EGFP
BOP1pro:EGFP
B
CR-Vector Fd
NPR4pro:EGFP Fd
CaMV35Spro:EGFP Fd
CaMV35Spro:EGFP Fo
BOP2pro:EGFP Fd
BOP1pro:EGFP Fd
NPR3pro:EGFP Fd
NPR3pro:EGFP Fo
C
CaMV35Spro:EGFP apical meristem
CaMV35Spro:EGFP root
NPR4pro:EGFP silique
NPR4pro:EGFP root
BOP2pro:EGFP hypocotyl
BOP2pro:EGFP petiole
BOP2pro:EGFP apical meristem
BOP1pro:EGFP silique
D
Molecules of NPR3 mRNA/ng of total RNA
120000
100000
80000
60000
40000
20000
0
Radicle
Root
Developing Flower
Opened Flower
Leaf
Senescent Leaf

Figure 2-2 EGFP localization of NPR promoters:EGFP transgenic plants and Q-PCR analysis of *NPR3* expression. EGFP expression in leaves (A) and petals (B) of *Arabidopsis* plants transformed with $NPR1_{pro}$:EGFP through $BOP1_{pro}$:EGFP observed by confocal microscope (A) and epi-fluorescence microscope (B). Scale bars represent 10 μm (A) and 100 μm (B). Fd, developing flower; Fo, opened flowers. C, EGFP expression in various tissues, including apical meristems, roots, siliques, hypocotyls and petioles, of plants transformed with promoter:EGFP fusions observed by stereomicroscope. Scale bars represent 250 μm. D, Q-PCR analysis of *NPR3* expression in six different tissues. Absolute copy number of *NPR3* transcripts was determined for each ng of total RNA. Each data point represents the mean ± SE of three biological replicates, each containing tissue from five individual plants.

Identification of *npr3* insertional mutants

To further explore the function of *NPR3*, we utilized the homozygous T-DNA insertion mutants of *NPR3* SALK_043055 and SALK_009990, also known as *npr3-2* and *npr3-3*. The *npr3-2* and *npr3-3* lines each carry a single insertion at slightly different positions in the third exon of *NPR3* (Figure 3A). After three backcrosses to wild-type Col-0, plants with T-DNA insertions were allowed to self to obtain homozygous mutant lines. Q-PCR analysis showed that the levels of *NPR3* transcript in the developing flowers of *npr3-2* and *npr3-3* mutants were approximately five-fold less and 17-fold less (Figure 3B), respectively, than in wild-type flowers, suggesting that the *npr3-2* is a weak allele.

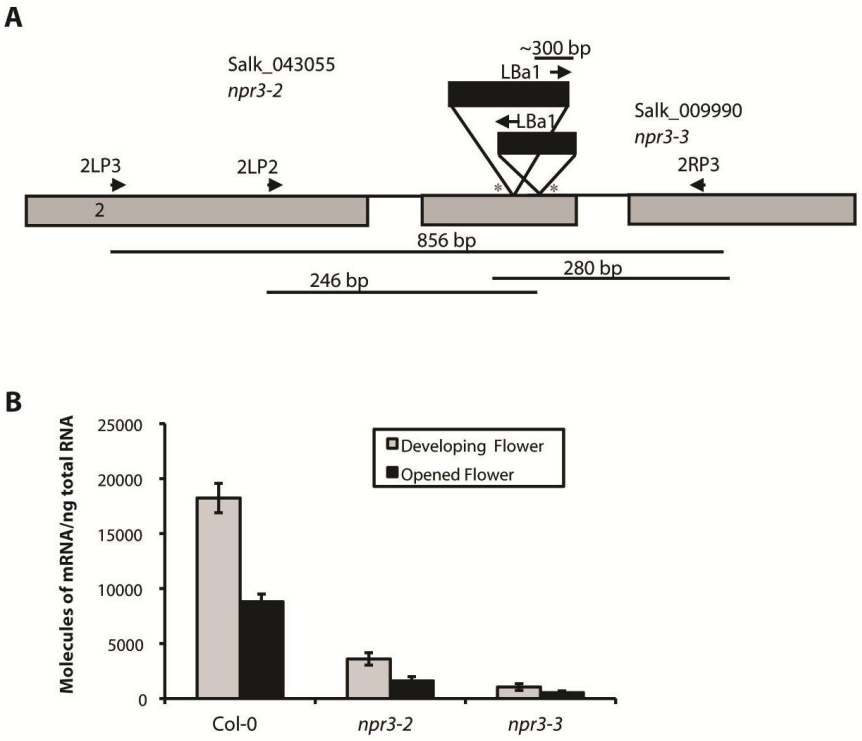


Figure 2-3 Characterization of the *npr3* mutants. A, Schematic diagram showing the insertion in the *npr3-2* (SALK_043055) and *npr3-3* (SALK_009990) mutants. Exons are represented by grey boxes and introns by black lines. "2" signifies the second exon of *NPR3* gene. Black boxes represent the T-DNA insertions and the gray asterisks show the sites of insertion. The two pairs of primers used for genotyping are indicated by arrows, 2RP3/2LP3 for wild-type copy, LBa1/2RP3 and LBa1/2LP2 for T-DNA insertion copy in *npr3-2* and *npr3-3*, respectively. B, Q-PCR analysis showing that *NPR3* transcripts level dramatically decrease in developing and opened flowers of both *npr3-2* and *npr3-3* mutants. Expression levels are presented as absolute copy numbers of *NPR3*/ng of total RNA. Means ± SE of three biological replicates, each consisting of flower tissues from five individual plants are presented.

The *npr3* mutant exhibits enhanced resistance to bacterial infection during early flower development

Considering the close evolutionary relatedness between *NPR1* and *NPR3*, we speculated that *NPR3* may play a role in defense response. To confirm this, we infiltrated leaves of Col-

0 and *npr3* mutants with *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*), but no significant difference was found in bacterial growth between Col-0 and two *npr3* mutant lines (data not shown). Further considering the strong expression of the *NPR3* gene in developing flower buds, we hypothesized that *NPR3* may play a role in the defense response during flower development. To test this, we developed a floral pathogen infection assay to challenge homozygous T-DNA insertion mutants with *P.s.t.*. When plants were grown without pathogens, we did not observe any obvious phenotypic differences between Col-0 and the two alleles of homozygous *npr3* mutants. Seven days after inoculation, Col-0 and *npr1-3* mutants showed impaired flower development and produced short siliques (Figure 4A-B). When challenged with pathogenic bacteria however, the *npr3-3* mutant flowers developed normally (Figure 4C). Remarkably, flower and silique development of inoculated *npr3-3* mutants was indistinguishable from that of non-inoculated plants. A similar phenotype was observed for the *npr3-2* allele (data not shown) and these phenotypes were reproducibly demonstrated in five independent experiments. To quantify this difference, the length of the fourth silique from the bottom of each infected inflorescence was measured (Figure 4D). The mean silique length of the *npr3* mutants after infection was about 12 mm, which was not statistically different from that of uninfected plants, whereas silique lengths on infected *npr1-3* and Col-0 plants were dramatically reduced to approximately 2 mm (Figure 4D). Interestingly, there was no significant difference in silique growth between Col-0 and the *npr3* mutants if the flowers were challenged after fully opening (data now shown). Thus, developing Col-0 flowers are more susceptible to *P.s.t.* infection than are opened flowers.

To test whether *npr3* mutant alleles also reduced bacterial growth on developing flowers, we quantified bacterial titers in floral extracts. The *npr3* mutant plants supported about 30-fold less bacterial growth than did the wild type and the *npr1-3* mutant, which were statistically indistinguishable from each other (Figure 4E).

NPR3 acts as a repressor of the NPR1-dependent signal transduction pathway

To test if NPR3 acts on the NPR1 defense signaling pathway, the double mutant *npr1-3 npr3-3* was obtained by crossing. Growth of *P.s.t.* pathogen on developing flowers was measured in the *npr1-3 npr3-3* double mutant and control plants. Again, we observed that the Col-0 and *npr1-3* controls supported about 30 times more bacteria then the *npr3-3* mutant. The double mutant showed an intermediate level of bacterial growth (Figure 4H), suggesting that the NPR3 repression occurs in part through the NPR1-dependent pathway but might also act on an NPR1-independent pathway(s) (Clarke et al., 2000).

Basal and induced levels of *PR1* are elevated in flowers of the *npr3-3* mutant

The activation of the *PR1* gene has long been regarded as a major hallmark of NPR1-mediated transcriptional defense response activation (Dong, 2004; Loake and Grant, 2007). As a further test of the role of NPR3 in repression of this pathway, we examined the level of *PR1* transcript in the *npr3-3* mutant as a proxy for the activation status of the NPR1-dependent pathway. Six-week-old plants were inoculated with *P.s.t.*, and *PR1* transcript levels were measured 5 days after infection. In developing flowers of the *npr3-3* mutant, the basal level of *PR1* was approximately two-fold higher than in Col-0 (Figure 4F) and after infection with *P.s.t.* it was induced an additional 3-fold, indicating NPR1-dependent defense signaling is present in flowers. In older opened flowers, while there was no significant difference in the level of *PR1* between mock-inoculated Col-0 and *npr-3-3* mutants, the *PR1* gene was slightly up-regulated in the *npr3* mutant in response to inoculation with *P.s.t.* (Figure 4G). We found that the basal and induced expression levels of *NPR3* were higher in developing flowers compared to opened flowers, however, the *NPR1* gene was expressed at a similar level in both treatments and tissues. In addition, *PR1* expression was overall higher in opened Col-0 flowers compared to developing Col-0 flowers and was greatest in developing, inoculated *npr3-3* flowers. In summary, there is a negative correlation between *NPR3* and *PR1* expression. These data, along with the infection phenotype (Figure 4D and 4E), strongly suggest that NPR1 and NPR3 interact during flower development in the regulation of the defense response.

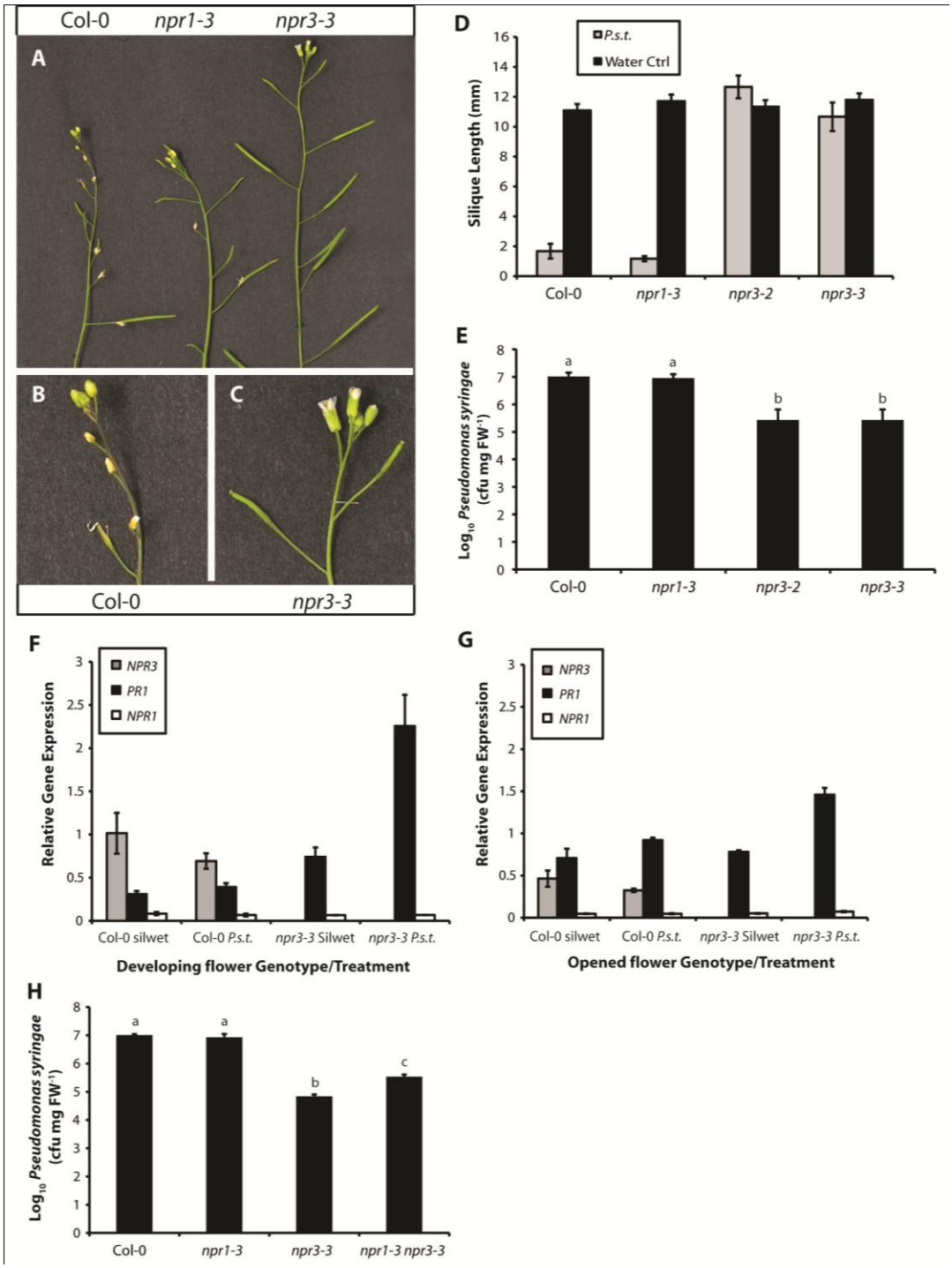


Figure 2-4 Increased resistance of *npr3* mutant to virulent strain *P.s.t.* DC3000. A, Representative image of siliques from infected inflorescences of Col-0, *npr1-3,* and *npr3-3* seven days after inoculation with a virulent bacteria *P.s.t.* DC3000 suspension of OD_{600}=0.2. B and C, Close-up views of the WT and *npr3-3* mutant, respectively, shown in A. D, Average

silique length of the four genotypes seven days after inoculation (grey bars) compared to water controls (black bars). Data represent means ± SE of treated siliques from four inflorescences per treatment. E, Bacterial populations in flowers of six-week-old plants (genotypes indicated on x-axis) five days after inoculation with virulent *P.s.t.* DC3000. Data represent mean ± SE of three replicates, each consisting of four infected inflorescences. Significant differences among the genotypes ($P<0.05$) were determined by Fisher's PLSD analysis. cfu, colony forming units. F and G, Dot-blot analysis of *NPR3, PR1,* and *NPR1* expression in developing and opened flowers from Col-0 and *npr3-3* mutant plants treated with *P.s.t.* and compared to flowers from control plants treated with Silwet L-77 only. Six-week-old soil-grown wild-type Col-0 and *npr3* mutant plants were treated with *P.s.t.* DC3000 and samples were collected 24 h after treatment. Relative expression levels were normalized to the expression of *ubiquitin*. Data represent means ± SE of three biological replicates, each containing flower tissues from five individually treated plants. H, Bacterial populations in flowers of six-week-old Col-0, *npr1-3*, *npr3-3* and *npr1-3 npr3-3* double mutant five days after inoculation of virulent *P.s.t.* DC3000 at $OD_{600}=0.2$. Data represent means ± SE of three replicates, each containing four infected inflorescences. Significant differences among the genotypes ($P<0.05$) determined by single factor ANOVA analysis.

Overexpression of *NPR3* confers hypersensitivity to bacterial infection

NPR3 was constitutively over-expressed in transgenic Col-0 plants and pathogen resistance evaluated seven days after six-week-old plants were inoculated with *P.s.t..* Developing flowers of NPR3 over expressing plants were more susceptible than Col-0 to infection but only when the bacterial solution was diluted to one-tenth of the concentration normally used (Figure 5A). Whereas the *NPR3* overexpressor displayed the susceptible phenotype, namely short siliques and compromised flower development, under the conditions of this experiment (lowered inoculum concentration $OD_{600}=0.02$) Col-0 exhibited normal development seven days after inoculation.

We reasoned that since NPR3 appears to act as a repressor of defense response in developing flowers based on its high expression level, over-expression in leaves might

result in repression of NPR1-mediated pathogen resistance as well. We performed a *P.s.t.* infection assay on leaves and measured bacterial titers of five independent *NPR3* overexpression lines. Symptoms were more severe in the transgenic plants, indicating that constitutive overexpression of *NPR3* resulted in increased susceptibility (Figure 5B). All five lines supported about 300-fold more bacterial growth than Col-0 (Figure 5C). We plotted bacterial levels (cfu) against the level of *NPR3* transcript in various transgenic lines (Figure 5D). The results show a strong positive correlation of bacterial growth (R^2 0.9011) with the level of *NPR3* expression. These results further confirm that *NPR3* plays a role in the negative regulation of the defense response and demonstrate that NPR3 can act as a repressor in both floral and leaf tissues.

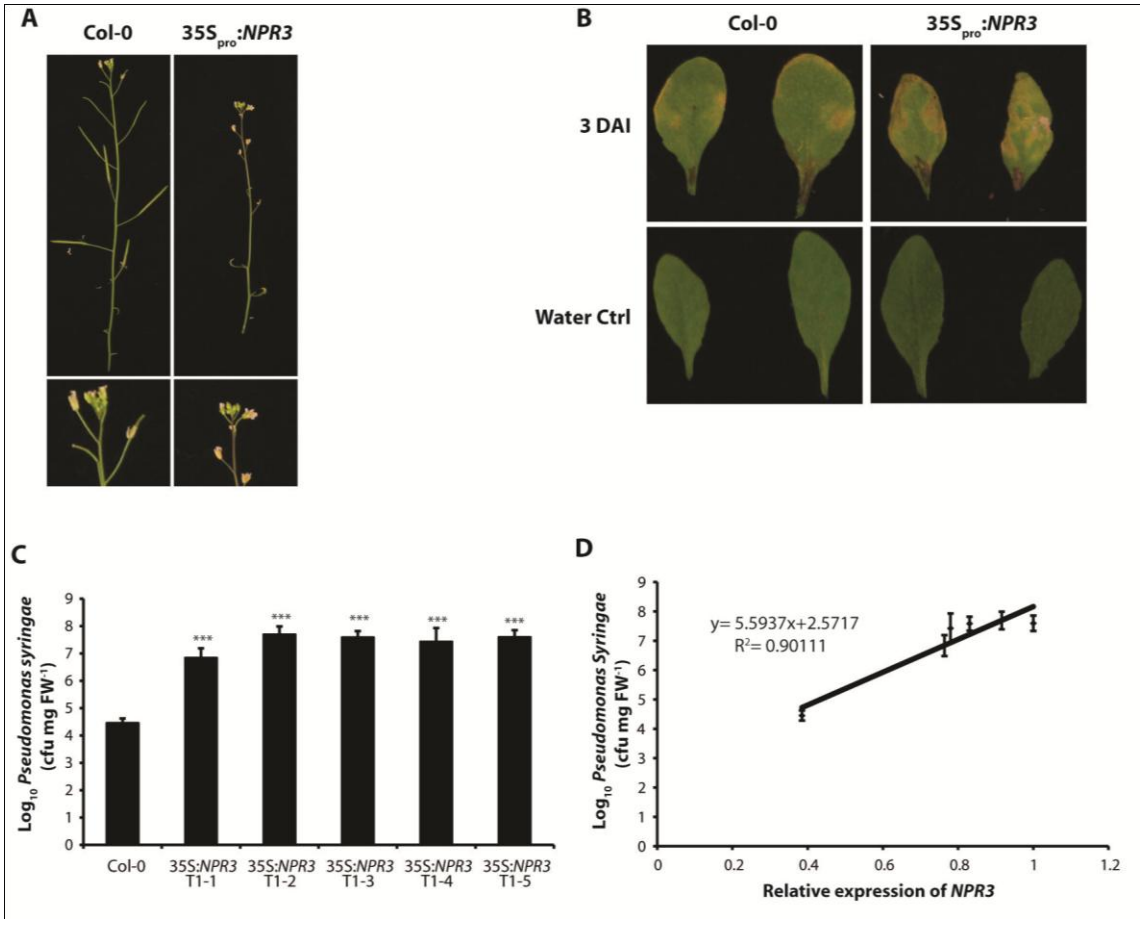


Figure 2-5 Increased susceptibility to *P.s.t.* DC3000 infection of transgenic plants constitutively overexpressing *NPR3*. A, Representative images of inflorescences of Col-0 and CaMV35S$_{pro}$:*NPR3* plants seven days after drop-inoculation with a low concentration

(OD_{600}=0.02) of virulent *P.s.t.* DC3000. B, Representative images of leaves of Col-0 and CaMV35S_{pro}:*NPR3* plants three days after infiltration (DAI) with a lower concentration (OD_{600}=0.002) of *P.s.t.* bacteria. C, Bacterial populations in infected leaves of wild-type Col-0 and CaMV35S_{pro}:*NPR3* transgenic plants. Leaves of four-week-old plants were infiltrated with a bacterial suspension of OD_{600}=0.002. Samples were collected three days after inoculation. Data represent means ± SE of three biological replicates, each containing two individually infected plants. Asterisks show the significant differences (P<0.001) determined by single factor ANOVA. D, Correlation between *NPR3* expression level and bacterial growth in CaMV35S_{pro}:*NPR3* leaves. Total leaf RNA was extracted from Col-0 and five CaMV35S_{pro}:*NPR3* transgenic lines and cDNA prepared from the individual samples. *NPR3* expression levels were determined by dotblot analysis and relative expression values were normalized to the expression of *ubiquitin*.

NPR3 interacts with TGA2 and NPR1 in transient expression assays

To investigate the possibility of physical interactions among NPR3, TGA2, and NPR1 proteins *in vivo*, we performed bimolecular fluorescence complementation (BiFC) (Walter et al., 2004) in onion epidermal cells. Co-bombardment of the two halves of YFP served as the negative control, and resulted in no fluorescence (Figure 6B), whereas transient expression of CaMV35S_{pro}:EGFP served as the positive control, and resulted in green fluorescence in both the cytoplasm and the nucleus (Figure 6C). We first co-bombarded onion epidermal cells with a translational fusion of NPR1 fused to the N-terminal half of cyan fluorescent protein (NPR1:CFP^N), and a transcriptional fusion of TGA2 fused to the C-terminal portion of yellow fluorescent protein (TGA2:YFP^C). Consistent with previous reports (Zhou et al., 2000; Fan and Dong, 2002), we observed cyan fluorescence in the cytoplasm and nucleus (Figure 6D). Furthermore, we confirmed a previous report (Zhang et al., 2006) showing that NPR3 associates with TGA2 (Figure 6E) in both the cytoplasm and nucleus. However, in our assay, this interaction occurred primarily in the cytoplasm, whereas Zhang et al. (2006) described it as occurring predominantly in the nucleus.

NPR1 and NPR3 contain BTB/POZ domains in their N-termini, and the BTB/POZ domain mediates the formation of inactive NPR1 oligomers in uninduced cells (Mou et al., 2003). We thus hypothesized that the BTB/POZ domains of NPR1 and NPR3 might also interact with each other to form NPR1-NPR3 hetero-dimers *in vivo*. To test this, NPR1:YFP^{C} and NPR3:YFP^{N} constructs were co-bombarded into onion epidermal cells. Confocal microscopy revealed yellow fluorescence in the cytoplasm (Figure 6F), indicating that NPR1 and NPR3 interact *in vivo* in the onion epidermal peel transient assay.

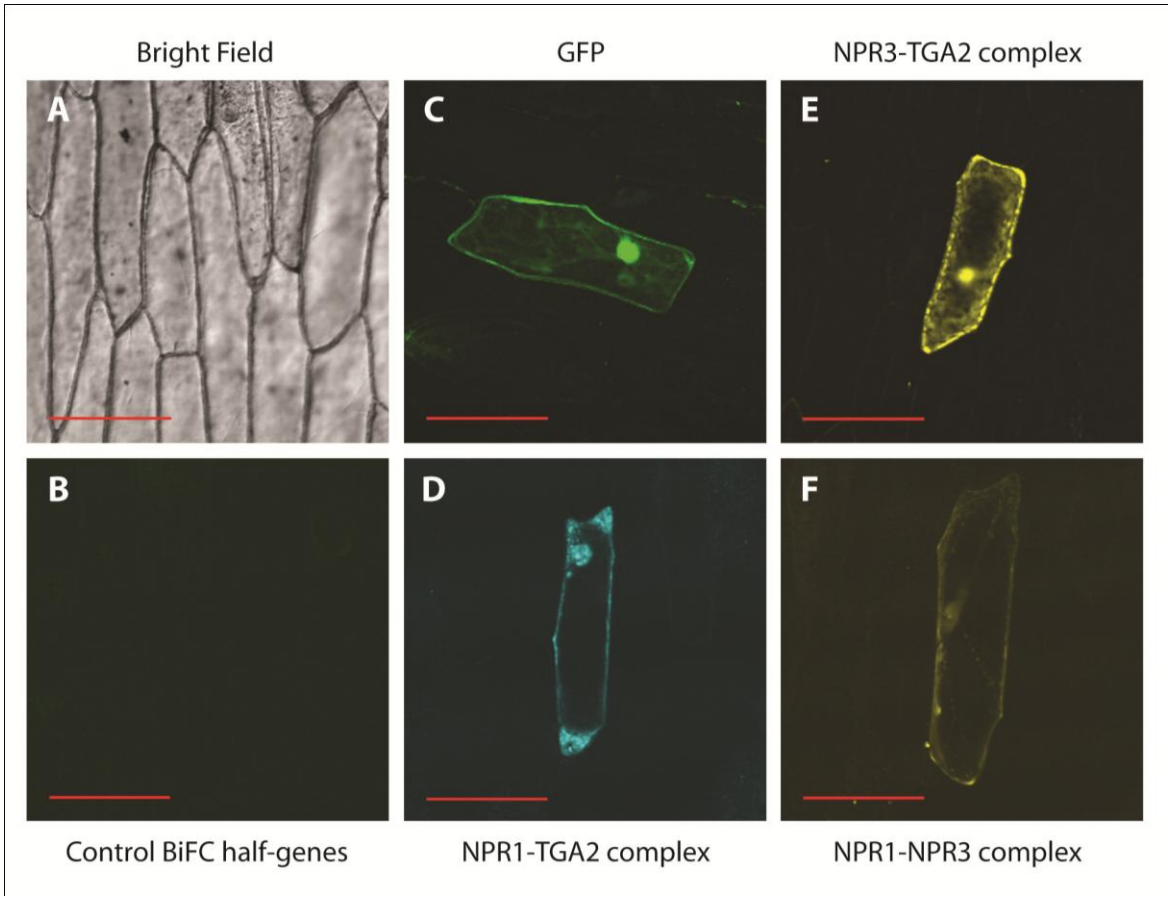


Figure 2-6 Bimolecular fluorescence complementation (BiFC) analysis of interactions among NPR3, TGA2, and NPR1 in transiently transformed onion epidermal cells. A, Representative bright field image. B through F, Representative confocal images of onion epidermal cells co-bombarded with constructs expressing B, BiFC control half-genes, which served as the negative control, C, CaMV35S$_{pro}$:EGFP, which served as the positive control, D, NPR1:CFP^{N} and TGA2:YFP^{C}, which formed NPR1:TGA2 complex E, NPR3:YFP^{N} and

TGA2:YFP^{C}, which formed NPR3-TGA2 complex and F, NPR1:YFP^{C} and NPR3:YFP^{N}, which formed NPR1-NPR3 complex. Scale bars represent 150 μm.

NPR3, NPR1 and TGA2 interactions inhibit NPR1-TGA2 complex nuclear translocation and increase susceptibility to *P.s.t.* DC3000

Based on our results, we developed a working model in which NPR3 interacts with NPR1 and/or TGA2 in the cytoplasm and this interaction restricts the nuclear translocation of the NPR1-TGA2 complex into the nucleus, resulting in repression of the NPR1 defense pathway. To test this model, we used multi-color, bi-molecular fluorescence complementation interaction analysis with stably transformed plants. Col-0 plants were transformed with TGA2:YFP^{C} and NPR3:YFP^{N} transgenes and developing flower petals were imaged by confocal microscopy. We observed a yellow fluorescence signal in the cytoplasm of transgenic petal cells, indicating that NPR3 and TGA2 associate with each other in the cytoplasm of stably transformed plants under uninfected conditions (Figure 7A). 24 h after *P.s.t.* inoculation, NPR3-TGA2 complexes were also observed to be primarily localized to the cytoplasm, suggesting that the NPR3-TGA2 complex, unlike the NPR1-TGA2 complexes, was not translocated to the nucleus after bacterial infection (see below). We did not detect significant yellow fluorescence in transgenic plants transformed with the YFP^{C} and YFP^{N} transgenes lacking NPR1, NPR3 and TGA2 translational fusions (control BiFC half-genes)(Figure 7A), although minor punctate regions of fluorescence were observed primarily in cell wall junctions.

Although NPR1 and TGA2 are known to interact in leaves (Despres et al., 2000; Zhou et al., 2000; Fan and Dong, 2002), this interaction has not been demonstrated in flowers. We therefore examined the petals from young flowers of transgenic plants co-transformed with TGA2:YFP^{C} and NPR1:CFP^{N} by confocal microscopy. NPR1 and TGA2 were found to interact in the cytoplasm of uninfected petal cells (Figure 7B). Furthermore, the NPR1-TGA2 interaction was also observed in the nucleus twenty-four hours after *P.s.t.* infection, as previously reported in leaves (Despres et al., 2000; Zhou et al., 2000; Fan and Dong, 2002; Pieterse and Van Loon, 2004). Thus, the mechanism of NPR1-TGA2 mediated

induced defense may be the same in the flower and the leaf. No fluorescence was detected in control BiFC half-gene transgenic flowers (Figure 7B).

If NPR1 and NPR3 compete for TGA2, we would expect the two complexes to primarily co-localize in cells, thus we imaged the interaction among the three proteins using multicolor BiFC. We crossed transgenic plants expressing NPR1:CFP^N to transgenic plants expressing TGA2:YFP^C and NPR3:YFP^N to obtain plants expressing all three fusion proteins. Both yellow fluorescence (indicating NPR3-TGA22 interaction) and cyan fluorescence (indicating NPR1-TGA2 interaction) were detected in the cytoplasm of petal cells of non-inoculated developing transgenic flowers (Figure 7C). Although the yellow fluorescence was brighter than cyan, a yellowish green color resulted in the merged image, indicating that the NPR1-TGA2 and NPR3-TGA2 complexes primarily co-localize. Consistent with our model of NPR3 repression, the NPR1-TGA2 complex did not translocate to the nucleus when flowers of the triple expressor were inoculated with *P.s.t.*, indicating that the interactions between NPR3-TGA2 and/or NPR3-NPR1 may impair the nuclear localization of the NPR1-TGA2 complex after bacterial infection. It is possible that NPR1, NPR3 and TGA2 form a stable tertiary complex, but our data cannot discern this possibility.

To investigate the correlation between the nuclear localization of NPR1-TGA2 and resistance to *P.s.t.* DC3000, we inoculated developing flowers of five different genotypes, namely, Col-0, *npr3-3* mutant, and transgenic plants co-expressing TGA2:YFP^C and NPR3: YFP^N, TGA2:YFP^C and NPR1: CFP^N, and the triple transgenic NPR1:CFP^N, TGA2:YFP^C, and NPR3:YFP^N. Consistent with the results presented in Figure 4E, we found that the *npr3-3* mutant supported about 30-fold less bacteria than Col-0 (Figure 7D). The bacterial growth on transgenic plants expressing TGA2:YFP^C and NPR1:CFP^N was lower than that on Col-0, suggesting that over-expression of these proteins and their nuclear translocation conferred higher resistance. Although bacterial growth in the triple transgenic plant was slightly lower than that in Col-0 (Figure 7D), the triple transgenic plant was more susceptible to *P.s.t.* than transgenic plant co-expressing TGA2:YFP^C and NPR1:CFP^N. Furthermore, transgenic plants co-expressing TGA2:YFP^C and NPR3:YFP^N were more susceptible than Col-0 to *P.s.t.*. All of the plants transformed with BiFC constructs were more susceptible

than the *npr3-3* mutant, indicating that pathways in addition to the NPR1-dependent signaling pathway may be involved in resistance to *P.s.t.* DC3000 during early flower development. Overall, these bacterial infection data are in good agreement with the confocal microscopy observations (Figure 7A-C) and further confirm that *NPR3* is a negative regulator of the defense response in developing flowers.

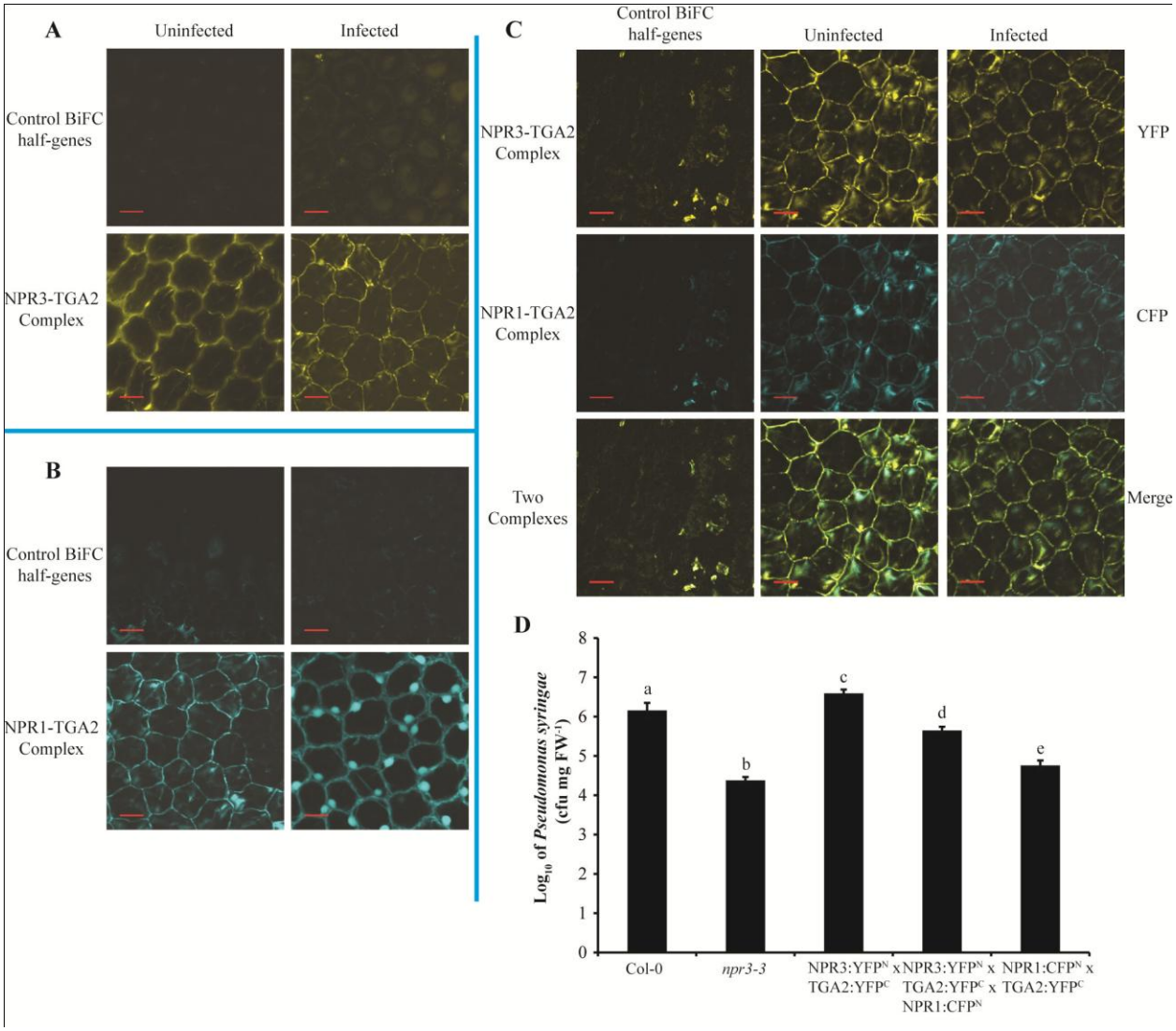


Figure 2-7 *In planta* bimolecular fluorescence complementation (BiFC) analysis of interactions among NPR1, TGA2, and NPR3. A, B and C, Representative confocal microscope images of stably transformed *Arabidopsis* floral petals. Uninfected petals were harvested from inflorescences inoculated with only Silwet L-77 and infected petals were collected from inflorescences infected with a suspension of *P.s.t.* DC3000 (OD_{600}=0.2 (10^8cfu/ml)) 24

h prior to observation. Scale bars represent 10 μm. A, Petal cells from plants co-transformed with NPR3:YFP^N and TGA2:YFP^C constructs to form NPR3-TGA2 complex. Co-transformation of BiFC control half-genes served as the negative control. B, Petals of transgenic plants co-transformed with NPR1:CFP^N and TGA2:YFP^C constructs to form NPR1-TGA2 complex. Petals of transgenic plants co-transformed with CFP^N and YFP^C alone vectors served as the negative control. C, Petals of transgenic plants co-expressing all three NPR3:YFP^N, NPR1:CFP^N, and TGA2:YFP^C fusion proteins. The petals of plants co-transformed with YFP^N, CFP^N, and YFP^C vectors served as empty vector negative controls. NPR3-TGA2 complex was observed using YFP settings and NPR1-TGA2 complex was observed under CFP settings. Merged images were obtained by overlaying images from the YFP and CFP channels. D, *P.s.t.* DC3000 bacterial populations in Col-0, *npr3-3* mutant, and transgenic plants expressing various combinations of BiFC constructs. Whole inflorescences of six-week-old soil-grown plants were inoculated with a bacterial suspension of OD_{600}=0.2. Bacterial growth was assayed five days after inoculation. Data represent means ± SE of three replicates, each containing four infected inflorescences. Significant differences among the five genotypes were determined by single factor ANOVA.

The *npr3* mutant exhibits lower fitness in the absence of pathogen

Our results led to the hypothesis that de-repression of the NPR1 mediated defense response pathway in the *npr3* mutant would reduce plant fitness due to the reduction of carbon allocation to vegetative growth and seed production. As measures of plant fitness, we recorded the primary root length (PRL), relative growth rate (RGR), and seed production in *npr3* mutant and Col-0 plants. For root length, 16 individuals of each genotype were grown on vertical petri dishes, and root lengths measured one week after germination. For relative growth rate and seed production, flats of plants were grown under highly controlled conditions and at least 8 plants were pooled for each measurement. We found that the primary roots of *npr3* mutants were reproducibly about 14% shorter than those of Col-0 (Figure 8A, B). Moreover, roots of *npr1-3* and CaMV35S$_{pro}$:*NPR3* transgenic plants were significantly longer than those of Col-0. Consistent with our

expression results that suggest *npr3-2* is a weak allele (Figure 3B), the roots of *npr3-2* plants were intermediate in size compared to Col-0 and *npr3-3*.

In the absence of pathogen, RGR was the same in wild-type Col-0 and the *npr3* mutants (data not shown). Furthermore, while there was statistically no difference in seed production between Col-0 and the *npr1-3* mutant, both *npr3* mutant alleles produced about 25% fewer seeds than Col-0 (Figure 8C). Pollination efficiency was the same in Col-0 and *npr3* plants, as neither the number of siliques produced nor the number of seeds within a single silique differed significantly between Col-0 and *npr3* plants. To determine whether the mass per seed of *npr3* was lower than that of the wild type, we counted the number of seeds present in 10 mg of seed. The two *npr3* mutant lines produced seeds which weighed about 22% less than Col-0 seeds (Figure 8D), indicating that seeds from *npr*3 mutants were significantly lighter than those from Col-0.

The *npr3-3* mutation is semi-dominant

To investigate the heritability of the *npr3* mutant allele, we inoculated the inflorescences of plants in an F_2 segregating population derived from a cross between Col-0 and the *npr3-3* parental genotype with *P.s.t*. at concentration of OD_{600}=0.2. The bacterial titers of 94 plants were assayed five days after inoculation and the genotypes of each plant were determined. The response of the heterozygotes to *P.s.t.* infection was almost perfectly intermediate between that of Col-0 and *npr3-3* homozygous mutant plants (Figure 8E), and the heterozygotes harbored significantly less bacteria than Col-0 and significantly more bacteria than *npr3-3*. Moreover, the wild-type susceptible phenotype, the intermediate phenotype, and the *npr3-3* resistant phenotype were found roughly in a 1:2:1 ratio within the population. Thus, the *npr3-3* mutation appears to be semi-dominant. This is consistent with a mechanism of competitive interactions whereby the concentration of NPR3 is critical in the degree of repression of NPR1 mediated defense.

To rule out the possibility that the difference in primary root length between Col-0 and *npr3* is due to maternal effects on seed development and subsequent seedling growth,

and to further confirm the semi-dominance of the *npr3-3* allele, we measured the primary root length of 94 plants from the F_2 segregation population. As expected, the roots of heterozygous NPR3/*npr3-3* plants were statistically shorter than those of Col-0, but longer than those of *npr3-3* homozygous mutants (Figure 8F). The root lengths of the heterozygote were closer in size to the Col-0 than to *npr3-3* suggesting that the NPR3 gene acts as an incompletely dominant gene in roots. Thus, the difference in root length between Col-0 and *npr3* is not the result of maternal effects, because all seeds were produced on the same mother plant. Our data indicates that the reduced fitness of the *npr3* mutant is instead dependent on the dose of the *NPR3* functional allele.

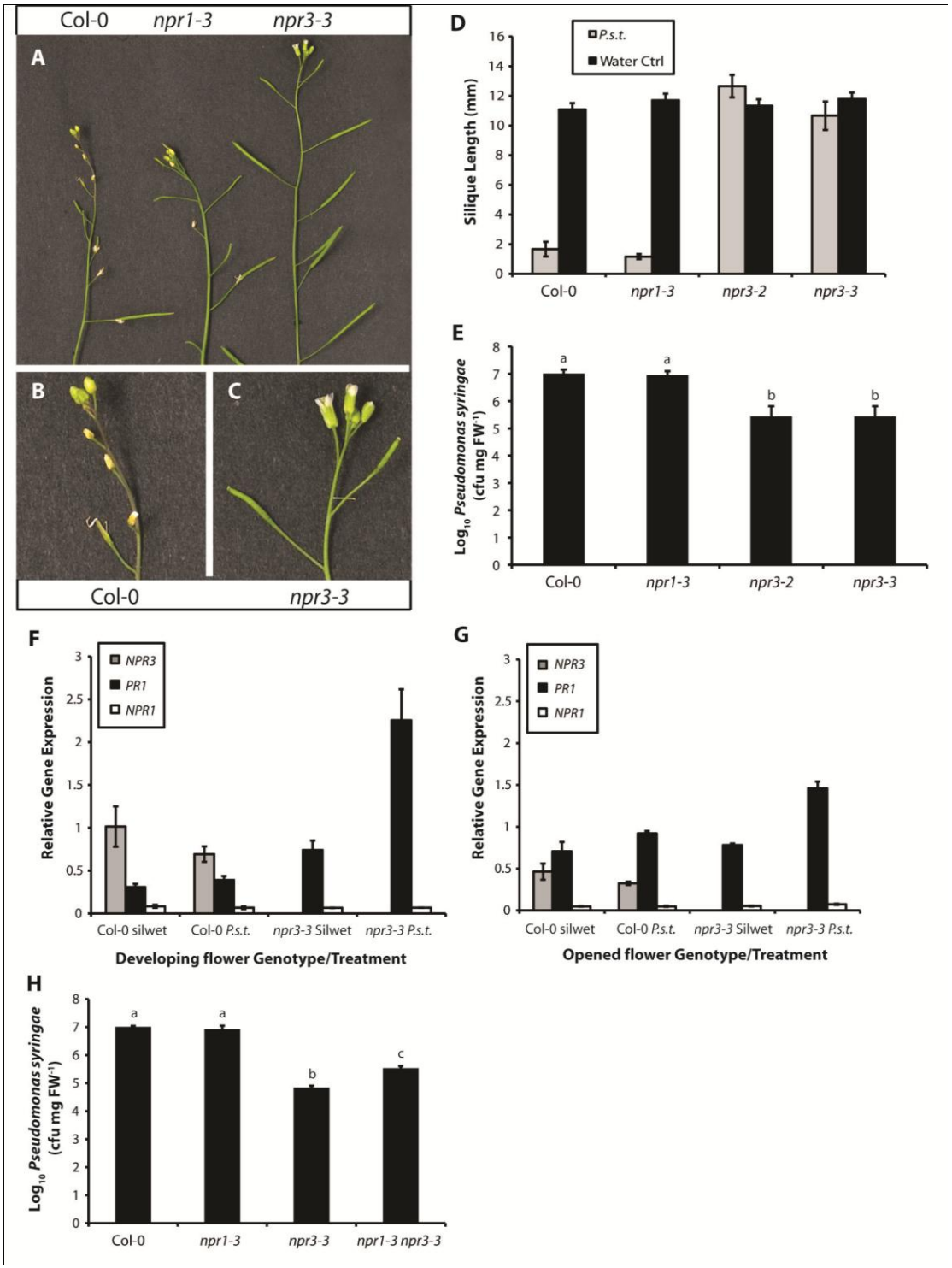


Figure 2-8 The *npr3* mutant exhibited reduced fitness in the absence of pathogen and the *npr3-3* mutation is semi-dominant. A, Representative images of seedlings of Col-0, *npr1-3*, *npr3-2*, and *npr3-3* mutants, and CaMV35S$_{pro}$:*NPR3* transgenic plants one week after germination. B, Root lengths of plate-grown seedlings one week after germination. Means ±

SE of 16 individual plants from each genotype are presented. Significant differences among the five genotypes ($P<0.05$) were determined by single factor ANOVA. C, Seed production (g) per Col-0, *npr1-3*, *npr3-2,* and *npr3-3* plant in the absence of pathogen. Bar chart represents means ±SE (n=20). Significant differences ($P<0.05$) were determined by single factor ANOVA. D, Average weight per seed (mg) of Col-0, *npr1-3*, *npr3-2,* and *npr3-3* mutants. Three biological replicates of 10 mg seed were weighted, seed numbers were counted per replicate and average weight per seed was calculated. Data represent means ± SE of the three biological replicates and statistically significant differences among four genotypes ($P<0.05$) were determined by single factor ANOVA. E, Bacterial growth in an F_2 (Col-0 X *npr3-3*) segregating population of 94 plants. Inflorescences of six-week-old plants were inoculated with 0.2 OD_{600} (10^8cfu/ml) of virulent *P.s.t.* DC3000 by drop inoculation. Bacterial growth was assayed five days after inoculation. Data represent means ± SE of 24 Col-0, 49 heterozygous *NPR3/npr3-3,* and 21 homozygous *npr3* mutants. Significant differences among the three genotypes ($P<0.05$) were determined by single factor ANOVA. F, Primary root length of an F_2 (Col-0 X *npr3-3*) segregating population of 94 plants. Root lengths of plate-grown plants were measured one week after germination. Bar chart represents means ± SE of 25 Col-0, 50 heterozygous *NPR3/npr3-3,* and 19 homozygous *npr3* mutants and significance at $P<0.05$ was determine by single factor ANOVA.

2.4 Discussion

In this study, we sought to define the role of NPR3 in *Arabidopsis*. Given that NPR1 is a key component of the plant defense response (Cao et al., 1997; Dong, 2004) and that NPR1-NPR4 are highly similar to each other and share a common ancestor (Figure 1A and B), we hypothesized that NPR3 is also involved in the plant's physiological response to pathogen infection. Since *NPR3* is strongly expressed in young flowers and at a higher level than any other *NPR* family members in this tissue, we reasoned that it might play an important role during flower development.

Our observations suggest that NPR3 is a negative regulator of the plant defense response in early flower development. Firstly, silique development was inhibited in wild-

type plants challenged with *Pseudomonas syringae* pv. tomato DC3000, but not in challenged *npr3* mutants (Figure 4A), and secondly, the inflorescence of *npr3* mutants supported 30-fold less bacteria than wild type five days after *P.s.t.* infection (Figure 4E). Furthermore, the observations that *NPR3* levels declined as flowers opened and that the level of *PR1* expression was inversely related to that of *NPR3* (Figure 4F and 4G) suggest that the negative regulation of the defense response conferred by *NPR3* is dependent on developmental stage and affects *PR1* gene expression, the hallmark of NPR1-mediated defense response. Although *PR1* transcripts accumulated significantly in developing *npr3-3* flowers infected with *P.s.t.*, neither the mutation of *NPR3* nor pathogen challenge affected *NPR1* levels (Figure 4F and 4G). This indicates that the regulation of *NPR1* after *P.s.t.* infection differs in flowers and leaves, since there is 2-4 fold induction of *NPR1* in leaf tissues after infection (Dong, 2004). However, there was only about a three-fold increase in *PR1* expression upon infection of unopened *npr3* flowers, which was less than the 10-fold induction previously observed in leaf tissue of infected wild-type plants (Cao et al., 1994; Cao et al., 1997; Ryals et al., 1997; Shah et al., 1997; Liu et al., 2005; Zhang et al., 2006). The intermediate bacterial susceptibility of the *npr1 npr3* double mutant (Figure 4H) suggests that while NPR3 indeed is a repressor of the NPR1-dependent defense response, it may also act on NPR1-independent pathways (Clarke et al., 2000)as well. This suggests that downstream defense components in addition to *PR1* might function in the *NPR3*-dependent defense pathway in flowers.

The interaction between NPR1 and TGA2 is known to be essential for the induction of defense genes, such as *PR1*, in leaf tissue (Zhou et al., 2000; Fan and Dong, 2002; Rochon et al., 2006). Our BiFC results show that the NPR1-TGA2 protein complex forms in Col-0 petal cells and is translocated into the nucleus upon bacterial infection (Figure 7B). Furthermore, consistent with previous study (Zhang et al., 2006) NPR3 was shown to interact with TGA2 in both cytoplasm and nucleus in transiently transformed onion epidermal cells (Figure 6E). However, it was demonstrated that NPR3 interacts with TGA2 only in the cytoplasm of both infected and uninfected stable transgenic *Arabidopsis* plants (Figure 7A). It is possible that expression in heterologous system might not represent the real localization due to high amount of DNA injected and cell wounding, thus we speculate

that the cytoplasmic interaction of NPR3-TGA2 observed in petals of stable transgenic *Arabidopsis* is more actual. In addition, our multicolor BiFC results show that in this experimental system, the induced nuclear translocation of the NPR1-TGA2 complex is blocked by *NPR3* over-expression (Figure 7C), and that the NPR1-TGA2 complex and NPR3-TGA2 complex co-localized in the cytoplasm of both uninfected and infected cells over-expressing all three proteins (Figure 7C). These results are totally consistent with the *npr1-3* phenotype in floral infection (Figure 4D and E), where *npr1-3* mutant showed no effect on bacterial growth compared to Col-0. Since NPR3 negatively regulates NPR1-dependent signaling, no *npr1* susceptible phenotype will be observed if NPR3 is present. Interestingly, we found that NPR1 and NPR3 can also associate with each other in the cytoplasm of transiently transformed onion cells, indicating that NPR3 might regulate the defense response in flowers by directly binding to cytoplasmic NPR1. BiFC approach has been used to identify interactions between a number of different proteins in plant cells (Bracha-Drori et al., 2004; Zhang et al., 2006; Zhang and McCormick, 2007; Feng et al., 2008). These results should be taken with caution however, because the concentrations of protein are likely to be much higher than that of the native proteins *in vivo,* and this could cause artifactual associations or abnormal localization. To further confirm our findings, co-immunoprecipitation of subcellular protein fractions using TGA2 or NPR antibodies could be used.

Based on previous literature, and data presented here, we propose a possible mechanistic model in which NPR3 functions as a negative regulator of the defense response in developing flowers. The evidence presented in Figures 4 and 7 strongly suggests that in developing flowers, NPR3 is a repressor of the NPR1-dependent defense response, reducing *PR1* gene expression and increasing susceptibility to bacterial infection. Our data demonstrate that in onion epidermal cells (Figure 6) and in transgenic *Arabidopsis* (Figure 7A-C), NPR3 interacts with both NPR1 and TGA2. Additionally, we demonstrated that NPR1 and NPR3 can also interact predominantly in the cytoplasm of onion epidermal cells (Figure 6). It was previously demonstrated that in *Arabidopsis* leaf cells, NPR1 interacts with TGA2 through its ankyrin repeat domain and with itself to form homodimers through its BTB/POZ domain (Zhang et al., 1999). Based on the high conservation of these domains

between NPR1 and NPR3, we propose that similar interactions between NPR3-TGA2 and NPR3-NPR1 most likely occur through the same functional domains to form a protein complex consisting of NPR1, NPR3 and two TGA2 molecules. TGA2 is a bZip transcription factor that exists as homo or heterodimers *in vivo* and has been shown to also interact with TGA5 and TGA6 (Blanco et al., 2009). It is likely that the hypothesized multimeric protein complexes would contain mixtures of all of these proteins and perhaps others such as NPR4. These interactions are hypothesized to prevent the NPR1-TGA2 complex from translocating into the nucleus upon *Pseudomonas* infection, thus repressing the expression of defense genes and diminishing disease resistance. Further investigation, such as *in vitro* protein complex analysis, is required to identify all the components in this protein complex and to clarify the mechanism by which NPR3 interacts with other defense-related proteins to inhibit the defense response.

Our results showed that developing flower buds were more susceptible than open flowers to *P.s.t.* infection, indicating that plant defense mechanisms are repressed in young flowers. Since pollination is completed as flowers fully open in *Arabidopsis* (Smyth et al., 1990; Scott et al., 2004), it is possible that plants "sacrifice" young, un-pollinated flowers during pathogen infection to save energy for seed production by older, pollinated flowers. By attacking pathogens only in pollinated flowers, plants increase their chance of transferring their genetic information to the next generation.

We have also shown that, in the absence of pathogens, *npr3* mutants exhibit lower fitness, which manifests as reduced primary root length and seed mass (Figure 8B and 8D). The resources might be allocated differently in the *npr3* mutant, with more carbon being used for over-expression of defense genes, and less for development and reproduction. Thus, it can be concluded that the physiological cost of constitutive direct defense in the *npr3* mutant is a burden for plants during normal development in the absence of pathogen. We have shown that NPR3 negative regulation is only partially dependent on NPR1 (Figure 4H) and we assumed that other defensive genes in NPR1-independent pathway (Clarke et al., 2000) are also de-repressed in *npr3* mutant. Thus, the de-repression of both NPR1-

dependent and independent defense responses in *npr3* mutant contributed to the reduction in whole plant fitness. Additionally, the fact that *npr3* mutation primarily affected the seed production (Figure 8C) in the absence of pathogen but not RGR suggested that the de-repression of defensive genes in flowers majorly lowered whole plant fitness but not in leaves. This cost and benefit balance may underlie the evolution of a negative regulator of the defense response, such as NPR3. When the plant is not being attacked by a pathogen, a negative regulator suppresses the defense pathways conserving energy for normal physiological processes. Upon exposure to a pathogen and/or achieving certain developmental stages, the repression is removed and the defense pathways can be induced. Taken together, our data clearly suggests that a strong evolutionary pressure resulted in the neofunctionalization of the *NPR3* gene after the duplication of the ancestral NPR progenitor, resulting in a mechanism capable of repressing the defense response in young flowers and acting to maintain an equilibrium between disease protection and a high level of fitness.

We showed that the phenotype of the heterozygous *NPR3*/*npr3* mutants was intermediate between that of Col-0 and the homozygous mutant (Figure 8E and 8F), suggesting that NPR3 function depends on the dose of NPR3 functional allele. This is consistent with a mechanism by which NPR3 acts as a competitive inhibitor of NPR1-mediated defense response signaling which is capable of fine tuning the degree of defense response through alteration in the concentrations of NPR1 and NPR3.

A recent study showed that nuclear NPR1 turnover plays a vital role in regulating plant immunity (Spoel et al., 2009). In the absence of a pathogen, proteasomal degradation of NPR1 prevents the constitutive expression of defense genes. Upon induction of SAR, the recruitment of phosphorylated NPR1 to a Cullin3-based ubiquitin ligase facilitates the turnover of NPR1 and triggers the transcription of genes that regulate the defense response. Another study demonstrated that some BTB domain-containing proteins can interact with Cullin3 *in vivo* (Figueroa et al., 2005). If the NPR3-NPR1 and/or TGA2 complexes associate with Cullin3 or similar proteins, it may target specific proteins for degradation or itself be targeted for degradation by Cullin3-based ubiquitin ligase.

We found that *NPR4* expression was relatively high in roots (Figure 2C); thus, it is reasonable to speculate that *NPR4* has a unique function in root disease resistance. Silique growth of the *npr4* mutant was reduced upon *P.s.t.* infection, and growth was intermediate between that of infected Col-0 and the *npr3* mutant (Figure S4). This result suggests that *NPR4* might be partially redundant to *NPR3*. The observation that *NPR3* and *NPR4* genes are evolutionarily closely related (Figure 1B), and that the leaves of the *npr3 npr4* double mutant are significantly more resistant to bacterial inoculation than either of the single mutants (Zhang et al., 2006) suggests functional redundancy between NPR3 and NPR4. Further studies are needed to clarify the function of NPR4 and to decipher the interactions between NPR3 and NPR4.

2.5 Materials and Methods

Phylogenetic Analysis

The six proteins predicted from the gene sequences of the *Arabidopsis* NPR family were aligned using ClustalW v1.8 (Thompson et al., 1994). The sequences were analyzed for potential functional sites by querying the Prosite database (Expasy, http://expasy.org/prosite/).

Using the TBLASTN program, full length *Arabidopsis* NPR1 protein sequence was used to search the Phytozome database (http://www. phytozome.net) to obtain the NPR1-like genes in poplar, Medicagao, grape and rice (E-value cutoff $1e^{-20}$) (Altschul et al., 1997). Six *NPR1-like* genes were identified in poplar four in Mecicago, three in grape and six in rice. Multiple protein sequence alignment of 25 sequences was performed by MUSCLE software (Edgar, 2004). The phylogenetic tree was constructed by MEGA 4.1 software using the neighbor-joining algorithm with Poisson correction model and option of pairwise deletion (Kumar et al., 2004).2000 replicates were performed to obtain the bootstrap value.

Data Mining

Data mining was performed on the Genevestigator website (https://www.genevestigator.com/gv/index.jsp; accessed on 10/14/2009). To examine expression in different organs, the gene atlas tool was used. We examined only the ATH 22K arrays hybridized with cDNA from Col-0.

Identification of *npr3* Knockout Mutants

Mutant seeds were obtained from the Arabidopsis Biological Resource Center at The Ohio State University and screened by PCR genotyping with wild type and T-DNA-specific primer pairs (Table S1). The plants tagged with T-DNA were backcrossed to wild-type Col-0 three times to eliminate extra T-DNA insertions. Homozygous plants from the F_2 generation were evaluated by DNA southern blot analysis to determine T-DNA copy number. Transgenic lines carrying a single T-DNA insertion were used in the expression analysis and to characterize the *npr3* mutant phenotype.

Vector Constructs and Plant Transformation

For each *NPR* promoter construct, an approximately 1.5 kb fragment directly upstream of the ATG translation initiation codon was PCR amplified from BAC clones provided by the Arabidopsis Biological Resource Center at The Ohio State University (see Table S2 for primer sequences and BAC clones IDs). Each promoter fragment was cloned into a pGEM®-T Easy vector (Promega Corp., Madison, WI) and the sequences were verified. Final gene cassettes were assembled in a pUC19 cloning vector such as each individual construct contained one of six *NPR* gene promoter fragments, EGFP coding sequence and 35S terminator. The junction sequences between the promoters and the EGFP coding sequence were verified and the constructs were transferred into binary vector pCAMBIA1300 (http://www.cambia.org/daisy/cambia/585.html). *NPR1, NPR3, NPR4*, and *BOP1* promoter constructs were ligated into the binary vector as *Hind*III and *Eco*RI fragments, the *NPR2* promoter construct was inserted as a *Hind*III and *Pvu*II fragment, and the *BOP2* promoter construct as a *Sal*I and *Pvu*II fragment. Each binary vector was mobilized into

Agrobacterium tumefacians AGLI and transformed into wild-type *Arabidopsis* plants accession Columbia (Col-0) by the floral dip method (Clough and Bent, 1998). Control transformations were performed with pCAMBIA1300 and pCAMBIA1300 with $35S_{pro}$:EGFP fragment. Hygromycin B-resistant plants were selected and transformation was verified by PCR with EGFP specific primers listed in Table S2.

EGFP Fluorescence Visualization by Stereomicroscopy

Bright field and fluorescent images were acquired with a Nikon SMZ-4 stereomicroscope equipped with an epi-fluorescence attachment, a 100 W mercury light source and a 3 CCD video camera system (Optronics Engineering, Goleta, CA). The microscope filter set included a 450- to 490-nm excitation and a 520- to 560-nm emission filters. EGFP fluorescence was recorded in various organs and tissues of randomly selected plants from four to six different transgenic lines per construct and compared to control pCAMBIA1300 transgenic plants.

EGFP, YFP and CFP Fluorescence Visualization by Confocal Microscopy

Fluorescence was observed in onion epidermal cell, flower petals and whole leaves of transgenic *Arabidopsis* plants. In all cases, the tissue was placed in a drop of water on a standard glass slide and covered with a glass cover slip. The slides were then positioned on the inverted platform of an Olympus FV1000 Laser Scanning Confocal Microscope (Olympus America Inc., Melville, NY). Water objectives (10X for the onion cells and 40X for the *Arabidopsis* tissues) with a numerical aperture of 1.15 were used. Excitation filters were applied as follow: for EGFP-488 nm, for CFP-458 nm, and for YFP-515 nm. Emission wavelengths of 500-600 nm were detected through a variable bandpass filter positioned in front of photomultiplier tube. The OLYMPUS FLUOVIEW software version 1.6 was used to collect images, select slices, and export data as 24-bit tiff files.

EGFP Fluorescence Visualization by Epi-Fluorescence Microscopy

Flowers were placed in a drop of water on a slide and covered with a cover slip. The slides were imaged using an Olympus BX61Epi-Fluorescence Microscope (Olympus America Inc., Melville, NY) using a 10x objective with a numerical aperture of 0.3. A FITC cube (Chroma filter cube #41001) was used to observe GFP. Images were acquired with a Hamamatsu Orca-ER camera and processed with SlideBook 4.1 software.

Q-PCR Analysis

Total RNA was extracted from radicles, roots, developing flowers, opened flowers, leaves, and senescent leaves. Radicle samples were collected from Col-0 plants grown in liquid MS medium for four days and root samples were harvested from Col-0 plants grown on solid MS medium for two weeks. Leaf samples were collected from four-week-old soil-grown plants. Opened and developing flower samples were collected from six-week-old plants and senescent leaf samples were collected from eight-week-old plants. Gene specific primers and probes used are presented in Table S3. The primers were synthesized at the Penn State Nucleic Acid Facility with a MerMade12 automated DNA synthesizer (Bioautomation, Plano TX) and the specific fluorescent probes were synthesized by Biosearch Technologies (Novato, CA). Q-PCR analysis was performed at the Penn State Nucleic Acid Facility using the ABI 7300 Sequence Detection System (Applied Biosystems Inc, Foster City, CA). Absolute copy number was determined for each transcript using recombinant RNA calibration curves for each gene generated from full-length cDNA fragments.

Pathogenesis Assay

Arabidopsis plants were grown in a Conviron growth chamber at 22ºC, with 16 h light/8 h dark cycle. Light intensity was maintained at 200 μM/m^2 with Octron 4100K Ecologic bulbs (Sylvania, http://www.sylvania.com). The *Pseudomonas syringae* pv. tomato DC3000 was activated on pseudomonas agar supplemented with kanamycin (25 ng/μl) and rifampicin

(100 ng/µl) at 28ºC for two days. The bacteria were suspended in sterile water and the concentration was adjusted to OD_{600}=0.2 (10^8 cfu/ml). Surfactant Silwet L-77 (0.02%) was added to the suspension. Treatment with sterile water with 0.02% Silwet L-77 was used as a negative control. Whole inflorescences, still attached to the plant, were inoculated using pipette with 100 µl of bacterial solution. Plants were covered with a plastic dome to maintain high humidity levels and grown as described above. Seven days after inoculation, phenotypes were assessed and silique lengths were measured and recorded. For the bacterial growth assay, the inflorescences were collected five days after inoculation. Three samples, each containing four individual inflorescences, were weighed and placed in 1.5-ml centrifuge tubes, containing 100µl sterile water. Samples were ground using a polypropylene pestle that was driven by a Pellet Pestle Motor (Kontes, Vineland NJ) and 900 µl of sterile water was added to each sample. Each sample was then serially diluted in 1:10 steps to 10^{-7} in water. Ten microliters of each dilution were spotted on Pseudomonas Agar and incubated at 28ºC. After two days, colony numbers were recorded. Bacterial growth is presented as colony forming unit/fresh weight of inflorescence sample. Data were statistically analyzed using Fisher's PLSD test.

Dot-blot Gene Expression analysis

Col-0 and *npr3-3* mutant plants were grown in a pest-free growth chamber at 22ºC. Six weeks after germination, inflorescences were challenged as described above. Opened and developing flowers were collected 24 h after treatments. Three biological replicates, each containing five individually infected plants, were collected. RNA was extracted using the RNeasy Plant Mini Kit (Qiagen) and was reverse-transcribed to cDNA by MMLV Reverse Transcriptase (New England BioLabs). RT-PCR was performed with cDNA templates and intron-spanning primers for *PR1*, *NPR1,* and *NPR3,* and for *ubiquitin,* as a control (see Table S4 for primer sequences) in 96 well PCR plates. PCR was terminated after 18 cycles and 0.4 M NaOH and 20 mM EDTA (final concentration) were added to each reaction. Complete reaction solutions were then transferred onto Hybond-N^+ Nylon membrane (Amersham, Pittsburgh PA) using a dot-blot apparatus, according to the user's manual (Cat. 170-6542,

Biorad, Hercules CA). After air drying overnight, membranes were UV cross-linked by GS gene Linker UV chamber (Biorad, Hercules CA) and hybridized using probes generated from PCR products of the target genes labeled with P^{32} using the MEGA Labeling Kit (GE Healthcare, Piscataway NJ). Hybridization product intensity were detected and quantified by phosphorimaging using STORM 860 scanner and ImageQuant software (GE Healthcare, Pittsburgh PA) half an hour after exposure. The pixel intensity of the target genes was normalized by dividing their intensity values to these of *ubiquitin* control. The mean and standard error of expression values relative to *ubiquitin* for three biological replicates were calculated and plotted.

Overexpression of *NPR3*

The full-length cDNA of *NPR3* was amplified by PCR and assembled into the pUC19 cloning vector downstream of the modified CaMV35S E12Ω promoter (Mitsuhara et al., 1996). The cassette was cloned into the pCAMBIA1300 SmaI site. The *NPR3* overexpression vector was transferred into *Agrobacterium* AGlI and subsequently used for transformation of Col-0 plants by floral dip method (Clough and Bent, 1998). Transgenic plants were selected on MS medium supplemented with 25 ng/µl hygromycin B. The inflorescences of five independent transgenic lines were inoculated with *P.s.t.* as described above, using a bacterial suspension at 0.02 OD_{600}. Disease evaluation in leaves (infection and bacterial growth assays) was carried out as previously described (Katagiri et al., 2002). Expression levels of *NPR3* in five individual lines were determined by dot-blot analysis as described above.

Bimolecular Fluorescence Complementation (BiFC) Assay in Transiently Transformed Onion Epidermal Cells

To evaluate the interaction between NPR3 and TGA2, full-length cDNA of *TGA2* was amplified by PCR from cDNA clone U50303 and cloned into pSAT4A-cEYFP-N1 to obtain a cassette containing the 2x35S promoter, the *TGA2* coding sequence, and the C-terminal

fragment of YFP. The full-length *NPR3* cDNA was amplified by PCR from cDNA clone U22208 and cloned into pSAT4A-nEYFP-N1 to obtain a cassette containing the 2x35S promoter, the *NPR3* coding sequence, and the N-terminal fragment of YFP (see Table S5 for primer sequences and cloning sites). Both cassettes were confirmed by sequencing and then transferred into pCAMBIA1300 binary vectors. For transient expression assay, 2µg each of two vectors (*NPR3:YFP^N* and *TGA2:YFP^C*) were mixed at a ratio of 1:1 and co-bombarded into onion epidermal cells as previously described (von Arnim and Deng, 1994). Model PDS-1000/He Biolistic particle delivery system (Biorad, Hercules CA) was used for bombardment with tungsten M-10 (0.7 micron) as microcarrier and 1100 psi rupture disk.

Similarly, to evaluate the interaction between NPR1 and TGA2, full length *NPR1* cDNA was amplified by PCR from cDNA clone U13446 and cloned into pSAT4A-nCerulean-N to obtain a cassette containing the 2x35S promoter, the *NPR1* coding sequence, and the N-terminal fragment of CFP. The *NPR1:CFP^N* and *TGA2:YFP^C* vectors were co-bombarded in onion cells.

The interaction between NPR1 and NPR3 was also evaluated in onion cells. For this, *NPR1* full-length cDNA was amplified by PCR from cDNA clone U13446 and cloned into pSAT4A-cEYFP-N1 to obtain a cassette containing the 35S promoter, the *NPR1* coding sequence, and the C-terminal fragment of YFP. Vectors *NPR3:YFP^N* and *NPR1:YFP^C* were co-bombarded. Transient expression of CaMV35S$_{pro}$:EGFP served as a positive control and co-bombardment of 35S$_{pro}$:YFP^N and 35S$_{pro}$:YFP^C served as a negative control. In all experiments transient expression of CFP and YFP in onion cells was documented by confocal microscopy as described above.

Bimolecular Fluorescence Complementation in Stably Transformed Plants

Binary vectors *NPR3:YFP^N, TGA2:YFP^C,* and *NPR1:CFP^N* (described above) were used to transform Col-0 plants by floral dip method (Clough and Bent, 1998). Developing flowers from transgenic plants co-expressing NPR3:YFP^N and TGA2:YFP^C fusion proteins were inoculated with a *P.s.t.* DC3000 suspension of OD_{600}=0.2 in 0.02% Silwet L-77. Control

samples were treated with a 0.02% solution of Silwet L-77 in water. Twenty-four hours after infection, petals were collected from control and induced plants and observed with confocal microscope as described. Transgenic negative control plants were generated by co-transformation with vectors containing only the YFP^N or YFP^C fragments (BiFC control half-genes). Flowers from transgenic plants co-expressing NPR1:CFP^N and TGA2:YFP^C were also challenged and evaluated by confocal microscopy. Transgenic plants co-expressing both CFP^N and YFP^C fragments only served as negative controls (BiFC control half-genes).

For a bimolecular fluorescence competition assays, flowers from transgenic plants co-expressing NPR3:YFP^N, TGA2:YFP^C, and NPR1:CFP^N underwent the same infection and observation process, but were sequentially excited at 515 nm and 458 nm to image YFP and CFP, respectively. Transgenic plants co-expressing all three fluorescent protein halves YFP^N, YFP^C, and CFP^N without the genes of interest served as the half-gene negative controls. Merged images were obtained by overlaying YFP and CFP channels.

Whole inflorescences of six-week-old, soil-grown Col-0, *npr3-3* mutant, BiFC transgenic plants and BiFC competition plants (transgenic plants co-expressing NPR3:YFP^N and TGA2:YFP^C fusion proteins, plants co-expressing NPR1:CFP^N and TGA2:YFP^C and plants co-expressing NPR3:YFP^N, TGA2:YFP^C, and NPR1:CFP^N three fusion proteins) were inoculated with *P.s.t.* DC3000 at concentration of 0.2 OD_{600} and bacterial growth was assayed as described above.

Fitness test

Col-0, *npr1-3*, *npr3-2*, *npr3-3,* and CAMV35S_{pro}:*NPR3* seeds were cultured on MS medium containing 4.3 g MS/L, 2% sucrose, and 0.7% agar. Plants were grown vertically in a growth chamber (22°C; 16 h light/8 h dark). One week after germination, the primary root length of 16 plants grown on individual plates was measured for each genotype.

Additionally, Col-0, *npr1-3*, *npr3-2,* and *npr3-3* plants were grown in soil in a pest-free growth chamber (Convrion Model No. MTPS144, 22°C; 16 h light/8 h dark). Starting at week three, 10 plants were randomly selected at weekly intervals up to week seven, the

above-ground portion was removed from soil, dried for 48 h at 70°C, and used to measure whole plant dry weight. Relative growth rate (RGR) ($g \cdot g^{-1} \cdot day^{-1}$) was calculated according to the formula RGR= (InW2-InW1)/t2-t1, where W1 and W2 were the dry weights of plants at time points t1 and t2, respectively. Plants were regularly watered until they reached senescence, seeds were collected, and the total weight of seeds per plant was determined. The total number of seeds in 10 mg samples per plant was counted under a dissecting microscope for each sample. Data was collected from three biological replicates.

Dominance of *NPR3*

An F_2 segregation population of 94 plants from a cross between Col-0 and *npr3-3* was planted in soil. The flowers of six-week-old plants were infected with *P.s.t.* and the bacterial growth assay was performed as described above. After the infected flowers were collected, the leaves were also collected for genomic DNA extraction. The genotype of each plant was determined by PCR analysis, as described in the section on mutant identification.

For the root assay, an F_2 segregation population of 94 plants from a cross between Col-0 and *npr3-3* was cultured on MS solid medium and grown vertically. One week after germination, the primary root length was measured for each plant and genomic DNA was extracted from seedlings to determine the genotype of each plant.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: At1g64280 (*NPR1*), At4g26120 (*NPR2*), At5g45110 (*NPR3*), At4g19660 (*NPR4*), At2g41370 (*BOP2*), At3g57130 (*BOP1*), At2g14610 (*PR1*), At3g52590 (*ubiquitin*) and At2g37620 (*Actin*).

Authors' contributions

ZS performed most of the experiments, *ie,* promoter-GFP plants construction and observation, gene expression analysis, bacterial growth assays, BIFC plant construction and confocal microscopy observations, fitness tests and drafted the manuscript. SNM

participated in the design of the study, directed the transformation vector construction and transgenic lines generation, and participated in drafting of the manuscript. YL participated in RNA isolation for gene expression analysis and helped to analyze the sequence. JV initiated the experiments with sequence analysis of six genes in the family. MJG conceived the study, drafted the manuscript and gave advice on experimental design, data analysis and execution. All authors read and approved the final manuscript.

Acknowledgement

We would like to thank Dr. Stanton Gelvin at Purdue University for providing the BiFC vectors, Nicole Zembower at Penn State Cytometry Facility for providing help with microscopy, Ann Young and Sharon Pishak for their technical assistance, and Jonathon Arias for critical review. This work was supported in part by The Pennsylvania State University, College of Agricultural Sciences and The Huck Institutes of Life Sciences.

2.6 Supplemental Data and Figures

Supplemental Table 1. Primer sequences used in this study.

Table 1. Primer sequences used for *npr3-3* and *npr3-2* genotyping.

Primer Name	Primer Sequences
2LP3	5'-AGTGATCCGAAAGTTGTTGCC-3'
2RP3	5'-CCCTTCTCAAAATGGAAATCG-3'
2LP2	5'-TGTGTTCCATTTCTATGCCTG-3'
LBa1	5'-TGGTTCACGTAGTGGGCCATCG-3'

Supplemental Table 2. Primer sequences used to amplify promoters of *NPR* genes and cDNA clones used.

Gene Name	Primer Sequences	BAC Clone ID
NPR1	5'-CGAAGCTTTTTGGCAATTGGGTGTTT-3' 5'-ATCCATGGAATCGGCGAATCCATCAA-3'	BAC F15H21
NPR2	5'-GCAAGCTTAAGAAAGCAAGACCAAAA-3' 5'-CGCCATGGGCATCACGAGATCGGCG-3'	BAC F20B18
NPR3	5'-GCAAGCTTTAGCCTTGTAGGTAGAGT-3' 5'-CTCCATGGCCTTCTTGTAGAATTATCA-3'	BAC K17O22
NPR4	5'-CGAAGCTTCTTGGCTTGAGAGGTTTG-3' 5'-CGCCATGGTTTGTTCCCCAGAAGTAT-3'	BAC T16H5
BOP2	5'-CGAAGCTTAACCCATAGTATTTCTTA-3' 5'-CGCCATGGAGTTGCCTCCTTTTATTG-3'	BAC F13H10
BOP1	5'-GCAAGCTTTTCTGCGATAAATGTGGT-3' 5'-GCCCATGGAGCTCCTTTGTTGATTTC-3'	BAC F24I13

Supplemental Table 3. Primer and probe sequences for TaqMan® Assays of *NPR* genes in different plant tissues.

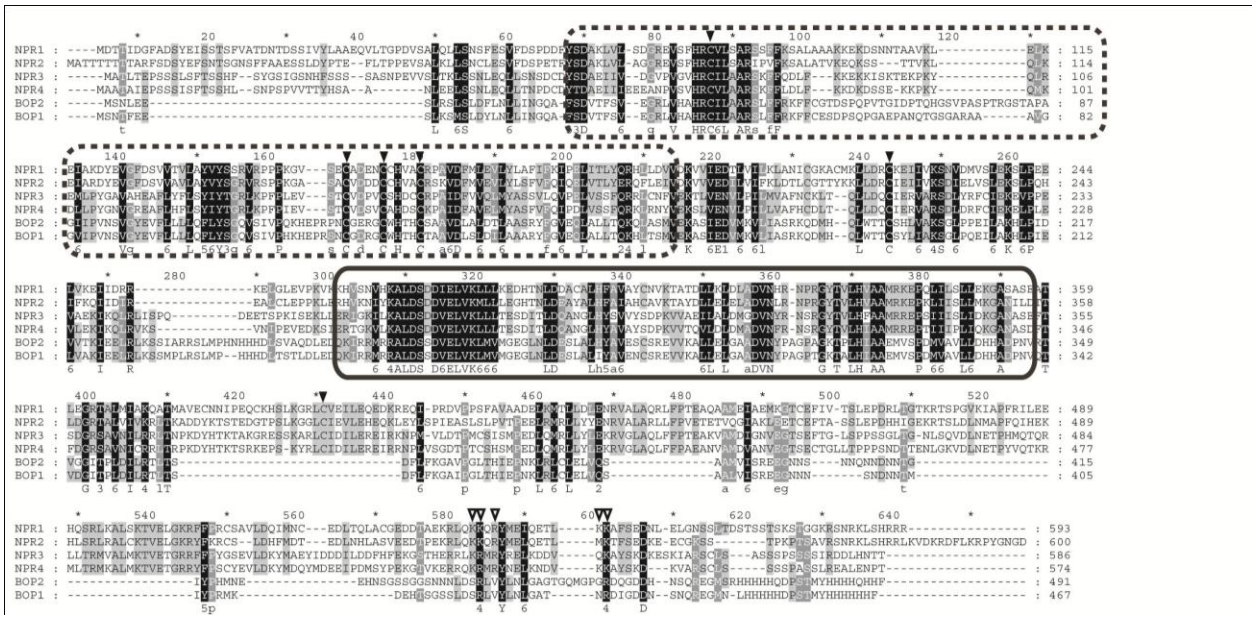
Gene Name	Primer Sequences	Probe Sequences
NPR1	5'-GTCACGTGCTGTCTGACCTTGT-3' 5'-TCACAGTTCATAATCTGGTCGAGC-3'	5'-TYCCGCGCTGTTCGGCAGT-3'
NPR2	5'-TTGCGTAGACGACGATTGTTG-3' 5'-AAAGAACCTCCACCATGAAATCC-3'	5'-CACGTGGCTTGCCGGTCAAAG G-3'
NPR3	5'-AAGGATGATGTCCAAAAGGCATA-3' 5'-AAGACAAGACCGCGCAATCT-3'	5'-AGCAAAGACAAAGAGTCT-3'
NPR4	5'-CACCCACTTGTTCCCATTCG-3' 5'-CTGAGCAAGTCCCACTCGCT-3'	5'-TGCCCGAGGATCTCCAAATGAGGTT-3'
BOP2	5'-GAGCTGCCGATGTGAATTATCC-3' 5'-CAGCCGCGATGTGCAA-3'	5'-CGGGTCCGGCAGGGAAAACAC-3'
BOP1	5'-GGCCGTCTAGTCCACGCT C-3' 5'-GAATTTGCGGAAGAAGAGGCT-3'	5'-CCGTTGCATCCTCGCTGCTCG-3'
Ubiquitin	5'-AGGCCTCAACTGGTTGCTGT-3' 5'-ACCGGCAAGACCATCACTCT-3'	5'-CGAGAGCAGCGACACCCATCGACA-3'
Actin	5'-GATTCAGATGCCCAGAAGTCTTG-3' 5'-TCTCGTGGATTCCAGCAGCT-3'	5'-CCAGCCCTCGTTTGTGGGAAAGG-3'

Supplemental Table 4. Primer sequences used for amplifying cDNA fragments for dotblot analysis.

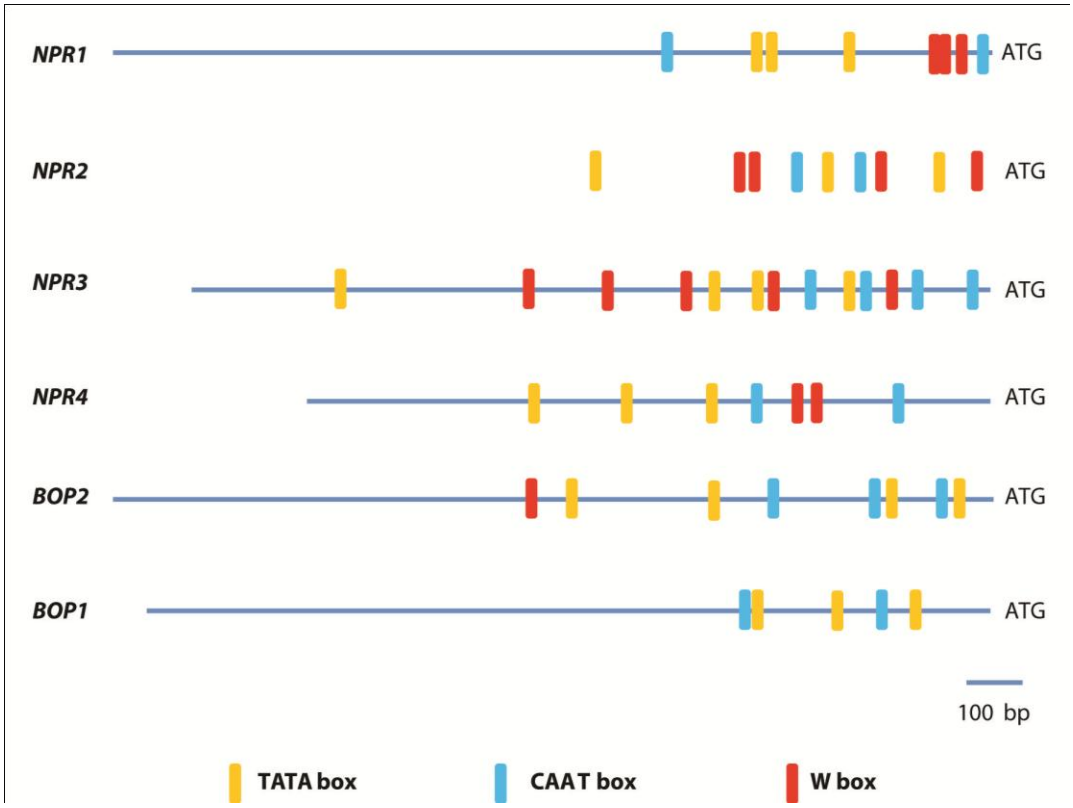
Gene Name	Primer Sequences
NPR1	5'- ATGCGTGTGCTCTTCATTTCGCTG-3' 5'- ATCGTTTCCCGAGTTCCACGGTTT-3'
NPR3	5'- AGTGATCCGAAAGTTGTTGCCGAG-3' 5'- ACCTTCGACCAGTCTCAACTGT-3'
PR1	5'- CTCGAAAGCTCAAGATAGCCCACA-3' 5'- CTTCTCGTTCACATAATTCCCACG-3'
Ubiquitin	5'- ACCGGCAAGACCATCACTCT-3' 5'- AGGCCTCAACTGGTTGCTGT-3'

Supplemental Table 5. Primer sequences used for BiFC vector construction and vector names.

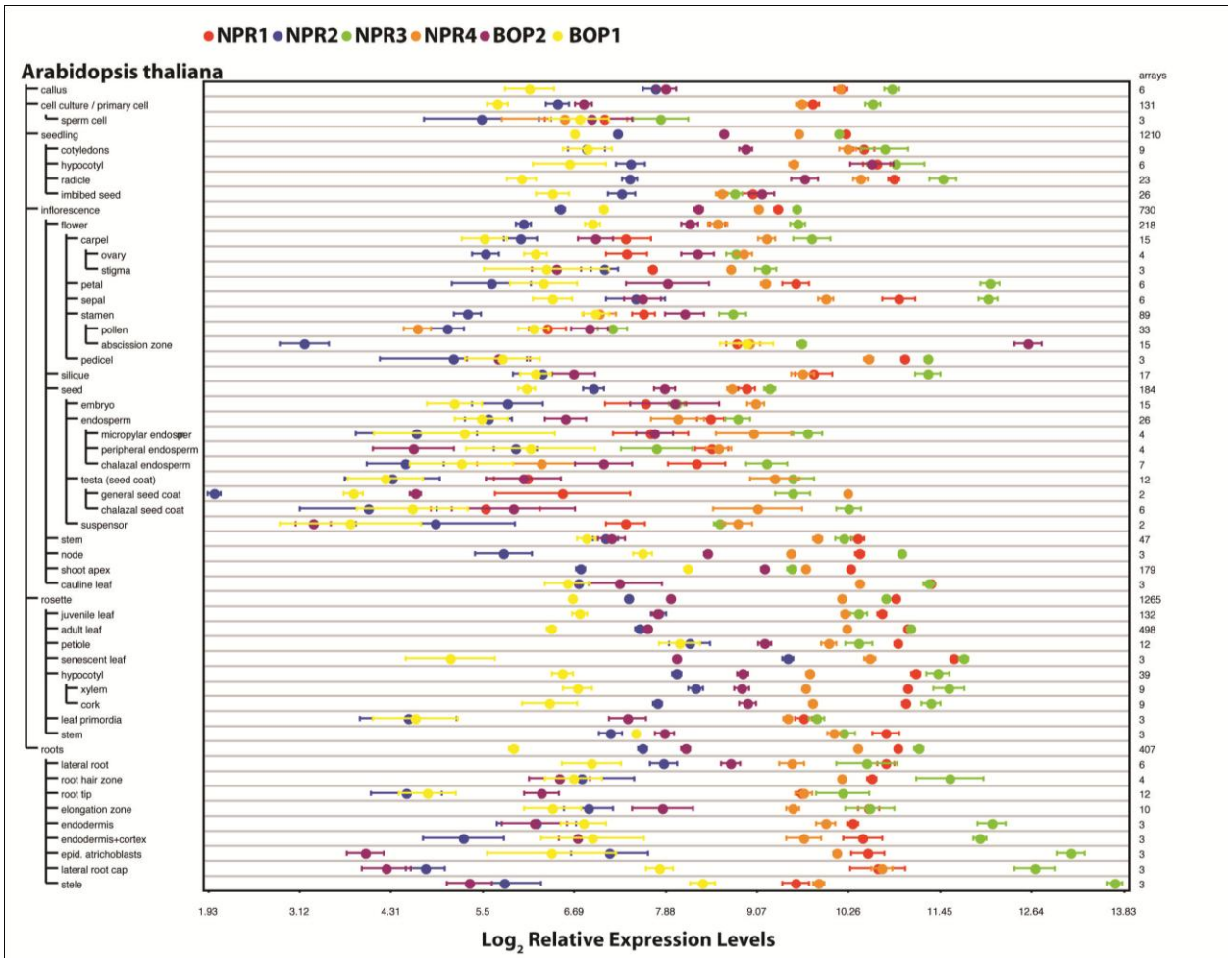
Gene Name	Primer Sequences and cloning site	BiFC Vector Name
NPR1	5'-GCGAGCTCCCATGGACACCACCAT (SacI) 5'-GCCCGCGGCACCGACGACGATGAG-3' (SacII)	pSAT4A-nCerulean-N (pE3247)
NPR1	5'-GCGAGCTCCCATGGACACCACCAT-3' (SacI) 5'-GCCCGCGGCACCGACGACGATGAG-3'(SacII)	pSAT4A-cEYFP-N1 (pE3323)
TGA2	5'-GCGAATTCCATGGCTGATACCAGT-3' (EcoRI) 5'-ATGGTACCACTCTCTGGGTCGAGC-3' (KpnI)	pSAT4A-cEYFP-N1 (pE3323)
NPR3	5'-GCGAGCTCCCATGGCTACTTTGAC-3' (SacI) 5'-ATCCGCGGCATGTTGTGTTGTGC-3' (SacII)	pSAT4A-nEYFP-N1 (pE3085)



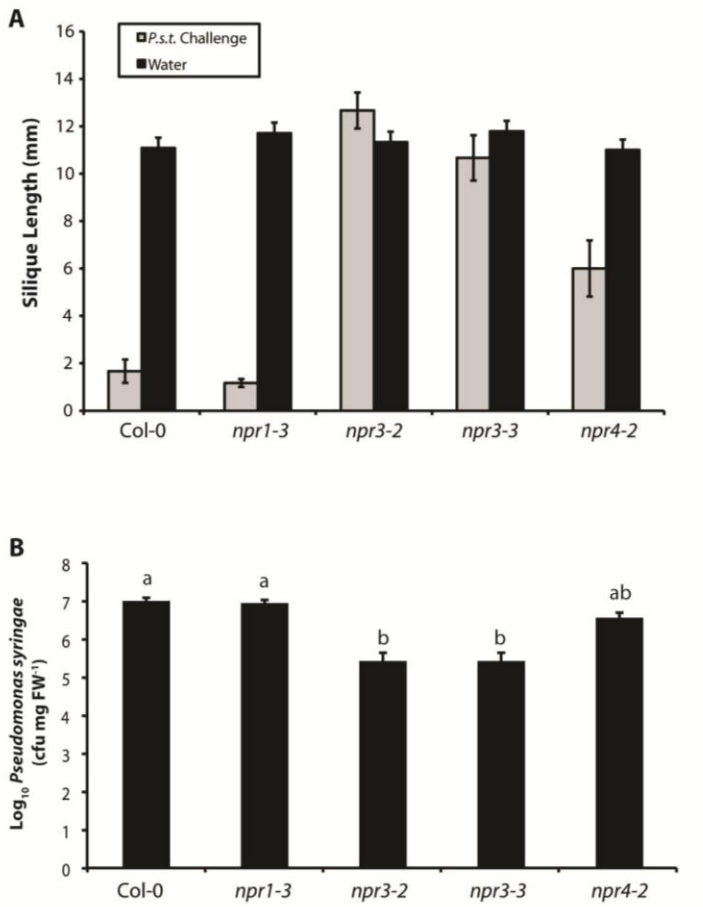
Supplemental Figure 2-1 Protein sequence alignment of the six members of the *NPR* family. Protein alignment was generated using MUSCLE software. Amino acid residues highlighted in back share similar chemical properties in all six members and residues highlighted in grey are common for some of the members. Numbers refer to the amino acid position on NPR1. BTB/POZ and ankyrin repeats domains are represented by dashed line box and solid line box, respectively. Conserved cysteines are indicated by solid black arrows. The five basic amino acids facilitating the nuclear localization of NPPR1 are indicated by empty black arrows.



Supplemental Figure 2-2 Schematic representation of cis-acting element in the promoters of *NPR* family genes. The blue lines represent the promoter regions of the six members (about 1.5kb upstream of ATG start codon). Putative TATA boxes (yellow blocks), CAAT boxes (blue blocks) and W boxes (red blocks) were identified by querying the PLACE database (http://www.dna.affrc.go.jp/PLACE/).



Supplemental Figure 2-3 Expression profiles of NPR gene family members from microarray database resource Genevestigator. Summary of relative expression levels (Log_2) of *NPR* family genes in various tissues of *Arabidopsis*, generated by Genevestigator software. The expression of each gene is represented by different color.



Supplemental Figure 2-4 Analysis of silique length and bacterial growth in inflorescences from Col-0 and mutants *npr1-3*, *npr3-2*, *npr3-3* and *npr4-2* plants after inoculation with virulent *P.s.t.* DC3000. (A) Silique lengths of all siliques from four inflorescences per genotype per treatment (inoculated and control) were measured seven days after inoculation and mean lengths ± SE were calculated. (B) Bacterial populations were determined in the flowers of the five genotypes at five days after infection. Means ± SE of three biological replicates per genotype, each containing four infected inflorescences were calculated. Significant differences among the genotypes ($P<0.05$) were determined by Fisher's PLSD analysis.

2.7 References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389-3402

Baldwin AS, Jr. (1996) The NF-kappa B and I kappa B proteins: new discoveries and insights. Annu Rev Immunol **14:** 649-683

Beckers GJ, Spoel SH (2006) Fine-tuning plant defence signalling: Salicylate versus jasmonate. Plant Biol **8:** 1-10

Blanco F, Salinas P, Cecchini NM, Jordana X, Van Hummelen P, Alvarez ME, Holuigue L (2009) Early genomic responses to salicylic acid in *Arabidopsis*. Plant Mol Biol **70:** 79-102

Bracha-Drori K, Shichrur K, Katz A, Oliva M, Angelovici R, Yalovsky S, Ohad N (2004) Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. Plant J **40:** 419-427

Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an Arabidopsis mutant that is nonresponsive to inducers of systemic acquired resistance. Plant Cell **6:** 1583-1592

Cao H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The *Arabidopsis NPR1* gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell **88:** 57-63

Clarke JD, Volko SM, Ledford H, Ausubel FM, Dong X (2000) Roles of salicylic acid, jasmonic acid, and ethylene in cpr-induced resistance in Arabidopsis. Plant Cell **12:** 2175-2190

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735-743

Delaney TP, Friedrich L, Ryals JA (1995) *Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance. Proc Natl Acad Sci U S A **92:** 6602-6606

Despres C, DeLong C, Glaze S, Liu E, Fobert PR (2000) The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. Plant Cell **12:** 279-290

Dong X (2004) NPR1, all things considered. Curr Opin Plant Biol **7:** 547-552

Fan W, Dong X (2002) In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis. Plant Cell **14:** 1377-1389

Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, Chen L, Yu L, Iglesias-Pedraz JM, Kircher S, Schafer E, Fu X, Fan LM, Deng XW (2008) Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. Nature **451:** 475-479

Figueroa P, Gusmaroli G, Serino G, Habashi J, Ma L, Shen Y, Feng S, Bostick M, Callis J, Hellmann H, Deng XW (2005) Arabidopsis has two redundant Cullin3 proteins that are essential for embryo development and that interact with RBX1 and BTB proteins to form multisubunit E3 Ubiquitin ligase complexes in vivo. Plant Cell **17:** 1180-1195

Ghosh S, May MJ, Kopp EB (1998) NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol **16:** 225-260

Glazebrook J, Rogers EE, Ausubel FM (1996) Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening. Genetics **143:** 973-982
Grant M, Lamb C (2006) Systemic immunity. Curr Opin Plant Biol **9:** 414-420
Ha CM, Jun JH, Nam HG, Fletcher JC (2004) *BLADE-ON-PETIOLE1* encodes a BTB/POZ domain protein required for leaf morphogenesis in *Arabidopsis thaliana*. Plant Cell Physiol **45:** 1361-1370
Hepworth SR, Zhang Y, McKim S, Li X, Haughn GW (2005) BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. Plant Cell **17:** 1434-1448
Johnson C, Boden E, Arias J (2003) Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. Plant Cell **15:** 1846-1858
Johnson C, Mhatre A, Arias J (2008) NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2. Biochim Biophys Acta **1779:** 583-589
Jun JH, Ha CM, Fletcher JC (2010) BLADE-ON-PETIOLE1 coordinates organ determinacy and axial polarity in Arabidopsis by directly activating ASYMMETRIC LEAVES2. Plant Cell **22:** 62-76
Katagiri F, Thilmony R, He S (2002) The *Arabidopsis thaliana-Pseudomonas syringae* interaction. *In* The Arabidopsis Book. The American Society of Plant Biologists
Kinkema M, Fan W, Dong X (2000) Nuclear localization of NPR1 is required for activation of *PR* gene expression. Plant Cell **12:** 2339-2350
Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief Bioinform **5:** 150-163
Liu G, Holub EB, Alonso JM, Ecker JR, Fobert PR (2005) An *Arabidopsis NPR1-like* gene, *NPR4*, is required for disease resistance. Plant J **41:** 304-318
Loake G, Grant M (2007) Salicylic acid in plant defence--the players and protagonists. Curr Opin Plant Biol **10:** 466-472
McKim SM, Stenvik GE, Butenko MA, Kristiansen W, Cho SK, Hepworth SR, Aalen RB, Haughn GW (2008) The *BLADE-ON-PETIOLE* genes are essential for abscission zone formation in *Arabidopsis*. Development **135:** 1537-1546
Mitsuhara I, Ugaki M, Hirochika H, Ohshima M, Murakami T, Gotoh Y, Katayose Y, Nakamura S, Honkura R, Nishimiya S, Ueno K, Mochizuki A, Tanimoto H, Tsugawa H, Otsuki Y, Ohashi Y (1996) Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. Plant Cell Physiol **37:** 49-59
Mou Z, Fan W, Dong X (2003) Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. Cell **113:** 935-944
Mukhtar MS, Nishimura MT, Dangl J (2009) NPR1 in plant defense: It's not over 'til It's turned over. Cell **137:** 804-806
Ndamukong I, Abdallat AA, 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
Norberg M, Holmlund M, Nilsson O (2005) The *BLADE ON PETIOLE* genes act redundantly to control the growth and development of lateral organs. Development **132:** 2203-2213

Pieterse CM, Van Loon LC (2004) NPR1: the spider in the web of induced resistance signaling pathways. Curr Opin Plant Biol **7:** 456-464

Rochon A, Boyle P, Wignes T, Fobert PR, Despres C (2006) The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. Plant Cell **18:** 3670-3685

Ryals J, Weymann K, Lawton K, Friedrich L, Ellis D, Steiner HY, Johnson J, Delaney TP, Jesse T, Vos P, Uknes S (1997) The Arabidopsis NIM1 protein shows homology to the mammalian transcription factor inhibitor I kappa B. Plant Cell **9:** 425-439

Scott RJ, Spielman M, Dickinson HG (2004) Stamen structure and function. Plant Cell **16 Suppl:** S46-60

Shah J, Tsui F, Klessig DF (1997) Characterization of a salicylic acid-insensitive mutant (*sai1*) of *Arabidopsis thaliana*, identified in a selective screen utilizing the SA-inducible expression of the *tms2* gene. Mol Plant Microbe Interact **10:** 69-78

Smyth DR, Bowman JL, Meyerowitz EM (1990) Early flower development in Arabidopsis. Plant Cell **2:** 755-767

Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P, Dong X (2009) Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. Cell **137:** 860-872

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG (2005) Sequence and structural analysis of BTB domain proteins. Genome Biol **6:** R82

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:** 4673-4680

von Arnim AG, Deng XW (1994) Light inactivation of *Arabidopsis* photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning. Cell **79:** 1035-1045

Walter M, Chaban C, Schutze K, Batistic O, Weckermann K, Nake C, Blazevic D, Grefen C, Schumacher K, Oecking C, Harter K, Kudla J (2004) Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant J **40:** 428-438

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 Pathog **2:** e123

Yang P, Chen C, Wang Z, Fan B, Chen Z (1999) A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. Plant J **18:** 141 - 149

Zhang Y, Cheng YT, Qu N, Zhao Q, Bi D, Li X (2006) Negative regulation of defense responses in *Arabidopsis* by two *NPR1* paralogs. Plant J **48:** 647-656

Zhang Y, Fan W, Kinkema M, Li X, Dong X (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene. Proc Natl Acad Sci U S A **96:** 6523-6528

Zhang Y, McCormick S (2007) A distinct mechanism regulating a pollen-specific guanine nucleotide exchange factor for the small GTPase Rop in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A **104:** 18830-18835

Zhou JM, Trifa Y, Silva H, Pontier D, Lam E, Shah J, Klessig DF (2000) NPR1 differentially interacts with members of the TGA/OBF family of transcription factors

that bind an element of the *PR-1* gene required for induction by salicylic acid. Mol Plant Microbe Interact **13:** 191-202

Zimmermann P, Hennig L, Gruissem W (2005) Gene-expression analysis and network discovery using Genevestigator. Trends Plant Sci **10:** 407-409

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

CHAPTER 3: FUNCTIONAL ANALYSIS OF THE *THEOBROMA CACAO NPR1* GENE IN *ARABIDOPSIS*

Zi Shi, Siela N. Maximova , Yi Liu, Joseph Verica and Mark J. Guiltinan

3.1 Abstract

Background: The *Arabidopsis thaliana NPR1* gene encodes a transcription coactivator (NPR1) that plays a major role in the mechanisms regulating plant defense response. After pathogen infection and in response to salicylic acid (SA) accumulation, NPR1 translocates from the cytoplasm into the nucleus where it interacts with other transcription factors resulting in increased expression of over 2000 plant defense genes contributing to a pathogen resistance response.

Results: A putative *Theobroma cacao NPR1* cDNA was isolated by RT-PCR using degenerate primers based on homologous sequences from *Brassica*, *Arabidopsis* and *Carica papaya*. The cDNA was used to isolate a genomic clone from *Theobroma cacao* containing a putative *TcNPR1* gene. DNA sequencing revealed the presence of a 4.5 kb coding region containing three introns and encoding a polypeptide of 591 amino acids. The predicted TcNPR1 protein shares 55% identity and 78% similarity to *Arabidopsis* NPR1, and contains each of the highly conserved functional domains indicative of this class of transcription factors (BTB/POZ and ankyrin repeat protein-protein interaction domains and a nuclear localization sequence (NLS)). To functionally define the *TcNPR1* gene, we transferred *TcNPR1* into an *Arabidopsis npr1* mutant that is highly susceptible to infection by the plant pathogen *Pseudomonas syringae* pv. tomato DC3000. Driven by the constitutive CaMV35S promoter, the cacao *TcNPR1* gene partially complemented the *npr1* mutation in transgenic *Arabidopsis* plants, resulting in 100 fold less bacterial growth in a leaf infection assay. Upon induction with SA, *TcNPR1* was shown to translocate into the nucleus of leaf and root cells in a manner identical to *Arabidopsis* NPR1. Cacao NPR1 was also capable of participating in SA-JA signaling crosstalk, as evidenced by the suppression of JA responsive gene expression in *TcNPR1* overexpressing transgenic plants.

Conclusion: Our data indicate that the *TcNPR1* is a functional ortholog of *Arabidopsis NPR1*, and is likely to play a major role in defense response in cacao. This fundamental knowledge can contribute to breeding of disease resistant cacao varieties through the application of molecular markers or the use of transgenic strategies.

3.2 Background

Plants have evolved a complex network of defense responses, often associated with a response local to the site of infection (Heath, 2000; Durrant and Dong, 2004; Mur et al., 2008; Attaran et al., 2009). In addition, defenses are also systemically induced in remote parts of the plant in a process known as systemic acquired resistance (SAR) (Dong, 2004; Durrant and Dong, 2004; Mukhtar et al., 2009). Induction of the SAR pathway leads to heightened broad-spectrum resistance to secondary pathogen attacks by a variety of pathogens. Multiple studies in both monocots and dicots have shown that salicylic acid (SA) plays a central role as a signaling molecule in SAR (Cao et al., 1994; Liu et al., 2002; Fitzgerald et al., 2004; Makandar et al., 2006; Yuan et al., 2007; Vlot et al., 2008; Chen et al., 2009; Le Henanff et al., 2009) . Following pathogen attack, SA levels increase both locally and systemically in infected plants. In addition, SA is required for the induced expression of a set of pathogenesis-related (*PR*) genes (Gaffney et al., 1993; Cao et al., 1994; Rochon et al., 2006; Loake and Grant, 2007).

NPR1 was originally identified by screening for mutants that were insensitive to SA (or its chemical analogs, 2,6-dichloroisonicotic acid (INA) or benzothiadiazole (BTH)) in *Arabidopsis* (Cao et al., 1994; Delaney et al., 1995; Glazebrook et al., 1996; Shah et al., 1997). These screens identified a mutation designated as *Non-Expressor of PR1* (*NPR1*). Studies that followed further documented that *npr1* mutants displayed reduced expression of *PR* genes upon SA treatment and were more susceptible to pathogens (Cao et al., 1994; Glazebrook et al., 1996; Ryals et al., 1997; Shah et al., 1997). Conversely, when *NPR1* was overexpressed, the resulting transgenic plants displayed increased resistance to pathogens, and were able to induce increased levels of *PR* genes in a dose-dependent fashion (Cao et al., 1998).

NPR1 encodes a protein containing ankyrin repeats and a BTB/POZ domain, both of which mediate protein-protein interactions in animals (Stogios et al., 2005). NPR1 shares homology with IκBα transcription inhibitors, which regulate the innate immunity response (Baldwin, 1996; Ryals et al., 1997). Recent work has shed light onto the mechanisms of NPR1 function (Feys and Parker, 2000; Mou et al., 2003; Dong, 2004; Pieterse and Van Loon, 2004; Loake and Grant, 2007; Vlot et al., 2008; Mukhtar et al., 2009). *NPR1* is constitutively expressed, and NPR1 protein is present as inactive oligomers in the cytoplasm of the cell. Upon SAR induction, the redox state of the cell is altered, resulting in the reduction of NPR1 to its active monomeric form. Monomeric NPR1 moves into the nucleus where it can affect the induction of *PR* genes. Although NPR1 itself has no DNA binding domains, it participates in the regulation of defense gene transcription via interactions with TGA transcription factors (Despres et al., 2000; Zhou et al., 2000; Fan and Dong, 2002; Rochon et al., 2006; Ndamukong et al., 2007; Johnson et al., 2008; Boyle et al., 2009). In *Arabidopsis*, two conserved cysteine residues (C82 and C216) have been shown to be essential to the oligomerization and cytoplasmic localization of AtNPR1 (Mou et al., 2003). Mutation of these residues results in constitutive monomerization and nuclear localization of NPR1.

It is believed that NPR1 also plays a role in the jasmonic acid (JA) signaling pathway and mediates the crosstalk between SA-JA defense pathways to fine-tune defense responses (Feys and Parker, 2000; Spoel et al., 2003; Ndamukong et al., 2007; Koornneef and Pieterse, 2008; Leon-Reyes et al., 2009). SA-mediated defenses are mainly effective against biotrophic pathogens, whereas JA-mediated defenses are predominantly efficient against necrotrophic pathogens and herbivorous insects. NPR1 mediates the antagonistic effect of SA on JA signaling by suppressing the expression of JA-responsive genes upon combined treatment of SA and methyl jasmonate (MeJA) (Spoel et al., 2003).

A growing body of evidence has revealed that the salicylic acid dependent, NPR1-mediated defense pathway is also conserved in other plant species across wide phylogenetic distances. Two *NPR1-like* genes have been characterized from *Vitis vinifera* (grapevine) (Le Henanff et al., 2009). When translational fusions of the proteins encoded

by the two genes with GFP were transiently expressed in *Nicotiana benthamiana* leaves, the proteins were localized predominantly to the nucleus and triggered the accumulation of pathogenesis-related proteins PR1 and PR2. In addition, the silencing of a tomato *NPR1-like* gene leads to increased bacterial growth upon *Ralstonia solanacearum* infection in tomato (Chen et al., 2009). In tobacco, the suppression of *NPR1-like* gene leads to increased susceptibility to tobacco mosaic virus (Liu et al., 2002). Similarly, overexpression of the apple *MpNPR1* gene in transgenic apple plants resulted in the up-regulation of *PR* genes and enhanced resistance to bacterial and fungal pathogens (Malnoy et al., 2007). In wheat, the expression of *Arabidopsis NPR1* confers resistance to Fusarium head blight in susceptible cultivar Bobwhite (Makandar et al., 2006). Major efforts have been made to study the SA and NPR1-dependent pathway in rice, the model monocot plant. Treatment of rice plants with the salicylic acid analog probenazole results in enhanced resistance against rice blast fungus (Sakamoto et al., 1999). In addition, rice plants expressing bacterial salicylate hydrolase (*nahG*) are unable to accumulate salicylic acid and display increased susceptibility to rice blast (Yang et al., 2004). Overexpression of the *Arabidopsis NPR1* gene in rice leads to enhanced resistance to the bacterial pathogen *Xanthomonas oryzae* pv. *oryzae* (Fitzgerald et al., 2004). An ortholog of *NPR1* has been isolated from rice (*OsNPR1*/*NH1*), and the overexpression of *OsNPR1* in rice leads to enhanced resistance to both bacterial and oomycete pathogens (Chern et al., 2005). Moreover, *OsNPR1* is able to complement the *Arabidopsis npr1-1* mutant (Yuan et al., 2007). Like AtNPR1, OsNPR1 is also constitutively expressed and localizes to the cytoplasm. Treatment of rice cells with a reducing agent resulted in the movement of OsNPR1 into the nucleus. Similar to *Arabidopsis* NPR1, mutation of the corresponding cysteines (C82 and C216) in OsNPR1 also resulted in constitutive nuclear localization (Yuan et al., 2007). Thus, it appears that the mechanisms of SA-dependent, NPR1-mediated defense response likely evolved very early in the emergence of the plant kingdom.

Theobroma cacao L, (cacao) is a small tropical tree species endemic to the Amazon rainforest of South America. Cacao seeds are harvested and processed into cocoa beans and chocolate, providing an income for millions of small-holder farmers in West Africa, Central and South America, the Caribbean, Malaysia, Indonesia and other tropical areas.

Pathogens are a major problem for cacao production, causing annual crop losses estimated at 30-40% (Hebbar, 2007). In its center of diversity, the Amazon basin, cacao is susceptible to several potentially devastating pathogens, such as *Moniliophthora perniciosa*, the causal agent of witches' broom disease, *Moniliophthora roreri*, the causal agent of frosty pod rot (Wood and Lass, 1985; Evans and Priori, 1987; Purdy and Schmidt, 1996; Hebbar, 2007; Phillips-Mora and Wilkinson, 2007) and several *Phytophthora* spp., the causal agent of black pod disease (Bailey et al., 2005; Bailey et al., 2005). Outside this region, cacao is susceptible to a number of opportunistic pathogens (Buddenhagen, 1977; Schnell et al., 2007; Guiltinan et al., 2008) .

Several defense-related genes in *Theobroma cacao* have been identified through gene expression analyses after hormone treatments (Bailey et al., 2005; Bailey et al., 2005; Bae et al., 2008). An endo-1,4-β-glucanase is induced by the application of ethylene, and a type III peroxidase and a class VII chitinase are induced by methyl jasmonate treatment in mature cacao leaves. Those genes are responsible for induced resistance to pests in cacao, though the responses to hormone induction are different depending on developmental stages. In addition, transgenic overexpression of a class I chitinase gene in cacao enhances foliar resistance against the fungal pathogen, *Colletotrichum gloeosporioides* (Maximova et al., 2006). Moreover, ESTs sharing sequence homology to known *PR* genes have been isolated from cacao (Jones et al., 2002; Verica et al., 2004; Argout et al., 2008). Several of these genes have been shown to be up-regulated by treatment of plants with benzothiadiazole (BTH), the salicylic acid analog (Verica et al., 2004). All together, recent evidence suggests that cacao may utilize SAR pathway during the defense response; however, the extent of conservation of the pathway in cacao is presently unknown.

In this paper, we report the isolation and characterization of an *NPR1* homologue from the tropical tree, *Theobroma cacao*. We show that *Theobroma cacao NPR1* (*TcNPR1*) shares similar functions as *Arabidopsis NPR1*. It is able to partially complement the *Arabidopsis npr1-2* mutation in transgenic *Arabidopsis* plants in a leaf infection assay and translocate into nucleus upon SA induction in the same manner as the endogenous *Arabidopsis* NPR1 protein.

3.3 Material and Methods

Full-length cDNA Cloning by Degenerate PCR

NPR1 cDNA sequences from *Arabidopsis* (U76707), *Brassica napus* (AF527176), and *Carica papaya* (AY550242) were aligned using the ClustalW program v1.8 (Thompson et al., 1994). Degenerate primers (TcNPR1dg-5', TATTGTCAARTCTRATGTAGAT; TcNPR1dg-3', GAARAAYCGTTTCCCKAGTTCCAC) were designed to regions highly conserved among all three sequences.

Total RNA was isolated from cacao leaves from variety Scavina6 as previously described (Verica et al., 2004). Cacao leaf cDNA was synthesized using the SMART RACE cDNA Amplification Kit (Clontech Laboratories Inc., Mountain view, CA http://www.clontech.com/) according to the manufacturer's instructions.

PCR reactions were performed using cacao leaf 2.5 µl cDNA from first strand synthesis from SMART RACE cDNA Amplification Kit, 10 µl Redi-prime PCR mix (GeneChoice, Inc., Frederick, MD) and 5µM of the above degenerate primers. Following denaturation (94º for 5 min.), PCR was performed for 32 cycles using the following condition (94º for 30 sec., 45 º for 30 sec, 72º for 1 min.), followed by a 5min. extension at 72º. PCR products were resolved on 1% agarose gels, purified with the GENECLEAN II Kit (Q-Biogene Inc., Solon OH) and cloned into the pGEM-T-Easy vector (Promega Corporation, Madison WI) according to the manufacturer's instructions. DNA sequencing was performed at the Penn State Genomics Core Facility using an ABI Hitachi 3730XL DNA Analyzer. The resulting clone was designated as pGEM-TcNPR1.

Genomic DNA cloning by BAC library screening

Theobroma cacao BAC filter arrays constructed using genomic DNA from genotype LCT-EEN 37 were purchased from the Clemson University Genomic Institute (http://www.genome.clemson.edu/). Filter arrays were blocked for 4 hours at 60 ºC in a

solution containing 1% BSA, 1 mM EDTA, 7% SDS, and 0.25M sodium phosphate. PCR generated *TcNPR1* cDNA fragment labeled with ^{32}P dCTP using the MegaPrimer DNA Labeling System (GE Healthcare, Buckinghamshire, UK) according to the manufacturer's instructions was added and hybridized overnight at 60 ºC. The next day, the filter arrays were washed twice in 2X sodium chloride/sodium citrate (SSC), 0.5% sodium dodecyl sulfate (SDS) for 20 minutes at 60 ºC. Radiographic imaging was performed via storage phosphor imaging (Molecular Dynamics, http://www.mdyn.com/). After filter alignment and clone number identification, a BAC clone (2K13) containing a putative *TcNPR1* fragment was obtained from a frozen stock. The sequence of *TcNPR1* genomic fragment was acquired by series sequencing of the BAC clone from ATG start codon. Sequencing primer was designed based on the *TcNPR1* cDNA at the first round and following series primers were designed based on the known sequence resulting from previous sequencing. Introns were identified by aligning the genomic sequence and full length cDNA using SPIDEY software (http://www.ncbi.nlm.nih.gov/spidey/). The same strategy was applied to clone the 1.1kb promoter region upstream of the ATG. Forward and reverse sequencing was also performed to validate the sequence.

For sequence verification the *Arabidopsis* NPR1 protein sequence (At1g64280) and putative cacao NPR1 protein sequences (genbank accession HM117159) were aligned using the ClustalW program v1.83 (Thompson et al., 1994). The TcNPR1 protein sequence was analyzed for potential functional sites by querying the Simple Modular Architecture Research Tool (SMART) database (http://smart.embl-heidelberg.de/).

Semi-quantitative RT-PCR analysis of *TcNPR1* expression in cacao tissues

Total RNA was isolated from Scavina6 leaves stages A, C and E (corresponding to stages YR, IG, MG respectively, as described in (Melnick et al., 2008)), open flowers, unopened flowers, roots, exocarps and seeds as previously described (Verica et al., 2004). For each tissue, three biological replicates were collected and analyzed. Cacao cDNA was synthesized in a final volume of 25µl from 2µg of total cacao RNA using M-MLV reverse transcriptase (New England Biolabs, Inc., Ipswich, MA). RNA and 0.5µg oligo(dT) were added to sterile water to

final volume of 18µl. The mixture was then incubated at 70°C for 5min, chilled on ice, which was followed by adding 10x reverse transcription buffer (New England Biolabs, Inc., Ipswich, MA), 0.1M fresh made DTT and 10mM dNTP. The mixture was further incubated at 42°C for 2min, followed by incubation at 42°C for 1hr with 10 units of reverse transcriptase MMLV (New England Biolabs, Inc., Ipswich, MA). The reaction was terminated at 70°C for 15min.

Semi-quantitative RT-PCR was performed using intron-spanning primers for *TcNPR1* (TcNPR1RT-5': ATGGATTCCCGTCTGGAACTTGGT; TcNPR1RT-3': TCTGGAGTGTCATTTCCTCCGCAT) and *TcActin* (CL33contig2 in Esttik Database http://esttik.cirad.fr/) used as an internal normalization and cDNA loading control (TcActinRT-5': AGCTGAGAGATTCCGTTGTCCAGA and TcActinRT-3': CCCACATCAACCAGACTTTGAGTTC). RT-PCR reactions were set up using 1 µl of ½ diluted cDNA and 5µM of the *TcNPR1* or *TcActin* primers. Titration of cycles was carried out and it was determined that the PCR amplification of *TcNPR1* was within its linear range at 27 cycles using the following condition: 94ºC for 30 sec., 56 ºC for 30 sec, 72ºC for 1 min. Similarly, PCR of *TcActin* was performed under non-saturation conditions within the linear range (22 cycles at 94ºC for 30 sec., 55 ºC for 30 sec, 72ºC for 1 min). *TcActin* served as a cDNA loading control.

SA treatment of cacao seedlings

The leaves of two to three-month old cacao plants generated by rooted cuttings from two different genotypes (ICS1 and Scavina6) were sprayed with salicylic acid (SA) dissolved in water at three different concentrations, 1mM, 2mM and 4mM. Control plants were treated with water. Plants were grown in a greenhouse under conditions previously described (Verica et al., 2004) and leaf tissue from fully expanded young leaves (developmental stage C, corresponding to stage IG in (Bailey et al., 2005)) was harvested at 24hrs after treatment and frozen in liquid nitrogen. Total RNA was isolated and cDNA was synthesized as described above. For each genotype and each treatment, three biological replicates were collected. Semi-quantitative RT-PCR and expression analysis were performed to assay the

levels of *TcNPR1* expression as described above. The PCR products were analyzed on 1% agarose gel, stained with ethidium bromide. The expression values of *TcNPR1* and *TcActin* were quantified using ImageQuant software (Molecular Dynamics, Amersham Bioscience) as described in (Brechenmacher et al., 2004) and relative expression of *TcNPR1* was calculated by comparing with the expression of *TcActin*.

Transgenic *Arabidopsis* mutant complementation assay

All binary plant transformation vectors were constructed by incorporating the genes of interest into pCAMBIA-1300 binary transformation vector containing plant selectable marker for hygromycin resistance (Hajdukiewicz et al., 1994).

Binary Vector p35S:TcNPR1 **-** The *TcNPR1* coding sequence fragment was generated by PCR using pGEM-TcNPR1 as described above and included *Xma*I and *Not*I restriction sites at the 5'- and 3'-ends respectively (TcNPR1-5'-*Xma*I, CCCGGGATGGATAACAGAAATGGCTT; TcNPR1-3'-*Not*I, GCGGCCGCTTGCATTAGGCCTATGGTCTA). The fragment was cloned into pGEM T-Easy (Promega Corporation, Madison WI) according to the manufacturer's instructions and sequenced for integrity. The *TcNPR1* coding sequence was then cloned into the *Xma*I and *Not*I sites of an intermediate cloning vector (pE2113) between E12-Ω promoter (Mitsuhara et al., 1996) and 35SCaMV terminator. A 3 kb restriction fragment containing *TcNPR1* gene cassette was excised from pE2113 using *Pvu*II and ligated into the *Sma*I site of pCAMBIA-1300.

Ligations were performed overnight at 16º with 3Units of T4 DNA ligase (Promega Corporation, Madison WI).

Binary vector p35S:AtNPR1 - The *AtNPR1* coding sequence fragment was generated by PCR using the *AtNPR1* cDNA clone U13446 from Arabidopsis Biological Resource Center (http://www.biosci.ohio-state.edu/~plantbio/Facilities/abrc/abrchome.htm) and

included *Nco*I sites at the 5'- and 3'-ends (AtNPR1-5'-*Nco*I, CCATGGACACCACCATTGATGGATTC; AtNPR1-3'-*Nco*I, CCATGGTCCGACGACGATGAGAGAGTTTACG) . The PCR fragment was cloned into pGEM T-Easy and sequenced. The resulting intermediate plasmid was designated pGEM-AtNPR1. The *AtNPR1* coding sequence was then excised by *Nco*I from pGEM-AtNPR1, and blunt-end cloned into pE2113 between E12-Ω promoter (Mitsuhara et al., 1996) and 35SCaMV terminator as *Xma*I and *Not*I fragment. Contently 3.1 kb fragment containing the *AtNPR1* gene cassette was obtained by digestion with *Pvu*II, and blunt-end ligated into the *Sma*I site of pCAMBIA-1300.

Binary Vector p35S:TcNPR1:EGFP – The cassette of E12-Ω promoter and EGFP on the intermediate cloning vector pE2113 was cloned into *EcoR*I and *Hind*III sites of pCambia1300. The resulting binary vector was designated pXCGH. PCR generated *TcNPR1*fragment, including *Sma*I and *KpnI* sites at the 5'- and 3'-ends (TcNPR1-5'-*Sma*I, CCCGGGATGGATAACAGAAATGGCTT; TcNPR1_3'-*Kpn*I, GGTACCGACCGCCCCTACCACTACCAGTTAG) was first cloned into pGEM T-Easy (pGEM-TcNPR1-EGFP). The sequence was verified, the DNA fragment was excised with *SmaI* and *KpnI* and blunt ends ligated into the blunt-ended *NcoI* site of pXCGH positioned between the E12-Ω promoter and at the 5'end of the EGFP coding sequence to generate the binary vector p35S:TcNPR1:EGFP.

Binary vector p35S:AtNPR1:EGFP - The pGEM-AtNPR1 containing *AtNPR1* coding sequence was digested with *Nco*I and the fragment was ligated into the *Nco*I site of pE2113 as described above. The 3.6 kb fragment containing the *AtNPR1*- EGFP fusion gene cassette was digested with *Sal*I and *Eco*RI and cloned into the *Sal*I and *Eco*RI sites of pCAMBIA-1300.

Arabidopsis Transformation

The binary vectors described above were introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation, as previously described in (Lin, 1995). *Arabidopsis* Col-0 plants were grown in a Conviron growth chamber (Model No. MTPS144) maintained at 22ºC, under a 16:8::L:D cycle. Light intensity was maintained at 200 $\mu M/m^2 \cdot s$ with Octron 4100K Ecologic bulbs (Sylvania, Danvers MA). To increase the number of inflorescences, plants were cut back after bolting, and allowed to re-grow. The floral dip method was used to transform *Arabidopsis* as described previously (Clough and Bent, 1998). Briefly, Agrobacterium cultures were grown at 25º on a platform shaker (200 rpm) to an OD_{600}=1.2. Cells were centrifuged at 1,500 X *g* for 6 minutes and re-suspended in 300mls of a solution containing 2.15 g L^{-1} MS salts, 5% sucrose, 0.02% Silwet-77. The flowers were dipped in the solution for three seconds, domed to remain humidity and covered with black cloth. The cloth was removed the next day and plants were regularly watered until seed maturation.

Following seed set, seeds were collected from nine plants for each independent transgenic event. Seeds from 5 individual lines were soaked in 0.1% Tween-20 for 2 minutes and sterilized with 50% bleach for 10 minutes at room temperature. Seeds were then washed five times with 1 ml of sterile water. To select for transformants, seeds were planted on ½ MS media, agar plates (pH 5.7) supplemented with 25 µg ml^{-1} hygromycin B. Plates were place in a Conviron growth chamber under the same light and temperature conditions as above. After 10 days, germinated seedlings were examined for leaf development and root elongation. Those seedlings that showed root elongation were transferred to soil and allowed to grow. Transformations were performed with the following vectors: p35S:TcNPR1, p35S:AtNPR1, p35S:TcNPR1:EGFP, and p35S:AtNPR1:EGFP constructed as described above, and control vectors p35S: EGFP (pGH00.0126, (Maximova et al., 2003))and pCambia 1300.

Salicylic acid (SA) *Arabidopsis* induction assay

Four week-old wild type *Arabidopsis* Col-0 and *npr1-2* mutants and five independent transgenic lines growing in soil were sprayed with 1mM SA, along with water-treated control plants. Three biological replicates, each containing leaves from 5 individual plants were collected 24hrs after treatment. Total RNA was isolated from treated and control samples using RNeasy plant mini kit (QIAGEN, Valencia CA). cDNA was generated as described above. Semi-quantitative RT-PCR was performed as described above to measure the expression of *TcNPR1* and *AtPR1*. *Arabidopsis Ubiquitin* served as cDNA loading and normalization control. The following primer sets and conditions were employed:

TcNPR1-5': ATGGATTCCCGTCTGGAACTTGGT; TcNPR1-3': TCTGGAGTGTCATTTCCTCCGCAT (27cycles of 94ºC for 30 sec., 56 ºC for 30 sec., 72ºC for 1 min). AtPR1-5': CTCGAAAGCTCAAGATAGCCCACA; AtPR1-3': CTTCTCGTTCACATAATTCCCACG (25 cycles of 94ºC for 30 sec., 54 ºC for 30 sec., 72ºC for 1 min). Ubiquitin-5': ACCGGCAAGACCATCACTCT; Ubiquitin-3': AGGCCTCAACTGGTTGCTGT (22 cycles of 94ºC for 30 sec., 54 ºC for 30 sec., 72ºC for 1 min). The conditions of PCR were determined by cycle titration to avoid saturating conditions. The relative expression levels were determined as described above.

Pseudomonas syringae **infection assay of** ***Arabidopsis*** **transgenic plants overexpressing** ***NPR1*** **genes**

Pseudomonas syringae pv. tomato DC3000 (*P.s.t.*) was grown on Difco Pseudomonas agar (PA) (Becton, Dickinson and Company, http://www.bdbioscience.com/) supplemented with rifampicin (100 µg ml-1) and kanamycin (25 µg ml-1) at 25oC for 48 hrs. Cells were scraped from plates using a bacterial inoculating loop and re-suspended in water. Plant infection assays and bacterial growth assays were carried out as described previously in (Katagiri et al., 2002). Five individual transgenic *npr1-2* mutant overexpressing *TcNPR1* coding sequence were infected with *P.s.t.* at OD_{600}=0.002. Briefly, three days after inoculation leaf disks from treated leaves of 2 independent replicate plants were pooled for a single sample. Data represents means ±SE (cfu/mg FW) of three biological replicates per treatment and statistical differences were determined by Single factor ANOVA analysis.

Nuclear translocation of TcNPR1 in transgenic *Arabidopsis* plants

For observations of green leaves, four week-old soil-grown transgenic plants containing one of transgenes 35S:AtNPR1:EGFP, 35S:TcNPR1:EGFP, 35S:EGFP and plants transformed with empty binary vector pCambia 1300 were sprayed with either a 1mM solution of SA in water or water. For root observations, control and transgenic seed were germinated on MS agar or MS agar supplemented with 0.5 mM SA (Kinkema et al., 2000) and seedlings were grown for 10 days . Leaves and roots were placed in a drop of water on a standard microscope glass slide and overlaid with a cover slip. The samples were imaged with an inverted Olympus FV1000 Laser Scanning Confocal Microscope (Olympus America Inc., Melville, NY). For imaging EGFP, tissues were excited with a blue argon laser (488nm) and emission wavelengths of 500-600 nm were detected through a variable bandpass filter positioned in front of the photomultiplier tube. Tissues were observed using 40x and 10x objectives for leaf cells and root cells, respectively, each with numerical apertures and 1.15. FV10-ASW version 1.6 software (OLYMPUS, Pittsburgh, PA) was used to collect images, select slices, and create intensity projections over the Z axis.

SA and JA combination treatment of *Arabidopsis* transgenic plants overexpressing *TcNPR*1

Four weeks old soil-grown wild type *Arabidopsi*s Col-0, *npr1-2* mutants and five independent transgenic lines containing p35S:TcNPR1 were sprayed with a combination of 1mM SA and 0.1mM MeJA in 0.015% Silwet L-77. Plants treated with 1mM SA alone in water, 0.1mM MeJA alone in 0.015% Silwet L-77 and water with 0.015% Silwet L-77 served as control treatment. Three biological replicates each consisting of leaves from 5 individual plants were collected at 48hrs after treatment, total RNA was isolated, cDNA was synthesized and semi-quantitative RT-PCR was performed as described above to determine the transcripts level of *TcNPR1* and *AtPR1*. For expression analysis of *VSP2* and *PDF1.2*,

following primer sets and conditions were used to maintain the reaction in its linear amplification range. *VSP2* Forward: TACGGTCTCGGCATCCGTTC; *VSP2* Reverse: CCTCAAGTTCGAACCATTAGGCT (21cycles of 94ºC for 30 sec., 58 ºC for 30 sec., 72ºC for 1 min). *PDF1.2* Forward: TCATCATGGCTAAGTTTGCTTCCATC; *PDF1.2* Reverse: TGTCATAAAGTTACTCATAGAGTGAC (27cycles of 94ºC for 30 sec., 60 ºC for 30 sec., 72ºC for 1 min). The PCR products were analyzed on 1% agarose gel, stained with ethidium bromide. The expression values of *AtVSP2* and *AtPDF1.2* were quantified using ImageQuant software (Molecular Dynamics, Amersham Bioscience) as described in (Brechenmacher et al., 2004) and relative expression of two genes was calculated by comparing with the expression of *AtUbiquitin*.

3.4 Results

Isolation of a putative *TcNPR1* gene

Degenerate PCR was utilized to clone the full length cDNA of *Theobroma cacao NPR1* (*TcNPR1*). The degenerate primers were designed based on the alignment of NPR1 homologs from *Arabidopsis*, *Brassica* and *Carica papaya* and cDNA from cacao genotype Scavina6 (SCA6) leaf was used as template. A fragment of 1776 bp was isolated, cloned into pGEM sequencing vector and sequenced to reveal an intact coding sequence of the expected length and with high homology to the *Arabidopsis NPR1* gene.

A genomic fragment containing a putative *TcNPR1* gene was obtained by screening Clemson University Genomics Institute (CUGI) cacao BAC library using the putative cacao *TcNPR1* cDNA clone as probe. Two BAC clones were found to contain the *TcNPR1* gene: 2K13 and 11K17. The genomic sequence of *TcNPR1* was isolated by primer walking sequencing from known sequence using clone 2K13. A similar strategy was performed to sequence a region of 1.1kb containing the promoter sequence upstream of ATG start codon. The full sequence consisted of a 4.5kb genomic region of *TcNPR1* containing 1.1kb promoter, four exons and three introns (depicted in Figure 1A), which is similar to the genomic structure of *AtNPR1*.

Arabidopsis and cacao NPR1 protein sequences are highly similar

Conceptual translation of the cacao NPR1 protein revealed that it consists of 591 amino acid residues, only two amino acids shorter than AtNPR1. Alignment of the AtNPR1 and TcNPR1 protein sequences revealed that they are highly similar to each other (55% identity, 74% similarity). Both the *Arabidopsis* and cacao *NPR1* genes encode predicted proteins that share a number of structural features (Figure 1B). Each has a BTB/POZ domain near its N-terminal end (dashed line box) which shares 65% identity. Similarly, an ankyrin repeat region (solid line box) is present in both proteins which shares about 72% identity. In other ankyrin containing proteins, these domains have been shown to play roles in protein-protein interactions (Albagli et al., 1995; Stogios et al., 2005; Li et al., 2006; Rochon et al., 2006). In the AtNPR1 protein, the BTB/POZ domain has been shown to function in homo-dimerization of NPR1, and the ankyrin repeat region mediates interactions with TGA transcription factors (Zhang et al., 1999). In addition, two cysteine residues (C82 and C216 in AtNPR1), which have been shown to play a role in the redox regulated activation and nuclear localization (Mou et al., 2003), are also conserved in TcNPR1 (Figure 1B. grey triangles). In fact, the AtNPR1 and TcNPR1 proteins share eleven conserved cysteine residues, suggesting that they share a similar structural conformation. The C-terminal region of AtNPR1 has been shown to contain a nuclear localization signal (NLS) that directs NPR1 monomers into the nucleus upon induction (Kinkema et al., 2000). Five basic amino acids in this region function directly in this role (Figure 1B, black arrows). Four out of five of these basic amino acids are identical in TcNPR1, suggesting that TcNPR1 may also contain functional nuclear localization sequences. These similarities in protein structure suggest that *TcNPR1* gene may also share the same function as *AtNPR1* during plant defense response.

Cacao NPR1 gene promoter contains putative SA regulatory elements

We analyzed the 1.1kb promoter region of the TcNPR1 gene (Figure 1C) using plant cis-acting regulatory elements databases PLACE (http://www.dna.affrc.go.jp/PLACE/)(Higo et al., 1999) and PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)

(Rombauts et al., 1999; Lescot et al., 2002). Although a potential CAAT box was found 290bp and 140bp upstream of the ATG start codon, we did not observe an element resembling a TATA box. This is not surprising, as recent studies of core promoter regions in both plants and animals suggest that only 24%-29% of genes contain TATA-like elements (Molina and Grotewold, 2005; Yang et al., 2007). A variety of other regulatory elements were also found. Several elements known to regulate inducibility by salicylic acid were found, such as the AS-1 element (TGACG). TGACG motifs were found involved in transcription activation by SA and this element was previously shown to be required for the SA-induced expression of *PR1* (Lebel et al., 1998). In addition, there were multiple copies of the W-box (TTGAC), an element similar to the AS-1 element, which was also found in promoter of *AtNPR1*. W-box was shown to be the binding site for SA-induced WRKY DNA binding proteins (Yu et al., 2001), and was required for the SA induction of the tobacco (*Nicotiana tabacum*) class I chitinase gene (Yang et al., 1999). All of the information suggests that the *TcNPR1* gene might be regulated by SA in a manner similar to *AtNPR1*. Interestingly, several cis-elements involved in light responsiveness and circadian control are also presented in the *TcNPR1* promoter, suggesting that *TcNPR1* might be also regulated by light.

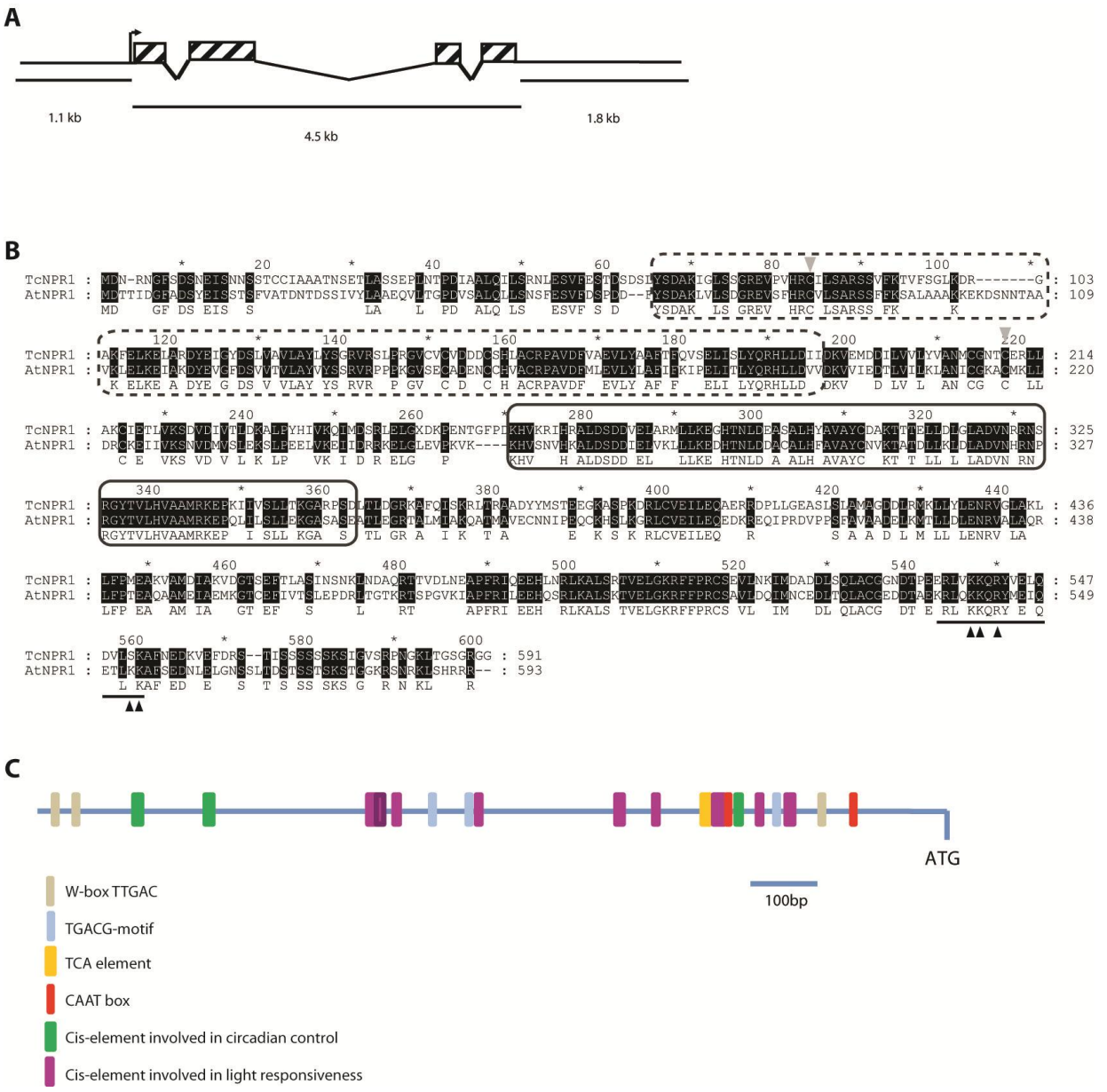


Figure 3-1 Gene and protein structures of *Theobroma cacao NPR1.* **A.** Diagram of *TcNPR1* gene structure. Boxes with diagonal stripes represent exons. The slanted lines represent introns. The arrow represents the start site of transcription. The sizes of promoter region sequenced, the coding region and the 3'-untranslated region (UTR) of *TcNPR1* are indicated. **B.** Alignment of AtNPR1 and TcNPR1 proteins. Protein alignment was carried out by ClustalW. Residues blocked in black are identical in both sequences. Numbers refer to the amino acid position in AtNPR1 protein. BTB/POZ and ankyrin repeats domains are highlighted by dashed line box and solid line box, respectively. Two of the conserved

cysteines (C82 and C216 in AtNPR1) are shown with grey triangles. The potential nuclear localization signal identified in *Arabidopsis* is underlined. Amino acids demonstrated to be critical for AtNPR1 nuclear translocation are indicated with black triangles. **C.** Schematic representation of cis-acting regulatory DNA element in cacao *TcNPR1* promoter region. The 1.1kb DNA fragment upstream of start codon was analyzed. The cis-acting regulatory elements were identified by querying the PLACE and PlantCare databases. The colored blocks represent different cis-elements as indicated.

Basal and induced expression of *TcNPR1* in cacao tissues

Semi-quantitative RT-PCR was performed to illustrate the basal expression level of *TcNPR1* in various cacao tissues of Scavina6, including leaves from stage A (young/expanding), C (expanded/soft), E (mature/hardened), open flowers, unopened flowers, roots, seeds and fruit exocarps. *TcNPR1* transcript was detected in all tissues tested (Figure 2A), an expression pattern similar to the *Arabidopsis* gene, however, the basal level of expression varies among different tissues. The expression of *TcNPR1* was relatively high in the younger leaves (stage A and C) and lower in the later stages of development (stage E). The lowest expression of *TcNPR1* in all tested tissue was observed in seeds whereas the expression was relatively high in fruit exocarps. In flowers, expression of *TcNPR1* was higher in open flowers than in unopened ones. The expression of *TcNPR1* in roots was at a moderate level, comparable to that in flowers and younger leaves.

Induction by SA

Since it is well-characterized that *NPR1* transcript accumulation can be increased by SA treatment of *Arabidopsis* leaves, we tested if *TcNPR1* can respond to exogenous SA in the same manner. We applied various concentrations of SA to stage C leaves of two genotypes, Scavina6 and ICS1, which differ in their resistance to witches' broom disease (Scavina6 is more resistant) (Faleiro et al., 2006). In *Arabidopsis*, the *NPR1* gene is induced approximately 2-3 fold 24hrs after treatment of leaves with 1mM SA (Cao et al., 1997; Durrant and Dong, 2004). Semi-quantitative RT-PCR was employed to demonstrate the

induced level of *TcNPR1* 24 hours after SA application (Figure 2B). To quantify the expression of *TcNPR1* after SA treatment, we measured the fluorescence intensity of ethidium bromide stained DNA fragments irradiated with UV light using a high-sensitivity camera and ImageQuant software. Data were normalized to the expression level of an *actin* control. The results presented in Figure 2B showed that there was no significant change of *TcNPR1* expression upon 1mM, 2mM and 4mM SA treatment in ICS1 compared to water control. However, in the Scavina6, there was a statistically-significant 2-fold increase of *TcNPR1* at 4mM SA induction, though there was no change upon 1mM and 2mM SA treatment.

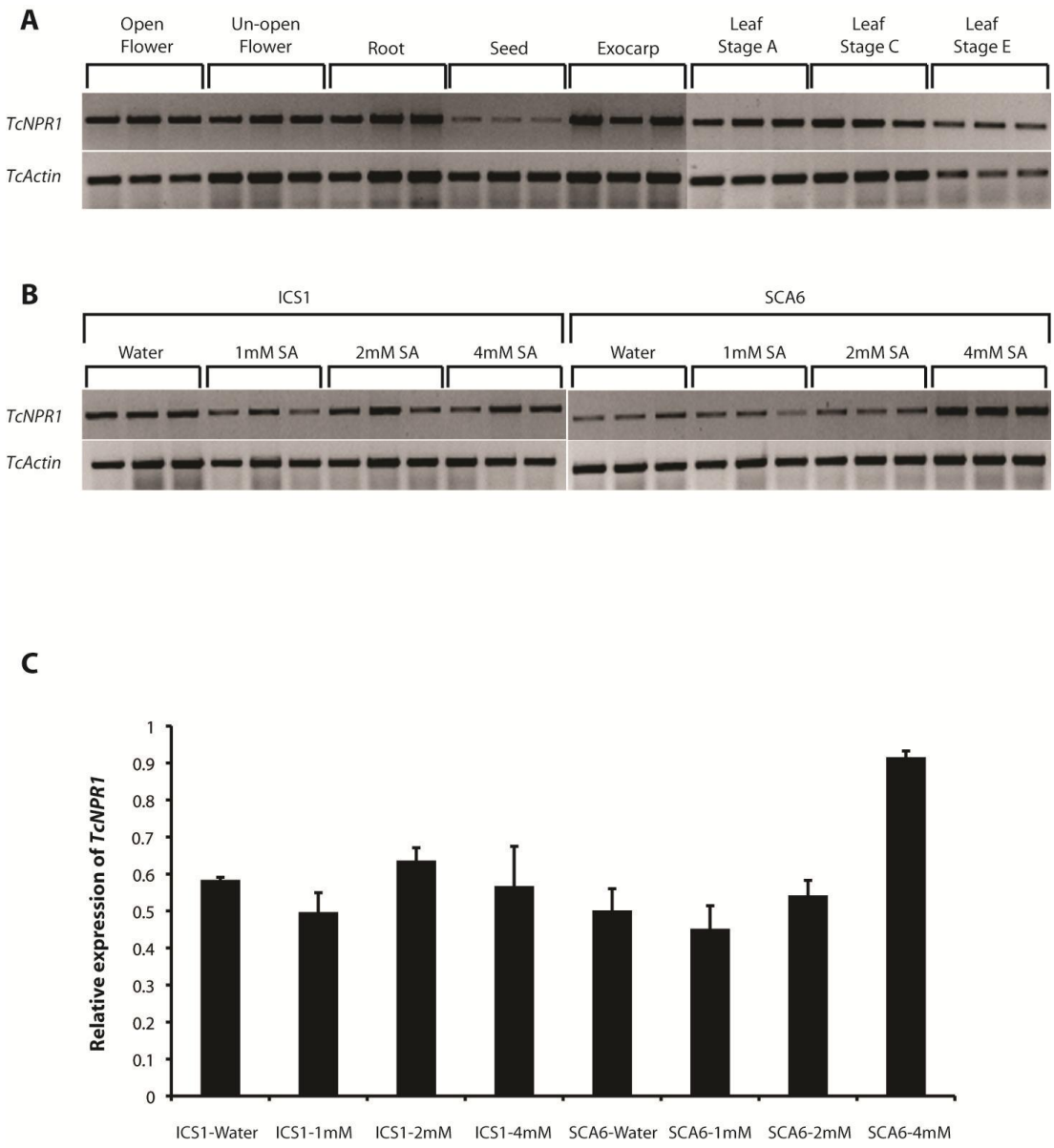


Figure 3-2 Gene expression analysis of *TcNPR1* in cacao. **A.** Expression of *TcNPR1* in various cacao tissues. Total RNA samples were collected from open flower, unopened flower, root, seed, exocarp and three different leaf developmental stages from youngest to oldest (A, C and E) from cacao genotype Scavina6 (SCA6). Semi- quantitative RT-PCR was performed and cacao *actin* (*TcActin*) was used as cDNA loading control. **B.** Expression of *TcNPR1* in cacao leaf tissue after salicylic acid (SA) treatment. Semi-quantitative RT-PCR was performed with cDNA from stage C leaves of two different cacao genotypes ICS1 (left panel) and Scavina6 (right panel), sampled 24hrs after SA treatment in three different

concentrations (1mM, 2mM and 4mM). Water-treated samples served as a control and *TcActin* was used as cDNA normalization control. **C.** Calculated average relative gene expression levels from B. Gel images were quantified by ImageQuant and expression of *TcNPR1* was normalized to *TcActin*. Expression levels are presented as the means ± standard errors of three biological replicates.

Complementation of *Arabidopsis npr1-2* mutant

To assess the function of TcNPR1, we placed the cacao *TcNPR1* gene under the control of the E-12 omega promoter and introduced it into the *Arabidopsis npr1-2* mutant to test if it can restore the mutant phenotype. One of the well characterized phenotypes of this mutant is the lack of SA-dependent activation of the *PR1* gene (Ryals et al., 1997; Shah et al., 1997). The *PR1* gene is thought to encode a protein active in defense response and has been used as a marker of SA pathway activation in many studies and in different plant species.

Five independent *TcNPR1* transgenic lines, wild type *Arabidopsis* Col-0 along with the *npr1-2* mutant were sprayed with 1mM SA, and the expression of *TcNPR1* and *AtPR1* was determined by semi-quantitative RT-PCR 24hr after induction. Five transgenic lines all showed heterologous *TcNPR1* expression with varied expression levels (Figure 3). As expected, there was no significant up-regulation of the transgene after SA treatment because *TcNPR1* was expressed constitutively from the E12-Ω promoter. The *Arabidopsis PR1* gene showed a very large induction after SA treatment in wild type *Arabidopsis* Col-0 (Figure 3), but there was no up regulation in the *npr1-2* mutant, which is consistent with previous report (Cao et al., 1997). There was a small increase in *PR1* expression in the mutant treated with water, which could be expected from plant to plant biological variation. However, the *PR1* gene expression level did not change after SA treatment, as expected for the *npr1-2* mutant. We observed a moderate induction of the *PR1* gene in 3 out of 5 transgenic lines (Line 2, 3 and 4), though the level of induction was not as high as in wild type Col-0. No *PR1* gene induction was observed for transgenic lines 1 and 5. These results suggest that the *TcNPR1* gene can at least partially complement the *Arabidopsis*

*npr1*mutant and act to mediate SA dependent *PR1* gene expression in *Arabidopsis* leaves but it may not act as efficiently as the endogenous NPR1 itself.

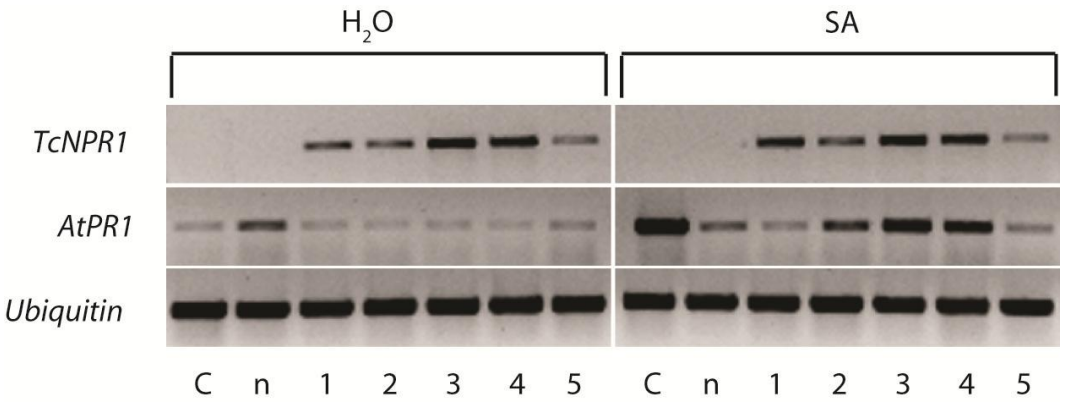


Figure 3-3 Gene expression of *TcNPR1* and *AtPR1* in transgenic *Arabidopsis npr1-2* mutant lines. Semi-quantitative RT-PCR was performed with cDNA prepared from the leaves of 4-week-old plants of wild type (C), *npr1-2* (n) and 5 independent transgenic *npr1-2* mutant expressing *TcNPR1* (1-5). *TcNPR1* and *AtPR1* expression were evaluated 24hrs after 1mM SA treatment. Water-treated control leaves (left panel) from each genotype were also analyzed. *Arabidopsis Ubiquitin* (*AtUbiquitin*) expression was assayed as a non SA-induced, RNA loading control.

Another phenotype of the *Arabidopsis npr1* mutation is increased pathogen growth after bacterial infection of leaves (Cao et al., 1997; Ryals et al., 1997; Shah et al., 1997). To test if *TcNPR1* overexpression in *npr1-2* mutant can complement the mutant disease susceptible phenotype, we infected leaves from 5 transgenic lines with *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*) by syringe infiltration. The results indicated that the *npr1-2* mutant was more susceptible than Col-0 (Figure 4A) three days after inoculation, exhibiting yellow necrosis similar to previous results (Cao et al., 1997). Three transgenic lines overexpressing the *TcNPR1* gene and exhibiting SA dependent *PR1* up-regulation partially restored induced resistance compared to the control *npr1-2* mutant (Figure 4A). Although several yellow necrotic spots were displayed on leaves of the transgenic plants, they did not exhibit severe necrosis or senescence. However, the other two transgenic lines,

line 1 and 5, showed necrosis all over the leaves and the tissues were wilted. Water infiltration served as a control to demonstrate that the injection of water alone did not damage the tissues.

To quantify the disease symptom, bacterial assays were carried out to measure the titer of bacterial on infected leaves. The levels of bacterial in infected *npr1-2* mutant leaf disks increased more than 250 fold as compared to Col-0 controls (Figure 4B). The three transgenic lines overexpressing the *TcNPR1* gene (Line 2, 3 and 4), which exhibited significant up-regulation of the *PR1* after SA treatment, showed a 30 to 100 fold reduction of bacterial growth compared to the *npr1-2* mutant. There was no significant change in bacterial growth rates in leaf disks of the other two transgenic lines tested (Line 1 and 5). To assess the relationship between the level of SA-dependent induction of *PR1* and the degree of bacterial growth in the transgenic lines, we plotted the values as depicted in Figure 4C. A significant negative correlation between SA dependent gene induction and bacterial growth was observed (R^2=0.88), suggesting that the resistance conferred by *TcNPR1* is via the SA dependent resistance pathway and further supports our hypothesis that TcNPR1 plays a similar function to *Arabidopsis* NPR1 in plant defense response.

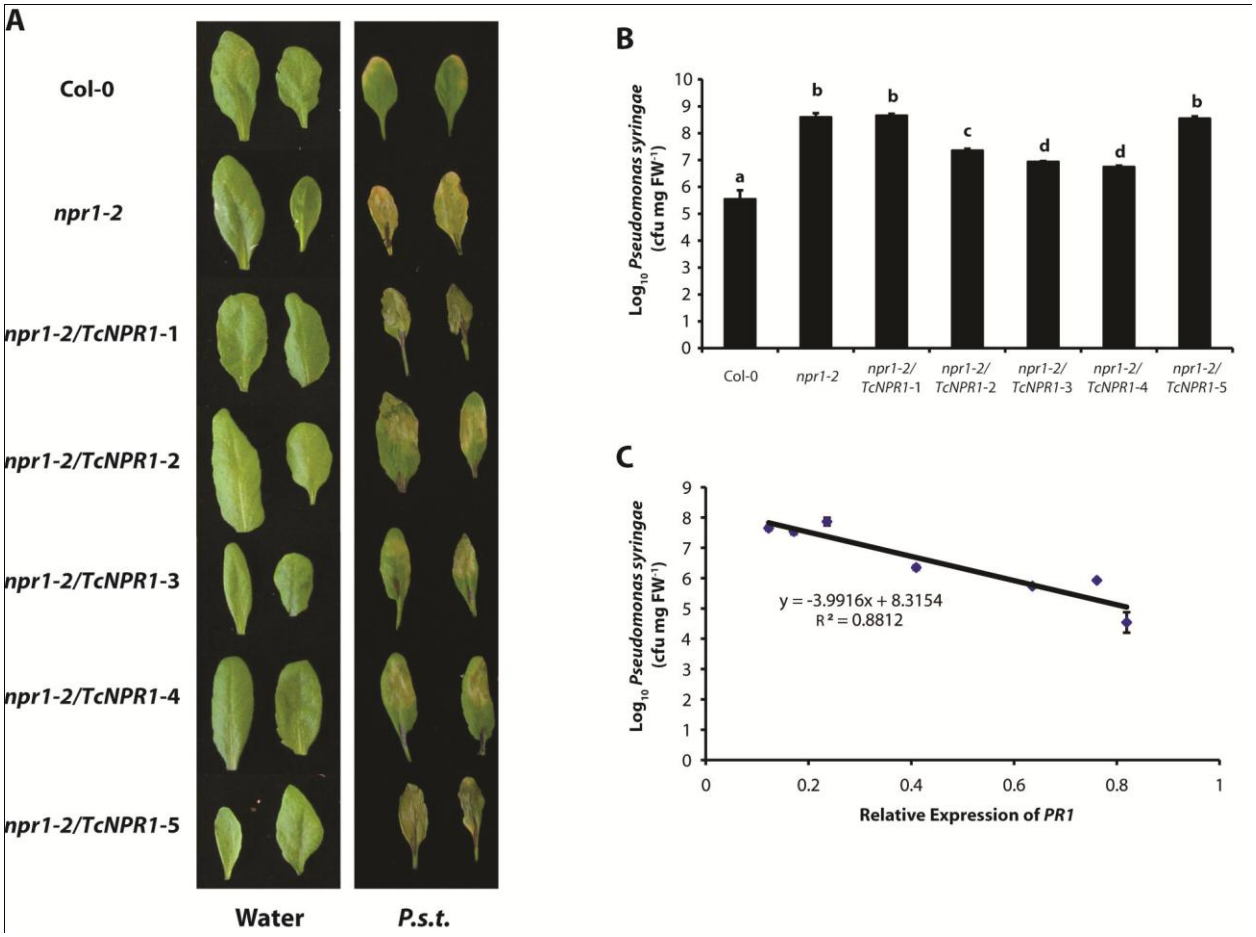


Figure 3-4 *Pseudomonas syringae* infection assay of transgenic *Arabidopsis npr1-2* mutant overexpressing *TcNPR1*. **A.** Disease symptoms on leaves of Col-0, *npr1-2* and five independent lines of *npr1-2* plants transformed with *TcNPR1* (*npr1-2*/*TcNPR1*) inoculated with *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*) (OD_{600}=0.002) at three days post inoculation. Additionally, leaves from those seven genotypes were infiltrated with water as a control treatment. **B.** Growth of *P.s.t.* in leaves from Col-0, *npr1-2* and five individual transgenic lines (*npr1-2*/*TcNPR1*). Three days after inoculation, leaf disks were collected and bacterial titers were measured. Data represents the means ± standard errors of three replicates, each containing three leaf disks from three individual plants. Letters above the histogram indicate statistically significant differences among genotypes ($P<0.01$) using the single factor ANOVA. **C.** Correlation of bacterial growth with relative *AtPR1* expression level. Average growth of *Pseudomonas syringae* pv. tomato DC3000 (Figure 3-4B) and average

AtPR1 gene expression (Figure 3-3) were evaluated in leaf tissue of Col-0, *npr1-2* mutant and five transgenic lines expressing *TcNPR1* and correlations were established.

Nuclear translocation of *TcNPR1* after SA induction

Another hallmark of AtNPR1 function is its nuclear localization in response to treatment with SA (Kinkema et al., 2000; Mou et al., 2003; Durrant and Dong, 2004; Spoel et al., 2009; Zhang et al., 2010). To determine if TcNPR1 can also translocate into the nucleus in response to SA in a manner similar to *Arabidopsis* NPR1, we created transgenic *Arabidopsis* plants containing a TcNPR1-EGFP translational fusion and observed the subcellular localization of the fusion protein using confocal microscopy (Figure 5). This construct (35S*:TcNPR1:EGFP)* was stably transformed into the *npr1-2* mutant and we observed the localization of EGFP fusion protein before and 24hrs after SA treatment in both leaf and root cells of four independent transgenic lines. We observed no EGFP fluorescence in negative control plants transformed with the identical vector lacking the TcNPR1-EGFP fusion gene (Figure 5A and 5B). As an additional control, transgenic plants overexpressing EGFP without a fusion to TcNPR1 were imaged, and we observed strong fluorescence in both cytoplasm and nucleus with no localization changes after SA treatment. A final control consisted of a construct designed for the overexpression of the *Arabidopsis* NPR1 protein translationally fused to EGFP (35S:*AtNPR1*:*EGFP*). Consistent with the findings of others (Kinkema et al., 2000; Mou et al., 2003), we observed an extremely strong nuclear translocation of the fusion protein in leaf guard cells and in root cells 24hrs after SA treatment.

The TcNPR1-EGFP fusion protein appeared to be evenly distributed in cytoplasm of leaf guard cells from water-treated 4-week-old soil grown plants, however, the protein accumulated moderately in guard cell nucleus 24 hours after SA application (Figure 5A, red arrow). Similarly, a modest level of nuclear translocation could also be observed in the root cells from 10-day-old seedlings grown on MS medium supplemented with 0.5mM SA (Figure 5B). Although protein translocation of TcNPR1 is of lesser extent than observed with the *Arabidopsis* NPR1-EGFP protein based on reduced nuclear fluorescence observed

in TcNPR1-EGFP transgenic plants, our results taken together indicate that TcNPR1, like *Arabidopsis* NPR1, can translocate into nucleus after SA induction and participate in the induction of defense related gene expression.

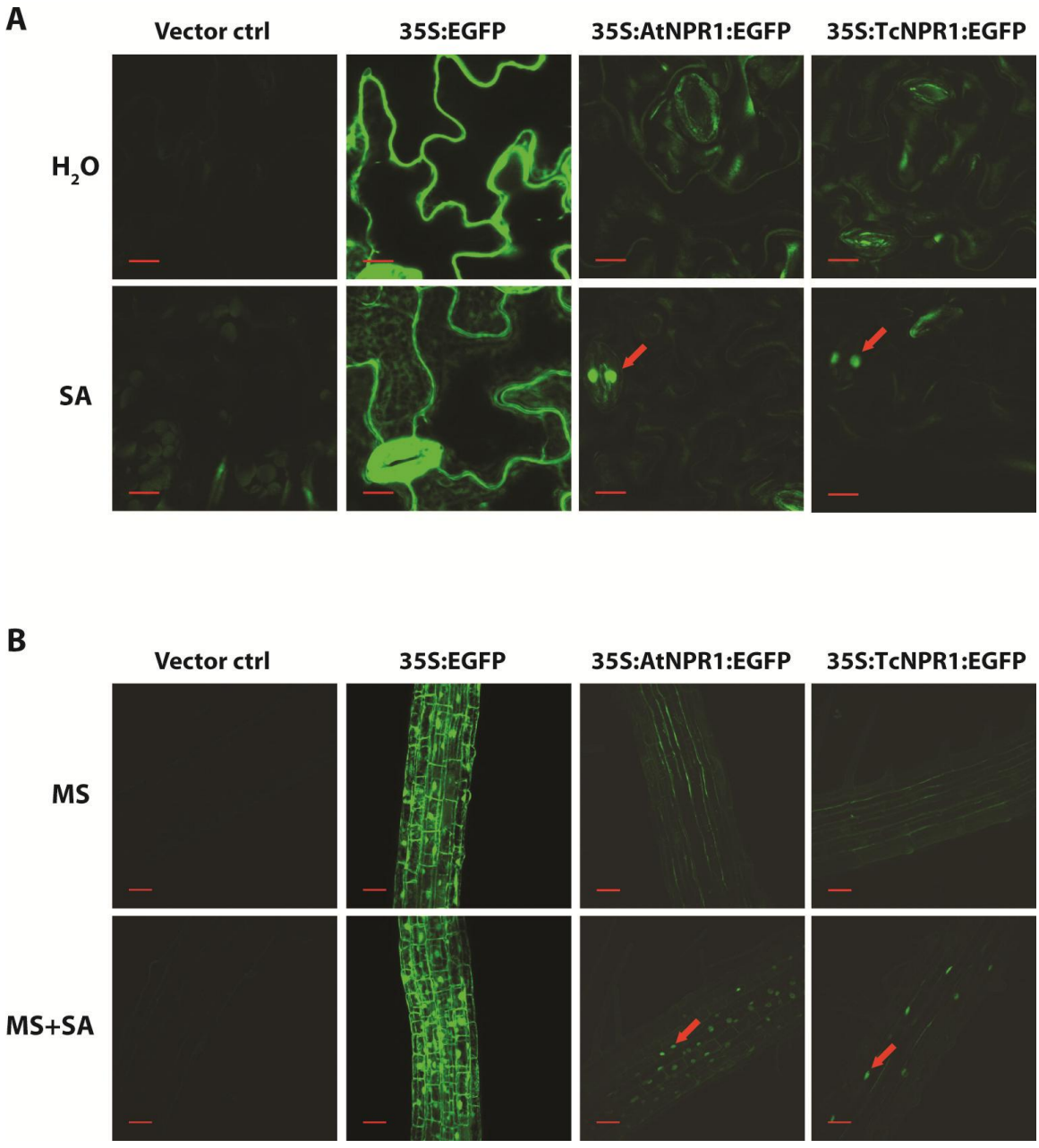


Figure 3-5 Nuclear localization of TcNPR1-EGFP in transgenic *Arabidopsis* plants in response to SA induction. **A.** Confocal images of EGFP fluorescence in *Arabidopsis* leaves of

4-week-old soil-grown plants 24hrs after H_2O (upper images) or 1mM SA (lower images) treatment. All images were taken at the same magnification and exposure. Arrows indicate the accumulation of green fluorescence in the guard cell nucleus after SA treatment. Scale bar, 10μm. **B.** Confocal images of EGFP fluorescence in *Arabidopsis* roots from 10-day-old seedlings grown on MS (upper images) or MS supplemented with 0.5mM SA (lower images). All images were captured using the same exposure settings. Arrows indicate the accumulation of EGFP in the nucleus of root cells after SA treatment. Scale bar, 30μm. Samples from transgenic plants generated with pCAMBIA1300 (vector ctrl) was used as negative control and samples from transgenic plants expressing 35S:EGFP served as positive control in **A** and **B**.

TcNPR1 and SA-JA crosstalk

It has been previously demonstrated that *Arabidopsis* NPR1 can mediate the antagonism between SA and jasmonic acid (JA) by suppressing JA-responsive genes (Feys and Parker, 2000; Spoel et al., 2003; Leon-Reyes et al., 2009), suggesting that it plays an important role in fine tuning the cross-talk between different regulatory pathways. To explore the role of TcNPR1 in cross-talk, we tested the effect of SA and JA treatments on defense gene expression in wild type Col-0, *npr1-2* mutant and five independent 35S:TcNPR1 transgenic *Arabidopsis* lines. Semi-quantitative RT-PCR showed that all five lines carrying the cacao transgene expressed *TcNPR1* at moderate levels, and these did not change much during hormone treatments (Figure 6A). Exogenous application of 1mM SA activated *PR1* in Col-0 and three transgenic lines, but not in *npr1-2* mutant. Additionally, 48hrs after treatment with 0.1mM methyl jasmonate (MeJA) in 0.015% Silwet L-77, two well established MeJA inducible genes (*VSP2* and *PDF1.2)* were up-regulated in wild-type plant and in *npr1-2* mutant, consistent with previous reports (Penninckx et al., 1998; Spoel et al., 2003). Two DNA bands were detected in some of the *PDF1.2* PCR products, and we determined that the smaller molecular weight band resulted from cDNA amplification and the large fragment resulted from amplification of genomic DNA (data not shown). As predicted, all five 35S:TcNPR1 transgenic lines exhibited levels of increased *VSP2* and *PDF1.2* that were similar to those seen in Col-0 plants. Upon treatment with a combination of 1mM SA and 0.1mM MeJA in 0.015% Silwet L-77, *PR1* was expressed at a level similar to seen when

plants were treated with SA alone, indicating that MeJA had no effects on SA-responsive *PR1* expression. Both *VSP2* and *PDF1.2* expressed at significantly lower levels in Col-0 compared to that in *npr1-2* mutant after SA and MeJA combined treatment, demonstrating the function of AtNPR1 in antagonistic repression of JA-responsive genes. All five transgenic lines containing *TcNPR1* gene displayed reduced expression levels of JA-responsive gene expression upon SA and JA combined treatment compared to *npr1-2* mutant, suggesting that TcNPR1 can also mediate SA-JA cross-talk in a manner similar to AtNPR1.

To quantify the expression of *VSP2* and *PDF1.2* after the treatment of the combination of SA and MeJA, we measured the band intensity as above (Figure 6B). The data was normalized to an *Ubiquitin* control for loading effects. The relative expression levels of *VSP2* and *PDF1.2* were significantly decreased in *TcNPR1* expressing transgenic lines compared to *npr1-2* mutant ($P<0.05$), a pattern similar to wild-type Col-0, suggesting that *TcNPR1* restored the *npr1* phenotype. These data support our hypothesis that TcNPR1 may play a role in mediating SA-JA cross talk as does *Arabidopsis* NPR1.

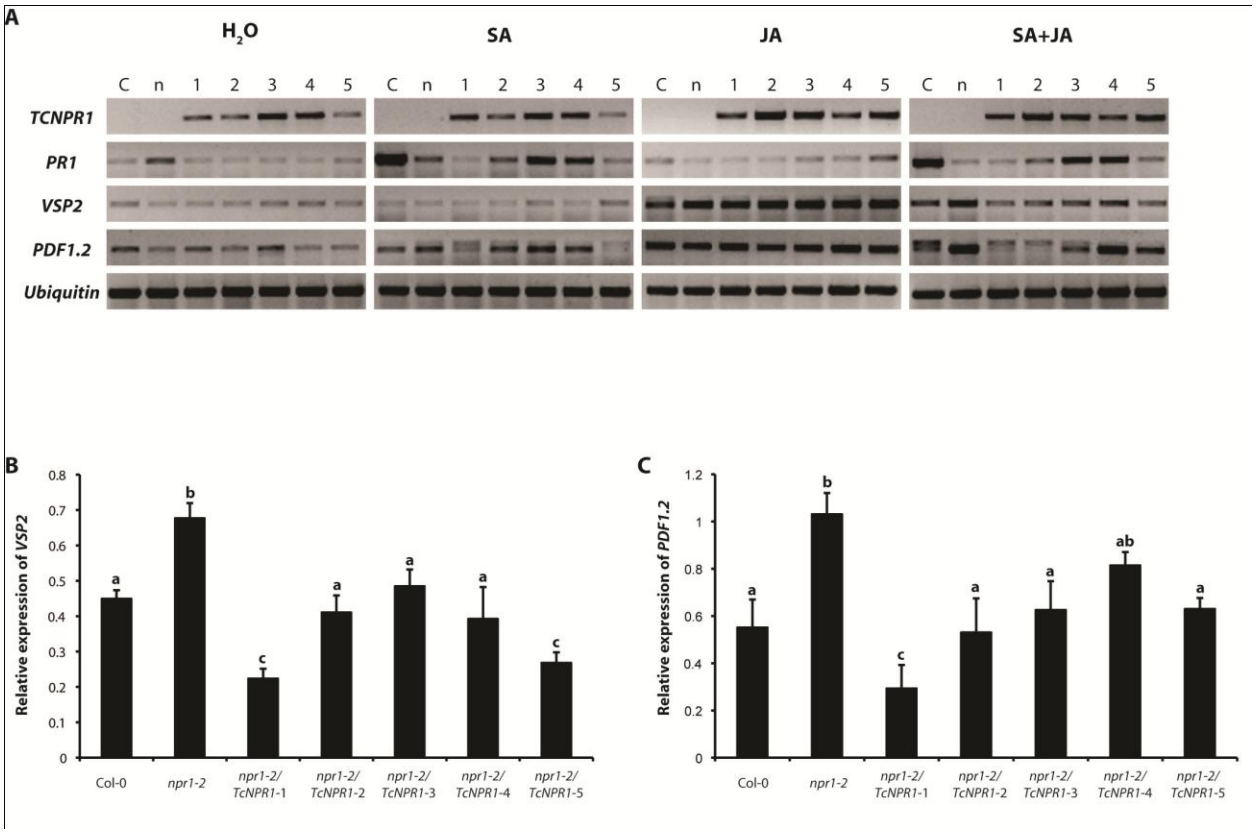


Figure 3-6 Gene expression of *TcNPR1*, SA- and JA-responsive genes in transgenic *Arabidopsis npr1-2* mutants. **A.** Semi-quantitative RT-PCR was performed with cDNA prepared from leaves of 4-week-old plants of wild type(C), *npr1-2* (n) and 5 independent transgenic *npr1-2* mutant expressing *TcNPR1* (1-5). The expression of *TcNPR1*, *AtPR1*, *AtVSP2* and *AtPDF1.2* was evaluated 48hrs after treatment with water control, 1mM SA water solution alone, 0.1mM MeJA alone in 0.015% Silwet L-77 and the combination of 1mM SA and 0.1mM MeJA in 0.015% Silwet L-77. *AtUbiquitin* was used as a cDNA loading and normalization control. **B.** The intensity of *AtVSP2* and *AtPDF1.2* RT-PCR gel bands in Figure 3-6**A** were quantified by ImageQuant software for total pixel intensity and the expression of JA-responsive genes *AtVSP2* and *AtPDF1.2* was normalized by *AtUbiquitin*. The bar charts represent the means ± standard errors of relative expression value of *AtVSP2* and *AtPDF1.2* following 48hrs treatment of SA-MeJA combination of three biological replicates. Letters above the bar chart indicate statistically significant differences among genotypes ($P<0.05$) determined by single factor ANOVA.

3.5 Discussion

We have isolated an *NPR1* homologous gene from the tropical tree, *Theobroma cacao*, and have generated transgenic *Arabidopsis npr1-2* mutant lines overexpressing *TcNPR1*. All of our results are consistent with the hypothesis that *TcNPR1* is a functional ortholog of the well characterized *Arabidopsis* gene. *TcNPR1* complemented each of the major *Arabidopsis npr1-2* mutant phenotypes that were tested. Over-expression of *TcNPR1* in the *npr1-2* mutant conferred *PR1* up-regulation after SA treatment and increased resistance to *Pseudomonas syringae* pv. tomato DC3000 (Figure 3 and 4 A, B). *TcNPR1* was shown to be translocated into the nucleus in response to SA and to participate in SA-JA cross talk regulation (Figure 5 and 6). In our data, we found that transgenic lines 1 Line exhibited reduced complementation in SA-induced *PR1* expression and disease resistance (Figure 3 and 4), while at the same time same two lines efficiently mediated crosstalk between SA and JA (Figure 6). In previous studies, the activation of defense related genes was shown to involve the nuclear translocation of NPR1 (Kinkema et al., 2000) while the crosstalk between SA and JA signaling was shown to be mediated by cytosolic NPR1 (Spoel et al., 2003), thus it appears that very different mechanisms exist for these two functions of NPR1. The differential efficiencies of complementation of TcNPR1 we observed may reflect these different mechanisms. It is well known that positional effects (differential transgene transcription levels due to different genomic insertion sites in individual transgenic events) can have a large effect on protein expression levels. As suggested by RNA expression levels of the different TcNPR1 expressing transgenic lines (Figure 3), lines 1 and 5 may have lower protein expression than lines 2-4. It seems plausible that the differential complementation of the two NPR1 functions resulted from the differences in expression levels, potentially as a result of different protein accumulation levels in the cytosol vs nuclear compartments. Consistent with this idea, only the higher levels of expression seen in lines 2-4 was sufficient to complement the nuclear gene induction function, but the levels of expression were high enough in all lines to complement the cytosolic SA/JA crosstalk function.

In all, our results demonstrate a high degree of evolutionary and functional conservation of NPR1 from the Brassicales to the Malvales. NPR1 is also conserved in

species as diverse as grapevine (Le Henanff et al., 2009), tomato (Chen et al., 2009), apple (Malnoy et al., 2007), banana (Endah et al., 2008), cotton (Zhang et al., 2008), tobacco (Liu et al., 2002) and rice (Yuan et al., 2007). This high degree of functional conservation suggests that NPR1 function evolved very early in the development of higher plants and that it plays a very critical role in plant development and reproductive success.

Little is known about the mechanisms of defense signaling in cacao. Our data suggests that the central mechanisms operative in *Arabidopsis* are likely to be conserved in cacao. At a minimum, our data suggests that the mechanisms and molecules that interact with NPR1 during SA and JA signaling and nuclear translocation are also conserved in cacao. If this were not the case, we would not expect the cacao NPR1 protein to function normally in *Arabidopsis*. However, the cacao protein in some cases only partially restored function of the *npr1* mutant, which is likely the result of transgene expression level differences compared to the endogenous gene and/or partial molecular incompatibility with its interacting protein partners. It is possible that the binding affinities between the cacao NPR1-interacting proteins are reduced as compared to the endogenous *Arabidopsis* coevolved partners. Partial complementation has commonly been observed in heterologous complementation analysis in many other systems (Irish and Yamamoto, 1995; Dong et al., 2001; Maizel et al., 2005).

Further investigation is needed to explore the entire defense response pathway in *Theobroma cacao* and to understand the similarities and differences with *Arabidopsis* overall. For example, our expression data shows that *TcNPR1* can be up-regulated only at 4mM SA treatment but not at lower concentrations, which is higher than the optimal level of 1mM in *Arabidopsis* as previously indicated (Cao et al., 1997). It would be interesting to test the endogenous SA level of cacao and to determine dose response dynamics in various tissues and during different stages of development. Another area of interest is to identify and characterize the downstream targets of *TcNPR1* and to compare them to the approximately 2,248 genes that are regulated by NPR1 during systemic acquired resistance in *Arabidopsis* (Wang et al., 2006). Surveying these genes in cacao could reveal interesting differences in the defense responses unique to this tropical tree relative to *Arabidopsis*. Furthermore, *Arabidopsis* NPR1 has been shown to interact with several different proteins

such as the TGA transcription factors (Despres et al., 2000; Johnson et al., 2003; Rochon et al., 2006; Boyle et al., 2009; Spoel et al., 2009). Thus another area of interest is to isolate *TcNPR1* interacting cacao proteins, which will further enhance our knowledge of this pathway in cacao. We are also interested in studying other NPR1-like genes of cacao and the recent completion of a draft cacao genome sequence has led to the identification of three additional NPR1-like cacao genes (*Argout et al., 2010).

Plant diseases, especially pathogenic fungi, are estimated to cause about 30-40% yield loss on cacao annually (Hebbar, 2007; Ploetz, 2007), and thus disease resistance is of substantial interest to cacao breeders. Our findings can be utilized in several approaches to help develop varieties of cacao with enhanced disease resistance. The sequence of the *TcNPR1* gene could possibly be used to develop molecular markers and probes that can be employed to select disease resistant varieties with specific allelic variations. Interestingly, the major quantitative trait locus (QTLs) for witches' broom disease resistance is tightly linked to the *TcNPR1* gene (*Argout et al., 2010), thus the *TcNPR1* gene serves as a key candidate gene for generation of molecular markers that can be used for marker assisted selection of new disease resistant varieties. In addition, *TcNPR1* expression levels could be modified in transgenic cacao varieties to develop broad-spectrum disease resistance. This approach has already been successful in several species but to our knowledge, has not yet been deployed in commercial production for any species. However, consumer and industry reluctance to accept transgenic plant technology remains a formidable barrier to development of any transgenic cacao varieties for commercialization.

3.6 Conclusion

The isolation of the *TcNPR1* gene and its heterologous complementation in *Arabidopsis* allowed us to rapidly characterize the function of this defense-related gene. The up-regulation of *PR1* and increased bacterial resistance in transgenic *Arabidopsis npr1-2* mutants strongly supported that *TcNPR1* is a functional ortholog of Arabidopsis *Npr1,* and vital component in SA-dependent signaling pathway in *Theobroma cacao*. Our results provide potential opportunities to enhance disease resistance in this crop species through

conventional breeding or biotechnological approaches. Further investigation is needed to identify the TcNPR1 interacting transcription factors and their downstream targets in cacao and to reveal further details of the molecular mechanisms of the role TcNPR1 plays as a central mediator of the plant defense response.

Accession numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative, GenBank/EMBL databases or Esttik database (http://esttik.cirad.fr/) under the following accession numbers: At1g64280 (*NPR1*), At2g14610 (*PR1*), At5g24770 (*VSP2*), At5g44420 (*PDF1.2*), At3g52590 (*ubiquitin*), HM117159 (*TcNPR1*) and CL33contig2 (*TcActin*).

Abbreviations

NPR1: non expressor of PR genes 1; SA: salicylic acid; INA: 2,6-dichloroisonicotic acid; BTH: benzothiadiazole; BTB/POZ: broad complex, tramtrack and bric a brac/pox virus and zinc finger; JA: jasmonic acid; PR: pathogenesis related; SAR: systemic acquired resistance; NLS: nuclear localization signal; MeJA: methyl jasmonate; VSP2: vegetative storage protein 2; PDF1.2: plant defensin 1.2; QTL: quantitative trait locus

Authors' contributions

ZS performed most of the experiments, *ie,* sequence analysis, gene expression studies, phenotypic analysis of transgenic *Arabidopsis* plants, confocal microscopy observations and drafted the manuscript. SNM participated in the design of the study, directed the transformation vector construction and transgenic lines generation, and participated in drafting of the manuscript. YL participated in transgenic *Arabidopsis* plants analysis and helped to analyze the sequence. JV cloned the *TcNPR1* gene. MJG conceived the study, drafted the manuscript and gave advice on experimental design, data analysis and execution. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank Nicole Zembower at cytometry facility for providing help for confocal microscopy imaging. Thanks to the people in the Guiltinan lab, especially Ann Young and Sharon Pishak for their technical assistance in construction the transformation vectors and generation of transgenic plants. This work is supported in part by The Pennsylvania State University, The Huck Institutes of Life Sciences and American Research Institute Penn State Endowed Program in the Molecular Biology of Cacao.

3.7 References

1. Heath MC: **Hypersensitive response-related death**. *Plant Mol Biol* 2000, **44**(3):321-334.
2. Durrant WE, Dong X: **Systemic acquired resistance**. *Annu Rev Phytopathol* 2004, **42**:185-209.
3. Attaran E, Zeier TE, Griebel T, Zeier J: **Methyl salicylate production and jasmonate signaling are not essential for systemic acquired resistance in Arabidopsis**. *Plant Cell* 2009, **21**(3):954-971.
4. Mur LA, Kenton P, Lloyd AJ, Ougham H, Prats E: **The hypersensitive response; the centenary is upon us but how much do we know?** *J Exp Bot* 2008, **59**(3):501-520.
5. Dong X: **NPR1, all things considered**. *Curr Opin Plant Biol* 2004, **7**(5):547-552.
6. Mukhtar MS, Nishimura MT, Dangl J: **NPR1 in plant defense: It's not over 'til It's turned over**. *Cell* 2009, **137**(5):804-806.
7. Cao H, Bowling SA, Gordon AS, Dong X: **Characterization of an Arabidopsis mutant that is nonresponsive to inducers of systemic acquired resistance**. *Plant Cell* 1994, **6**(11):1583-1592.
8. Liu Y, Schiff M, Marathe R, Dinesh-Kumar SP: **Tobacco *Rar1*, *EDS1* and *NPR1/NIM1* like genes are required for N-mediated resistance to tobacco mosaic virus**. *Plant J* 2002, **30**(4):415-429.
9. Fitzgerald HA, Chern MS, Navarre R, Ronald PC: **Overexpression of *(At)NPR1* in rice leads to a BTH- and environment-induced lesion-mimic/cell death phenotype**. *Mol Plant Microbe Interact* 2004, **17**(2):140-151.
10. Vlot AC, Klessig DF, Park SW: **Systemic acquired resistance: the elusive signal(s)**. *Curr Opin Plant Biol* 2008, **11**(4):436-442.
11. Yuan Y, Zhong S, Li Q, Zhu Z, Lou Y, Wang L, Wang J, Wang M, Yang D, He Z: **Functional analysis of rice *NPR1-like* genes reveals that *OsNPR1/NH1* is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility**. *Plant Biotechnol J* 2007, **5**(2):313-324.
12. Chen YY, Lin YM, Chao TC, Wang JF, Liu AC, Ho FI, Cheng CP: **Virus-induced gene silencing reveals the involvement of ethylene-, salicylic acid- and mitogen-**

activated protein kinase-related defense pathways in the resistance of tomato to bacterial wilt. *Physiol Plant* 2009, **136**(3):324-335.

13. Makandar R, Essig JS, Schapaugh MA, Trick HN, Shah J: **Genetically engineered resistance to *Fusarium* head blight in wheat by expression of *Arabidopsis NPR1***. *Mol Plant Microbe Interact* 2006, **19**(2):123-129.
14. Le Henanff G, Heitz T, Mestre P, Mutterer J, Walter B, Chong J: **Characterization of *Vitis vinifera NPR1* homologs involved in the regulation of pathogenesis-related gene expression**. *BMC Plant Biol* 2009, **9**:54.
15. Gaffney T, Friedrich L, Vernooij B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H, Ryals J: **Requirement of salicylic acid for the induction of systemic acquired resistance**. *Science* 1993, **261**(5122):754-756.
16. Rochon A, Boyle P, Wignes T, Fobert PR, Despres C: **The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines**. *Plant Cell* 2006, **18**(12):3670-3685.
17. Loake G, Grant M: **Salicylic acid in plant defence--the players and protagonists**. *Curr Opin Plant Biol* 2007, **10**(5):466-472.
18. Shah J, Tsui F, Klessig DF: **Characterization of a salicylic acid-insensitive mutant (*sai1*) of *Arabidopsis thaliana*, identified in a selective screen utilizing the SA-inducible expression of the *tms2* gene**. *Mol Plant Microbe Interact* 1997, **10**(1):69-78.
19. Delaney TP, Friedrich L, Ryals JA: ***Arabidopsis* signal transduction mutant defective in chemically and biologically induced disease resistance**. *Proc Natl Acad Sci U S A* 1995, **92**(14):6602-6606.
20. Glazebrook J, Rogers EE, Ausubel FM: **Isolation of *Arabidopsis* mutants with enhanced disease susceptibility by direct screening**. *Genetics* 1996, **143**(2):973-982.
21. Ryals J, Weymann K, Lawton K, Friedrich L, Ellis D, Steiner HY, Johnson J, Delaney TP, Jesse T, Vos P *et al*: **The Arabidopsis NIM1 protein shows homology to the mammalian transcription factor inhibitor I kappa B**. *Plant Cell* 1997, **9**(3):425-439.
22. Cao H, Li X, Dong X: **Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance**. *Proc Natl Acad Sci U S A* 1998, **95**(11):6531-6536.
23. Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG: **Sequence and structural analysis of BTB domain proteins**. *Genome Biol* 2005, **6**(10):R82.
24. Baldwin AS, Jr.: **The NF-kappa B and I kappa B proteins: new discoveries and insights**. *Annu Rev Immunol* 1996, **14**:649-683.
25. Mou Z, Fan W, Dong X: **Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes**. *Cell* 2003, **113**(7):935-944.

26. Pieterse CM, Van Loon LC: **NPR1: the spider in the web of induced resistance signaling pathways**. *Curr Opin Plant Biol* 2004, **7**(4):456-464.
27. Feys BJ, Parker JE: **Interplay of signaling pathways in plant disease resistance**. *Trends Genet* 2000, **16**(10):449-455.
28. Despres C, DeLong C, Glaze S, Liu E, Fobert PR: **The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors**. *Plant Cell* 2000, **12**(2):279-290.
29. Zhou JM, Trifa Y, Silva H, Pontier D, Lam E, Shah J, Klessig DF: **NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the *PR-1* gene required for induction by salicylic acid**. *Mol Plant Microbe Interact* 2000, **13**(2):191-202.
30. Ndamukong I, Abdallat AA, Thurow C, Fode B, Zander M, Weigel R, Gatz C: **SA-inducible *Arabidopsis* glutaredoxin interacts with TGA factors and suppresses JA-responsive *PDF1.2* transcription**. *Plant J* 2007, **50**(1):128-139.
31. Johnson C, Mhatre A, Arias J: **NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2**. *Biochim Biophys Acta* 2008, **1779**(10):583-589.
32. Fan W, Dong X: **In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis**. *Plant Cell* 2002, **14**(6):1377-1389.
33. Boyle P, Le Su E, Rochon A, Shearer HL, Murmu J, Chu JY, Fobert PR, Despres C: **The BTB/POZ domain of the Arabidopsis disease resistance protein NPR1 interacts with the repression domain of TGA2 to negate its function**. *Plant Cell* 2009, **21**(11):3700-3713.
34. Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Metraux JP, Brown R, Kazan K *et al*: **NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol**. *Plant Cell* 2003, **15**(3):760-770.
35. Leon-Reyes A, Spoel SH, De Lange ES, Abe H, Kobayashi M, Tsuda S, Millenaar FF, Welschen RA, Ritsema T, Pieterse CM: **Ethylene modulates the role of NONEXPRESSOR OF PATHOGENESIS-RELATED GENES1 in cross talk between salicylate and jasmonate signaling**. *Plant Physiol* 2009, **149**(4):1797-1809.
36. Koornneef A, Pieterse CM: **Cross talk in defense signaling**. *Plant Physiol* 2008, **146**(3):839-844.
37. Malnoy M, Jin Q, Borejsza-Wysocka EE, He SY, Aldwinckle HS: **Overexpression of the apple *MpNPR1* gene confers increased disease resistance in *Malus* x *domestica***. *Mol Plant Microbe Interact* 2007, **20**(12):1568-1580.
38. Sakamoto K, Tada Y, Yokozeki Y, Akagi H, Hayashi N, Fujimura T, Ichikawa N: **Chemical induction of disease resistance in rice is correlated with the expression of a gene encoding a nucleotide binding site and leucine-rich repeats**. *Plant Mol Biol* 1999, **40**(5):847-855.

39. Yang Y, Qi M, Mei C: **Endogenous salicylic acid protects rice plants from oxidative damage caused by aging as well as biotic and abiotic stress**. *Plant J* 2004, **40**(6):909-919.
40. Chern M, Fitzgerald HA, Canlas PE, Navarre DA, Ronald PC: **Overexpression of a rice *NPR1* homolog leads to constitutive activation of defense response and hypersensitivity to light**. *Mol Plant Microbe Interact* 2005, **18**(6):511-520.
41. Hebbar PK: **Cacao diseases: a global perspective from an industry point of view**. *Phytopathology* 2007, **97**(12):1658-1663.
42. Purdy LH, Schmidt RA: **STATUS OF CACAO WITCHES' BROOM: biology, epidemiology, and management**. *Annu Rev Phytopathol* 1996, **34**:573-594.
43. Wood G, Lass R: **Cocoa**. New York: Longman Scientific & Technical Press; 1985.
44. Evans H, Priori C: **Cocoa pod diseases: causal agents and control**. *Outlook on Agriculture* 1987, **16**:35-41.
45. Phillips-Mora W, Wilkinson MJ: **Frosty pod of cacao: a disease with a limited geographic range but unlimited potential for damage**. *Phytopathology* 2007, **97**(12):1644-1647.
46. Bailey BA, Bae H, Strem MD, Antunez de Mayolo G, Guiltinan MJ, Verica JA, Maximova SN, Bowers JH: **Developmental expression of stress response genes in *Theobroma cacao* leaves and their response to Nep1 treatment and a compatible infection by *Phytophthora megakarya***. *Plant Physiol Biochem* 2005, **43**(6):611-622.
47. Bailey BA, Strem MD, Bae H, de Mayolo GA, Guiltinan MJ: **Gene expression in leaves of *Theobroma cacao* in response to mechanical wounding, ethylene, and/or methyl jasmonate**. *Plant Science* 2005, **168**(5):1247-1258.
48. Buddenhagen IW: **Resistance and Vulnerability of Tropical Crops in Relation to Their Evolution and Breeding**. *Annals of the New York Academy of Sciences* 1977, **287**(Feb25):309-326.
49. Guiltinan M, Verica J, Zhang D, Figueira A: **Genomics of *Theobroma cacao*, "The foods of the Gods"**. In: *Genomics of Tropical Crop Plants.* Edited by P.M. Moore, Ming R: New York: Springer; 2008.
50. Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W, Amores FM, Motamayor JC: **Development of a marker assisted selection program for cacao**. *Phytopathology* 2007, **97**(12):1664-1669.
51. Bae H, Kim SH, Kim MS, Sicher RC, Lary D, Strem MD, Natarajan S, Bailey BA: **The drought response of *Theobroma cacao* (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses**. *Plant Physiol Biochem* 2008, **46**(2):174-188.
52. Maximova SN, Marelli JP, Young A, Pishak S, Verica JA, Guiltinan MJ: **Over-expression of a cacao class I chitinase gene in *Theobroma cacao L.* enhances**

resistance against the pathogen, *Colletotrichum gloeosporioides*. *Planta* 2006, **224**(4):740-749.

53. Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ: **Isolation of ESTs from cacao (*Theobroma cacao L.*) leaves treated with inducers of the defense response**. *Plant Cell Rep* 2004, **23**(6):404-413.
54. Jones PG, Allaway D, Gilmour DM, Harris C, Rankin D, Retzel ER, Jones CA: **Gene discovery and microarray analysis of cacao (*Theobroma cacao L.*) varieties**. *Planta* 2002, **216**(2):255-264.
55. Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M *et al*: **Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of *Theobroma cacao L.* generated from various tissues and under various conditions**. *BMC Genomics* 2008, **9**:512.
56. Albagli O, Dhordain P, Deweindt C, Lecocq G, Leprince D: **The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins**. *Cell Growth Differ* 1995, **6**(9):1193-1198.
57. Li J, Mahajan A, Tsai MD: **Ankyrin repeat: a unique motif mediating protein-protein interactions**. *Biochemistry* 2006, **45**(51):15168-15178.
58. Zhang Y, Fan W, Kinkema M, Li X, Dong X: **Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene**. *Proc Natl Acad Sci U S A* 1999, **96**(11):6523-6528.
59. Kinkema M, Fan W, Dong X: **Nuclear localization of NPR1 is required for activation of *PR* gene expression**. *Plant Cell* 2000, **12**(12):2339-2350.
60. Higo K, Ugawa Y, Iwamoto M, Korenaga T: **Plant cis-acting regulatory DNA elements (PLACE) database: 1999**. *Nucleic Acids Res* 1999, **27**(1):297-300.
61. Lescot M, Dehais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouze P, Rombauts S: **PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences**. *Nucleic Acids Res* 2002, **30**(1):325-327.
62. Rombauts S, Dehais P, Van Montagu M, Rouze P: **PlantCARE, a plant cis-acting regulatory element database**. *Nucleic Acids Res* 1999, **27**(1):295-296.
63. Molina C, Grotewold E: **Genome wide analysis of *Arabidopsis* core promoters**. *BMC Genomics* 2005, **6**(1):25.
64. Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E: **Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters**. *Gene* 2007, **389**(1):52-65.
65. Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E: **Functional analysis of regulatory sequences controlling *PR-1* gene expression in Arabidopsis**. *Plant J* 1998, **16**(2):223-233.

66. Yu D, Chen C, Chen Z: **Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression**. *Plant Cell* 2001, **13**(7):1527-1540.
67. Yang P, Chen C, Wang Z, Fan B, Chen Z: **A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter**. *Plant J* 1999, **18**(2):141 - 149.
68. Faleiro F, Queiroz V, Lopes U, Guimarães C, Pires J, Yamada M, Araújo I, Pereira M, Schnell R, Filho G *et al*: **Mapping QTLs for witches' broom (*Crinipellis Perniciosa*) resistance in cacao (*Theobroma Cacao L.*)**. *Euphytica* 2006, **149**(1):227-235.
69. Cao H, Glazebrook J, Clarke JD, Volko S, Dong X: **The *Arabidopsis NPR1* gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats**. *Cell* 1997, **88**(1):57-63.
70. Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P, Dong X: **Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity**. *Cell* 2009, **137**(5):860-872.
71. Zhang X, Chen S, Mou Z: **Nuclear localization of NPR1 is required for regulation of salicylate tolerance, *isochorismate synthase 1* expression and salicylate accumulation in *Arabidopsis***. *J Plant Physiol* 2010, **167**(2):144-148.
72. Penninckx IA, Thomma BP, Buchala A, Metraux JP, Broekaert WF: **Concomitant activation of jasmonate and ethylene response pathways is required for induction of a plant defensin gene in Arabidopsis**. *Plant Cell* 1998, **10**(12):2103-2113.
73. Endah R, Beyene G, Kiggundu A, van den Berg N, Schluter U, Kunert K, Chikwamba R: **Elicitor and *Fusarium*-induced expression of *NPR1-like* genes in banana**. *Plant Physiol Biochem* 2008, **46**(11):1007-1014.
74. Zhang Y, Wang X, Cheng C, Gao Q, Liu J, Guo X: **Molecular cloning and characterization of *GhNPR1*, a gene implicated in pathogen responses from cotton (*Gossypium hirsutum L.*)**. *Biosci Rep* 2008, **28**(1):7-14.
75. Irish VF, Yamamoto YT: **Conservation of floral homeotic gene function between Arabidopsis and antirrhinum**. *Plant Cell* 1995, **7**(10):1635-1644.
76. Maizel A, Busch MA, Tanahashi T, Perkovic J, Kato M, Hasebe M, Weigel D: **The floral regulator LEAFY evolves by substitutions in the DNA binding domain**. *Science* 2005, **308**(5719):260-263.
77. Dong X, Braun EL, Grotewold E: **Functional conservation of plant secondary metabolic enzymes revealed by complementation of Arabidopsis flavonoid mutants with maize genes**. *Plant Physiol* 2001, **127**(1):46-57.
78. Wang D, Amornsiripanitch N, Dong X: **A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants**. *PLoS Pathog* 2006, **2**(11):e123.

79. Johnson C, Boden E, Arias J: **Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis**. *Plant Cell* 2003, **15**(8):1846-1858.
80. *Argout X, *Salse J, *Aury J, *Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S *et al*: (*These authors contributed equally to this work.) **Deciphering the genome structure and paleohistory of *Theobroma cacao***. *Nature Genetics* 2010, **In Press**.
81. Ploetz RC: **Cacao diseases: important threats to chocolate production worldwide**. *Phytopathology* 2007, **97**(12):1634-1639.
82. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice**. *Nucleic Acids Res* 1994, **22**(22):4673-4680.
83. Melnick RL, Zidack NK, Bailey BA, Maximova SN, Guiltinan M, Backman PA: **Bacterial endophytes: *Bacillus* spp. from annual crops as potential biological control agents of black pod rot of cacao**. *Biological Control* 2008, **46**(1):46-56.
84. Brechenmacher L, Weidmann S, van Tuinen D, Chatagnier O, Gianinazzi S, Franken P, Gianinazzi-Pearson V: **Expression profiling of up-regulated plant and fungal genes in early and late stages of *Medicago truncatula-Glomus mosseae* interactions**. *Mycorrhiza* 2004, **14**(4):253-262.
85. Hajdukiewicz P, Svab Z, Maliga P: **The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation**. *Plant Mol Biol* 1994, **25**(6):989-994.
86. Mitsuhara I, Ugaki M, Hirochika H, Ohshima M, Murakami T, Gotoh Y, Katayose Y, Nakamura S, Honkura R, Nishimiya S *et al*: **Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants**. *Plant Cell Physiol* 1996, **37**(1):49-59.
87. Lin JJ: **Electrotransformation of *Agrobacterium***. *Methods Mol Biol* 1995, **47**:171-178.
88. Clough SJ, Bent AF: **Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana***. *Plant J* 1998, **16**(6):735-743.
89. Maximova S, Miller C, Antunez de Mayolo G, Pishak S, Young A, Guiltinan MJ: **Stable transformation of *Theobroma cacao L.* and influence of matrix attachment regions on GFP expression**. *Plant Cell Rep* 2003, **21**(9):872-883.
90. Katagiri F, Thilmony R, He S: **The *Arabidopsis thaliana-Pseudomonas syringae* interaction**. In: *The Arabidopsis Book.* The American Society of Plant Biologists; 2002.

CHAPTER 4: IDENTIFICATION OF OTHER *NPR-LIKE* GENES IN *THEOBROMA CACAO* AND CHARACTERIZATION OF *THEOBROMA CACAO NPR3* GENE

4.1 Introduction

To effectively combat a wide range of pathogens, plants are able to induce a complex network of defense responses against invasions (Vlot et al., 2008). Upon activation of the defense response, resources are allocated to produce defensive proteins, resulting in an energetic and/or metabolic cost to plants (Heil, 2002; van Hulten et al., 2006). Therefore, defense responses are highly regulated by numbers of signal modules through a complicated network of positive and negative regulatory mechanisms and inter-connected communications among different pathways, which allows plants to react appropriately to infection and conserve metabolic energy for growth and reproduction needs (Glazebrook, 2001; Zhang and Klessig, 2001; Koornneef and Pieterse, 2008).

Intensive research over the past decades has begun to reveal the positive and negative regulators of the plant defense response. The most well studied positive regulator in systemic acquired resistance (SAR) is Non Expressor of PR1 (NPR1), a transcription coactivator that regulates *PR* gene expression in SA-dependent signaling pathway (Dong, 2004; Zhang and Cai, 2005). *Arabidopsis npr1* mutants exhibit higher susceptibility to pathogen infection and impaired expression of *PR* genes (Cao et al., 1997). Additional positive regulators have been identified in SA-dependent systemic signaling, such as Enhanced Disease susceptibility 1 (EDS1) (McDowell et al., 2000), SA Induction Deficient 1 (SID1) (Nawrath and Metraux, 1999) and TGA transcription factors (Zhang et al., 2003).

Although negative regulators suppress the defense response, they are of great importance in controlling the timing and efficiency of resistance (Bowling et al., 1994; Li et al., 1999; Journot-Catalino et al., 2006; Zhang et al., 2006). Several negative regulators have been isolated from mutant screening in *Arabidopsis*. The *cpr1* (constitutive Expresser of PR 1) mutation causes constitutively high SA levels, increased *PR1* expression in the absence of pathogen, and thus enhanced disease resistance (Bowling et al., 1994). The resistance phenotype of *cpr1* is only partially blocked by *npr1* mutation, indicating a SA-dependent

but NPR1-independent signaling mechanism (Clarke et al., 2000). Additionally, SNI1 (suppressor of *npr1-1* inducible) was identified as a repressor of genes targeted by NPR1 induction, thus dampening the basal expression levels of downstream *PR* genes (Li et al., 1999; Durrant et al., 2007). SNI1 is a nuclear protein and it may function via chromatin remodeling through histone modification. Moreover, mutant analyses revealed that the loss of function mutant of IId subfamily of WRKY transcription factor WRKY11 displays increased resistance towards *P.s.t.* infection and the resistance is even higher in a *wrky11 wrky17* double mutant (Journot-Catalino et al., 2006). Genome-wide gene expression studies in single and double mutants demonstrated that both WRKY11 and WRKY17 mediate transcriptional changes upon pathogen infection. Furthermore, two NPR1 paralogs, NPR3 and NPR4 have been shown to be negative regulators of defense responses (Zhang et al., 2006). The *npr3* single mutant has elevated basal *PR1* level in leaves and an *npr3 npr4* double mutant also exhibits enhanced resistance against virulent bacterial and oomycete pathogens in a leaf infection assay.

Other than *NPR1*, *NPR3* and *NPR4*, three other *NPR1-like* genes are present in *Arabidopsis* genome. All of six predicted proteins contain a BTB/POZ domain at N-terminus followed by ankyrin repeats region in the center of the protein and they all possess several conserved cysteine residues (Hepworth et al., 2005; Norberg et al., 2005). *NPR2*, which shares the highest similarity with *NPR1*, was suggested to act as a regulator of both NPR1-dependent and independent *PR* expression, but this information was only presented in an abstract in a non-peer reviewed publication (Racki et al., 2003) and has not been confirmed. NPR5 (also named BOP2) and NPR6 (also named BOP1) have been shown to be functionally redundant in determining proper vegetative development and growth symmetry in leaf morphogenesis (Ha et al., 2003; Ha et al., 2004; Hepworth et al., 2005; Norberg et al., 2005; Ha et al., 2007; McKim et al., 2008; Jun et al., 2010). Despite the growing evidence for the molecular function of each of the NPR family members in *Arabidopsis*, there is no comprehensive identification of this family in other species, including sequenced genomes (rice, grape, poplar etc). Furthermore, even in *Arabidopsis*, little is known about the function of NPR proteins in tissues other then leaves.

The work presented in Chapter 2 of this thesis was designed to further define the role of the NPR1-like proteins in regulation of the defense response in *Arabidopsis.* Surprisingly, I discovered that NPR3 negatively regulates the SA dependent defense responses during early flower development, as demonstrated by the analysis of an *npr3* mutant, which is highly resistant to bacterial infection. I propose a model describing a possible molecular mechanism by which NPR3 represses the defense response and discuss the evidence supporting this model.

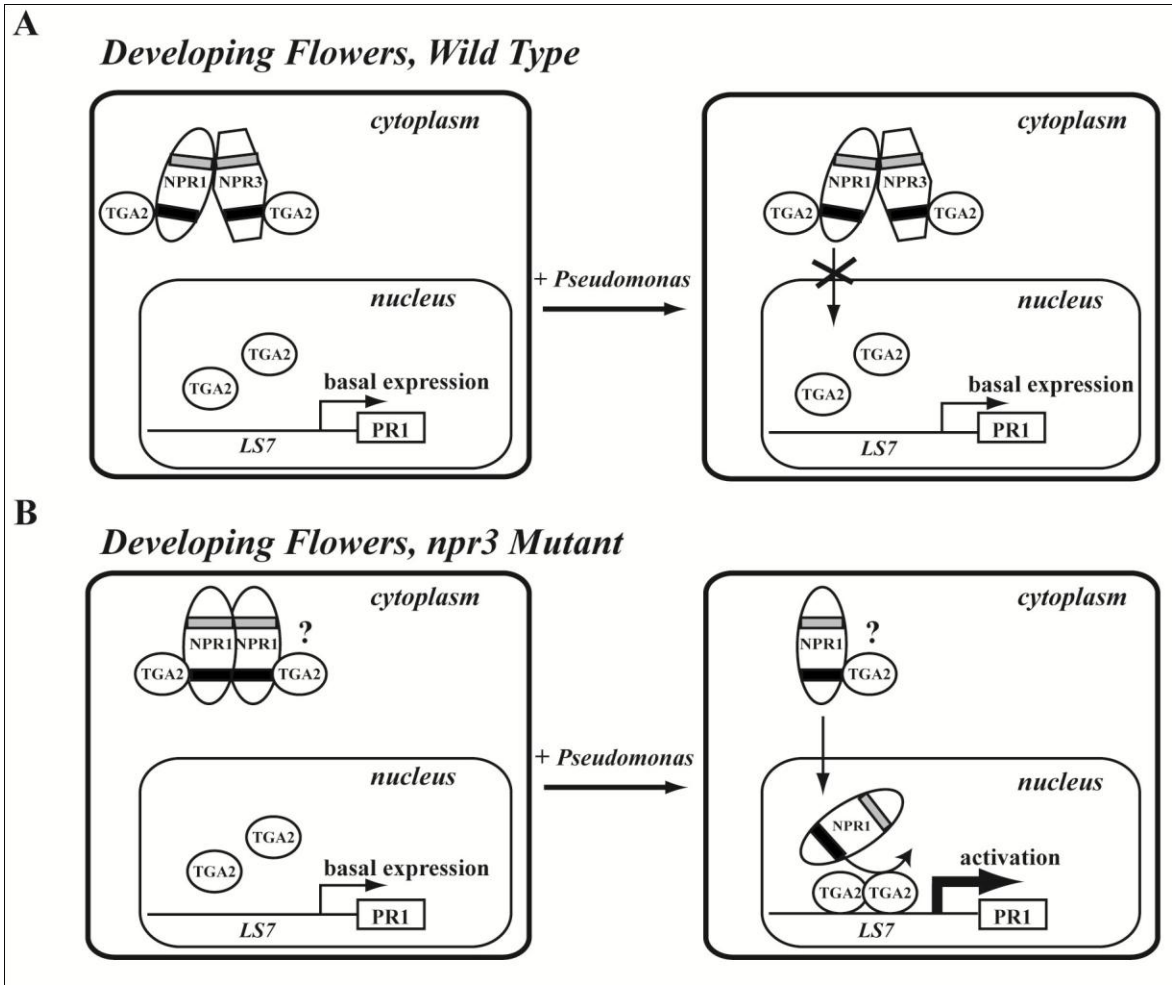


Figure 4-1 Proposed model for the mechanism of NPR3-mediated repression of the NPR1-dependent defense response. The model proposes that NPR3 acts to repress the NPR1-dependent defense response through protein-protein interactions with NPR1 and TGA2, thus preventing nuclear translocation of NPR1 and repressing defense gene activation (see Discussion for detailed explanation of the model and its supporting data). A, left panel: In

wild type developing flowers, interactions between NPR3, NPR1 and TGA2 inhibit the nuclear translocation and functional activity of NPR1. Upon *Pseudomonas* infection (A, right panel), NPR3 blocks the defense response. B, In the *npr3* mutant developing flower, NPR3 is not present. Upon *Pseudomonas* infection, NPR1 moves from the cytoplasm to the nucleus where it acts as a transcriptional coactivator and binds to TGA2 to facilitate the recruitment of TGA2 dimer at the *LS7* cis-element (Johnson et al., 2008), resulting in the activation of *PR1* gene and other defense related genes (B, left panel). Grey bars represent BTB/POZ domains in NPRs and black bars represent ankyrin repeats domains. The ? in panel B indicates interactions between TGA2 and NPR1 in the *npr3* mutant that are hypothesized without supporting data.

Understanding the mechanisms regulating the defense response has profound implications for agriculture. Such knowledge can be used to design breeding strategies, or to create transgenic plants with enhanced disease resistance against the plethora of plant pathogens. One such disease, witches' broom disease of cacao, is caused by the fungal pathogen *Moniliophthora perniciosa*, is a serious disease in Latin America. Indigenous to the Amazon rainforest, it now occurs in most cocoa growing regions in South America and has became a big threat to world cocoa production (Purdy and Schmidt, 1996; Meinhardt et al., 2008). Terminal and axillary buds, flower cushions and developing pods are highly susceptible to witches' broom infection, resulting in severe yield loss (Andebrhan et al., 1993). The genome of this pathogen was recently sequenced (Mondego et al., 2008) and a deeper understanding of the mechanisms of pathogenicity is beginning to emerge (Gesteira et al., 2007; Marelli et al., 2009; Lopes et al., 2010; Tiburcio et al., 2010). In an effort to advance breeding of disease resistant varieties of cacao, I conducted research to further our understanding of the defense response in this tropical tree species (Chapter 3).

Several homologs of *Arabidopsis* defensive proteins have been identified in cacao after SA/MeJA treatment, which suggests cacao and *Arabidopsis* may share similar defense signaling pathways (Verica et al., 2004; Bailey et al., 2005) and suggesting that a translational biology approach could be used to accelerate our knowledge of the cacao

defense pathways. In Chapter 3, I describe the isolation and functional analysis of the cacao *NPR1* gene. Using heterologous transgenic complementation, I showed that the cacao NPR1 gene is capable of restoring the *npr1-2* mutant phenotype. Cacao NPR1 protein was translocated into nuclei of transgenic *Arabidopsis* plants upon SA treatment, further indicating that cacao may share the same defense mechanisms as *Arabidopsis*. The fact that young flowers are of high susceptibility to witches' broom disease in cacao (Andebrhan et al., 1993), along with the evidence that NPR3 protein acts as a negative regulator in defense response during flower development in *Arabidopsis* (Chapter 2), led me to further investigate the possibility of a similar mechanism of defense repression in cacao. In this chapter, I describe the isolation and functional analysis of the *TcNPR3* gene. The *TcNPR3* gene restored the major phenotype of enhanced bacterial resistance to infection in *Arabidopsis npr3-3* mutant. Thus, I propose that TcNPR3 shares a similar function to *Arabidopsis* native NPR3 as a negative regulator of the defense response. This knowledge can be used in the future to develop varieties of cacao resistant to pathogens such as witches' broom disease.

4.2 Material and Methods

Genomic Resources Utilized

Esttik database (http://esttik.cirad.fr/): Fifty-six different cDNA libraries were constructed and sequenced from different organs, different genotypes and different environmental conditions (Argout et al., 2008). Expressed Sequence Tag Treatment and Investigation Kit is a tool dedicated to the analysis and annotation of cDNA raw data.

454 transcriptome: The cDNA library from six mixed tissues of Criollo genotype (B97-61/B2) was sequenced by Roche/454 technology (*Argout et al., 2010).

Draft Genome Assembly: The genome of *Theobroma cacao* was sequenced and analyzed from a Belizean Criollo genotype (B97-61/B2) (*Argout et al., 2010). The assembly consists of 4,792 scaffolds and it represents of 76% estimated genome size, which covers a large portion of the euchromatin of the *T. cacao* genome.

Full length cDNA cloning and Genomic DNA cloning of *TcNPR3*

A partial sequence of *TcNPR3* gene was identified and sequenced by screening the BAC filter arrays constructed using genomic DNA of genotype LCT-EEN 37 from the Clemson University Genomic Institute (http://www.genome.clemson.edu/) using known partial *TcNPR1* sequence as probe. Based on its partial sequence, primers were designed to clone the full length cDNA of *TcNPR3* from genotype Scavina6 (SCA6). The *TcNPR3* coding sequences was generated by RT-PCR using 1 µl of ½ diluted cDNA of SCA6 leaf stage C as template and 5µM of primers were used to include *Kpn*I and *Not*I restriction sites at the 5'- and 3'-ends respectively (TcNPR3-5'-*Kpn*I, GCGGTACCATGGCGTATTTATCTGAG; TcNPR3-3'-*Not*I, GCGCGGCCGCTCACAATTTTCTGAGC). PCR was performed using the following condition: 94ºC for 2 min., 32 cycles of 94ºC for 30 sec., 60ºC for 30 sec, followed by a 5min. final extension at 72ºC. PCR product was resolved on 1% agarose gels, purified with the GENECLEAN II Kit (Q-Biogene Inc., Solon OH) and cloned into the pGEM T-Easy vector (Promega Corporation, Madison WI). Forward and reverse sequencing was performed at the Penn State Genomics Core Facility to verify the sequence. The resulting clone was designated as pGEM-TcNPR3.

Full length *TcNPR3* coding sequence was used to blast against a cacao genome assembly database (paper under review) to obtain full length genomic sequence by performing blastn (Altschul et al., 1990).

Gene Structure and Protein Sequence Alignment

The full length CDS and full length genomic fragment of *TcNPR3* were uploaded onto SPIDEY (http://www.ncbi.nlm.nih.gov/spidey/), an online mRNA-to-genomic alignment program, and the gene structure was determined.

To verify the sequence, *Arabidopsis* NPR3 protein sequence was retrieved from TAIR (At5g45110) and TcNPR3 protein sequence was obtained by conceptual translation of its CDS. Protein sequence alignment was performed using MUSCLE software (Edgar, 2004). The TcNPR3 protein sequence was analyzed for potential functional sites by querying the

Simple Modular Architecture Research Tool (SMART) database (http://smart.embl-heidelberg.de/).

Identification of other *NPR-like* genes in *Theobroma cacao*

Full length protein sequences of 6 *Arabidopsis NPR* genes were used to blast against local cacao genome database and 454 EST database by performing tblastn with the cutoff E value of 1e-20 (Altschul et al., 1990; *Argout et al., 2010). Three scaffolds and five EST sequences were identified by performing four separate queries using NPR1, 2, 3 and 4 individually. Two of scaffolds and four ESTs were the ones containing *TcNPR1* and *TcNPR*3 (Table 1), and the novel scaffold was scaffold03937 and the novel EST sequence was named CL11980contig1. Another round of blast was carried out using CL11980contig1 as query in both 454 EST database and Esttik (Argout et al., 2008) to determine the full length CDS of this new cacao gene (named *TcNPR4*). The gene structure of *TcNPR4* was obtained by comparing all EST sequences and scaffold sequence in SPIDEY software. One scaffold and one EST sequence were identified using BOP2 and BOP1 as queries. The gene structure of *TcBOP2* was carried out by aligning EST and scaffold in SPIDEY. Protein sequences of TcNPR4 and TcBOP2 were obtained by conceptual translation from their CDS.

Table 1 *NPR1-like* genes in cacao and their corresponding scaffolds and EST sequences.

	scaffolds	EST sequences	
TcNPR1	scaffold00937	CL13863Contig1	CL4922Contig1
TcNPR3	scaffold00078	CL1Contig114	CL1Contig252
TcNPR4	scaffold03937	CL11980contig1	
TcBOP2	scaffold02275	CL7684Contig1	

To verify the sequences, putative protein sequences of *TcNPR4* and *TcBOP2* were aligned with their *Arabidopsis* counterpart respectively using MUSCLE software (Edgar, 2004). Potential functional sites were obtained by querying the Simple Modular Architecture Research Tool (SMART) database (http://smart.embl-heidelberg.de/).

Phylogenetic analysis

Multiple protein alignment of six *Arabidopsis* NPRs and four cacao NPRs was carried out by MUSCLE with default parameters (Edgar, 2004). The phylogenetic tree was constructed by MEGA 4.0 (Tamura et al., 2007) using neighbor-joining (NJ) method with the option of pairwise deletion. The bootstrap values were obtained from 2000 replicates.

Semi-quantitative RT-PCR analysis of *TcNPR3* expression in cacao tissues

Semi-quantitative RT-RCR was performed as described in Chapter 3. Basically, total RNA was extracted from leaf stages A, C and E, open flowers, un-open flowers, roots, pod seeds and pod exocarps. Three biological replicates were collected for each tissue. Cacao cDNA was synthesized in a final volume of 25µl from 2µg of total cacao RNA using M-MLV reverse transcriptase (New England Biolabs, Inc., Ipswich, MA) following the protocol given in Chapter 3.

Semi-quantitative RT-PCR was performed using primers for *TcNPR3* (TcNPR3-RT5': TGCTTGTCGACCCGCCATCAATTT; TcNPR3-RT3': AGGTTGTCTCAGCATGTGCTATGTCC) and *TcActin* used as cDNA loading and internal normalization controls (TcActinRT-5': AGCTGAGAGATTCCGTTGTCCAGA and TcActinRT-3': CCCACATCAACCAGACTTTGAGTTC). As for RT-PCR of *TcNPR1*, PCR was set up with 1 µl of ½ diluted cDNA and 5µM of the *TcNPR3* or *TcActin* primers. Titration of cycles was also performed to determine the amplification of *TcNPR3* is within linear range at cycle with following condition 94ºC for 2 min, 27 cycles of 94ºC for 30 sec., 56 ºC for 30 sec, 72ºC for 1 min, followed by the final extension of 72ºC for 5min. RT-PCR of *TcActin* was performed exactly the same as described in Chapter 3. PCR products were analyzed on 1% agrose gel stained with ethidium bromide. The intensity of PCR bands were quantified by ImageQuant software (Molecular Dynamics, Amersham Bioscience) and the relative expression values of *TcNPR3* were calculated by dividing the intensity values of corresponding *TcActin*.

Transgenic *Arabidopsis* mutant complementation assay

A T-DNA binary vector designed for overexpression of the *TcNPR3* coding sequence (p35S: TcNPR3) was created as follows. The *TcNPR3* coding sequence was excised from pGEM-TcNPR3 vector and cloned into *Kpn*I and *Not*I sites of an intermediate cloning vector

(pE2113) between the very strong E12-Ω promoter (Mitsuhara et al., 1996)and a 35SCaMV terminator. A 3 kb restriction fragment containing *TcNPR3* gene cassette was excised from pE2113 using *EcoR*I and *Pvu*II and ligated into *EcoR*I and *Sma*I site of pCAMBIA-1300 (Hajdukiewicz et al., 1994). Ligations were performed 1 hour at room temperature with 3Units of T4 DNA ligase (Promega Corporation, Madison WI) resulting in p35S: TcNPR3.

Arabidopsis transformation and transgenic plant identification

The binary vector of p35S: TcNPR3 containing a plant selectable marker for hygromycin resistance was introduced into *Agrobacterium tumefaciens* strain AGL1 by electroporation as previously described in (Lin, 1995). The floral dip method was used to transform *Arabidopsis npr3-3* mutants with p35S:TcNPR3 as described in Chapter 3 (Clough and Bent, 1998). Seeds were collected for 5 individual transgenic lines and the screening of positive transformants was conducted as described in Chapter 3. Flower tissues from six-week old soil-grown Col-0, *npr3-3* mutant and five individual transgenic lines were collected. RNA was extracted from each genotype using RNeasy plant mini kit (QIAGEN, Valencia CA). cDNA was synthesized using M-MLV reverse transcriptase (New England Biolabs, Inc., Ipswich, MA), as described in Chapter 3. Semi-quantative RT-PCR was performed to identify the heterologous *TcNPR3* expression. *AtUbiquitin* served as a cDNA loading and normalization control. Following primers and conditions were employed: TcNPR3-RT5': TGCTTGTCGACCCGCCATCAATTT; TcNPR3-RT3': AGGTTGTCTCAGCATGTGCTATGTCC (27cycles of 94ºC for 30 sec., 56 ºC for 30 sec., 72ºC for 1 min). Ubiquitin-5': ACCGGCAAGACCATCACTCT; Ubiquitin-3': AGGCCTCAACTGGTTGCTGT (22 cycles of 94ºC for 30 sec., 54 ºC for 30 sec., 72ºC for 1 min). The conditions of PCR were determined by cycle titration as described above to make sure the reaction was in the linear amplification range. The PCR products were resolved on 1% agarose gels, and the relative expression level of *TcNPR3* was calculated by comparing the level of *AtUnibquitin* using ImageQuant software (Molecular Dynamics, Amersham Bioscience) as described in (Brechenmacher et al., 2004).

Pseudomonas syringae infection assay of *Arabidopsis* transgenic plants overexpressing *TcNPR3* genes

Positive transgenic plants, along with wild-type Col-0 and *npr3-3* mutants, were grown in a Conviron growth chamber at 22ºC, with 16 h light/8 h dark cycle. Humidity was maintained at 60% and light intensity was 200 μM/m^2 using Octron 4100K Ecologic bulbs. The *Pseudomonas syringae* pv. tomato DC3000 was activated on pseudomonas agar with kanamycin (25 ng/μl) and rifampicin (100 ng/μl) at 28ºC for two days. The bacteria were suspended in sterile water to the concentration of OD_{600}=0.2 and surfactant Silwet L-77 (0.02%) was added to the suspension. Sterile water with 0.02% Silwet L-77 was used for treatments as negative control. The *Pseudomonas syringae* inflorescence infection and bacteria assay on Col-0, *npr3-3* mutants and transgenic *npr3-3* overexpressing *TcNPR3* were conducted as described in Chapter 2. The statistical analysis was performed using single factor ANOVA and/or Fisher's PLSD.

4.3 Results

Isolation of putative *TcNPR3* gene

Initially, a partial *TcNPR3* gene was identified by screening a BAC library by hybridization with partial *TcNPR1* sequence. Based on this sequence, PCR primers were designed and used to amplify cDNA isolated from cacao genotype Scavina6 (SCA6) stage C leaves. A fragment of 1764bp was isolated, cloned into pGEM vector and sequenced to reveal an intact coding sequence of the expected length and with high homology to the *Arabidopsis NPR3* gene.

Subsequently, a genomic sequence containing a putative *TcNPR3* gene was identified by searching a cacao genome database (unpublished data) using the full length cacao NPR3 cDNA as a query and the Blast algorithm. The structure of the TcNPR3 gene was obtained by comparing the full length cDNA and genomic sequences using SPIDEY program, which reveals four exons and three introns, similar to the genomic structure of *Arabidopsis NPR3* (Figure 4-1A).

Arabidopsis and cacao NPR3 protein sequences are highly similar

Conceptual translation of cacao NPR3 gene showed that it encodes a putative protein consisting of 587 amino acid residues, only one amino acid longer than *Arabidopsis*

NPR3. Alignment of TcNPR3 and AtNPR3 protein sequences demonstrated they are highly similar to each other (60% identity and 77% similarity). Both *Arabidopsis* and cacao NPR3 encode predicated proteins sharing key structural features (Figure 4-1B). Both proteins have a BTB/POZ domain near their N-terminus (dashed line box) that shares 64% identity and an ankyrin repeat region (solid line box) is present in both proteins which share about 72% identity. It has been shown that both BTB/POZ domain and ankyrin repeats are involved in protein-protein interactions (Albagli et al., 1995; Stogios et al., 2005; Li et al., 2006; Rochon et al., 2006). In *Arabidopsis*, it has been hypothesized that the BTB/POZ domain may be involved in the interaction with NPR1 protein and ankyrin repeats may mediate the interaction with TGA2 transcription factor (Chapter 2). These interactions interrupt the nuclear localization of NPR1-TGA2 complex upon floral infection and thus NPR3 functions as a negative regulator during early flower development. These similarities in protein structure suggest that *TcNPR3* gene may also share the same function as *AtNPR3* during defense response in developing flowers.

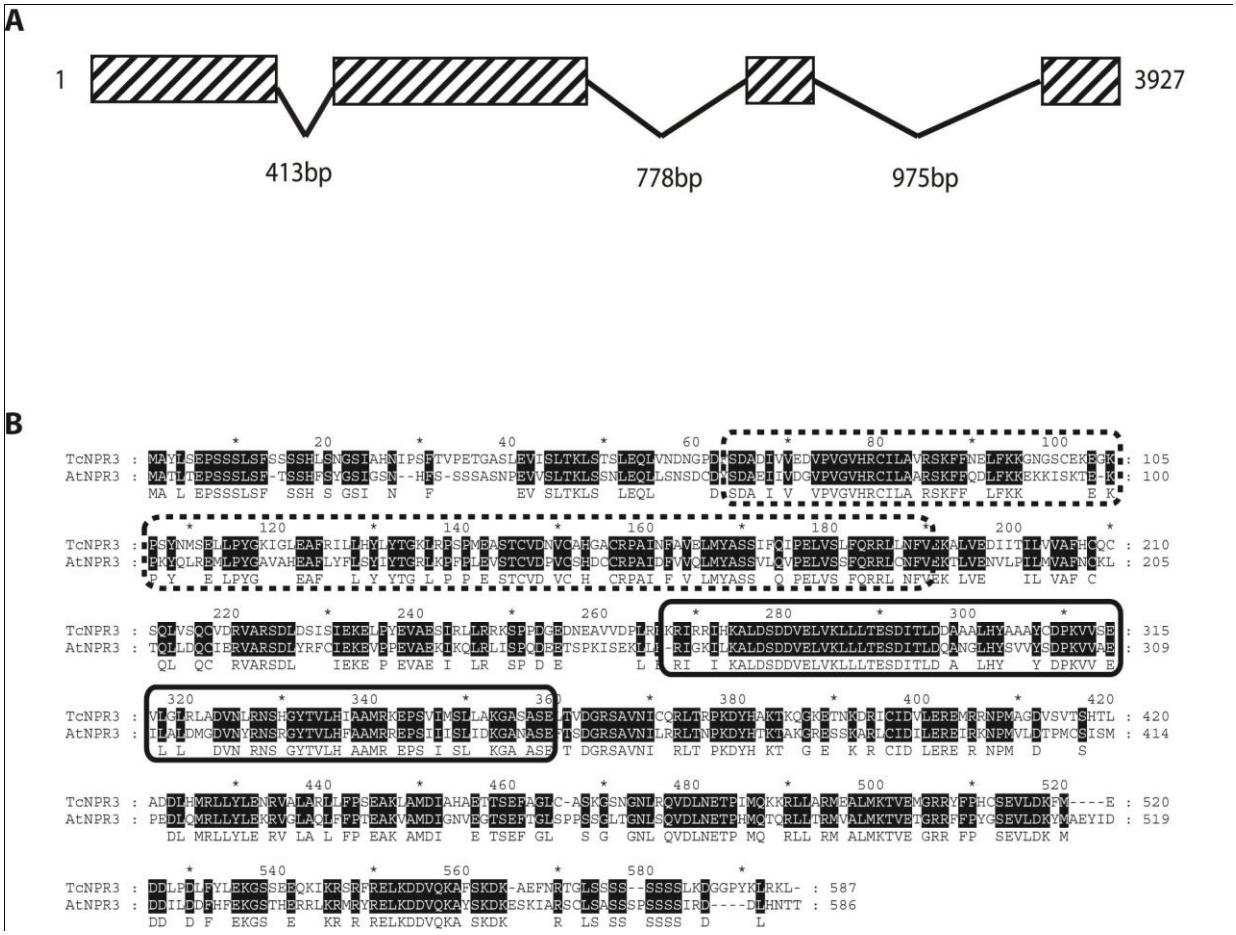


Figure 4-2 Gene and protein structure of *Theobroma cacao* NPR3. **A.** Diagram of *TcNPR3* gene structure. Boxes with diagonal stripes represent exons. The slanted lines represent introns. The sizes of the entire genomic fragment and each intron are indicated. **B.** Alignment of AtNPR3 and TcNPR3 proteins. Protein alignment was performed using MUSCLE software. Amino acids blocked in black indicate identical residues in both sequences and the amino acids below the blocks represent the consensus sequence. BTB/POZ and ankyrin repeat domains are highlighted by dashed line and solid line boxes, respectively.

Genomic sequence based identification of other *NPR-like* genes in cacao

To retrieve all the *NPR-like* genes in cacao, protein sequences of six *Arabidopsis NPR* genes were used to blast local cacao genome database and 454 EST database from

genotype B97-61/B2, a criollo type isolated in Belize (*Argout et al., 2010)(http://cocoagendb.cirad.fr/gbrowse/cgi-bin/gbrowse/theobroma/). Scaffold00937, scaffold00078 and scaffold03937 were identified with an E value cutoff of 1e-20 when using NPR1-4 as queries and two of them (00937 and 00078) contained *TcNPR1* and *TCNPR3*. Scaffold03937 contained a novel cacao gene, named *Theobroma cacao NPR4* (*TcNPR4*). Similarly, the blast results against 454 EST database revealed five EST sequences and only one, CL11980contig1, was from the novel cacao gene. Another round of blast analysis was conducted against 454 EST database and Esttik (http://esttik.cirad.fr/) using CL11980contig1 as query. Four more EST sequences were identified specific to *TcNPR4*, KZ0AAV3YB07FM1, KZ0ABF8YL04FM1, CL7174Contig2 and CL7174Contig1. A total of five ESTs and scaffold03937 were used to elucidate the predicted gene structure of *TcNPR4*. *TcNPR4* contains 4 exons and three introns (Figure 4-2A), the same structure as each of the *Arabidopsis* NPR1-4 genes, though its first intron is about 6.9kb long. When using *Arabidopsis BOP2* as queries, scaffold02275 and CL7684Contig1 were isolated from genome and EST databases with the E value of 1e-139 and 3e-98 respectively, and this new gene was named *Theobroma cacao BOP2* (*TcBOP2*). SPIDEY was performed to predict the gene structure by aligning genomic sequence and EST sequence. *TcBOP2* is composed of 2 exons and one intron (Figure 4-2B), the same gene structure as *Arabidopsis BOP2* and *BOP1*.

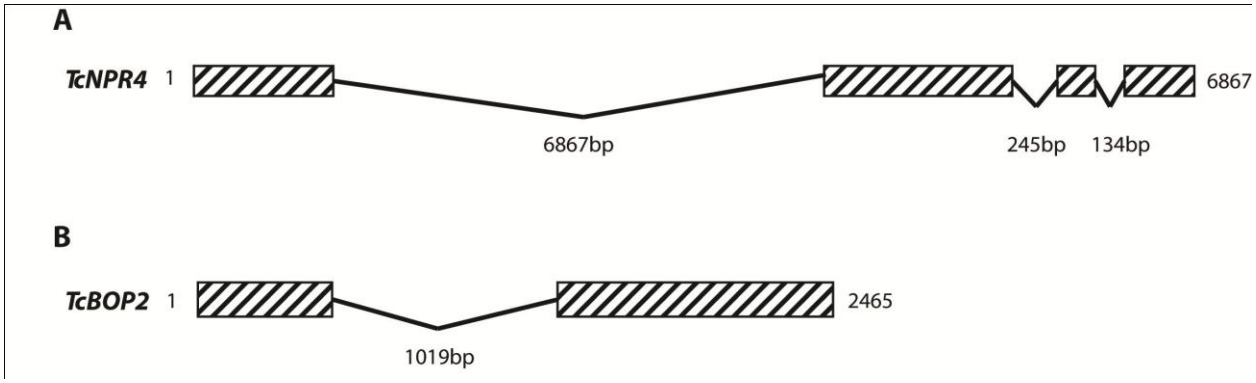


Figure 4-3 Diagrams of *TcNPR4* and *TcBOP2* gene structure. **A.** *TcNPR4* gene structure and **B.** *TcBOP2* gene structure. Boxes with diagonal stripes represent exons. The slanted lines represent introns. The sizes of the mRNA (from start codon to stop) and each intron are indicated.

Conceptual translation of *TcNPR4* revealed that it consists of 583 amino acid residues, nine amino acids longer than *Arabidopsis* NPR4. Protein alignment of TcNPR4 and AtNPR4 illustrated that they are similar to each other (47% identity and 65% similarity). Both proteins share similar function domains (Figure 4-3A). TcNPR4 and AtNPR4 have BTB/POZ domain near their N-terminal end (dashed line box), which shares 54% identity. An ankyrin repeat region (solid line box) is also present in both proteins that shares 60% identity. TcBOP2 protein is comprised of 481 amino acid residues, ten amino acids shorter than *Arabidopsis* BOP2. Protein alignment of TcBOP2 and AtBOP2 demonstrated that they are extremely similar to each other with 81% identity and 86% similarity. TcBOP2 protein also possesses BTB/POZ domain and ankyrin repeats region, like all other cacao NPRs (Figure 4-3B). Both TcBOP2 and AtBOP2 have BTB/POZ domain near N-terminus of the protein (dashed line box), which shares 81% identity. They both have an ankyrin repeats region (solid line box) at the center of the protein, which share 94% identity. The high similarities in protein functional domains suggest they may share the same function in vegetative development (Norberg et al., 2005).

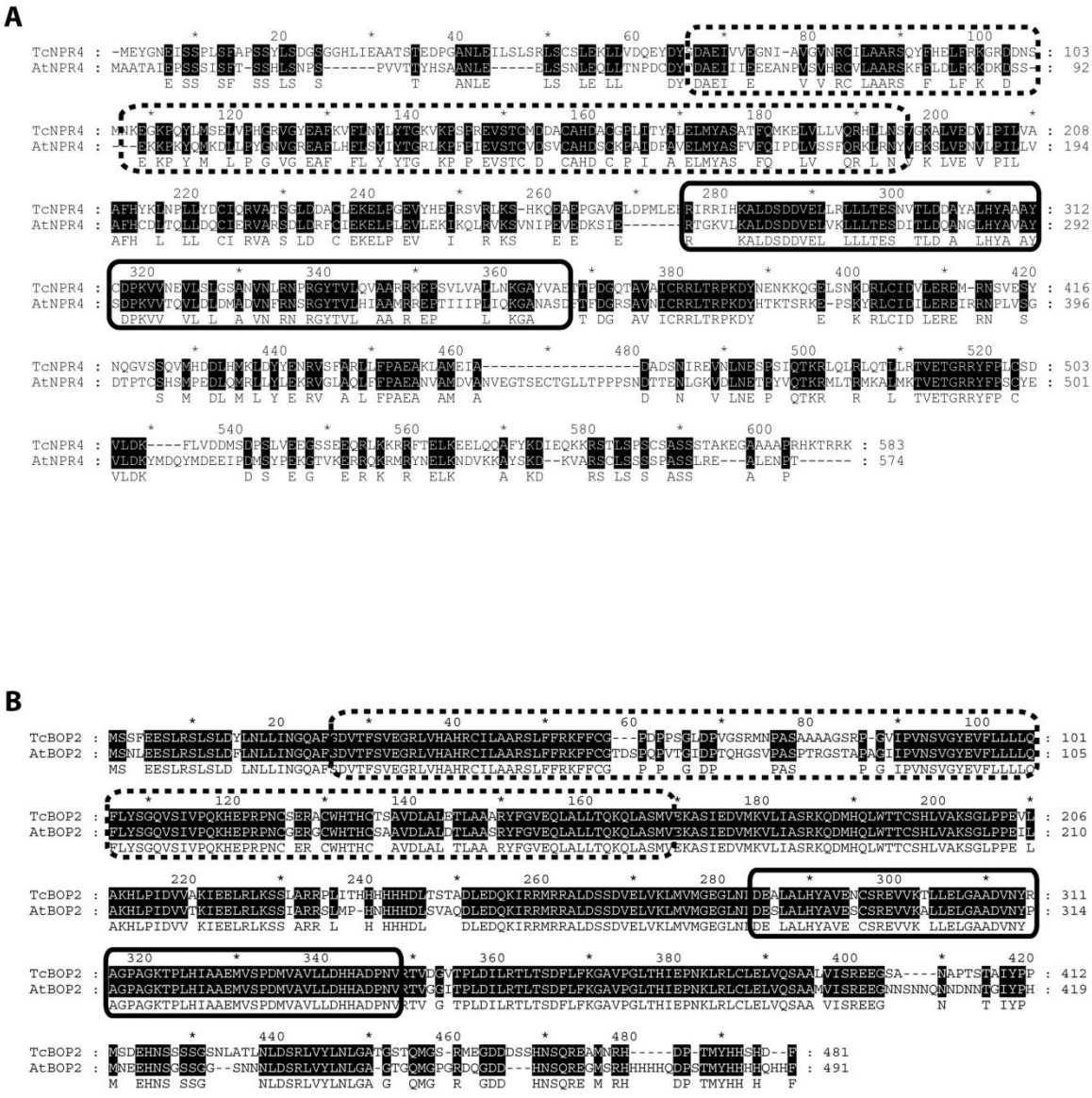


Figure 4-4 Protein alignment of TcNPR4/TcBOP2 and their *Arabidopsis* orthologs. The alignments were performed by MUSCLE software with default parameters. **A.** Alignment of TcNPR4 and AtNPR4. **B.** Alignment of TcBOP2 and AtBOP2. Amino acid residues blocked in black indicate identical residues in both sequences and the amino acids below the blocks represent consensus sequences. Dashed line boxes and solid line boxes designate BTB/POZ domain and ankyrin repeats region, respectively.

To investigate the evolutionary history of NPR- like gene family, phylogenetic analysis was performed based on protein sequences of four *NPR-like* genes in cacao and the

six *Arabidopsis NPRs* (Figure 4-4). A neighbor-joining tree was constructed using MEGA 4.0 with pairwise deletion. It suggests that the duplication of NPR1-4 and BOPs occurred before the divergence of *Arabidopsis* and cacao. The pairs of AtNPR1 and AtNPR2, as well as AtNPR3 and AtNPR4, appear to have been duplicated after the split of *Arabidopsis* and cacao. The analysis suggests that TcNPR3 and TcNPR4 are co-ortholog of AtNPR3/4, though TcNPR3 is closer to its *Arabidopsis* homologues. TcNPR1 is the co-ortholog of AtNPR1/2 and TcBOP2 is the ortholog of AtBOP2.

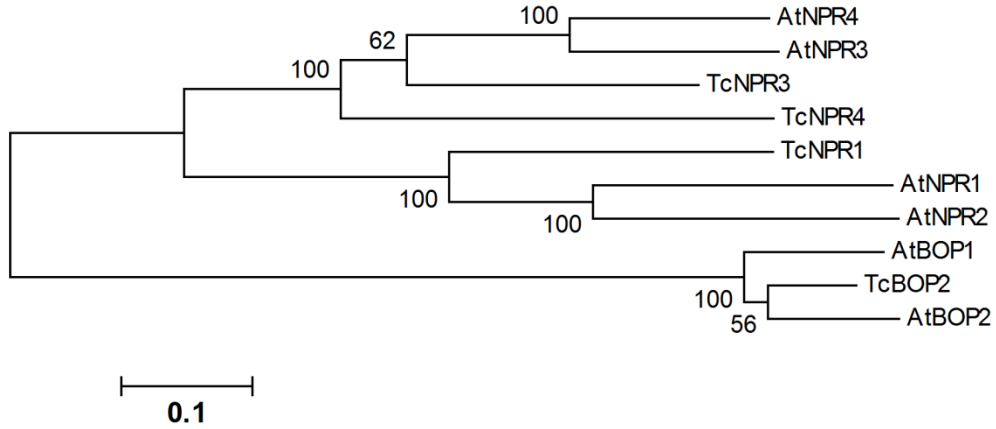


Figure 4-5 Phylogenetic tree of *Arabidopsis* and cacao NPR family members. Ten full length protein sequences from Arabidopsis and cacao were aligned using MUSCLE with default parameters. Neighbor-Joining (NJ) tree was constructed using MEGA 4.0 with pairwise deletion. The scale bar represents 0.1 substitutions per site and the values next to the nodes are bootstrap values from 2000 replicates.

Expression of *TcNPR3* in cacao tissues

Semi-quantative RT-PCR was performed to illustrate the basal expression level of *TcNPR3* in various cacao tissues of Scavina6, including leaves from sequential developmental stages A, C, E, representing young, mid and mature stages of development. In addition, RNA was isolated from open flowers, un-opened flowers, roots, seeds and fruit exocarp. *TcNPR3* is constitutively expressed in all tissues tested (Figure 4-2A), similar to its ortholog in *Arabidopsis* (Chapter 2). However, the basal expression level varies among

different tissues. To quantify the relative expression levels of *TcNPR3* in different tissues, we measured the fluorescence intensity of ethidium bromide stained DNA fragments irradiated with UV light using a high-sensitivity camera (shown in Figure 4-2A) and ImageQuant software (semi quantitative PCR). The results represented in Figure 4-2B showed that the expression in un-open flowers was relatively high among all the tissues and was about 2-fold higher than open flowers, a profile similar to that observed for *AtNPR3* during *Arabidopsis* development (Chapter 2). *TcNPR3* basal expression levels in root, pod seed and exocarp were moderate and comparable to that in open flowers. The expression of *TcNPR3* was lowest in young leaves (Stage A), moderate in stage C and highest in stage E, suggesting that the expression level of *TcNPR3* increases during leaf maturation. The expression pattern of *TcNPR3* is similar to *AtNPR3* in *Arabidopsis*, where *AtNPR3* expressed highest in unopened flowers followed by senescent leaves (Chapter 2).

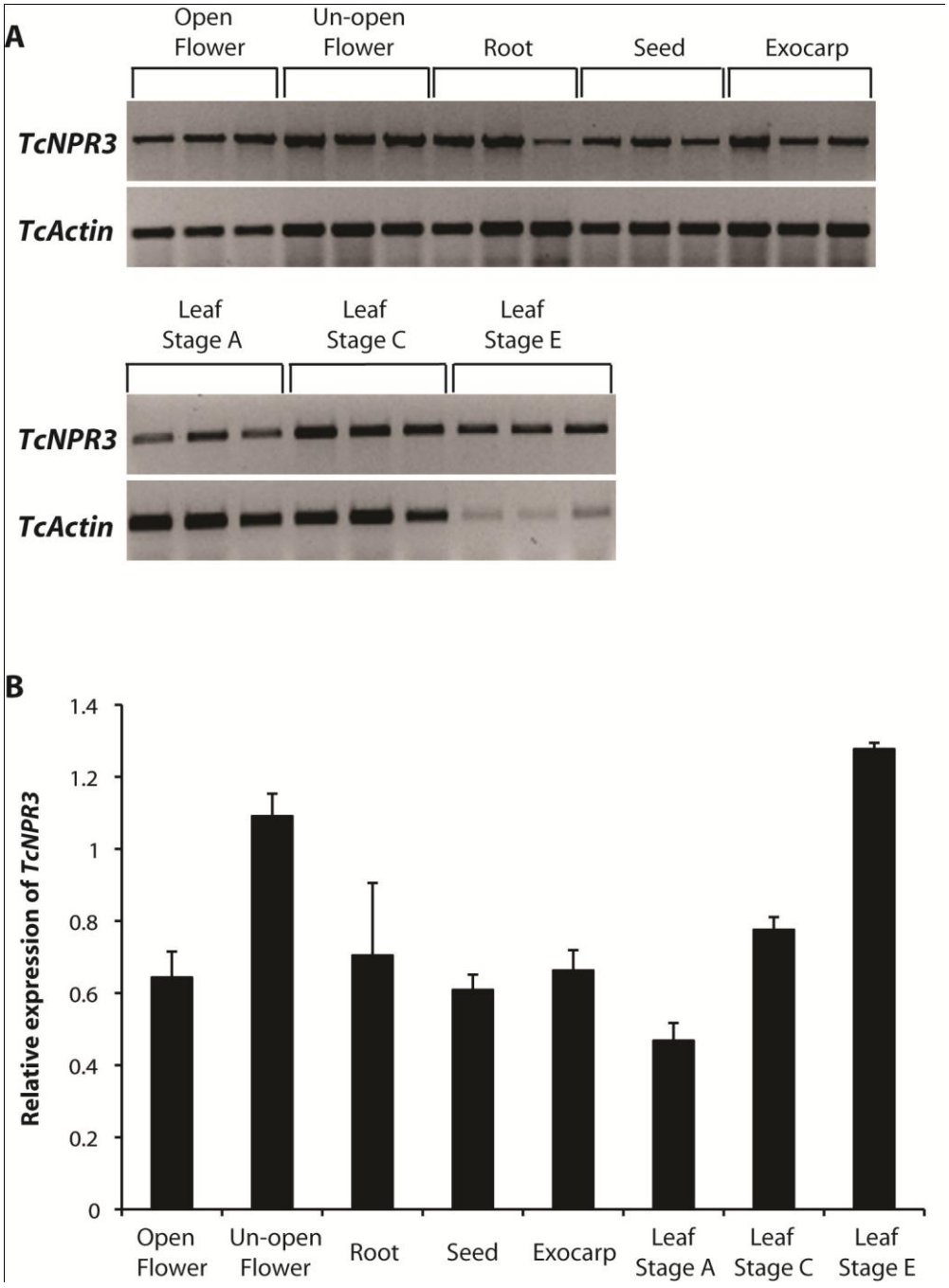


Figure 4-6 Gene expression analysis of *TcNPR3* in cacao. **A.** Expression of *TcNPR3* in various cacao tissues. Total RNA samples were collected from open flower, unopened flower, root, seed, exocarp and three different leaf developmental stages from youngest to oldest (A, C and E) from cacao genotype Scavina6 (SCA6), three biological replicates for each tissue. Semi-quantitative RT-PCR was performed and *TcActin* served as cDNA loading and normalization controls. **B.** Relative expression of *TcNPR3* from **A.** The intensity of PCR

band on gel images was quantified by ImageQuant and expression of *TcNPR3* was normalized to *TcActin*. Expression levels are presented as the means ± standard errors of three biological replicates.

Complementation of *Arabidopsis npr3-3* mutant

The work described above demonstrated that AtNPR3 acts as negative regulator during early flower development and the *npr3-3* mutant reduces whole plant fitness (Chapter 2). To evaluate the function of TcNPR3, we introduced the cacao *NPR3* CDS under the control of a constitutive promoter into *Arabidopsis npr3-3* mutant to examine if it can restore the mutant phenotype. To evaluate the expression of the transgene, semi-quantitative RT-PCR was performed using flowers of five independent transgenic *npr3-3* lines overexpressing *TcNPR3*, along with wild-type *Arabidopsis* Col-0 and the *npr3-3* mutant. As expected, *TcNPR3* expression was not detected in either Col-0 or the *npr3-3* mutant, but five independent transgenic lines all showed heterologous expression of *TcNPR3* with varied expression levels (Figure 4-6). We observed that line 1, 3 and 4 exhibited more *TcNPR3* expression than the other two lines.

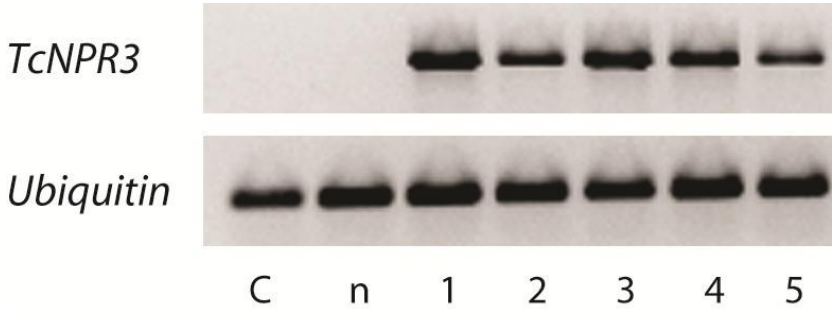


Figure 4-7 Gene expression of *TcNPR3* in transgenic *Araabidopsis npr3-3* lines. Semi-quantative RT-PCR was performed with cDNA prepared from flowers of 6-week-old plants of wild type (C), *npr3-3* mutant (n) and five individual transgenic *npr3-3* mutant overexpressing *TcNPR3* (1-5). *Arabidopsis Ubiquitin* (*AtUbiquitin*) was assayed as an internal and cDNA loading control.

The major phenotype of *Arabidopsis npr3-3* mutant is normal silique development after virulent *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*) inoculation compared to

short siliques in Col-0, ie. enhanced bacterial infection resistance (Chapter 2). To test if *TcNPR3* overexpression in the *npr3-3* mutant can complement the mutant resistant phenotype, the floral infection assay was carried out with five individual *TcNPR3* transgenic lines, wild-type Col-0 and *npr3-3* mutant. We inoculated the top of inflorescences with *P.s.t.* (OD_{600}=0.2 with 0.02% Silwet L-77) and disease symptoms and bacterial growth were assayed. In water-treated inflorescences, we did not observe any phenotypic differences among Col-0, *npr3-3* mutant and five transgenic lines (Figure 4-7A). Seven days after inoculation, the *P.s.t.* –treated samples indicated that Col-0 was more susceptible than *npr3-3* mutant, resulting in impaired flower development and shorter siliques similar to previous results (Chapter 2). However, flower and silique development of inoculated *npr3-3* mutant were comparable to mock inoculated plants. All five individual transgenic lines possessed intermediate silique development after inoculation. To quantify the difference, the length of fourth silique from the bottom of each infected inflorescence was measured from four biological replicates (Figure 4-7B). The mean silique length of *npr3-3* mutant was about 11mm after inoculation, which is the same as uninfected plants, but the silique length of infected Col-0 was significantly reduced to 1mm. All five independent transgenic *npr3-3* lines overexpressing *TcNPR3* exhibited medium silique length, ranging from 3mm to about 8mm. The inoculated silique length of line 1 and line 5 is statistically the same as Col-0 and *npr3-3* mutant, respectively, but the other three lines are significantly longer than Col-0 while shorter than *npr3-3* mutant. Taken together, I conclude that the cacao *NPR3* gene was capable of restoring *Arabidopsis npr3-3* mutant normal silique phenotype to wild-type short silique phenotype after inflorescence inoculation.

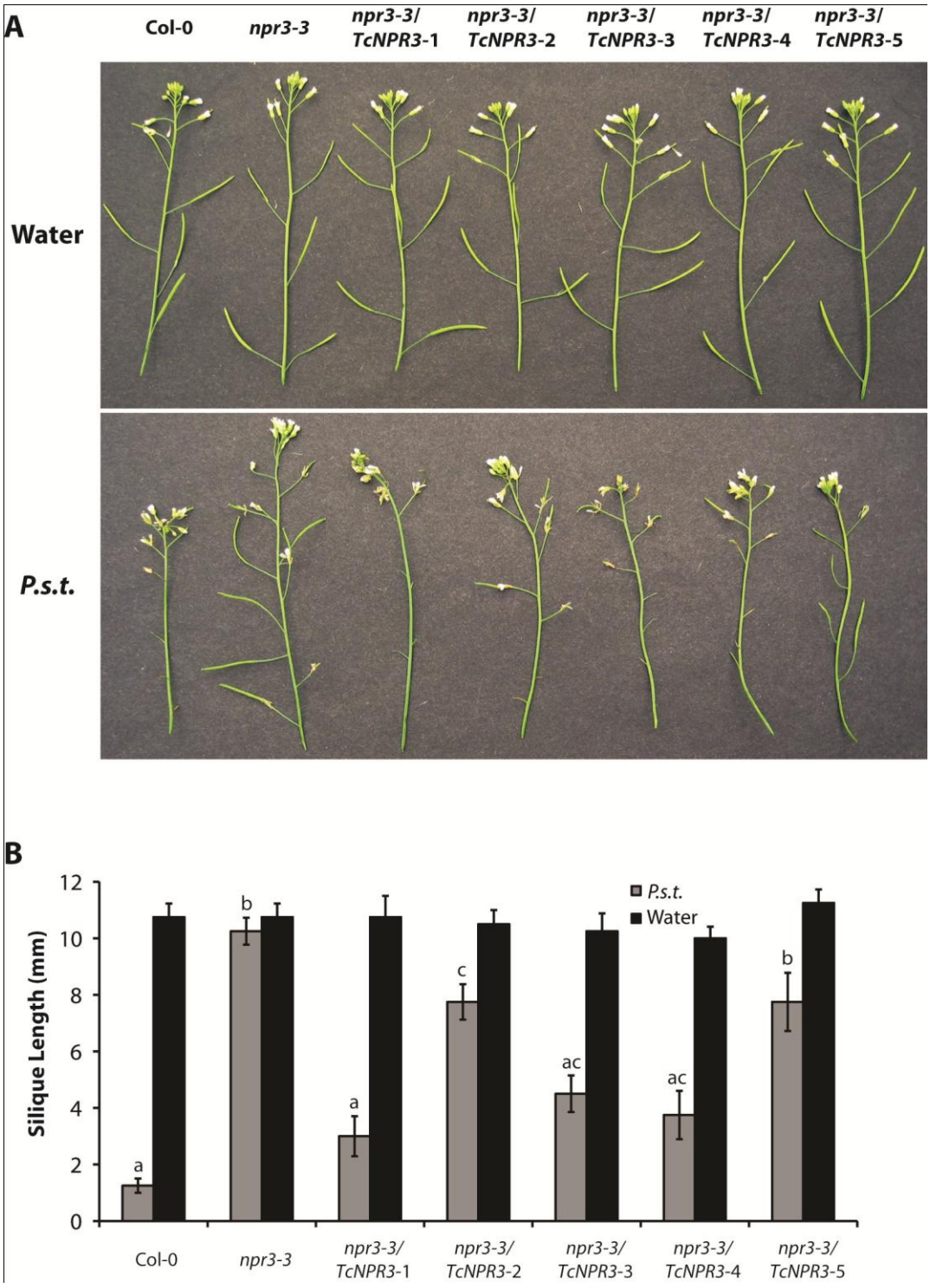


Figure 4-8 *Pseudomonas syringae* pv. tomato DC3000 infection assay of transgenic *npr3-3* mutant overexpressing *TcNPR3*. **A.** Representative images of siliques from infected Col-0, *npr3-3* mutant and five independent lines of *npr3-3* plants transformed with *TcNPR3* (*npr3-*

3/*TcNPR3*) seven days after inoculation with virulent bacteria *Pseudomonas syringae* pv. tomato DC3000 (*P.s.t.*) (OD_{600}=0.2 with 0.02% Silwet L-77) on top of the inflorescence. Additionally, water-treated (0.02% Silwet L-77) inflorescences of seven genotypes served as mock inoculation. **B.** Average length of the fourth silique from the bottom of infected inflorescence seven days after inoculation (gray bars) compared to water mock inoculation (black bars). Bar chart represents means ± SE of siliques from four biological replicates per treatment. Letters above the histogram indicate statistically significant differences among genotypes ($P<0.05$) determined by single factor ANOVA.

I previously demonstrated that the *Arabidopsis npr3-3* mutant can support 30-fold fewer bacteria than Col-0 (Chapter 2). Thus, to test whether the heterologous expression of *TcNPR3* can increase the bacterial growth in *npr3-3* mutant, bacterial growth was quantified in floral extracts. Similar to what was observed previously, the amount of bacteria in infected *npr3-3* inflorescence was decreased about 30-fold as compared to Col-0 (Figure 4-8A). Although line 5 exhibited the same statistical level as *npr3-3* mutant, two of the transgenic lines (Line 3 and 4) showed a similar level of bacterial growth to Col-0. The other two lines (line 1 and line 2) also supported significantly more bacteria than *npr3-3* mutant though the level was not as high as in Col-0. In all, four out of five transgenic lines exhibited significantly higher bacteria levels than did the *npr3-3* mutant, suggesting that *TcNPR3* can at least partially complement *npr3-3* mutant phenotype. To evaluate the relationship between bacterial growth level and heterologous *TcNPR3* expression in transgenic lines, we plotted $\log_{10}$ of bacterial growth (cfu mg FW^{-1}) against the relative expression level of *TcNPR3* in *npr3-3* mutant and five individual transgenic lines (Figure 4-8B). A strong positive correlation between those two variables was observed (R^2=0.8241), indicating that TcNPR3 plays a role in the negative regulation of plant disease resistance during flower development, the same function as *Arabidopsis* native NPR3 protein.

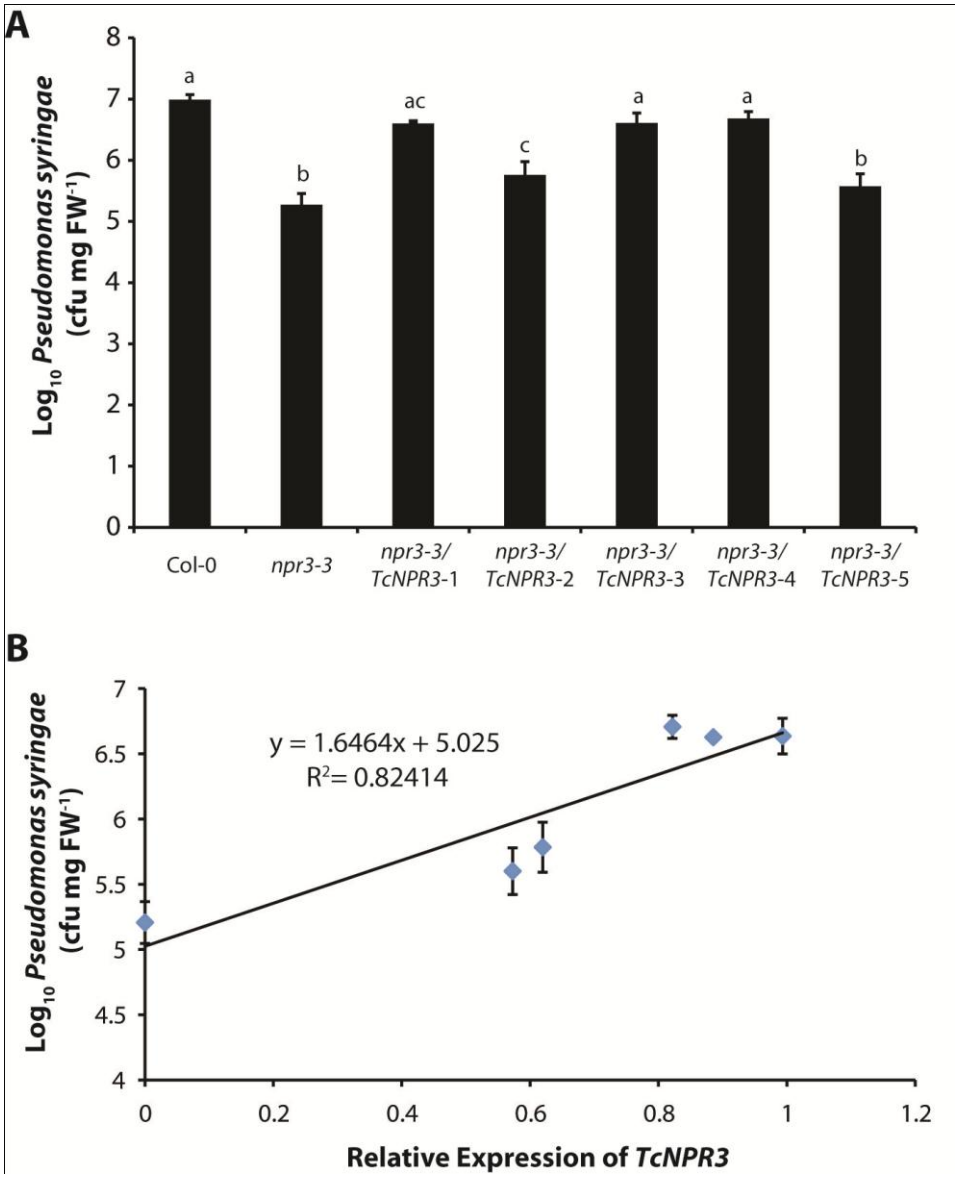


Figure 4-9 Bacterial titration on transgenic *npr3-3* mutant overexpressing *TcNPR3* and its correlation to heterologous *TcNPR3* expression. **A.** Bacterial populations of *P.s.t.* from infected flowers of Col-0, *npr3-3* mutant and five individual transgenic lines (*npr3-3*/*TcNPR3*). The whole inflorescences of six-week old plants (genotypes as indicated on x-axis) were inoculated with virulent *P.s.t.* DC3000 (OD_{600}=0.2 with 0.02% Silwet L-77) and bacterial growth was determined five days after inoculation. Data represent means ± SE of four replicates, each containing two infected inflorescences from two individual plants. Letters above the histogram indicate the statistical differences among different genotypes ($P<0.05$) determined by Fisher's PLSD analysis. cfu, colony forming units. **B.** Correlation

between bacterial growth on infected flower (Figure 4-8A) and relative expression of *TcNPR3* in *npr3-3* mutant and transgenic lines (Figure 4-6). The relative expression of *TcNPR3* from semi-quantative RT-PCR was obtained by comparing *TcNPR3* band intensity to *AtUbiquitin* control intensity.

4.4 Discussion

The complete genome sequence of *Theobroma cacao* allowed us to take advantage of the genomic resources of this tropical tree to identify all the *NPR-like* genes in the genome (*Argout et al., 2010). There are a total of four *NPR-like* genes present in the cacao genome. Besides *TcNPR1* and *TcNPR3*, which were isolated from traditional cDNA-based gene isolation methods, two more genes, TcNPR4 and TcBOP2, were been identified via bioinformatics analysis (Figure 4-2 and Figure 4-3). In phylogenetic analysis, all three clades of NPRs (NPR1/2, NPR3/4 and BOP1/2), contain at least one cacao gene (Figure 4-4), suggesting that ancient duplications occurred before the divergence of *Arabidopsis* and cacao, which is also supported by the genome-wide evolution data (*Argout et al., 2010). This is not surprising considering the importance of NPR1/NPR3 clades in plant defense response (Cao et al., 1997; Dong, 2004; Zhang et al., 2006; Le Henanff et al., 2009) and the BOP clade in vegetative development (Hepworth et al., 2005; Ha et al., 2007; McKim et al., 2008; Jun et al., 2010). Although TcBOP2 and AtBOP2 are grouped together on the tree, it is possible that TcBOP2 is the co-ortholog of AtBOP1 and AtBOP2 because of the low bootstrap value (56) at the node of TcBOP2 and AtBOP2. Similarly, it is also likely that TcNPR3 and TcNPR4 share the same evolutionary relationship with AtNPR3/AtNPR4 since the bootstrap value is only 62 at the node of TcNPR3 and its *Arabidopsis* counterparts. Evidently, phylogenetic analysis indicates that no further duplications of cacao *NPR-like* genes took place in cacao genome after the divergence of *Arabidopsis* and cacao.

We have isolated an NPR3 homolog from the tropical tree, *Theobroma cacao*, and have generated *Arabidopsis npr3-3* mutant transgenic lines over expressing TcNPR3. Our results are consistent with the hypothesis that *TcNPR3* is a functional ortholog of *Arabidopsis NPR3*. TcNPR3 is able to complement the resistance phenotype of *npr3-3*

mutant after floral inoculation. Overexpression of *TcNPR3* in *npr3-3* mutant conferred impaired silique development (Figure 4-7) and increased bacterial growth upon *P.s.t.* treatment (Figure 4-8A), indicating that TcNPR3 acts as a negative regulator in defense response during flower development, a function identical to native *Arabidopsis* NPR3 (Chapter 2). These results suggest that NPR3 function evolved before the divergence of *Arabidopsis* and cacao and it plays a critical role in defense response and reproductive success. Sequence alignment of TcNPR3 and AtNPR3 revealed they are highly similar to each other (60% identity and 77% similarity) and both of them have BTB/POZ domain and ankyrin repeat regions that share even higher similarity (Figure 4-1B). Thus, the mechanism of *TcNPR3* function may also through the interaction with NPR1 and TGA2 transcription factors, as proposed for *Arabidopsis* (Chapter 2). *In vitro* biochemical tests and/or in vivo bimolecular fluorescence complementation assays are in needed to demonstrate the mechanism of TcNPR3. Though NPR3 has not been identified in species other than *Arabidopsis* and cacao, it is reasonable to speculate that this gene is conserved in many plant species to negate defense response in the absence of pathogens and to balance the costs and benefits associated with defense, as we have demonstrated for *Arabidopsis* (Chapter 2).

Pathogenic fungi, including witches' broom, is considered to cause about 30% yield loss on cacao annually (Hebbar, 2007; Ploetz, 2007), and thus the resistance of witches' broom is of great interest to cacao breeders and chocolate industries. Since terminal and axillary buds, flower cushions are highly susceptible (Andebrhan et al., 1993), our results that TcNPR3 acts as a negative regulator of defense response during early flower development in transgenic *Arabidopsis* will provide new molecular marker information to select the varieties with optimal *TcNPR3* level to obtain higher resistance as well as favorable whole plant fitness.

It has been shown that *TcNPR1* is the functional ortholog of *AtNPR1*, because it can partially complement the phenotype of *Arabidopsis npr1-2* mutant (Chapter 3), along with the evidence that *TcNPR3* can restore *npr3-3* mutant phenotype, indicating high degree of evolutionary and functional conservation of *NPR* gene family. It is logical to hypothesize that *TcNPR4* and *TcBOP2* are also functional ortholog of *Arabidopsis* counterparts, because

of the high similarity between cacao and *Arabidopsis* genes, but further investigation is still needed to illustrate their physiological function and signal transduction in defense response and vegetative development in cacao.

4.5 References

***Argout X, *Salse J, *Aury J, *Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Sidibe-Bocs S, Lanaud C** (2010) (*These authors contributed equally to this work.) Deciphering the genome structure and paleohistory of *Theobroma cacao*. Nature Genetics **In Press**

Albagli O, Dhordain P, Deweindt C, Lecocq G, Leprince D (1995) The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. Cell Growth Differ **6:** 1193-1198

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Andebrhan T, Maddison AC, Rudgard SA, International Office of Cocoa Chocolate and Sugar Confectionary. (1993) Disease management in cocoa : comparative epidemiology of witches' broom. Chapman & Hall, London

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M, Kuhn D, Verica J, Courtois B, Loor G, Babin R, Sounigo O, Ducamp M, Guiltinan MJ, Ruiz M, Alemanno L, Machado R, Phillips W, Schnell R, Gilmour M, Rosenquist E, Butler D, Maximova S, Lanaud C (2008) Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of Theobroma cacao L. generated from various tissues and under various conditions. BMC Genomics **9:** -

Bailey BA, Strem MD, Bae H, de Mayolo GA, Guiltinan MJ (2005) Gene expression in leaves of Theobroma cacao in response to mechanical wounding, ethylene, and/or methyl jasmonate. Plant Science **168:** 1247-1258

Bowling SA, Guo A, Cao H, Gordon AS, Klessig DF, Dong X (1994) A mutation in Arabidopsis that leads to constitutive expression of systemic acquired resistance. Plant Cell **6:** 1845-1857

Brechenmacher L, Weidmann S, van Tuinen D, Chatagnier O, Gianinazzi S, Franken P, Gianinazzi-Pearson V (2004) Expression profiling of up-regulated plant and fungal genes in early and late stages of Medicago truncatula-Glomus mosseae interactions. Mycorrhiza **14:** 253-262

Cao H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell **88:** 57-63

Clarke JD, Volko SM, Ledford H, Ausubel FM, Dong X (2000) Roles of salicylic acid, jasmonic acid, and ethylene in cpr-induced resistance in arabidopsis. Plant Cell **12:** 2175-2190

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735-743

Dong X (2004) NPR1, all things considered. Curr Opin Plant Biol **7:** 547-552
Durrant WE, Wang S, Dong X (2007) Arabidopsis SNI1 and RAD51D regulate both gene transcription and DNA recombination during the defense response. Proc Natl Acad Sci U S A **104:** 4223-4227
Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:** 1792-1797
Gesteira AS, Micheli F, Carels N, Da Silva AC, Gramacho KP, Schuster I, Macedo JN, Pereira GAG, Cascardo JCM (2007) Comparative analysis of expressed genes from cacao meristems infected by Moniliophthora perniciosa. Annals of Botany **100:** 129-140
Glazebrook J (2001) Genes controlling expression of defense responses in Arabidopsis--2001 status. Curr Opin Plant Biol **4:** 301-308
Ha CM, Jun JH, Nam HG, Fletcher JC (2004) BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana. Plant Cell Physiol **45:** 1361-1370
Ha CM, Jun JH, Nam HG, Fletcher JC (2007) BLADE-ON-PETIOLE1 and 2 Control Arabidopsis Lateral Organ Fate through Regulation of LOB Domain and Adaxial-Abaxial Polarity Genes. Plant Cell
Ha CM, Kim GT, Kim BC, Jun JH, Soh MS, Ueno Y, Machida Y, Tsukaya H, Nam HG (2003) The BLADE-ON-PETIOLE 1 gene controls leaf pattern formation through the modulation of meristematic activity in Arabidopsis. Development **130:** 161-172
Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of Agrobacterium binary vectors for plant transformation. Plant Mol Biol **25:** 989-994
Hebbar PK (2007) Cacao diseases: a global perspective from an industry point of view. Phytopathology **97:** 1658-1663
Heil M (2002) Ecological costs of induced resistance. Curr Opin Plant Biol **5:** 345-350
Hepworth SR, Zhang Y, McKim S, Li X, Haughn GW (2005) BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. Plant Cell **17:** 1434-1448
Johnson C, Mhatre A, Arias J (2008) NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2. Biochim Biophys Acta **1779:** 583-589
Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in Arabidopsis thaliana. Plant Cell **18:** 3289-3302
Jun JH, Ha CM, Fletcher JC (2010) BLADE-ON-PETIOLE1 Coordinates Organ Determinacy and Axial Polarity in Arabidopsis by Directly Activating ASYMMETRIC LEAVES2. Plant Cell **22:** 62-76
Koornneef A, Pieterse CM (2008) Cross talk in defense signaling. Plant Physiol **146:** 839-844
Le Henanff G, Heitz T, Mestre P, Mutterer J, Walter B, Chong J (2009) Characterization of Vitis vinifera NPR1 homologs involved in the regulation of pathogenesis-related gene expression. BMC Plant Biol **9:** 54
Li J, Mahajan A, Tsai MD (2006) Ankyrin repeat: a unique motif mediating protein-protein interactions. Biochemistry **45:** 15168-15178

Li X, Zhang Y, Clarke JD, Li Y, Dong X (1999) Identification and cloning of a negative regulator of systemic acquired resistance, SNI1, through a screen for suppressors of npr1-1. Cell **98:** 329-339

Lin JJ (1995) Electrotransformation of Agrobacterium. Methods Mol Biol **47:** 171-178

Lopes MA, Hora Junior BT, Dias CV, Santos GC, Gramacho KP, Cascardo JCM, Gesteira AS, Micheli F (2010) Expression analysis of transcription factors from the interaction between cacao and Moniliophthora perniciosa (Tricholomataceae). Genetics and Molecular Research **9:** 1279-1297

Marelli J-P, Maximova S, Gramacho K, Kang S, Guiltinan M (2009) Infection Biology of Moniliophthora perniciosa on Theobroma cacao and Alternate Solanaceous Hosts. Tropical Plant Biology **2:** 149-160

McDowell JM, Cuzick A, Can C, Beynon J, Dangl JL, Holub EB (2000) Downy mildew (Peronospora parasitica) resistance genes in Arabidopsis vary in functional requirements for NDR1, EDS1, NPR1 and salicylic acid accumulation. Plant J **22:** 523-529

McKim SM, Stenvik GE, Butenko MA, Kristiansen W, Cho SK, Hepworth SR, Aalen RB, Haughn GW (2008) The BLADE-ON-PETIOLE genes are essential for abscission zone formation in Arabidopsis. Development **135:** 1537-1546

Meinhardt LW, Rincones J, Bailey BA, Aime MC, Griffith GW, Zhang D, Pereira GA (2008) Moniliophthora perniciosa, the causal agent of witches' broom disease of cacao: what's new from this old foe? Mol Plant Pathol **9:** 577-588

Mitsuhara I, Ugaki M, Hirochika H, Ohshima M, Murakami T, Gotoh Y, Katayose Y, Nakamura S, Honkura R, Nishimiya S, Ueno K, Mochizuki A, Tanimoto H, Tsugawa H, Otsuki Y, Ohashi Y (1996) Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. Plant Cell Physiol **37:** 49-59

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, Garcia O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araujo MR, de Moraes MH, Castro LA, Gramacho KP, Goncalves MS, Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA (2008) A genome survey of Moniliophthora perniciosa gives new insights into Witches' Broom Disease of cacao. BMC Genomics **9:** 548

Nawrath C, Metraux JP (1999) Salicylic acid induction-deficient mutants of Arabidopsis express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen inoculation. Plant Cell **11:** 1393-1404

Norberg M, Holmlund M, Nilsson O (2005) The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. Development **132:** 2203-2213

Ploetz RC (2007) Cacao diseases: important threats to chocolate production worldwide. Phytopathology **97:** 1634-1639

Purdy LH, Schmidt RA (1996) STATUS OF CACAO WITCHES' BROOM: biology, epidemiology, and management. Annu Rev Phytopathol **34:** 573-594

Racki LR, Yorgey P, Dewdney J, Ausubel FM (2003) NPR2: A Novel Arabidopsis Defense Response Gene. Poster, Plant Biology 2003, Honolulu, Hawaii USA

Rochon A, Boyle P, Wignes T, Fobert PR, Despres C (2006) The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. Plant Cell **18:** 3670-3685

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG (2005) Sequence and structural analysis of BTB domain proteins. Genome Biol **6:** R82

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:** 1596-1599

Tiburcio RA, Costa GG, Carazzolle MF, Mondego JM, Schuster SC, Carlson JE, Guiltinan MJ, Bailey BA, Mieczkowski P, Meinhardt LW, Pereira GA (2010) Genes acquired by horizontal transfer are potentially involved in the evolution of phytopathogenicity in Moniliophthora perniciosa and Moniliophthora roreri, two of the major pathogens of cacao. J Mol Evol **70:** 85-97

van Hulten M, Pelser M, van Loon LC, Pieterse CM, Ton J (2006) Costs and benefits of priming for defense in Arabidopsis. Proc Natl Acad Sci U S A **103:** 5602-5607

Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ (2004) Isolation of ESTs from cacao (Theobroma cacao L.) leaves treated with inducers of the defense response. Plant Cell Rep **23:** 404-413

Vlot AC, Klessig DF, Park SW (2008) Systemic acquired resistance: the elusive signal(s). Curr Opin Plant Biol **11:** 436-442

Zhang HZ, Cai XZ (2005) Nonexpressor of pathogenesis-related genes 1 (NPR1): a key node of plant disease resistance signalling network. Sheng Wu Gong Cheng Xue Bao **21:** 511-515

Zhang S, Klessig DF (2001) MAPK cascades in plant defense signaling. Trends Plant Sci **6:** 520-527

Zhang Y, Cheng YT, Qu N, Zhao Q, Bi D, Li X (2006) Negative regulation of defense responses in Arabidopsis by two NPR1 paralogs. Plant J **48:** 647-656

Zhang Y, Tessaro MJ, 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

CHAPTER 5: GENOME WIDE GENE EXPRESSION ANALYSIS OF *THEOBROMA CACAO* IN RESPONSE TO SALICYLIC ACID TREATMENT

5.1 Introduction

Theobroma cacao, the seeds of which are used to make chocolate and cocoa powder, is an important economic crop providing income to small-scale farmers in tropical regions all over the world (Guiltinan et al., 2008). However, about 30%-40% of the world cacao bean production is lost to pathogens due to its very high susceptibility (Hebbar, 2007; Argout et al., 2008). Thus, the study of host-pathogen interactions and characterization of genes involved in pathogen infections could contribute to better understanding of the cacao plant susceptibility and resistance responses, which could potentially be used in cacao breeding programs for selection of cultivars with higher resistance.

The plant hormone salicylic acid (SA) and the SA-dependent signaling pathway play a major role in combating pathogen infection (Shah, 2003; Durrant and Dong, 2004; Loake and Grant, 2007; Vlot et al., 2008). During the last fifteen years many genes in the SA-dependent pathway have been isolated and characterized from the model plant *Arabidopsis*, (Cao et al., 1994; Dong, 2004; Uquillas et al., 2004; Grant and Lamb, 2006; Lee et al., 2007; Loake and Grant, 2007; Attaran et al., 2009). In addition, comprehensive *Arabidopsis* genomics research has provided insights into the changes in gene regulation during SA treatment (Wang et al., 2006; Blanco et al., 2009). To date, research on gene expression and regulation upon pathogen infection in cacao has been focused on only one or limited numbers of genes (Borrone et al., 2004; Bailey et al., 2005; Bailey et al., 2005; Maximova et al., 2006) and represents limited knowledge of the gene expression changes in general and also more specifically of the SA signaling pathway in cacao.

The application of microarray methods enables investigation of the expression of a large number of genes and/or a global gene expression. Microarrays have been used for the analysis of signal transduction and transcriptional regulation of fiber cell development in cotton (Lee et al., 2007), evaluating the global gene expression of phosphorus stressed *Arabidopsis* plants and legumes (Valdes-Lopez and Hernandez, 2008), investigating the

early transcriptional changes associated with bacterial infection in different grape varieties using Combimatrix microarray platform (Polesani et al., 2010), analyzing whole life expression pattern of 84 peptide transporter family members in rice (Zhao et al., 2010), profiling the transcription patterns of 1,536 pollen cDNAs from lily during pollen germination and pollen tube growth (Huang et al., 2006), examining genome-wide early transcriptional responses to elevated temperature and salt stress in *Arabidopsis* (Matsuura et al., 2010), testing differential gene expression among each floral whorls corresponding to ABC model of the genus Aquilegia by EST sequences based microarray (Voelckel et al., 2010), and many others.

Microarrays have also been utilized to study gene expression patterns upon SA treatment in *Arabidopsis* (at the whole genome level) and in other species (Wang et al., 2006; Lee et al., 2009). NPR1 is required for the up/down regulation of 2,248 genes in response to 8h SA analog, BTH treatment (Wang et al., 2006). To expand the knowledge of SA function in defense, the Complete Arabidopsis Transcriptome MicroArray (CATMAv2) chip has been used to study early defense responses to 2.5 h SA treatment in *Arabidopsis* seedlings (Blanco et al., 2009), which revealed that 193 genes are rapidly induced in a SA- and NPR1-dependent manner. In addition, microarray analysis demonstrated that 177 of 7,900 cDNA clones exhibited more than 4-fold transcription accumulation in SA-treated unripe pepper fruit (Lee et al., 2009). In rice, microarray analysis was applied to identify SA-inducible WRKY transcription factors involved in rice blast resistance (Shimono et al., 2007). Taken together, microarray analysis is a powerful tool to study genome-wide signal transduction and transcriptional regulation in SA-induced plants, which could be utilized to further identify new defense-related genes and illustrate the signaling pathways. Thus for a large-scale gene expression analysis of cacao we utilized a Nimblegen custom cacao expression array containing 17,427 unigenes and evaluated differential gene expression in cacao leaves in response to SA treatment and among different leaf developmental stages and between two different cacao genotypes.

5.2 Material and Methods

Plant material and Salicylic acid treatment

Microarray Construction

The microarray was designed based on EST sequences containing 6,853 contigs and 12,959 singlets from mixed cacao tissues ((Argout et al., 2008) and unpublished data), and 2,781 unigenes resulting from 6,572 ESTs generated from several previous cacao EST sequencing projects (Jones, 2002, Verica, 2004, and Pugh, 2004). The arrays were synthesized by Nimblegen (http://www.nimblegen.com/products/exp/index.html) in a 4x72K format representing a total of 17,427 unigenes. 72,000 probes were included on each array, 4 probes for each gene. The probes were designed based on the EST collection as 50-60mers in length.

RNA extraction and hybridization

Three to four month old cacao rooted cutting plants of genotypes SCAVINA 6 and ICS1 were treated with 2mM SA dissolved in water under greenhouse conditions as described in (Swanson et al., 2008). Plants treated with only water served as negative controls. Twenty-four hours after treatment, leaf samples from different developmental stages A, C and E (corresponding to stages YR, IG, MG respectively, as described in (Melnick et al., 2008)) were collected, three biological replicates for each genotype, each treatment and each developmental stage. Thus, 36 samples were collected in total. RNA from each sample was isolated as previously described in (Verica et al., 2004). Microarray hybridizations were performed by the Genomics core facility at PSU using established protocols (http://hils.psu.edu/stf/dnama/hybridization.html). Sample labeling was performed using the Amino Allyl MessageAmp II aRNA Amplification Kit according to the manufacturer's instructions (Applied Biosystems/Ambion - Amino Allyl MessageAmp II aRNA Amplificatio Kit - RNA Amplification for Array Analysis). Hybridization, washing, and scanning were performed according to the manufacturer's instruction (Roche/NimbleGen NimbleGen Arrays User's Guide - Gene Expression Analysis) with the exceptions that one Cy3 and one Cy5 labeled sample were hybridized to each subarray, and following washing, arrays were coated with DyeSaver (Genispher Inc) to inhibit the degradation of fluorescent signal.

Cacao EST microarray

To investigate SA-dependent transcriptome in *Theobroma cacao* L. we utilized custom cacao microarray representing 17,427 unigenes. Two cacao genotypes Scavina6 and ICS1 were treated with either water or 2mM salicylic acid. We collected samples from three different leaf developmental stages, A, C and E separately, since it is known that the defense gene expression varies a lot depending on the developmental stages (Figure 5-1). Thus, we examined 12 samples in total and three biological replicates for each sample. ArrayStar software was performed to normalize and analyze the data. Since we have three variables (genotypes, developmental stages and treatments), differentially expressed genes were only identified from the comparisons with two fixed variables and one different, for example, water-treated ICS1 stage A leaf and SA-treated ICS1 stage A leaf. Therefore, we have 24 comparisons in total. The ICS1 genotype, water treated stage A leaf sample was used as a reference control. Differentially expressed genes were defined as genes differing in expression levels with P-values less than 0.05 and fold changes greater than 1.5.

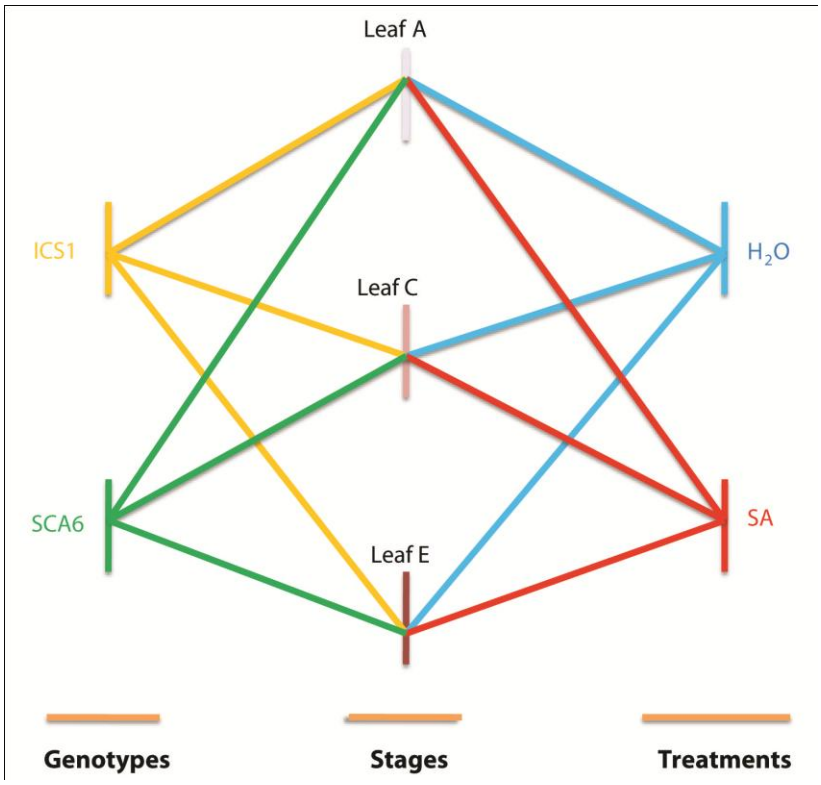


Figure 5-1. An illustration of Cacao EST microarray design. Vertical lines represent three variables in our microarray, genotypes, stages and treatments. Three leaf developmental stages (A, C and E) from two genotypes ICS1 and SCA6 were treated with either water or SA. Samples from yellow/green lines plus blue/red lines were collected.

Microarray data analysis

The data were analyzed with ArrayStar software (DNASTAR Inc, Madison WI). RMA preprocessing method was applied for background correction and quantile method was used for normalization. Student's t-test without multiple testing corrections was performed for statistical analysis with P=0.05 as cutoff for each comparisons. Scatter plots were generated and differentially expressed genes (fold change>1.5 and P<0.05) were identified for each comparison. Statistically significant changed genes were annotated based on the tblastx results against *Arabidopsis* database. Cacao genes with good matches to *Arabidopsis* were further analyzed for their GO functions using TAIR (http://www.arabidopsis.org).

Semi-quantitative RT-PCR validation of microarray results

The microarray analysis results were validated with semi-quantitative RT-PCR of the samples used for the hybridizations. cDNA was synthesized as described in Chapter 3 and semi-quantitative RT-PCR was performed using primers designed for Lite/contig/2252 (RT-5': AGCTTGCAGATTACCGTGGGTACA; RT-3': TGTATATTCCGGAGCTGGTGCCAA), CL160/contig1 (RT-5': CCTGGTCGACAATGGCAACAACAA; RT-3': TTGCAGAGAAGGGCACCAAGTAGT), Lite/contig/3309 (RT-5': GGGAGATCAAAGCGTACTCCC; RT-3': ACATTAATACTTTCTTCGCAACATCG)and Lite/contig/2028 (RT-5': ACCATGCTGCACCATCATCAAAGG; RT-3': TATTCAAGATGAAGGCCAGCTCCC). *TcActin* was used as normalization control as described above. The PCR products were analyzed on 1% agarose gel, stained with ethidium bromide. The expression level of Lite/contig/2252, CL160/contig1and TcActin in all 36 samples were quantified using ImageQuant software (Molecular Dynamics, Amersham Bioscience) as described in (Brechenmacher et al., 2004) and relative expression of two contigs was calculated by comparing with the expression of *TcActin*. The fold changes of two contigs in different samples were calculated by comparing the relative expression values. The fold changes from microarray were plotted against the fold changes in semi-quantitative RT-PCR.

GO annotation analysis

All genes with available *A. thaliana* loci accession numbers were classified according to gene ontology (GO) terms using the tools for GO annotations at The Arabidopsis Information Resource (http://arabidopsis.org/tools/bulk/go/index.jsp). Genes were evaluated for their predicted involvement in biological processes, molecular functions, and cellular localization. An annotation count list of imported genes was retrieved from cellular component category based on their functional GO, and pie charts were generated to present the percentage of each category. GO annotation analyses were performed for significantly up-regulated genes in response to SA treatment of stage C leaves of ICS1 and SCA6. The percentage difference of chloroplast and mitochondrial annotations between whole array and SA-induced genes were calculated and plotted. Chi-square analysis was applied to confirm the statistical significance.

5.3 Results

Overview of microarray results

Twenty-four comparisons were made using the results of the hybridizations of 12 individual samples. Four examples of scatter plots generated using ArrayStar software are presented (Figure 5-2 A to D). The data analysis indicated that that SA treatment induced 107 genes in leaf tissue from stage C leaves of ICS1 and 255 genes in SCA6 and that only 88 and 41 genes were down-regulated in ICS1 and SCA6, respectively (Table1, Figure 5-2 A and B). These results suggested that SA could induce changes in gene expression in cacao. It appears that expression is affected in significantly lower number of cacao genes under our experimental conditions. Much greater numbers of differentially expressed genes were recorded when comparing different developmental stages of cacao leaf tissue (Figure 5-2 C), suggesting that a great deal of gene expression change as developmental stages shift. In addition, we compared the gene expression profile in stage C leaves from ICS1 and SCA6 at both basal and SA-induced conditions (Figure 5-2D). Moderate numbers of differentially expressed gene were found in the comparison of different genotypes.

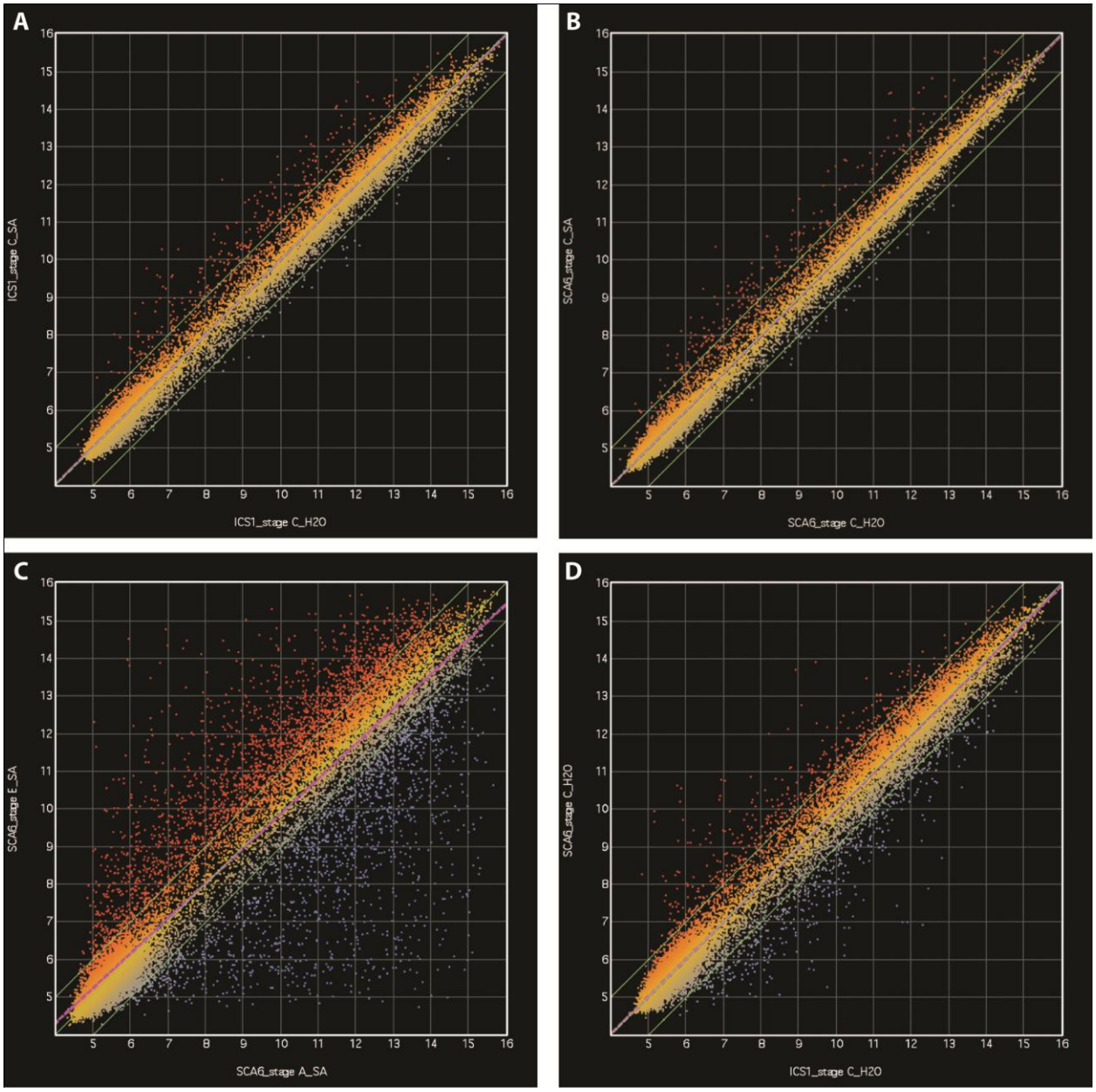


Figure 5-2 Scatter plots of four of selected comparisons generated by ArrayStar software. Three biological replicates for each treatment were grouped together after normalization step and scatter plots of the $\log_2$ of the normalized mean intensities were generated for each comparison. Orange dots represent $\log_2$ of mean intensity values of genes expressed higher in y-axis treatment and blue dots represent $\log_2$ of mean intensity values of genes expressed higher in x-axis treatment. Green lines represent 2-fold change threshold. **A.** Comparison of water- (H_2O) and SA-treated ICS1 stage C leaves. **B.** Comparison of water

(H_2O) - and SA-treated SCA6 stage C leaves. **C.** Comparison of stage A and stage C leaves of SA-treated SCA6. **D.** Comparison of ICS1 and SCA6 stage C water (H_2O)-treated leaves.

The results of all 24 comparisons (Figure 5-1) are represented in Table 1 including the numbers of significantly changed genes in each comparison. The largest number of significantly induced genes in response to SA appeared in different stages according to different genotypes, stage A in ICS1 and stage E in SCA6.

Table 2. Numbers of differentially expressed genes in 24 comparisons.

Comparison	UP	DOWN
ICS1_A_H2O vs. SA	345	154
ICS1_C_H2O vs. SA	107	88
ICS1_E_H2O vs. SA	44	34
SCA6_A_H2O vs. SA	69	46
SCA6_C_H2O vs. SA	255	41
SCA6_E_H2O vs. SA	233	74
ICS1_H2O_A vs. C	1056	1134
ICS1_H2O_A vs. E	2084	2023
ICS1_H2O_C vs. E	1142	1195
ICS1_SA_A vs. C	777	695
ICS1_SA_A vs. E	1763	1824
ICS1_SA_C vs. E	1451	1494
SCA6_H2O_A vs. C	670	719
SCA6_H2O_A vs. E	2246	2170
SCA6_H2O_C vs. E	1813	1700
SCA6_SA_A vs. C	699	870
SCA6_SA_A vs. E	2380	2224
SCA6_SA_C vs. E	1951	1635
A_H2O_ICS1 vs. SCA6	348	509
C_H2O_ICS1 vs. SCA6	556	617
E_H2O_ICS1 vs. SCA6	1585	1671
A_SA_ICS1 vs. SCA6	485	526
C_SA_ICS1 vs. SCA6	496	538
E_SA_ICS1 vs. SCA6	1679	1620

The comparisons between different leaf developmental stages manifested the largest number of differentially expressed genes. The comparisons of different genotypes displayed a moderate change in younger stages (A and C, 348 genes), but about three

thousand genes showed differential expression at stage E, indicating gene expression profiles of ICS1 and SCA6 varies substantially in the oldest developmental stage.

To identify potential overlap of SA-induced genes among different leaf developmental stages, we generated Venn diagrams. The results indicated that only several up and down regulated genes concurred in stage A and C in response to SA in genotype ICS1 and no genes were identified to be differentially regulated in all developmental stages (Figure 5-4 A). For genotype SCA6, nine genes were identified to be up-regulated in all developmental stages by SA, from which five genes were annotated with *Arabidopsis* locus identifier. We discovered two putative PHD finger transcription factor in nine overlapping genes, which is shown to be one of the transcription factor with major regulatory effects under changing environmental conditions in *Arabidopsis* (Carrera et al., 2009), indicating they might induce defensive gene transcription upon exogenous SA treatment. Only 3 down-regulated genes were common between stages A and E (1 gene) and stages C and E (2 genes). The fact that only few genes overlapped among the different developmental stages suggests that SA-induced defense response in cacao leaf tissue could dependent on the leaf developmental stage.

Table 3 Nine genes up regulated in all leaf developmental stages of SCA6 in response to SA treatment.

SEQ_ID	At Locus Identifier	E-value	Annotation
KAT11YG11FM1_putative	AT3G07160	1.00E-27	ATGSL10 (GLUCAN SYNTHASE-LIKE 10); 1,3-beta-glucan synthase
KCAF1YJ05FM1_Populus	AT4G14920	8.00E-06	PHD finger transcription factor, putative
KAT5YE13FM1_hypothetical	AT5G12400	7.00E-17	PHD finger transcription factor, putative
KAT3YF02FM1_hypothetical	AT5G24710	4.00E-08	similar to protein kinase family protein [Arabidopsis thaliana] (TAIR:AT5G38560.1); similar to Os01g0653800 [Oryza sativa (japonica cultivar-group)] (GB:NP_001043744.1); similar to WD-40 repeat family protein [Medicago sativa] (GB:AAZ31064.1); similar to Peptidase S8 and S53, subtilisin, kexin, sedolisin; WD40-like [Medicago truncatula] (GB:ABD32844.1); contains InterPro domain WD-40 repeat; (InterPro:IPR001680); contains InterPro domain WD40-like; (InterPro:IPR011046)
KAT10YL12FM1_gi_56806637_dbj_BAD83538_1_	ATMG00290	9.00E-81	encodes a mitochondrial ribosomal protein S4, a constituent of the small subunit of the ribosomal complex
KAQ3YN11FM1_	N/A	N/A	N/A

LKBE11YG18FM1_Vitis	N/A	N/A	N/A
KAV7YA04FM1_	N/A	N/A	N/A
LITE_CONTIG_3567	N/A	N/A	N/A

To identify genes that are differentially expressed in both genotypes ICS1 and SCA6 at basal level and after treatment with SA we also generated Venn diagram that illustrates the number of subset genes (Figure5-4 B). A number of differentially expressed genes were identified to concur in all three stages in four different conditions. Seventy genes were consistently up regulated and 72 down regulated in all developmental stages of ICS1 compared to SCA6, perhaps representing major genotypic differences at basal level. Similarly there were 81 up regulated and 59 down regulated genes that were identified to be differentially expressed in both genotypes, and 32 genes were found constitutively significantly higher in SCA6 in both basal and induced conditions (Table 3), some of which are shown to be associated with bacterial infection (Jones et al., 2006). Further functional analysis of these differentially expressed genes could contribute to characterization of the genotypic differences in SA response and potentially point to the mechanisms of disease resistance exhibited by SCA6.

Table 4 Genes that are significantly higher in all developmental stages of SCA6 compared to ICS1 in both basal and induced state.

SEQ_ID	At Locus Identifier	E-value	Annotation
LITE_CONTIG_6154	AT3G27890	2E-37	NQR (NADPH:QUINONE OXIDOREDUCTASE); FMN reductase
LITE_CONTIG_4240	AT5G67050	1E-11	lipase class 3 family protein
LITE_CONTIG_6216	AT5G55150	4E-10	F-box family protein
LITE_CONTIG_6331	AT5G54160	1E-09	ATOMT1 (O-METHYLTRANSFERASE 1)
LITE_CONTIG_1242	AT5G20620	8E-22	UBQ4 (ubiquitin 4); protein binding
LITE_CONTIG_1207	AT5G19150	1E-53	carbohydrate kinase family
KCAA6YL08FM1_hypothetical	AT5G07010	3E-14	sulfotransferase family protein
KCAA2YO18FM1_gi_125600653_gb_EAZ40229_1_	AT4G27300	5E-07	S-locus protein kinase, putative
1902_68416_m02826	AT3G22400	5E-15	lipoxygenase, putative

LITE_CONTIG_919	AT3G22320	1E-40	ATRPABC24.3 ("Arabidopsis thaliana RNA polymerase I, II and III 24.3 kDa subunit"); DNA binding / DNA-directed RNA polymerase
2452_68416_m01989	AT3G15690	3E-07	biotin carboxyl carrier protein of acetyl-CoA carboxylase-related
KAC10YP16FM1_hypothetical	AT3G01430	3E-09	similar to NHL repeat-containing protein [Arabidopsis thaliana] (TAIR:AT5G14890.1); similar to Os02g0743200 [Oryza sativa (japonica cultivar-group)] (GB:NP_001048094.1); similar to NHL repeat-containing protein-like [Oryza sativa (japonica cultivar-group)] (GB:BAD38163.1)
KBF10YF07FM1_Vitis	AT2G22370	1E-08	similar to Os02g0193900 [Oryza sativa (japonica cultivar-group)] (GB:NP_001046173.1); similar to mediator of RNA polymerase II transcription, subunit 18 homolog [Xenopus tropicalis] (GB:NP_001017086.1); contains domain Thiamin pyrophosphokinase, catalytic domain (SSF63999)
LITE_CONTIG_6151	AT1G78630	1E-30	EMB1473 (EMBRYO DEFECTIVE 1473); structural constituent of ribosome
KAV6YP12FM1_hypothetical	AT1G64350	4E-10	SEH1H (SEH1H); nucleotide binding
LITE_CONTIG_1007	AT1G07650	9E-06	leucine-rich repeat transmembrane protein kinase, putative
LKBE3YF17FM1_	N/A	N/A	N/A
LKBE5YP04FM1_	N/A	N/A	N/A
KCL4YH18FM1_	N/A	N/A	N/A
KCAK4YG22FM1_hypothetical	N/A	N/A	N/A
LKBE1YL12RM1_	N/A	N/A	N/A
KCAE1YG04FM1_	N/A	N/A	N/A
LITE_CONTIG_4627	N/A	N/A	N/A
KAQ7YM02FM1_	N/A	N/A	N/A
KAV5YN18FM1_	N/A	N/A	N/A
KAQ5YF07FM1_gi_125576048_gb_EAZ17270_1_	N/A	N/A	N/A
KAA14YD07FM1_	N/A	N/A	N/A
KCAK2YG13FM1_	N/A	N/A	N/A
LITE_CONTIG_4843	N/A	N/A	N/A
KAV6YE17FM1_gi_121916639_gb_EAY21418_1_	N/A	N/A	N/A
KAA5YG01FM1_	N/A	N/A	N/A
KAA9YG18FM1_	N/A	N/A	N/A

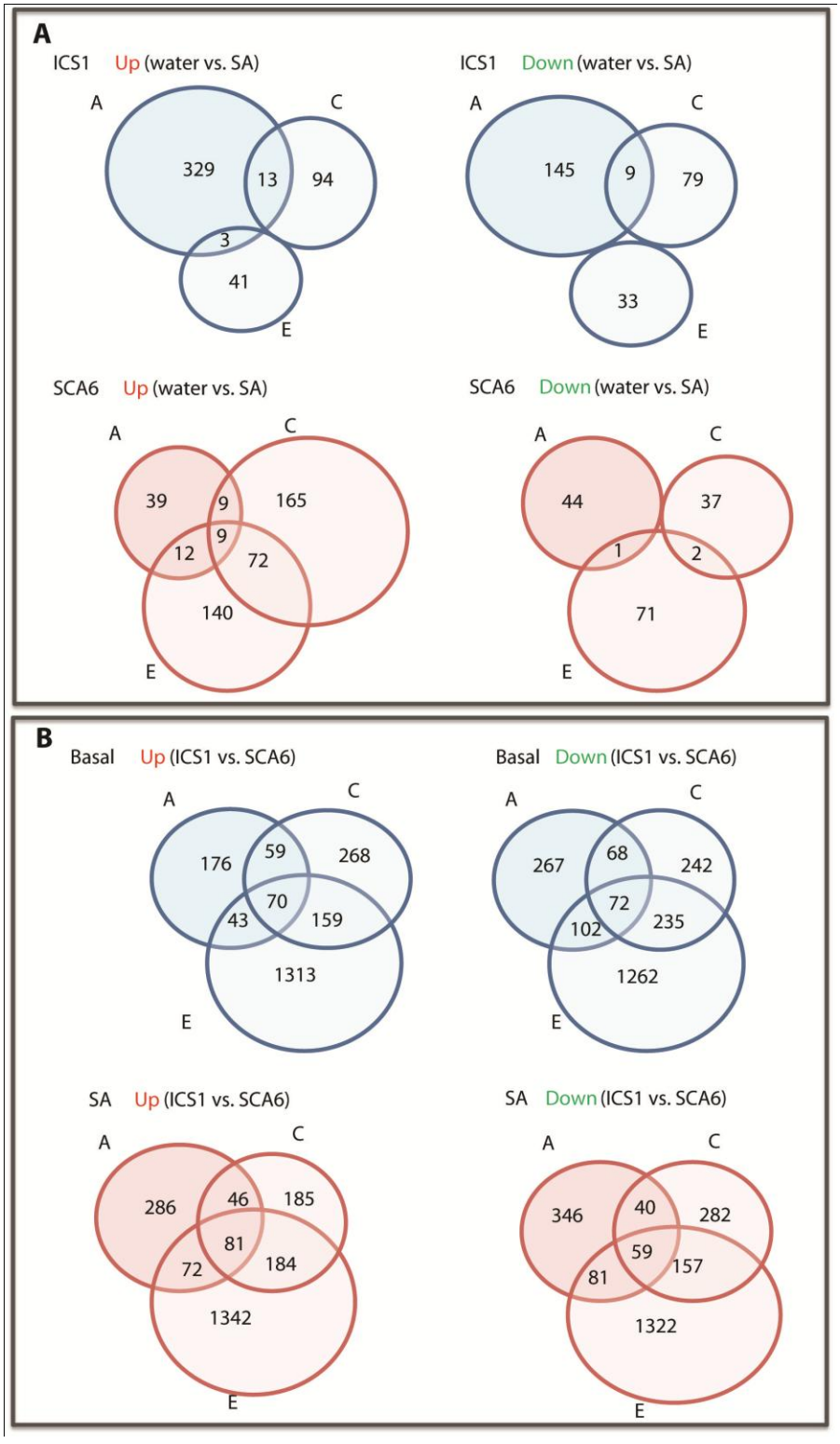


Figure 5-3. Venn diagrams of genes differentially expressed in different developmental stages and their intersections. **A.** The up- and down-regulated genes in the comparison of

water and SA-treated samples in both SCA6 and ICS1. **B.** The up- and down-regulated genes in the comparison of ICS1 and SCA6 samples at both basal and SA-induced conditions (UP: expression higher in SCA6; DOWN: expression higher in ICS1).

Comparison between SA-treated Scavina6 and ICS1 developmental stage C

It is well known that SCA6 is more tolerant than ICS1 to pathogen infection (Faleiro et al., 2006) and thus SCA6 has been used as a parent for different breeding programs. To investigate the molecular basis of this difference and to identify the gene expression changes in both genotypes in response to SA treatment, we further analyzed the microarray results comparing SA-treated ICS1 and SCA6 at developmental stage C. The analysis revealed that in ICS1 there were 107 genes up-regulated (Table 4) and 88 genes down-regulated in SA treatment compared to water-treated control, from which 76 up-regulated genes and 82 down-regulated genes were annotated and had *A. thaliana* loci accession numbers. The SCA6 analysis identified 255 genes up-regulated (Table 5) and 39 genes down-regulated in SA-treated SCA6 compared to water-treated controls, from these 175 and 30 genes respectively were annotated and had *A. thaliana* loci accession numbers.

Table 5 Transcripts elevated in SA-treated stage C leaf of ICS1 compared to water-treated samples.

SEQ_ID	Fold Induction	At Locus Identifier	Annotations
Lite_Contig_2028	up 8.088	AT1G75240	ATHB33 (ARABIDOPSIS THALIANA HOMEOBOX PROTEIN 33); DNA binding / transcription factor
Lite_Contig_2252	up 7.421	AT1G10070	ATBCAT-2; branched-chain-amino-acid transaminase/ catalytic
Lite_Contig_6344	up 6.118	AT3G12500	ATHCHIB (BASIC CHITINASE); chitinase
Lite_Contig_1168	up 5.853	N/A	N/A
KZ0ACAK3YI03FM1_	up 5.643	N/A	N/A
KZ0AAC7YC01FM1_gi_149944371_gb_ABR46228.1_	up 5.563	AT2G42110	similar to Os03g0128300 [Oryza sativa (japonica cultivar-group)] (GB:NP_001048839.1); contains domain SOX9B (PTHR10270:SF18); contains domain SOX TRANSCRIPTION FACTORS (PTHR10270)
KZ0ACL4YK11FM1_hypothetical	up 5.503	AT1G17020	SRG1 (SENESCENCE-RELATED GENE 1); oxidoreductase, acting on paired donors, with incorporation or reduction of molecular oxygen, 2-oxoglutarate as one donor, and incorporation of one atom each of oxygen into both donors
Lite_Contig_3482	up 5.386	AT1G50490	UBC20 (ubiquitin-conjugating enzyme 20); ubiquitin-protein ligase
Lite_Contig_686	up 5.380	AT4G25150	acid phosphatase, putative

Lite_Contig_6200	up 5.081	AT3G08770	LTP6 (Lipid transfer protein 6); lipid binding
Lite_Contig_6130	up 4.822	N/A	N/A
Lite_Contig_920	up 4.542	N/A	N/A
KZ0ACAE5YK13FM1_hypothetical	up 3.658	AT1G14600	DNA binding / transcription factor
KZ0AAV2YN20FM1_diphenol	up 3.656	AT5G03260	LAC11 (laccase 11); copper ion binding / oxidoreductase
KZ0AAC4YD19FM1_hypothetical	up 3.578	AT2G46410	CPC (CAPRICE); DNA binding / transcription factor
Lite_Contig_2515	up 3.556	AT5G40990	GLIP1 (GDSL LIPASE1); carboxylic ester hydrolase
KZ0AAP24YD22FM1_diphenol	up 3.541	AT5G09360	LAC14 (laccase 14); copper ion binding / oxidoreductase
Lite_Contig_938	up 3.500	AT5G48485	DIR1 (DEFECTIVE IN INDUCED RESISTANCE 1); lipid binding
KZ0ACAE2YM20FM1_gi_110743057_dbj_BAE99421.1_	up 3.416	AT3G04070	ANAC047 (Arabidopsis NAC domain containing protein 47); transcription factor
KZ0ACAF4YF18FM1_	up 3.347	N/A	N/A
KZ0AAC5YI07FM1_	up 3.195	N/A	N/A
Lite_Contig_4403	up 3.064	AT3G12870	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT5G56120.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:AAK09219.1); similar to Os03g0778400 [Oryza sativa (japonica cultivar-group)] (GB:NP_001051445.1)
KZ0AAC10YC12FM1_	up 2.949	N/A	N/A
CL404Contig1_	up 2.884	N/A	N/A
Lite_Contig_3863	up 2.811	AT1G20930	CDKB2;2 (CYCLIN-DEPENDENT KINASE B2;2); kinase
KZ0AAS3YP22FM1_jasmonic	up 2.772	AT1G19640	JMT (JASMONIC ACID CARBOXYL METHYLTRANSFERASE); jasmonate O-methyltransferase
Lite_Contig_5010	up 2.739	AT3G13960	AtGRF5 (GROWTH-REGULATING FACTOR 5)
KZ0ABB11YM19FM1_	up 2.663	N/A	N/A
KZ0AAT5YJ08FM1_gi_147816494_emb_CAN77349.1_	up 2.612	AT4G35100	PIP3 (PLASMA MEMBRANE INTRINSIC PROTEIN 3); water channel
Lite_Contig_1656	up 2.542	AT5G42330	unknown protein
Lite_Contig_4775	up 2.506	AT1G01370	HTR12 (Centromeric histone H3); DNA binding
Lite_Contig_2956	up 2.418	AT4G14130	XTR7 (XYLOGLUCAN ENDOTRANSGLYCOSYLASE 7); hydrolase, acting on glycosyl bonds
KZ0ABA6YK18FM1_CXE	up 2.385	AT1G47480	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G48700.1); similar to CXE carboxylesterase [Malus pumila] (GB:ABB89003.1); contains InterPro domain Esterase/lipase/thioesterase; (InterPro:IPR000379); contains InterPro domain Alpha/beta hydrolase fold-3; (InterPro:IPR013094)
KZ0ACAA5YD24FM1_gi_30023684_gb_AAP13375.1_	up 2.361	AT3G20600	NDR1 (NON RACE-SPECIFIC DISEASE RESISTANCE 1); signal transducer
Lite_Contig_3349	up 2.356	N/A	N/A
Lite_Contig_823	up 2.343	AT1G71695	peroxidase 12 (PER12) (P12) (PRXR6)

Lite_Contig_1193	up 2.339	AT5G48810	ATB5-B (Cytochrome b5 B)
Lite_Contig_2990	up 2.336	AT2G30550	lipase class 3 family protein
KZ0ACAE8YF05FM1_gi_94807644_gb_ABF47119.1_	up 2.332	AT1G31335	similar to Os03g0780200 [Oryza sativa (japonica cultivar-group)] (GB:NP_001051448.1)
KZ0AAQ5YE10FM1_	up 2.296	N/A	N/A
KZ0AAC12YD06FM1_	up 2.278	N/A	N/A
KZ0ABB10YN20FM1_	up 2.229	N/A	N/A
Lite_Contig_6380	up 2.224	AT2G22420	peroxidase 17 (PER17) (P17)
Lite_Contig_4285	up 2.189	AT1G78850	curculin-like (mannose-binding) lectin family protein
Lite_Contig_3899	up 2.163	AT3G07390	AIR12 (Auxin-Induced in Root cultures 12); extracellular matrix structural constituent
Lite_Contig_6147	up 2.049	AT3G46030	histone H2B, putative
KZ0AAV4YL15FM1_	up 2.022	N/A	N/A
Lite_Contig_1482	up 2.009	AT2G40435	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G56220.1); similar to Os03g0338400 [Oryza sativa (japonica cultivar-group)] (GB:NP_001050055.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:BAC10078.1); similar to Os05g0337200 [Oryza sativa (japonica cultivar-group)] (GB:NP_001055237.1); contains domain Regulatory domain in the aminoacid metabolism (SSF55021)
KZ0AAV11YM21FM1_	up 1.995	N/A	N/A
KZ0AAQ3YP13FM1_Vitis	up 1.953	N/A	N/A
KZ0AAC5YG21FM1_hypothetical	up 1.929	AT5G64620	C/VIF2 (CELL WALL / VACUOLAR INHIBITOR OF FRUCTOSIDASE 2); pectinesterase inhibitor
Lite_Contig_3635	up 1.924	N/A	N/A
Lite_Contig_3871	up 1.890	AT5G24105	AGP41
Lite_Contig_790	up 1.875	AT1G09815	DNA polymerase delta subunit 4 family
Lite_Contig_2978	up 1.870	AT3G55610	P5CS2 (DELTA 1-PYRROLINE-5-CARBOXYLATE SYNTHASE 2); catalytic/ glutamate 5-kinase/ glutamate-5-semialdehyde dehydrogenase
Lite_Contig_2491	up 1.843	N/A	N/A
Lite_Contig_4206	up 1.842	AT1G54690	histone H2A, putative
CL454Contig1_salicylic	up 1.824	AT1G17200	integral membrane family protein
Lite_Contig_3687	up 1.820	AT1G09310	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G56580.1); similar to Os12g0563600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001067038.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:BAC10818.1); similar to Os11g0683600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001068462.1); contains InterPro domain Protein of unknown function DUF538; (InterPro:IPR007493)
KZ0ABA4YO23FM1_	up 1.806	N/A	N/A
Lite_Contig_2161	up 1.746	AT1G78815	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G16910.1); similar to Protein of unknown function DUF640 [Medicago truncatula]

			(GB:ABE92798.1); contains InterPro domain Protein of unknown function DUF640; (InterPro:IPR006936)
Lite_Contig_4107	up 1.709	AT3G06680	60S ribosomal protein L29 (RPL29B)
KZ0AAT3YJ18FM1_histone	up 1.705	AT4G40030	histone H3.2
2499_68418.m06418	up 1.702	AT5G51760	protein phosphatase 2C, putative / PP2C, putative
Lite_Contig_130	up 1.692	N/A	N/A
KZ0AAS3YE04FM1_	up 1.687	N/A	N/A
KZ0AAS2YO07FM1_Engulfment	up 1.663	AT2G44770	phagocytosis and cell motility protein ELMO1-related
KZ0ACAA2YN06FM1_hypothetical	up 1.654	AT5G12080	mechanosensitive ion channel domain-containing protein / MS ion channel domain-containing protein
Lite_Contig_1964	up 1.652	AT5G42890	sterol carrier protein 2 (SCP-2) family protein
KZ0AAA1YH10FM1_	up 1.648	N/A	N/A
Lite_Contig_3905	up 1.648	N/A	N/A
KZ0ACAA3YO03FM1_gi_20196922_gb_AAC31837.2_	up 1.647	AT2G44870	similar to unknown protein [Oryza sativa (japonica cultivar-group)] (GB:AAL86463.1); contains InterPro domain Sigma factor, regions 3 and 4; (InterPro:IPR013324)
Lite_Contig_3578	up 1.625	AT4G21750	ATML1 (MERISTEM LAYER 1); DNA binding / transcription factor
Lite_Contig_1800	up 1.625	AT1G08510	FATB (FATTY ACYL-ACP THIOESTERASES B); acyl carrier/ acyl-ACP thioesterase
KZ0AAV3YA12FM1_Plant	up 1.619	AT5G22560	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT5G22550.2); similar to Os04g0505400 [Oryza sativa (japonica cultivar-group)] (GB:NP_001053253.1); similar to Plant protein of unknown function [Medicago truncatula] (GB:ABE89493.1); similar to Plant protein of unknown function [Medicago truncatula] (GB:ABE78995.1); contains InterPro domain Protein of unknown function DUF247, plant; (InterPro:IPR004158)
KZ0ABF3YO16FM1_hypothetical	up 1.619	AT2G19090	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT4G30130.1); similar to Protein of unknown function DUF630 [Medicago truncatula] (GB:ABE84030.1); contains InterPro domain Protein of unknown function DUF632; (InterPro:IPR006867); contains InterPro domain Protein of unknown function DUF630; (InterPro:IPR006868)
Lite_Contig_6761	up 1.612	AT1G08450	CRT3 (CALRETICULIN 3); calcium ion binding
KZ0AAT4YE11FM1_gi_2979554_gb_AAC06163.1_	up 1.609	AT2G45540	WD-40 repeat family protein / beige-related
Lite_Contig_2700	up 1.595	AT3G01130	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT5G15320.1); similar to conserved hypothetical protein [Medicago truncatula] (GB:ABE79949.1)
Lite_Contig_6246	up 1.592	AT4G14305	similar to PMP22 (peroxisomal membrane protein 22) [Arabidopsis thaliana] (TAIR:AT4G04470.1); similar to Os08g0566900 [Oryza sativa (japonica cultivar-group)] (GB:NP_001062544.1); similar to Os02g0226000 [Oryza sativa (japonica cultivar-group)] (GB:NP_001046343.1); similar to hypothetical protein DDBDRAFT_0188787 [Dictyostelium discoideum AX4] (GB:XP_635840.1); contains InterPro domain Mpv17/PMP22; (InterPro:IPR007248)

1726_	up 1.589	N/A	N/A
Lite_Contig_5901	up 1.584	AT1G16740	ribosomal protein L20 family protein
Lite_Contig_976	up 1.580	AT5G48485	DIR1 (DEFECTIVE IN INDUCED RESISTANCE 1); lipid binding
Lite_Contig_2577	up 1.580	N/A	N/A
KZ0ACAE4YG10FM1_hypothetical	up 1.579	AT1G72510	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT2G09970.1); similar to 4Fe-4S ferredoxin, iron-sulfur binding [Medicago truncatula] (GB:ABE81755.1); contains InterPro domain Protein of unknown function DUF1677, plant; (InterPro:IPR012876)
KZ0ABF6YH03FM1_	up 1.571	N/A	N/A
Lite_Contig_6222	up 1.570	AT4G14100	transferase, transferring glycosyl groups
KZ0ACAK5YE11FM1_gi_23296881_gb_AAN13194.1_	up 1.568	AT3G07090	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT5G25170.1); similar to expressed protein [Oryza sativa (japonica cultivar-group)] (GB:ABB47783.1); similar to unknown protein [Oryza sativa] (GB:AAG16855.1); similar to Os02g0814000 [Oryza sativa (japonica cultivar-group)] (GB:NP_001048488.1); contains InterPro domain Protein of unknown function DUF862, eukaryotic; (InterPro:IPR008580)
KZ0AAV5YG05FM1_gi_21689839_gb_AAM67563.1_	up 1.565	AT5G50650	WD-40 repeat family protein / St12p protein, putative
Lite_Contig_3649	up 1.563	AT5G59613	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G46430.1); similar to mitochondrial ATP synthase 6 KD subunit [Oryza sativa (japonica cultivar-group)] (GB:BAB21526.1)
Lite_Contig_2228	up 1.556	AT3G51030	ATTRX1 (Arabidopsis thaliana thioredoxin H-type 1); thiol-disulfide exchange intermediate
KZ0AAS1YB18_gi_113549197_dbj_BAF12640.1_	up 1.549	AT3G59420	ACR4 (ARABIDOPSIS CRINKLY4); kinase
KZ0AAS9YK15FM1_Transferase	up 1.547	AT4G24510	CER2 (ECERIFERUM 2); transferase
Lite_Contig_4146	up 1.546	AT1G58420	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G10140.1); similar to Os07g0123800 [Oryza sativa (japonica cultivar-group)] (GB:NP_001058800.1)
Lite_Contig_2817	up 1.546	N/A	N/A
KZ0AAT11YC07FM1_gi_90962932_gb_ABE02390.1_	up 1.543	AT1G73177	similar to conserved hypothetical protein [Medicago truncatula] (GB:ABD32847.1)
Lite_Contig_1120	up 1.543	AT1G20693	HMGB2 (HIGH MOBILITY GROUP B 2); transcription factor
Lite_Contig_1716	up 1.539	AT2G47730	ATGSTF8 (GLUTATHIONE S-TRANSFERASE 8); glutathione transferase
KZ0AAT3YK18FM1_gi_7269354_emb_CAB79413.1_	up 1.534	AT4G25040	integral membrane family protein
KZ0AAC10YD05FM1_	up 1.527	N/A	N/A
KZ0AAC4YC07FM1_cytochrome	up 1.517	AT1G26340	B5 #6 (cytochrome b5 family protein #6); heme binding / transition metal ion binding
Lite_Contig_4266	up 1.517	N/A	N/A
Lite_Contig_1161	up 1.513	AT1G14980	CPN10 (CHAPERONIN 10)
KZ0AAC11YD18FM1_Vitis	up 1.509	N/A	N/A

Lite_Contig_5161	up 1.508	N/A	N/A
Lite_Contig_5135	up 1.501	AT3G16080	60S ribosomal protein L37 (RPL37C)
Lite_Contig_6718	up 1.500	AT2G31880	leucine-rich repeat transmembrane protein kinase, putative

Table 6 Transcripts elevated in SA-treated stage C leaf of SCA6 compared to water-treated samples.

SEQ_ID	Fold Induction	At Locus Identifier	Annotations
LITE_CONTIG_2451	up 8.254	ATCG00680	encodes for CP47, subunit of the photosystem II reaction center.
KAA13YF15FM1_photosystem	up 7.961	ATCG00680	encodes for CP47, subunit of the photosystem II reaction center.
KAC3YC17FM1_gi_119224918_dbj_BAF41304_1_	up 6.994	ATCG01120	encodes a chloroplast ribosomal protein S15, a constituent of the small subunit of the ribosomal complex
LITE_CONTIG_2939	up 6.020	ATCG01240	30S chloroplast ribosomal protein S7
KAP12YH14FM1_NADH_ubiquinone	up 5.744	ATMG00285	encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from two precursors, NAD2A and NAD2B.
LITE_CONTIG_4986	up 5.549	ATCG01090	Encodes subunit of the chloroplast NAD(P)H dehydrogenase complex
LITE_CONTIG_5921	up 5.288	ATMG00560	encodes a mitochondrial ribosomal protein L2, a constituent of the large subunit of the ribosomal complex
LITE_CONTIG_6838	up 5.045	N/A	N/A
KBF4YD24FM1_HEAT	up 4.891	AT2G35630	MOR1 (MICROTUBULE ORGANIZATION 1)
LITE_CONTIG_2578	up 4.876	ATMG01320	encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from two precursors, NAD2A and NAD2B.
LITE_CONTIG_5208	up 4.852	ATCG00280	chloroplast gene encoding a CP43 subunit of the photosystem II reaction center. promoter contains a blue-light responsive element.
LITE_CONTIG_998	up 4.777	ATCG00040	Encodes a maturase located in the trnK intron in the chloroplast genome.
LITE_CONTIG_1235	up 4.480	ATCG00340	Encodes the D1 subunit of photosystem I and II reaction centers.
KAA8YD19FM1_	up 4.074	ATCG00170	RNA polymerase beta' subunit-2
LITE_CONTIG_2715	up 3.951	ATCG00190	Chloroplast DNA-dependent RNA polymerase B subunit. The transcription of this gene is regulated by a nuclear encoded RNA polymerase. This gene has been transferred to mitochondrial genome during crucifer evolution.
KBB14YA22FM1_hypothetical	up 3.925	AT1G04300	similar to meprin and TRAF homology domain-containing protein / MATH domain-containing protein [Arabidopsis thaliana] (TAIR:AT5G43560.2); similar to MATH domain, putative [Medicago truncatula] (GB:ABE87963.1); similar to Os12g0597100 [Oryza sativa (japonica cultivar-group)] (GB:NP_001067191.1); contains InterPro domain TRAF-like; (InterPro:IPR008974); contains InterPro domain MATH; (InterPro:IPR002083); contains InterPro domain TRAF-type; (InterPro:IPR013322)
LCL395Contig1_	up 3.924	N/A	N/A

LITE_CONTIG_2086	up 3.903	N/A	N/A
KAP10YP17FM1_gi_85687416_gb_ABC73628_1_	up 3.874	ATCG00350	Encodes psaA protein comprising the reaction center for photosystem I along with psaB protein; hydrophobic protein encoded by the chloroplast genome.
LITE_CONTIG_2510	up 3.720	ATMG00640	encodes a plant b subunit of mitochondrial ATP synthase based on structural similarity and the presence in the F(0) complex.
LITE_CONTIG_1962	up 3.699	AT2G38800	calmodulin-binding protein-related
KAS9YI17FM1_Gossypium	up 3.684	N/A	N/A
LITE_CONTIG_322	up 3.660	N/A	N/A
KCAA7YF10FM1_gi_125587847_gb_EAZ28511_1_	up 3.657	AT2G39810	HOS1 (High expression of osmotically responsive genes 1)
LITE_CONTIG_5476	up 3.654	ATCG00540	Encodes cytochrome f apoprotein; involved in photosynthetic electron transport chain; encoded by the chloroplast genome and is transcriptionally repressed by a nuclear gene HCF2.
LITE_CONTIG_6335	up 3.605	ATMG00580	NADH dehydrogenase subunit 4
LITE_CONTIG_4190	up 3.589	ATCG00490	large subunit of RUBISCO.
LKBE11YG18FM1_Vitis	up 3.550	N/A	N/A
LKBE7YP02FM1_	up 3.539	N/A	N/A
LCL399Contig1_	up 3.490	N/A	N/A
KAA9YK04FM1_gi_148491416_dbj_BAA99307_2_	up 3.476	ATMG00650	Encodes NADH dehydrogenase subunit 4L.
LITE_CONTIG_2432	up 3.453	ATMG00580	NADH dehydrogenase subunit 4
KAC1YI19RM1_	up 3.413	N/A	N/A
LITE_CONTIG_1504	up 3.385	AT4G20850	TPP2 (TRIPEPTIDYL PEPTIDASE II); subtilase
CL333Contig1_	up 3.383	N/A	N/A
KAP2YK05FM1_	up 3.377	N/A	N/A
KAA6YF21FM1_	up 3.330	N/A	N/A
KCAF7YG10FM1_	up 3.305	N/A	N/A
LKBE8YM20FM1_gi_80750940_dbj_BAE48016_1_	up 3.304	ATCG00550	PSII component
KAT5YB10FM1_	up 3.300	AT1G68650	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G25520.1); similar to Os08g0433100 [Oryza sativa (japonica cultivar-group)] (GB:NP_001061871.1); similar to PREDICTED: similar to conserved hypothetical protein [Strongylocentrotus purpuratus] (GB:XP_798718.2); contains InterPro domain Protein of unknown function UPF0016; (InterPro:IPR001727)
CL189Contig1_Solanum	up 3.282	N/A	N/A
LITE_CONTIG_5613	up 3.254	N/A	N/A
KAV4YA08FM1_	up 3.254	N/A	N/A

LITE_CONTIG_5346	up 3.190	AT1G22610	C2 domain-containing protein
KAQ5YJ19FM1_	up 3.175	N/A	N/A
KCAE8YH06FM1_	up 3.172	N/A	N/A
KAC5YK09FM1_	up 3.091	N/A	N/A
LKBE4YH22FM1_	up 3.054	N/A	N/A
KAA9YN18FM1_hypothetical	up 3.053	N/A	N/A
LITE_CONTIG_3770	up 3.047	ATCG00720	Encodes the cytochrome b(6) subunit of the cytochrome b6f complex.
LITE_CONTIG_2116	up 3.043	ATMG01275	Encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from three precursors, NAD1A, NAD1B, and NAD1C.
KAT4YI07FM1_gi_37591106_dbj_BAC98908_1_	up 3.036	ATMG00580	NADH dehydrogenase subunit 4
LITE_CONTIG_1453	up 3.010	AT4G02070	MSH6 (MUTS HOMOLOG 6-1)
KAT3YG24FM1_gi_85687429_gb_ABC73641_1_	up 3.004	ATCG00540	Encodes cytochrome f apoprotein; involved in photosynthetic electron transport chain; encoded by the chloroplast genome and is transcriptionally repressed by a nuclear gene HCF2.
LITE_CONTIG_4057	up 2.963	N/A	N/A
KAA6YM20FM1_hypothetical	up 2.951	AT1G13940	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G69360.1); similar to Os03g0758600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001051334.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:AAK16179.1); similar to expressed protein [Oryza sativa (japonica cultivar-group)] (GB:ABF98981.1); contains InterPro domain Protein of unknown function DUF863, plant; (InterPro:IPR008581)
LKBE4YO01FM1_	up 2.937	N/A	N/A
KAP4YH13FM1_	up 2.884	N/A	N/A
KCAF2YL24FM1_Gossypium	up 2.860	N/A	N/A
LITE_CONTIG_770	up 2.848	N/A	N/A
KCAK5YP14FM1_	up 2.845	N/A	N/A
LITE_CONTIG_1327	up 2.836	ATCG00180	RNA polymerase beta' subunit-1
KCAF5YA02FM1_	up 2.801	N/A	N/A
LITE_CONTIG_5334	up 2.798	ATCG01010	Chloroplast encoded NADH dehydrogenase unit.
KCAK1YE12FM1_	up 2.793	N/A	N/A
LITE_CONTIG_2703	up 2.788	ATCG00740	RNA polymerase alpha subunit
KCL5YP07FM1_Arabidopsis	up 2.787	N/A	N/A
KCAF3YO11FM1_	up 2.754	N/A	N/A
KCL2YI07FM1_	up 2.749	N/A	N/A
KCAK4YB08FM1_	up 2.728	N/A	N/A
LITE_CONTIG_986	up 2.725	ATCG00360	Encodes a protein required for photosystem I assembly and stability. In Chlamydomonas reinhardtii, this protein seems to act as a PSI specific chaperone facilitating the

			assembly of the complex by interacting with PsaA and PsaD. A loss of function mutation in tobacco leads to a loss of photosystem I.
KAA14YM09FM1_gi_49642980_emb_CAH00942_1_	up 2.713	N/A	N/A
KBF2YF10FM1_DDT	up 2.673	AT1G28420	HB-1 (homeobox-1); transcription factor
KAV6YJ11FM1_gi_85687401_gb_ABC73613_1_	up 2.641	ATCG00120	Encodes the ATPase alpha subunit, which is a subunit of ATP synthase and part of the CF1 portion which catalyzes the conversion of ADP to ATP using the proton motive force. This complex is located in the thylakoid membrane of the chloroplast.
LITE_CONTIG_5252	up 2.598	ATMG01360	cytochrome c oxidase subunit 1
LKBE8YO09FM1_	up 2.593	N/A	N/A
KAQ9YE03FM1_gi_58331773_gb_AAW70384_1_	up 2.581	AT1G62740	stress-inducible protein, putative
KBF10YK02FM1_	up 2.528	N/A	N/A
KAS3YB08FM1_	up 2.527	N/A	N/A
KAS1YN22_	up 2.515	N/A	N/A
CL3Contig4_hypothetical	up 2.511	N/A	N/A
KAP7YM15FM1_unnamed	up 2.509	AT5G53020	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G45900.1); similar to putative 200 kDa antigen p200 [Oryza sativa (japonica cultivar-group)] (GB:BAD88062.1); contains domain gb def: Emb
LITE_CONTIG_2938	up 2.507	AT5G02020	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT5G59080.1)
LITE_CONTIG_3430	up 2.497	AT5G19180	ECR1 (E1 C-TERMINAL RELATED 1); small protein activating enzyme
LITE_CONTIG_2025	up 2.487	AT3G52210	mRNA capping enzyme family protein
LITE_CONTIG_838	up 2.482	AT3G11910	ubiquitin-specific protease, putative
LKBE11YF17FM1_	up 2.479	N/A	N/A
LITE_CONTIG_5537	up 2.469	AT3G11670	DGD1 (DIGALACTOSYL DIACYLGLYCEROL DEFICIENT 1); galactolipid galactosyltransferase/ transferase, transferring glycosyl groups
KAV7YA04FM1_	up 2.456	N/A	N/A
KCL3YH24FM1_	up 2.456	N/A	N/A
KAT6YG07FM1_10	up 2.450	N/A	N/A
KAV2YI07FM1_	up 2.404	N/A	N/A
KAT10YO15FM1_gi_23197960_gb_AAN15507_1_	up 2.403	AT1G19110	inter-alpha-trypsin inhibitor heavy chain-related
LKBE9YF15FM1_	up 2.403	N/A	N/A
KBA3YN19FM1_	up 2.378	N/A	N/A
LITE_CONTIG_1233	up 2.370	N/A	N/A
KAT11YG11FM1_putative	up 2.368	AT3G07160	ATGSL10 (GLUCAN SYNTHASE-LIKE 10); 1,3-beta-glucan synthase
KBB2YF17FM1_	up 2.365	N/A	N/A

KAS1YG01FM1_	up 2.342	N/A	N/A
KCAE8YA11FM1_K_uniflora	up 2.340	N/A	N/A
LITE_CONTIG_5298	up 2.332	N/A	N/A
LITE_CONTIG_5917	up 2.329	N/A	N/A
KAV5YK05FM1_gi_110738680_dbj_BAF01265_1_	up 2.324	AT5G64220	calmodulin-binding protein
KCAF3YN07FM1_	up 2.321	N/A	N/A
LITE_CONTIG_1175	up 2.320	N/A	N/A
KAV9YO03FM1_	up 2.308	N/A	N/A
KAS1YA19RM1_gi_758363_emb_CAA37614_1_	up 2.307	N/A	N/A
LITE_CONTIG_528	up 2.305	N/A	N/A
LKBE7YF03FM1_subtilisin_like	up 2.294	AT5G59810	subtilase family protein
KAT9YD04FM1_hypothetical	up 2.291	AT4G32420	peptidyl-prolyl cis-trans isomerase cyclophilin-type family protein
LITE_CONTIG_6040	up 2.280	ATCG00500	Encodes the carboxytransferase beta subunit of the Acetyl-CoA carboxylase (ACCase) complex in plastids. This complex catalyzes the carboxylation of acetyl-CoA to produce malonyl-CoA, the first committed step in fatty acid synthesis.
KAA13YM11FM1_	up 2.260	N/A	N/A
KAP10YM17FM1_gi_113536665_dbj_BAF09048_1_	up 2.257	AT5G08630	DDT domain-containing protein
LITE_CONTIG_5559	up 2.251	ATCG00670	Encodes the only ClpP (caseinolytic protease) encoded within the plastid genome. Contains a highly conserved catalytic triad of Ser-type proteases (Ser-His-Asp). Part of the 350 kDa chloroplast Clp complex. The name reflects nomenclature described in Adam et. al (2001).
KAA9YH20FM1_hypothetical	up 2.245	N/A	N/A
CL1Contig8_gi_56806666_dbj_BAD83567_1_	up 2.236	ATMG00030	Identical to Hypothetical mitochondrial protein AtMg00030 (ORF107a) [Arabidopsis Thaliana] (GB:P93276); similar to hypothetical protein NitaMp027 [Nicotiana tabacum] (GB:YP_173374.1)
KAQ4YH02FM1_hypothetical	up 2.235	N/A	N/A
LITE_CONTIG_2564	up 2.235	N/A	N/A
LITE_CONTIG_4988	up 2.216	ATCG01040	Identical to Cytochrome c biogenesis protein ccsA (ccsA) [Arabidopsis Thaliana] (GB:P56770); similar to cytochrome c heme attachment protein [Gossypium hirsutum] (GB:YP_538991.1); contains InterPro domain Cytochrome c assembly protein; (InterPro:IPR002541)
KCAE6YG08FM1_unknown	up 2.212	AT1G69070	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G28770.1); similar to unknown protein [Oryza sativa (japonica cultivar-group)] (GB:AAN08225.2); similar to Nop14-like family protein, expressed [Oryza sativa (japonica cultivar-group)] (GB:ABF97462.1); similar to Nop14-like protein [Medicago truncatula] (GB:ABE90856.1); contains InterPro domain Nop14-like

			protein; (InterPro:IPR007276)
KCAE2YP03FM1_gi_119224884_dbj_BAF41270_1_	up 2.200	ATCG00660	encodes a chloroplast ribosomal protein L20, a constituent of the large subunit of the ribosomal complex
CL7Contig2_gi_56806666_dbj_BAD83567_1_	up 2.191	ATMG00030	Identical to Hypothetical mitochondrial protein AtMg00030 (ORF107a) [Arabidopsis Thaliana] (GB:P93276); similar to hypothetical protein NitaMp027 [Nicotiana tabacum] (GB:YP_173374.1)
CL3Contig2_gi_33086544_gb_AAP92584_1_	up 2.189	N/A	N/A
LITE_CONTIG_3729	up 2.177	AT4G33010	glycine dehydrogenase (decarboxylating), putative / glycine decarboxylase, putative / glycine cleavage system P-protein, putative
KAV1YD07_	up 2.173	N/A	N/A
LKBE4YK04FM1_	up 2.172	N/A	N/A
KAT3YF02FM1_hypothetical	up 2.167	AT5G24710	similar to protein kinase family protein [Arabidopsis thaliana] (TAIR:AT5G38560.1); similar to Os01g0653800 [Oryza sativa (japonica cultivar-group)] (GB:NP_001043744.1); similar to WD-40 repeat family protein [Medicago sativa] (GB:AAZ31064.1); similar to Peptidase S8 and S53, subtilisin, kexin, sedolisin; WD40-like [Medicago truncatula] (GB:ABD32844.1); contains InterPro domain WD-40 repeat; (InterPro:IPR001680); contains InterPro domain WD40-like; (InterPro:IPR011046)
LITE_CONTIG_5091	up 2.144	N/A	N/A
LITE_CONTIG_1531	up 2.134	ATCG00070	PSII K protein
KZOABB_CL4Contig4_hypothetical	up 2.128	N/A	N/A
LKBE8YJ13FM1_Beta	up 2.127	N/A	N/A
KAV12YP17FM1_putative	up 2.125	AT1G60200	splicing factor PWI domain-containing protein / RNA recognition motif (RRM)-containing protein
KBF2YP15FM1_	up 2.118	N/A	N/A
KBB3YO12FM1_gi_112296199_gb_ABI15136_1_	up 2.117	ATCG00770	chloroplast 30S ribosomal protein S8
KCAF5YA01FM1_	up 2.109	N/A	N/A
LITE_CONTIG_2934	up 2.106	AT3G16460	jacalin lectin family protein
CL1Contig23_hypothetical	up 2.106	N/A	N/A
KAS4YM24FM1_Strong	up 2.104	AT1G60070	clathrin binding
LITE_CONTIG_4292	up 2.104	N/A	N/A
LITE_CONTIG_3501	up 2.103	AT5G05170	CESA3 (CELLULASE SYNTHASE 3); cellulose synthase/ transferase, transferring glycosyl groups
KAA5YO04FM1_unknow	up 2.096	N/A	N/A
LITE_CONTIG_3116	up 2.092	AT5G67320	WD-40 repeat family protein
LITE_CONTIG_2184	up 2.088	N/A	N/A

LITE_CONTIG_1836	up 2.084	AT5G55040	DNA-binding bromodomain-containing protein
KCAA5YD24FM1_gi_30023684_gb_AAP13375_1_	up 2.074	AT3G20600	NDR1 (NON RACE-SPECIFIC DISEASE RESISTANCE 1); signal transducer
KCAF1YJ05FM1_Populus	up 2.064	AT4G14920	PHD finger transcription factor, putative
LITE_CONTIG_183	up 2.054	N/A	N/A
KAA3YL08FM1_	up 2.041	N/A	N/A
KCAF5YF24FM1_ABT1_associated	up 2.037	AT3G01160	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G28770.1); similar to conserved hypothetical protein [Medicago truncatula] (GB:ABE90093.1); contains domain RNA-binding domain, RBD (SSF54928); contains domain UNCHARACTERIZED (PTHR12202); contains domain UNCHARACTERIZED (PTHR12202:SF5)
LITE_CONTIG_1973	up 2.037	AT3G15220	protein kinase, putative
KAT10YF20FM1_hypothetical	up 2.035	AT2G29120	ATGLR2.7 (Arabidopsis thaliana glutamate receptor 2.7)
KBF3YD11FM1_Vitis	up 2.032	N/A	N/A
LITE_CONTIG_5462	up 2.028	ATCG01130	Identical to Putative membrane protein ycf1 (ycf1-B) [Arabidopsis Thaliana] (GB:P56785); similar to hypothetical chloroplast RF1 [Gossypium hirsutum] (GB:YP_538982.1); similar to Hypothetical protein [Gossypium barbadense] (GB:BAF41305.1); similar to hypothetical chloroplast RF1 [Eucalyptus globulus subsp. globulus] (GB:YP_636357.1); contains InterPro domain Ycf1; (InterPro:IPR008896)
LITE_CONTIG_6220	up 2.026	AT3G56150	EIF3C (EUKARYOTIC TRANSLATION INITIATION FACTOR 3)
KBB13YG10FM1_hypothetical	up 2.010	N/A	N/A
KAT5YE13FM1_hypothetical	up 2.009	AT5G12400	PHD finger transcription factor, putative
KAC13YP19FM1_probable	up 2.007	AT1G05320	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT2G32240.1); similar to unnamed protein product [Candida glabrata] (GB:XP_445454.1); similar to OSIGBa0126B18.8 [Oryza sativa (indica cultivar-group)] (GB:CAH66915.1); similar to OSJNBa0008M17.5 [Oryza sativa (japonica cultivar-group)] (GB:CAE04334.2); contains domain MYOSIN HEAVY CHAIN-RELATED (PTHR23160)
KCAE4YG02FM1_gi_2462838_gb_AAB72173_1_	up 2.007	AT1G22860	TGF beta receptor associated protein-related
LITE_CONTIG_1638	up 2.004	AT5G24470	APRR5 (PSEUDO-RESPONSE REGULATOR 5); transcription regulator
KAP24YN11FM1_	up 11.891	N/A	N/A
LKBE10YD11FM1_gi_66136_pir__DNOBU1	up 1.998	ATMG00516	Encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from three precursors, NAD1A, NAD1B, and NAD1C.
LITE_CONTIG_6689	up 1.996	AT4G14920	PHD finger transcription factor, putative
LITE_CONTIG_2769	up 1.985	N/A	N/A
KBB14YM14FM1_starch	up 1.980	AT3G29320	glucan phosphorylase, putative

LKBE9YG16FM1_gi_7270476_emb_CAB80241_1_	up 1.980	AT4G35240	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT2G17110.1); similar to Os09g0547300 [Oryza sativa (japonica cultivar-group)] (GB:NP_001063849.1); similar to bzip-related transcription factor -like [Oryza sativa (japonica cultivar-group)] (GB:BAD46467.1); contains InterPro domain Protein of unknown function DUF632; (InterPro:IPR006867); contains InterPro domain Protein of unknown function DUF630; (InterPro:IPR006868)
KAT10YL12FM1_gi_56806637_dbj_BAD83538_1_	up 1.970	ATMG00290	encodes a mitochondrial ribosomal protein S4, a constituent of the small subunit of the ribosomal complex
KBB1YE18_gi_58802853_gb_AAW82573_1_	up 1.968	N/A	N/A
LKBE5YE13FM1_	up 1.965	N/A	N/A
LITE_CONTIG_4751	up 1.959	AT5G11500	similar to unknown protein [Oryza sativa (japonica cultivar-group)] (GB:AAT07560.1); contains InterPro domain Protein of unknown function DUF814; (InterPro:IPR008532)
LITE_CONTIG_3567	up 1.959	N/A	N/A
KAP5YL24FM1_	up 1.956	N/A	N/A
KAQ3YN11FM1_	up 1.951	N/A	N/A
KAQ10YN20FM1_gi_2388582_gb_AAB71463_1_	up 1.947	AT1G05150	calcium-binding EF hand family protein
CL5Contig2_gi_56806578_dbj_BAD83479_1_	up 1.941	N/A	N/A
LITE_CONTIG_5674	up 1.923	AT1G64660	ATMGL; catalytic/ methionine gamma-lyase
KAV10YH16FM1_	up 1.920	N/A	N/A
CL260Contig1_gi_92898532_gb_ABE94226_1_	up 1.914	AT1G33060	ANAC014
LITE_CONTIG_5468	up 1.912	AT5G15930	PAM1 (PLANT ADHESION MOLECULE 1); RAB GTPase activator
KAC3YH21FM1_	up 1.906	N/A	N/A
KAQ11YM18FM1_pumilio	up 1.904	AT3G20250	APUM5 (ARABIDOPSIS PUMILIO 5); RNA binding
KAS7YM05FM1_gi_110737376_dbj_BAF00633_1_	up 1.901	AT1G14790	RDR1 (RNA-DEPENDENT RNA POLYMERASE 1); RNA-directed RNA polymerase/ nucleic acid binding
KAS10YG11FM1_gi_10177834_dbj_BAB11263_1_	up 1.893	AT5G56360	calmodulin-binding protein
KCAK3YB22FM1_	up 1.879	N/A	N/A
KAC1YK13FM1_gi_2664190_emb_CAA04679_1_	up 1.878	AT1G09570	PHYA (PHYTOCHROME A); G-protein coupled photoreceptor/ signal transducer
LITE_CONTIG_5985	up 1.876	AT3G48990	AMP-dependent synthetase and ligase family protein
LITE_CONTIG_5851	up 1.874	AT2G27170	TTN7 (TITAN7); ATP binding
LITE_CONTIG_2738	up 1.867	AT3G01470	ATHB-1 (Homeobox-leucine zipper protein HAT5); transcription factor

LITE_CONTIG_4346	up 1.861	AT1G25350	OVA9 (OVULE ABORTION 9); glutamine-tRNA ligase
LITE_CONTIG_1700	up 1.860	AT5G65630	DNA-binding bromodomain-containing protein
CL445Contig1_putative	up 1.852	AT1G58230	WD-40 repeat family protein / beige-related
LITE_CONTIG_4467	up 1.850	N/A	N/A
KCAK4YE11FM1_Vitis	up 1.848	N/A	N/A
LITE_CONTIG_6351	up 1.814	AT4G33740	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT2G22795.1); similar to conserved hypothetical protein [Medicago truncatula] (GB:ABE84763.1)
KAC2YK22FM1_	up 1.812	N/A	N/A
LKBE5YE10FM1_gi_87241360_gb_ABD33218_1_	up 1.808	AT2G42890	AML2; RNA binding
LITE_CONTIG_3592	up 1.799	ATMG00210	encodes a mitochondrial ribosomal protein L5, a constituent of the large subunit of the ribosomal complex
LKBE3YH11FM1_	up 1.790	N/A	N/A
KCAF3YI08FM1_	up 1.782	AT5G03450	zinc finger (C3HC4-type RING finger) family protein
LITE_CONTIG_1774	up 1.778	AT3G07770	ATP binding
KAT6YK22FM1_gi_51971230_dbj_BAD44307_1_	up 1.755	AT2G30600	BTB/POZ domain-containing protein
KCAK4YB13FM1_	up 1.745	N/A	N/A
KAV7YI18FM1_Ca2+_transporting	up 1.744	AT3G21180	ACA9 (autoinhibited Ca2+ -ATPase 9); calcium-transporting ATPase/ calmodulin binding
KAP3YP10FM1_	up 1.737	N/A	N/A
KAA11YH14FM1_serine_glyoxylate	up 1.729	AT2G13360	AGT (ALANINE:GLYOXYLATE AMINOTRANSFERASE)
LITE_CONTIG_3155	up 1.727	AT1G15740	leucine-rich repeat family protein
KBF3YN05FM1_gi_169128_gb_AAA33680_1_	up 1.726	AT5G50920	CLPC (HEAT SHOCK PROTEIN 93-V); ATP binding / ATPase
LITE_CONTIG_510	up 1.726	N/A	N/A
LKBE3YG04FM1_	up 1.723	N/A	N/A
KAA6YJ06FM1_hypothetical	up 1.714	N/A	N/A
KAS2YA17FM1_1,4_alpha_glucan	up 1.699	AT5G03650	SBE2.2 (STARCH BRANCHING ENZYME 2.2)
KCAE3YO13FM1_	up 1.688	N/A	N/A
KAV1YL24FM1_	up 1.687	N/A	N/A
LCL1Contig2_Sabia	up 1.661	N/A	N/A
LKBE1YL04FM1_unnamed	up 1.660	AT3G20720	similar to Os04g0628600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001053961.1); similar to OSJNBa0089N06.9 [Oryza sativa (japonica cultivar-group)] (GB:CAE04248.3); similar to unnamed protein product; gene_id:MOE17.2 unknown protein [Medicago truncatula] (GB:ABE93163.1); contains domain SUBFAMILY NOT NAMED (PTHR22774:SF4); contains domain UNCHARACTERIZED (PTHR22774)

KAQ7YM19FM1_gi_6041840_gb_AAF02149_1_AC009853_9	up 1.659	AT3G07400	lipase class 3 family protein
LITE_CONTIG_2495	up 1.646	AT4G30550	glutamine amidotransferase class-I domain-containing protein
2210_68414_m00950	up 1.643	AT1G08570	thioredoxin family protein
KAV13YC18FM1_unnamed	up 1.636	AT3G29390	RIK (RS2-INTERACTING KH PROTEIN)
LITE_CONTIG_2255	up 1.633	AT1G06950	ATTIC110/TIC110 (TRANSLOCON AT THE INNER ENVELOPE MEMBRANE OF CHLOROPLASTS 110)
LITE_CONTIG_3102	up 1.632	N/A	N/A
LITE_CONTIG_1738	up 1.625	ATMG01190	ATPase subunit 1
KAC1YD10FM1_gi_144923240_gb_ABP02416_1_	up 1.624	AT5G54280	ATM2 (ARABIDOPSIS THALIANA MYOSIN 4)
KBB14YC17FM1_cytochrome	up 1.623	ATMG00220	Mitochondrial apocytochrome b (cob) gene encodes a subunit of the ubiquinol-cytochrome c oxidoreductase and is part of a 5 kb transcript. The transcript also contains a pseudogene for ribosomal protein S14 called RPS15 and a tRNA(Ser) gene. Both the Cob and RPS15 genes are edited in the transcript.
KCL3YB16FM1_hypothetical	up 1.621	AT2G04240	XERICO; protein binding / zinc ion binding
LITE_CONTIG_166	up 1.616	N/A	N/A
LITE_CONTIG_6035	up 1.615	AT1G47900	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G19835.1); similar to Putative myosin-like protein [Oryza sativa (japonica cultivar-group)] (GB:AAL77142.1); similar to Os03g0246500 [Oryza sativa (japonica cultivar-group)] (GB:NP_001049544.1); contains InterPro domain Protein of unknown function DUF869, plant; (InterPro:IPR008587)
LCL405Contig1_	up 1.614	N/A	N/A
LITE_CONTIG_1752	up 1.610	AT3G45140	LOX2 (LIPOXYGENASE 2)
KAQ11YP13FM1_	up 1.608	N/A	N/A
LITE_CONTIG_1914	up 1.607	AT5G64550	loricrin-related
KBB12YK13FM1_	up 1.607	N/A	N/A
KAQ11YJ02FM1_	up 1.603	N/A	N/A
LITE_CONTIG_6400	up 1.591	AT1G68490	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G13390.2); similar to hypothetical protein [Cleome spinosa] (GB:ABD96883.1)
LITE_CONTIG_4730	up 1.590	N/A	N/A
KAP13YK17FM1_hypothetical	up 1.587	AT5G53890	leucine-rich repeat transmembrane protein kinase, putative
KBB12YG07FM1_RabGAP/TBC	up 1.583	AT4G29950	microtubule-associated protein
LITE_CONTIG_2669	up 1.583	ATCG01310	encodes a chloroplast ribosomal protein L2, a constituent of the large subunit of the ribosomal complex
KAQ4YI02FM1_	up 1.583	N/A	N/A
LITE_CONTIG_1604	up 1.580	AT4G03420	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G03610.1); similar to hypothetical protein

			[Oryza sativa (japonica cultivar-group)] (GB:AAO39871.1); similar to Os10g0494000 [Oryza sativa (japonica cultivar-group)] (GB:NP_001064942.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:AAP46235.1); contains InterPro domain Protein of unknown function DUF789; (InterPro:IPR008507)
LITE_CONTIG_2268	up 1.578	AT2G13360	AGT (ALANINE:GLYOXYLATE AMINOTRANSFERASE)
KAT11YJ19FM1_RelA_SpoT	up 1.574	AT4G02260	RSH1 (RELA-SPOT HOMOLOG); catalytic
LITE_CONTIG_1696	up 1.567	AT4G29060	EMB2726 (EMBRYO DEFECTIVE 2726); translation elongation factor
KBA3YB13FM1_hypothetical	up 1.566	AT1G19870	IQD32 (IQ-domain 32); calmodulin binding
KAS10YE10FM1_gi_110738020_dbj_BAF00945_1_	up 1.564	AT1G79090	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT3G22270.1); similar to Os02g0517300 [Oryza sativa (japonica cultivar-group)] (GB:NP_001046960.1); similar to Os01g0769000 [Oryza sativa (japonica cultivar-group)] (GB:NP_001044369.1); similar to unknown protein [Oryza sativa (japonica cultivar-group)] (GB:BAD52714.1); contains domain gb def: YUP8H12R.29 protein (PTHR21551:SF11); contains domain FAMILY NOT NAMED (PTHR21551)
LITE_CONTIG_6163	up 1.563	ATMG00640	encodes a plant b subunit of mitochondrial ATP synthase based on structural similarity and the presence in the F(0) complex.
LKBE3YF15FM1_unnamed	up 1.562	AT5G57990	UBP23 (UBIQUITIN-SPECIFIC PROTEASE 23); ubiquitin-specific protease
KAC5YF05FM1_	up 1.559	AT4G01510	ARV2
LITE_CONTIG_5331	up 1.554	ATCG01130	Identical to Putative membrane protein ycf1 (ycf1-B) [Arabidopsis Thaliana] (GB:P56785); similar to hypothetical chloroplast RF1 [Gossypium hirsutum] (GB:YP_538982.1); similar to Hypothetical protein [Gossypium barbadense] (GB:BAF41305.1); similar to hypothetical chloroplast RF1 [Eucalyptus globulus subsp. globulus] (GB:YP_636357.1); contains InterPro domain Ycf1; (InterPro:IPR008896)
LITE_CONTIG_3168	up 1.534	AT2G39930	ATISA1/ISA1 (ISOAMYLASE 1); alpha-amylase/ isoamylase
KCAF2YC17FM1_gi_62320777_dbj_BAD95440_1_	up 1.523	AT5G02770	similar to expressed protein [Oryza sativa (japonica cultivar-group)] (GB:ABF95330.1); similar to Os03g0283300 [Oryza sativa (japonica cultivar-group)] (GB:NP_001049753.1); contains domain FAMILY NOT NAMED (PTHR21594); contains domain gb def: Hypothetical protein Y53G8AR.6 (PTHR21594:SF7)
LITE_CONTIG_5751	up 1.522	ATCG00670	Encodes the only ClpP (caseinolytic protease) encoded within the plastid genome. Contains a highly conserved catalytic triad of Ser-type proteases (Ser-His-Asp). Part of the 350 kDa chloroplast Clp complex. The name reflects nomenclature described in Adam et. al (2001).
KBA3YM12FM1_	up 1.522	N/A	N/A
LITE_CONTIG_6390	up 1.512	ATCG00480	chloroplast-encoded gene for beta subunit of ATP synthase
KAQ4YG11FM1_Vitis	up 1.506	N/A	N/A
KAT10YM13FM1_hypothetical	up 1.503	AT3G14460	disease resistance protein (NBS-LRR class), putative

Analysis of the gene annotations revealed that most SCA6 top up-regulated genes were mitochondrial and chloroplast genes based on their homologous *Arabidopsis* locus IDs that encode proteins with functions in the electron transport chain (Table 5). About 50% of the top 40 up-regulated annotated genes in SCA6 were chloroplast genes (Table 6), including subunits of photosystem II and RUBISCO, 25% were mitochondrial genes, mainly NADH dehydrogenase subunit, and only 25% were nucleus genes. In contrast, no chloroplast and mitochondrial annotations were found in the top up-regulated genes in SA-treated ICS1. To further confirm the enrichment of chloroplast and mitochondrial genes in induced SCA6, we performed Gene Ontology (GO) annotation of three sets of genes using the *Arabidopsis* locus IDs: 1. All genes present in the whole array, 2. Genes significantly up-regulated by SA in SCA6 and 3. Genes significantly up-regulated by SA in ICS1. Among 17,427 unigenes on the whole genome array, 11,465 genes have *Arabidopsis* locus identifier and the cellular component categorization demonstrated that about 12% of the genes in the whole array gene set were chloroplast genes and 4% were mitochondrial genes (Figure 5-4 A). In SCA6, the chloroplast genes represented 17% and the mitochondrial genes 5% of up-regulated annotated genes. However, the upregulated genes in ICS1 included only 6% of chloroplast genes and 1% mitochondrial genes (Figure 5-4 B and C). Chi-square analysis showed that the enrichment of mitochondrial and chloroplast genes in SCA6 up-regulated set was statistically significant, as well as the reduction of chloroplast genes in ICS1 up-regulated set ($P<0.05$). Thus the results indicated enrichment of chloroplast and mitochondrial genes over the whole array in the SCA6 upregulated genes compared to reduced representation of these in ICS1 (Figure 5-4D), suggesting that those genes could play important roles in defense response in SCA6 and potentially confer higher resistance of SCA6 to pathogens.

Table 7 Top 40 up-regulated genes with *Arabidopsis* identifier SCA6 stage C leaf in response to SA.

SEQ_ID	Fold Induction	At Locus Identifier	Annotations
LITE_CONTIG_2451	up 8.254	ATCG00680	encodes for CP47, subunit of the photosystem II reaction center.
KAA13YF15FM1_photosystem	up 7.961	ATCG00680	encodes for CP47, subunit of the photosystem II reaction center.

KAC3YC17FM1_gi_119224918_dbj_BAF41304_1_	up 6.994	ATCG01120	encodes a chloroplast ribosomal protein S15, a constituent of the small subunit of the ribosomal complex
LITE_CONTIG_2939	up 6.020	ATCG01240	30S chloroplast ribosomal protein S7
KAP12YH14FM1_NADH_ubiquinone	up 5.744	ATMG00285	encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from two precursors, NAD2A and NAD2B.
LITE_CONTIG_4986	up 5.549	ATCG01090	Encodes subunit of the chloroplast NAD(P)H dehydrogenase complex
LITE_CONTIG_5921	up 5.288	ATMG00560	encodes a mitochondrial ribosomal protein L2, a constituent of the large subunit of the ribosomal complex
KBF4YD24FM1_HEAT	up 4.891	AT2G35630	MOR1 (MICROTUBULE ORGANIZATION 1)
LITE_CONTIG_2578	up 4.876	ATMG01320	encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from two precursors, NAD2A and NAD2B.
LITE_CONTIG_5208	up 4.852	ATCG00280	chloroplast gene encoding a CP43 subunit of the photosystem II reaction center. promoter contains a blue-light responsive element.
LITE_CONTIG_998	up 4.777	ATCG00040	Encodes a maturase located in the trnK intron in the chloroplast genome.
LITE_CONTIG_1235	up 4.480	ATCG00340	Encodes the D1 subunit of photosystem I and II reaction centers.
KAA8YD19FM1_	up 4.074	ATCG00170	RNA polymerase beta' subunit-2
LITE_CONTIG_2715	up 3.951	ATCG00190	Chloroplast DNA-dependent RNA polymerase B subunit. The transcription of this gene is regulated by a nuclear encoded RNA polymerase. This gene has been transferred to mitochondrial genome during crucifer evolution.
KBB14YA22FM1_hypothetical	up 3.925	AT1G04300	similar to meprin and TRAF homology domain-containing protein / MATH domain-containing protein [Arabidopsis thaliana] (TAIR:AT5G43560.2); similar to MATH domain, putative [Medicago truncatula] (GB:ABE87963.1); similar to Os12g0597100 [Oryza sativa (japonica cultivar-group)] (GB:NP_001067191.1); contains InterPro domain TRAF-like; (InterPro:IPR008974); contains InterPro domain MATH; (InterPro:IPR002083); contains InterPro domain TRAF-type; (InterPro:IPR013322)
KAP10YP17FM1_gi_85687416_gb_ABC73628_1_	up 3.874	ATCG00350	Encodes psaA protein comprising the reaction center for photosystem I along with psaB protein; hydrophobic protein encoded by the chloroplast genome.
LITE_CONTIG_2510	up 3.720	ATMG00640	encodes a plant b subunit of mitochondrial ATP synthase based on structural similarity and the presence in the F(0) complex.
LITE_CONTIG_1962	up 3.699	AT2G38800	calmodulin-binding protein-related
KCAA7YF10FM1_gi_125587847_gb_EAZ28511_1_	up 3.657	AT2G39810	HOS1 (High expression of osmotically responsive genes 1)
LITE_CONTIG_5476	up 3.654	ATCG00540	Encodes cytochrome f apoprotein; involved in photosynthetic electron transport chain; encoded by the chloroplast genome and is transcriptionally repressed by a nuclear gene HCF2.
LITE_CONTIG_6335	up 3.605	ATMG00580	NADH dehydrogenase subunit 4
LITE_CONTIG_4190	up 3.589	ATCG00490	large subunit of RUBISCO.
KAA9YK04FM1_gi	up 3.476	ATMG00650	Encodes NADH dehydrogenase subunit 4L.

_148491416_dbj_BAA99307_2_			
LITE_CONTIG_2432	up 3.453	ATMG00580	NADH dehydrogenase subunit 4
LITE_CONTIG_1504	up 3.385	AT4G20850	TPP2 (TRIPEPTIDYL PEPTIDASE II); subtilase
LKBE8YM20FM1_gi_80750940_dbj_BAE48016_1_	up 3.304	ATCG00550	PSII component
KAT5YB10FM1_	up 3.300	AT1G68650	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G25520.1); similar to Os08g0433100 [Oryza sativa (japonica cultivar-group)] (GB:NP_001061871.1); similar to PREDICTED: similar to conserved hypothetical protein [Strongylocentrotus purpuratus] (GB:XP_798718.2); contains InterPro domain Protein of unknown function UPF0016; (InterPro:IPR001727)
LITE_CONTIG_5346	up 3.190	AT1G22610	C2 domain-containing protein
LITE_CONTIG_3770	up 3.047	ATCG00720	Encodes the cytochrome b(6) subunit of the cytochrome b6f complex.
LITE_CONTIG_2116	up 3.043	ATMG01275	Encodes subunit of mitochondrial NAD(P)H dehydrogenase that is trans-spliced from three precursors, NAD1A, NAD1B, and NAD1C.
KAT4YI07FM1_gi_37591106_dbj_BAC98908_1_	up 3.036	ATMG00580	NADH dehydrogenase subunit 4
LITE_CONTIG_1453	up 3.010	AT4G02070	MSH6 (MUTS HOMOLOG 6-1)
KAT3YG24FM1_gi_85687429_gb_ABC73641_1_	up 3.004	ATCG00540	Encodes cytochrome f apoprotein; involved in photosynthetic electron transport chain; encoded by the chloroplast genome and is transcriptionally repressed by a nuclear gene HCF2.
KAA6YM20FM1_hypothetical	up 2.951	AT1G13940	similar to unknown protein [Arabidopsis thaliana] (TAIR:AT1G69360.1); similar to Os03g0758600 [Oryza sativa (japonica cultivar-group)] (GB:NP_001051334.1); similar to hypothetical protein [Oryza sativa (japonica cultivar-group)] (GB:AAK16179.1); similar to expressed protein [Oryza sativa (japonica cultivar-group)] (GB:ABF98981.1); contains InterPro domain Protein of unknown function DUF863, plant; (InterPro:IPR008581)
LITE_CONTIG_1327	up 2.836	ATCG00180	RNA polymerase beta' subunit-1
LITE_CONTIG_5334	up 2.798	ATCG01010	Chloroplast encoded NADH dehydrogenase unit.
LITE_CONTIG_2703	up 2.788	ATCG00740	RNA polymerase alpha subunit
LITE_CONTIG_986	up 2.725	ATCG00360	Encodes a protein required for photosystem I assembly and stability. In Chlamydomonas reinhardtii, this protein seems to act as a PSI specific chaperone facilitating the assembly of the complex by interacting with PsaA and PsaD. A loss of function mutation in tobacco leads to a loss of photosystem I.
KBF2YF10FM1_DDT	up 2.673	AT1G28420	HB-1 (homeobox-1); transcription factor
KAV6YJ11FM1_gi_85687401_gb_ABC73613_1_	up 2.641	ATCG00120	Encodes the ATPase alpha subunit, which is a subunit of ATP synthase and part of the CF1 portion which catalyzes the conversion of ADP to ATP using the proton motive force. This complex is located in the thylakoid membrane of the chloroplast.

LITE_CONTIG_5252	up 2.598	ATMG01360	cytochrome c oxidase subunit 1

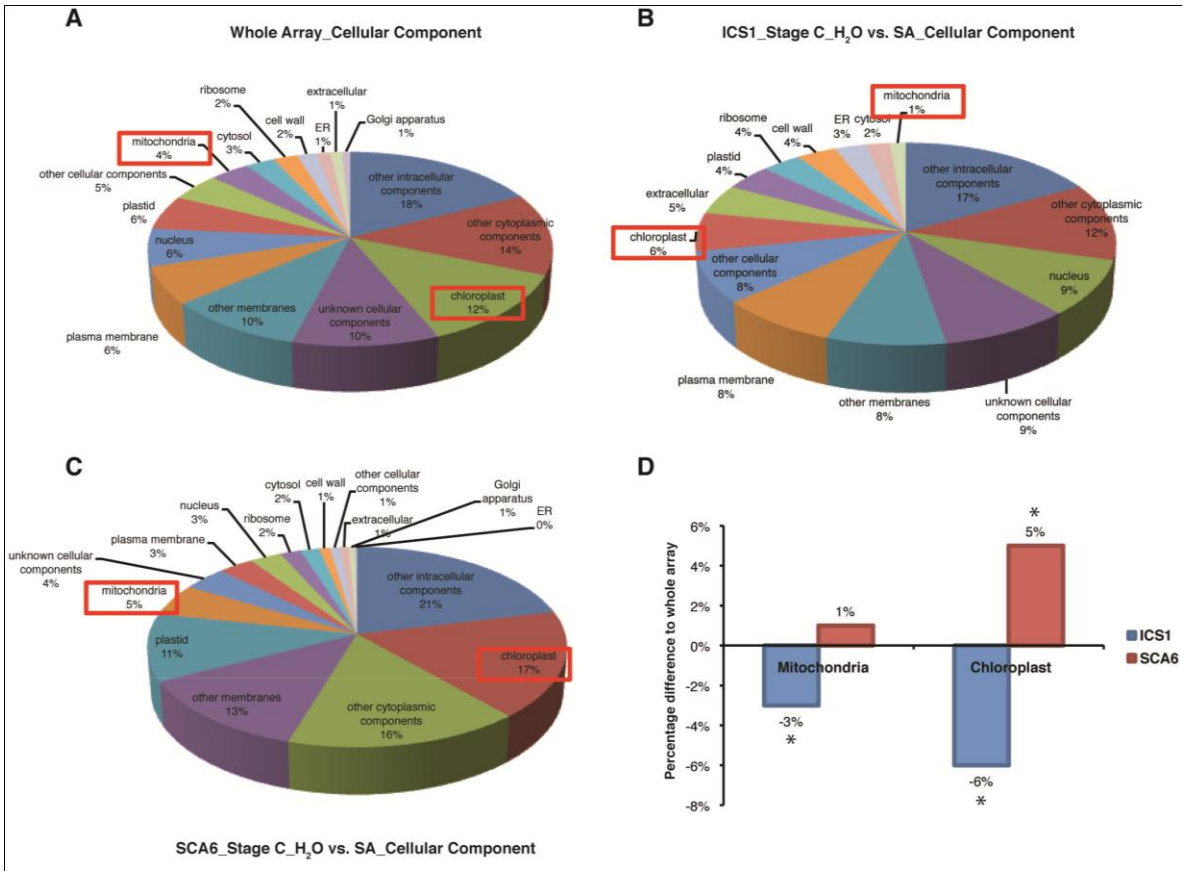


Figure 5-4 GO cellular component annotation analysis of significantly up-regulated genes in SA-treated stage C leaves of SCA6 and ICS1 with available with *Arabidopsis* locus IDs . **A.** GO cellular component analysis of all annotated genes on the whole array. **B.** GO cellular component analysis of annotated genes significantly up-regulated in SA-treated ICS1 genotype compared to control ICS1. . **C.** Pie chart of GO cellular component of annotated genes significantly up-regulated in SA-treated SCA6 genotype compared to control SCA6. The red boxes in **A**, **B** and **C** indicate the gene annotation percentages for mitochondria and chloroplast categories. **D.** Enrichment analysis of mitochondrial and chloroplast genes in ICS1 and SCA6 treated with SA. Bar chart represents the percentage difference of mitochondrial and chloroplast genes between gene annotations in the whole array and significantly SA-induced genes in ICS1 or SCA6. Stars represent statistical significance by Chi-square analysis ($P<0.05$).

Semi-quantitative RT-PCR validation

To validate the microarray results, semi-quantitative RT-PCR was performed on a subset of genes. Contig lite/contig/2252 with large up-regulation and CL160/contig1 with dramatic down-regulation in most of the comparisons, and two other probes lite/contig/2028 and lite/contig/3309 were chosen for RT-PCR. It has been shown that the relative expression value of both contigs varied among different samples (data not shown), similar to the microarray intensity values. In addition, all RT-PCR showed similar trends to microarray data when we plotted $\log_2$ fold change of all the comparisons that show significant changes in microarray (Figure 5-5 A-D). Although the fold changes in microarray are greater than in RT-PCR, which is probably due to the lower sensitivity of the gel intensity based quantification in RT-PCR, the microarray results are confirmed by semi-quantitative RT-PCR.

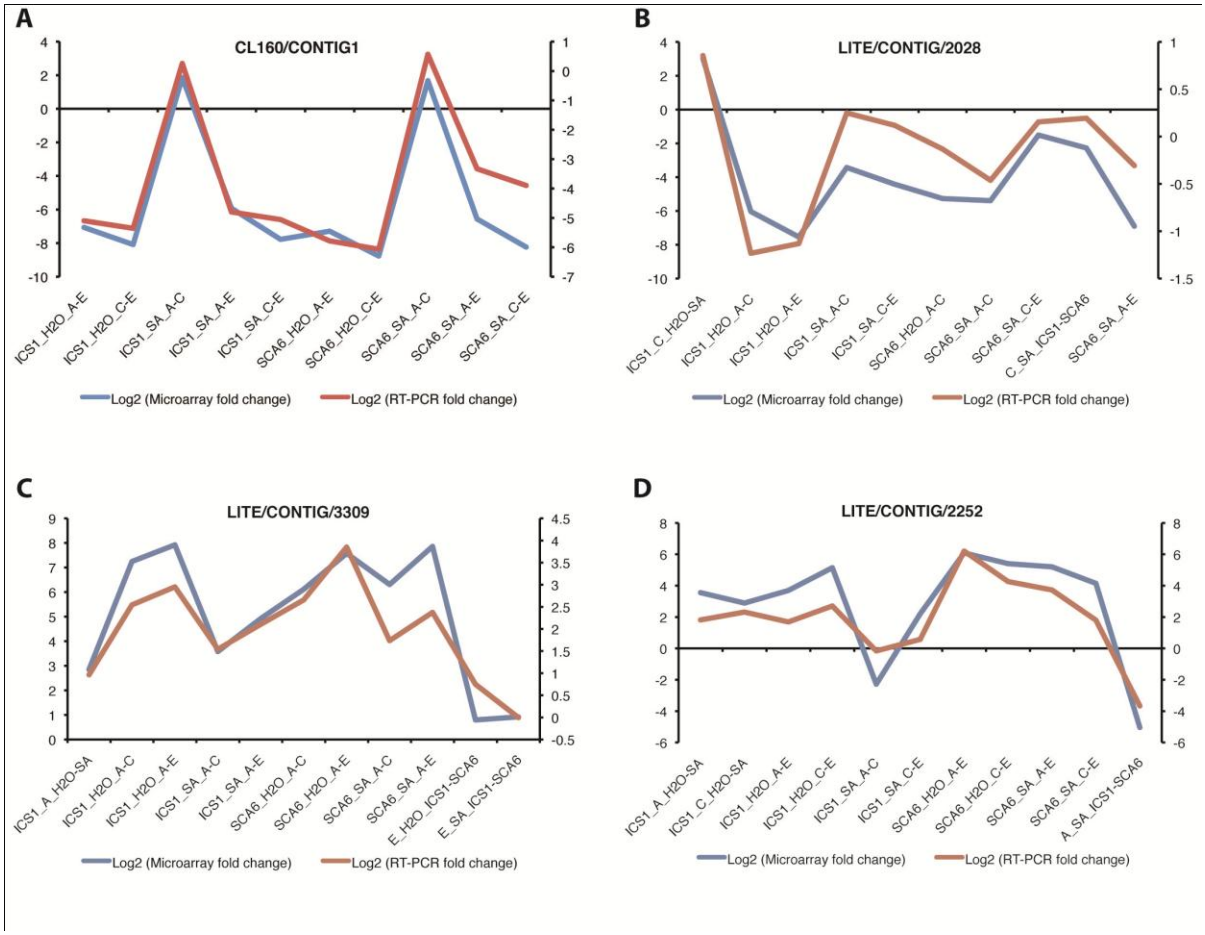


Figure 5-5 RT-PCR validation of microarray results. RT-PCR was performed with the same RNA samples as microarray using four probes, **A.** CL160/Contig1, **B.** Lite/contig/2028, **C.** Lite/contig/3309 and **D.** Lite/Contig/2252. Gel images were quantified by ImageQuant and relative expression of four contigs was normalized to *TcActin*. X-axis: All the comparisons that showed significant change in microarray. Blue line represents $\log_2$ of microarray fold change corresponding to the value on primary Y-axis and red line represents $\log_2$ of RT-PCR fold change corresponding to the value on secondary Y-axis.

5.4 Discussion

To characterize the molecular processes occurring during defense response in cacao, we used microarray analysis to compare the gene expression changes between SA- and control water- treated leaf samples of one susceptible (ICS1) and one disease-tolerant (SCA6) cacao genotypes (Faleiro et al., 2006). Since it has been previously demonstrated that defense response could vary among different leaf stages (Bailey et al., 2005), we also included analysis of genes expression among different developmental stages of the control and treated leaf tissue. This study represent an important step in cacao research and includes the first large-scale expression microarray analysis of this species in general and specifically in responsive to SA treatment. Although, the annotations of cacao ESTs are still limited, the microarray offered a remarkable way to study genome-wide defense response and development in cacao and we were able to identify large number of differentially expressed genes in our system and in all comparisons among treatments (Table 5-1),.

The two cacao genotypes included in this study, Scavina6 and ICS1 differ in their tolerance to witches' broom disease (Sca6 is more resistant) (Faleiro et al., 2006), but the molecular basis of this difference is not fully characterized. Our results indicated that chloroplast and mitochondrial genes were enriched in SA-induced SCA6 and the genes included NADPH dehydrogenases in mitochondria and genes encoding for proteins in the photosystem reaction center II. In contrast this enrichment is not observed in ICS1 treated with SA and furthermore the treatment contributed to reduce the percentage of the chloroplast and mitochondrial genes represented in the upregulated genes. The up-

regulation of those genes could be associated with oxidative burst (Rao et al., 1997; Torres and Dangl, 2005; Pogany et al., 2009; Dubreuil-Maurizi et al., 2010), because of their functions in electron transfer on the inner membrane of both organelles. It has been shown that the elevation of endogenous SA can induce the production of reactive oxygen species (ROS), which will in turn facilitate hypersensitive response (HR) and cell death at the site of infection (Apel and Hirt, 2004; Mur et al., 2008). Hypersensitive response is the first mechanism to combat pathogen infection, thus the difference of the induction of hypersensitive response between SCA6 and ICS1 may illuminate their distinct susceptibility to fungal pathogens. However, hypersensitive response has not been identified in cacao tissues. Therefore, further investigation of ROS production in ICS1 and SCA6 and the characterization of hypersensitive response in cacao tissues are of great importance to further confirm that SCA6 has the ability to induce HR while ICS1 lacks this capability.

The microarray, as a useful tool to study the global transcript changes in cacao, revealed possible existence of hypersensitive response in cacao tissue. In addition, it will help to identify new markers related to defense response and determine the developmental stages that cost the most upon defense response, which will be favored by both cacao farmers and chocolate industry.

5.5 References

***Argout X, *Salse J, *Aury J, *Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, Shi Z, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Sidibe-Bocs S, Lanaud C** (2010) (*These authors contributed equally to this work.) Deciphering the genome structure and paleohistory of *Theobroma cacao*. Nature Genetics **In Press**

Agrawal AA (1998) Induced responses to herbivory and increased plant performance. Science **279:** 1201-1202

Agrawal AA, Strauss SY, Stout MJ (1999) Costs of induced responses and tolerance to herbivory in male and female fitness components of wild radish. Evolution **53:** 1093-1104

Albagli O, Dhordain P, Deweindt C, Lecocq G, Leprince D (1995) The BTB/POZ domain: a new protein-protein interaction motif common to DNA- and actin-binding proteins. Cell Growth Differ **6:** 1193-1198

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403-410

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:** 3389-3402

Andebrhan T, Maddison AC, Rudgard SA, International Office of Cocoa Chocolate and Sugar Confectionary. (1993) Disease management in cocoa : comparative epidemiology of witches' broom. Chapman & Hall, London

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annu Rev Plant Biol **55:** 373-399

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M, Kuhn D, Verica J, Courtois B, Loor G, Babin R, Sounigo O, Ducamp M, Guiltinan MJ, Ruiz M, Alemanno L, Machado R, Phillips W, Schnell R, Gilmour M, Rosenquist E, Butler D, Maximova S, Lanaud C (2008) Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of Theobroma cacao L. generated from various tissues and under various conditions. BMC Genomics **9:** 512

Argout X, Fouet O, Wincker P, Gramacho K, Legavre T, Sabau X, Risterucci AM, Da Silva C, Cascardo J, Allegre M, Kuhn D, Verica J, Courtois B, Loor G, Babin R, Sounigo O, Ducamp M, Guiltinan MJ, Ruiz M, Alemanno L, Machado R, Phillips W, Schnell R, Gilmour M, Rosenquist E, Butler D, Maximova S, Lanaud C (2008) Towards the understanding of the cocoa transcriptome: Production and analysis of an exhaustive dataset of ESTs of Theobroma cacao L. generated from various tissues and under various conditions. BMC Genomics **9:** -

Attaran E, Zeier TE, Griebel T, Zeier J (2009) Methyl salicylate production and jasmonate signaling are not essential for systemic acquired resistance in Arabidopsis. Plant Cell **21:** 954-971

Bae H, Kim SH, Kim MS, Sicher RC, Lary D, Strem MD, Natarajan S, Bailey BA (2008) The drought response of *Theobroma cacao* (cacao) and the regulation of genes involved in polyamine biosynthesis by drought and other stresses. Plant Physiol Biochem **46:** 174-188

Bailey BA, Bae H, Strem MD, Antunez de Mayolo G, Guiltinan MJ, Verica JA, Maximova SN, Bowers JH (2005) Developmental expression of stress response genes in Theobroma cacao leaves and their

response to Nep1 treatment and a compatible infection by Phytophthora megakarya. Plant Physiol Biochem **43:** 611-622

Bailey BA, Strem MD, Bae H, de Mayolo GA, Guiltinan MJ (2005) Gene expression in leaves of Theobroma cacao in response to mechanical wounding, ethylene, and/or methyl jasmonate. Plant Science **168:** 1247-1258

Balbi V, Devoto A (2008) Jasmonate signalling network in Arabidopsis thaliana: crucial regulatory nodes and new physiological scenarios. New Phytol **177:** 301-318

Baldwin AS, Jr. (1996) The NF-kappa B and I kappa B proteins: new discoveries and insights. Annu Rev Immunol **14:** 649-683

Baldwin IT (1998) Jasmonate-induced responses are costly but benefit plants under attack in native populations. Proc Natl Acad Sci U S A **95:** 8113-8118

Beckers GJ, Spoel SH (2006) Fine-Tuning Plant Defence Signalling: Salicylate versus Jasmonate. Plant Biol (Stuttg) **8:** 1-10

Bennett AB (2003) Out of the Amazon: Theobroma cacao enters the genomic era. Trends Plant Sci **8:** 561-563

Blanco F, Salinas P, Cecchini NM, Jordana X, Van Hummelen P, Alvarez ME, Holuigue L (2009) Early genomic responses to salicylic acid in Arabidopsis. Plant Mol Biol **70:** 79-102

Borrone JW, Kuhn DN, Schnell RJ (2004) Isolation, characterization, and development of WRKY genes as useful genetic markers in Theobroma cacao. Theor Appl Genet **109:** 495-507

Bostock RM (2005) Signal crosstalk and induced resistance: straddling the line between cost and benefit. Annu Rev Phytopathol **43:** 545-580

Bowling SA, Guo A, Cao H, Gordon AS, Klessig DF, Dong X (1994) A mutation in Arabidopsis that leads to constitutive expression of systemic acquired resistance. Plant Cell **6:** 1845-1857

Boyle P, Le Su E, Rochon A, Shearer HL, Murmu J, Chu JY, Fobert PR, Despres C (2009) The BTB/POZ domain of the Arabidopsis disease resistance protein NPR1 interacts with the repression domain of TGA2 to negate its function. Plant Cell **21:** 3700-3713

Bracha-Drori K, Shichrur K, Katz A, Oliva M, Angelovici R, Yalovsky S, Ohad N (2004) Detection of protein-protein interactions in plants using bimolecular fluorescence complementation. Plant J **40:** 419-427

Brechenmacher L, Weidmann S, van Tuinen D, Chatagnier O, Gianinazzi S, Franken P, Gianinazzi-Pearson V (2004) Expression profiling of up-regulated plant and fungal genes in early and late stages of Medicago truncatula-Glomus mosseae interactions. Mycorrhiza **14:** 253-262

Buddenhagen IW (1977) Resistance and Vulnerability of Tropical Crops in Relation to Their Evolution and Breeding. Annals of the New York Academy of Sciences **287:** 309-326

Cao H, Bowling SA, Gordon AS, Dong X (1994) Characterization of an Arabidopsis mutant that is nonresponsive to inducers of systemic acquired resistance. Plant Cell **6:** 1583-1592

Cao H, Glazebrook J, Clarke JD, Volko S, Dong X (1997) The Arabidopsis NPR1 gene that controls systemic acquired resistance encodes a novel protein containing ankyrin repeats. Cell **88:** 57-63

Cao H, Li X, Dong X (1998) Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. Proc Natl Acad Sci U S A **95:** 6531-6536

Carrera J, Rodrigo G, Jaramillo A, Elena SF (2009) Reverse-engineering the Arabidopsis thaliana transcriptional network under changing environmental conditions. Genome Biology **10:** -

Chen YY, Lin YM, Chao TC, Wang JF, Liu AC, Ho FI, Cheng CP (2009) Virus-induced gene silencing reveals the involvement of ethylene-, salicylic acid- and mitogen-activated protein kinase-related defense pathways in the resistance of tomato to bacterial wilt. Physiol Plant **136:** 324-335

Chern M, Fitzgerald HA, Canlas PE, Navarre DA, Ronald PC (2005) Overexpression of a rice *NPR1* homolog leads to constitutive activation of defense response and hypersensitivity to light. Mol Plant Microbe Interact **18:** 511-520

Cipollini DF (2002) Does competition magnify the fitness costs of induced responses in Arabidopsis thaliana? A manipulative approach. Oecologia **131:** 514-520

Clarke JD, Volko SM, Ledford H, Ausubel FM, Dong X (2000) Roles of salicylic acid, jasmonic acid, and ethylene in cpr-induced resistance in arabidopsis. Plant Cell **12:** 2175-2190

Clough SJ, Bent AF (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J **16:** 735-743

Cronshaw DK, Evans HC (1978) Witches-Broom Disease of Cocoa Crinipellis-Perniciosa in Ecuador Part 2 Methods of Infection. Annals of Applied Biology **89:** 193-200

Delaney TP, Friedrich L, Ryals JA (1995) Arabidopsis signal transduction mutant defective in chemically and biologically induced disease resistance. Proc Natl Acad Sci U S A **92:** 6602-6606

Despres C, DeLong C, Glaze S, Liu E, Fobert PR (2000) The Arabidopsis NPR1/NIM1 protein enhances the DNA binding activity of a subgroup of the TGA family of bZIP transcription factors. Plant Cell **12:** 279-290

Dong X (2004) NPR1, all things considered. Curr Opin Plant Biol **7:** 547-552

Dong X, Braun EL, Grotewold E (2001) Functional conservation of plant secondary metabolic enzymes revealed by complementation of Arabidopsis flavonoid mutants with maize genes. Plant Physiol **127:** 46-57

Dubreuil-Maurizi C, Trouvelot S, Frettinger P, Pugin A, Wendehenne D, Poinssot B (2010) beta-Aminobutyric Acid Primes an NADPH Oxidase-Dependent Reactive Oxygen Species Production During Grapevine-Triggered Immunity. Mol Plant Microbe Interact **23:** 1012-1021

Durrant WE, Dong X (2004) Systemic acquired resistance. Annu Rev Phytopathol **42:** 185-209

Durrant WE, Wang S, Dong X (2007) Arabidopsis SNI1 and RAD51D regulate both gene transcription and DNA recombination during the defense response. Proc Natl Acad Sci U S A **104:** 4223-4227

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:** 1792-1797

Endah R, Beyene G, Kiggundu A, van den Berg N, Schluter U, Kunert K, Chikwamba R (2008) Elicitor and *Fusarium*-induced expression of *NPR1-like* genes in banana. Plant Physiol Biochem **46:** 1007-1014

Evans H, Priori C (1987) Cocoa pod diseases: causal agents and control. Outlook on Agriculture **16:** 35-41

Evans HC (1978) Witches-Broom Disease of Cocoa Crinipellis-Perniciosa in Ecuador Part 1 the Fungus. Annals of Applied Biology **89:** 185-192

Evans HC (2007) Cacao diseases-the trilogy revisited. Phytopathology **97:** 1640-1643

Faleiro F, Queiroz V, Lopes U, Guimarães C, Pires J, Yamada M, Araújo I, Pereira M, Schnell R, Filho G, Ferreira C, Barros E, Moreira M (2006) Mapping QTLs for witches' broom (*Crinipellis Perniciosa*) resistance in cacao (*Theobroma Cacao L.*). Euphytica **149:** 227-235

Fan W, Dong X (2002) In vivo interaction between NPR1 and transcription factor TGA2 leads to salicylic acid-mediated gene activation in Arabidopsis. Plant Cell **14:** 1377-1389

Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, Chen L, Yu L, Iglesias-Pedraz JM, Kircher S, Schafer E, Fu X, Fan LM, Deng XW (2008) Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. Nature **451:** 475-479

Feys BJ, Parker JE (2000) Interplay of signaling pathways in plant disease resistance. Trends Genet **16:** 449-455

Figueroa P, Gusmaroli G, Serino G, Habashi J, Ma L, Shen Y, Feng S, Bostick M, Callis J, Hellmann H, Deng XW (2005) Arabidopsis Has Two Redundant Cullin3 Proteins That Are Essential for Embryo

Development and That Interact with RBX1 and BTB Proteins to Form Multisubunit E3 Ubiquitin Ligase Complexes in Vivo
10.1105/tpc.105.031989. Plant Cell **17:** 1180-1195

Fitzgerald HA, Chern MS, Navarre R, Ronald PC (2004) Overexpression of *(At)NPR1* in rice leads to a BTH- and environment-induced lesion-mimic/cell death phenotype. Mol Plant Microbe Interact **17:** 140-151

Gaffney T, Friedrich L, Vernooij B, Negrotto D, Nye G, Uknes S, Ward E, Kessmann H, Ryals J (1993) Requirement of salicylic acid for the induction of systemic acquired resistance. Science **261:** 754-756

Gesteira AS, Micheli F, Carels N, Da Silva AC, Gramacho KP, Schuster I, Macedo JN, Pereira GAG, Cascardo JCM (2007) Comparative analysis of expressed genes from cacao meristems infected by Moniliophthora perniciosa. Annals of Botany **100:** 129-140

Ghosh S, May MJ, Kopp EB (1998) NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses. Annu Rev Immunol **16:** 225-260

Glazebrook J (2001) Genes controlling expression of defense responses in Arabidopsis--2001 status. Curr Opin Plant Biol **4:** 301-308

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

Glazebrook J, Rogers EE, Ausubel FM (1996) Isolation of Arabidopsis mutants with enhanced disease susceptibility by direct screening. Genetics **143:** 973-982

Grant M, Lamb C (2006) Systemic immunity. Curr Opin Plant Biol **9:** 414-420

Guiltinan M, Verica J, Zhang D, Figueira A (2008) Genomics of Theobroma cacao, "The foods of the Gods". *In* P.M. Moore, R Ming, eds, Genomics of Tropical Crop Plants. New York: Springer

Ha CM, Jun JH, Nam HG, Fletcher JC (2004) BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana. Plant Cell Physiol **45:** 1361-1370

Ha CM, Jun JH, Nam HG, Fletcher JC (2007) BLADE-ON-PETIOLE1 and 2 Control Arabidopsis Lateral Organ Fate through Regulation of LOB Domain and Adaxial-Abaxial Polarity Genes. Plant Cell

Ha CM, Kim GT, Kim BC, Jun JH, Soh MS, Ueno Y, Machida Y, Tsukaya H, Nam HG (2003) The BLADE-ON-PETIOLE 1 gene controls leaf pattern formation through the modulation of meristematic activity in Arabidopsis. Development **130:** 161-172

Hajdukiewicz P, Svab Z, Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. Plant Mol Biol **25:** 989-994

Heath MC (2000) Hypersensitive response-related death. Plant Mol Biol **44:** 321-334

Hebbar PK (2007) Cacao diseases: a global perspective from an industry point of view. Phytopathology **97:** 1658-1663

Heidel AJ, Clarke JD, Antonovics J, Dong X (2004) Fitness Costs of Mutations Affecting the Systemic Acquired Resistance Pathway in Arabidopsis thaliana. Genetics **168:** 2197-2206

Heil M (2002) Ecological costs of induced resistance. Curr Opin Plant Biol **5:** 345-350

Heil M, Hilpert A, Kaiser W, Linsenmair KE (2000) Reduced Growth and Seed Set Following Chemical Induction of Pathogen Defence: Does Systemic Acquired Resistance (SAR) Incur Allocation Costs? Journal of Ecology **88:** 645-654

Hepworth SR, Zhang Y, McKim S, Li X, Haughn GW (2005) BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. Plant Cell **17:** 1434-1448

Higo K, Ugawa Y, Iwamoto M, Korenaga T (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res **27:** 297-300

Huang J, Chen F, Del Casino C, Autino A, Shen M, Yuan S, Peng J, Shi H, Wang C, Cresti M, Li Y (2006) An ankyrin repeat-containing protein, characterized as a ubiquitin ligase, is closely associated with

membrane-enclosed organelles and required for pollen germination and pollen tube growth in lily. Plant Physiol **140:** 1374-1383

Irish VF, Yamamoto YT (1995) Conservation of floral homeotic gene function between Arabidopsis and antirrhinum. Plant Cell **7:** 1635-1644

Johnson C, Boden E, Arias J (2003) Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. Plant Cell **15:** 1846-1858

Johnson C, Mhatre A, Arias J (2008) NPR1 preferentially binds to the DNA-inactive form of *Arabidopsis* TGA2. Biochim Biophys Acta **1779:** 583-589

Jones AME, Thomas V, Bennett MH, Mansfield J, Grant M (2006) Modifications to the arabidopsis defense proteome occur prior to significant transcriptional change in response to inoculation with Pseudomonas syringae. Plant Physiology **142:** 1603-1620

Jones PG, Allaway D, Gilmour DM, Harris C, Rankin D, Retzel ER, Jones CA (2002) Gene discovery and microarray analysis of cacao (Theobroma cacao L.) varieties. Planta **216:** 255-264

Journot-Catalino N, Somssich IE, Roby D, Kroj T (2006) The transcription factors WRKY11 and WRKY17 act as negative regulators of basal resistance in Arabidopsis thaliana. Plant Cell **18:** 3289-3302

Jun JH, Ha CM, Fletcher JC (2010) BLADE-ON-PETIOLE1 Coordinates Organ Determinacy and Axial Polarity in Arabidopsis by Directly Activating ASYMMETRIC LEAVES2. Plant Cell **22:** 62-76

Katagiri F, Thilmony R, He S (2002) The Arabidopsis thaliana-Pseudomonas syringae interaction. *In* The Arabidopsis Book. The American Society of Plant Biologists

Kessler A, Baldwin IT (2002) Plant responses to insect herbivory: the emerging molecular analysis. Annu Rev Plant Biol **53:** 299-328

Kinkema M, Fan W, Dong X (2000) Nuclear localization of NPR1 is required for activation of *PR* gene expression. Plant Cell **12:** 2339-2350

Koornneef A, Pieterse CM (2008) Cross talk in defense signaling. Plant Physiol **146:** 839-844

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5:** 150-163

Kunkel BN, Brooks DM (2002) Cross talk between signaling pathways in pathogen defense. Curr Opin Plant Biol **5:** 325-331

Laurie-Berry N, Joardar V, Street IH, Kunkel BN (2006) The Arabidopsis thaliana JASMONATE INSENSITIVE 1 gene is required for suppression of salicylic acid-dependent defenses during infection by Pseudomonas syringae. Mol Plant Microbe Interact **19:** 789-800

Le Henanff G, Heitz T, Mestre P, Mutterer J, Walter B, Chong J (2009) Characterization of Vitis vinifera NPR1 homologs involved in the regulation of pathogenesis-related gene expression. BMC Plant Biol **9:** 54

Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E (1998) Functional analysis of regulatory sequences controlling *PR-1* gene expression in Arabidopsis. Plant J **16:** 223-233

Lee J, Nam J, Park HC, Na G, Miura K, Jin JB, Yoo CY, Baek D, Kim DH, Jeong JC, Kim D, Lee SY, Salt DE, Mengiste T, Gong Q, Ma S, Bohnert HJ, Kwak SS, Bressan RA, Hasegawa PM, Yun DJ (2007) Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. Plant J **49:** 79-90

Lee JJ, Woodward AW, Chen ZJ (2007) Gene expression changes and early events in cotton fibre development. Ann Bot **100:** 1391-1401

Lee S, Hong JC, Jeon WB, Chung YS, Sung S, Choi D, Joung YH, Oh BJ (2009) The salicylic acid-induced protection of non-climacteric unripe pepper fruit against Colletotrichum gloeosporioides is similar to the resistance of ripe fruit. Plant Cell Rep **28:** 1573-1580

Leon-Reyes A, Spoel SH, De Lange ES, Abe H, Kobayashi M, Tsuda S, Millenaar FF, Welschen RA, Ritsema T, Pieterse CM (2009) Ethylene modulates the role of NONEXPRESSOR OF

PATHOGENESIS-RELATED GENES1 in cross talk between salicylate and jasmonate signaling. Plant Physiol **149:** 1797-1809

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

Li J, Mahajan A, Tsai MD (2006) Ankyrin repeat: a unique motif mediating protein-protein interactions. Biochemistry **45:** 15168-15178

Li X, Zhang Y, Clarke JD, Li Y, Dong X (1999) Identification and cloning of a negative regulator of systemic acquired resistance, SNI1, through a screen for suppressors of npr1-1. Cell **98:** 329-339

Lin JJ (1995) Electrotransformation of Agrobacterium. Methods Mol Biol **47:** 171-178

Lin WC, Lu CF, Wu JW, Cheng ML, Lin YM, Yang NS, Black L, Green SK, Wang JF, Cheng CP (2004) Transgenic tomato plants expressing the Arabidopsis NPR1 gene display enhanced resistance to a spectrum of fungal and bacterial diseases. Transgenic Res **13:** 567-581

Liu G, Holub EB, Alonso JM, Ecker JR, Fobert PR (2005) An Arabidopsis NPR1-like gene, NPR4, is required for disease resistance. Plant J **41:** 304-318

Liu Y, Schiff M, Marathe R, Dinesh-Kumar SP (2002) Tobacco Rar1, EDS1 and NPR1/NIM1 like genes are required for N-mediated resistance to tobacco mosaic virus. Plant J **30:** 415-429

Loake G, Grant M (2007) Salicylic acid in plant defence--the players and protagonists. Curr Opin Plant Biol **10:** 466-472

Loon LCv, Strien EAv (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. Physiol.Mol.Plant Pathol. **55:** 85-97

Lopes MA, Hora Junior BT, Dias CV, Santos GC, Gramacho KP, Cascardo JCM, Gesteira AS, Micheli F (2010) Expression analysis of transcription factors from the interaction between cacao and Moniliophthora perniciosa (Tricholomataceae). Genetics and Molecular Research **9:** 1279-1297

Maizel A, Busch MA, Tanahashi T, Perkovic J, Kato M, Hasebe M, Weigel D (2005) The floral regulator LEAFY evolves by substitutions in the DNA binding domain. Science **308:** 260-263

Makandar R, Essig JS, Schapaugh MA, Trick HN, Shah J (2006) Genetically engineered resistance to Fusarium head blight in wheat by expression of Arabidopsis NPR1. Mol Plant Microbe Interact **19:** 123-129

Malnoy M, Jin Q, Borejsza-Wysocka EE, He SY, Aldwinckle HS (2007) Overexpression of the apple MpNPR1 gene confers increased disease resistance in Malus x domestica. Mol Plant Microbe Interact **20:** 1568-1580

Marelli J-P, Maximova S, Gramacho K, Kang S, Guiltinan M (2009) Infection Biology of Moniliophthora perniciosa on Theobroma cacao and Alternate Solanaceous Hosts. Tropical Plant Biology **2:** 149-160

Matsuura H, Ishibashi Y, Shinmyo A, Kanaya S, Kato K (2010) Genome-wide analyses of early translational responses to elevated temperature and high salinity in Arabidopsis thaliana. Plant Cell Physiol **51:** 448-462

Maximova S, Miller C, Antunez de Mayolo G, Pishak S, Young A, Guiltinan MJ (2003) Stable transformation of *Theobroma cacao L.* and influence of matrix attachment regions on GFP expression. Plant Cell Rep **21:** 872-883

Maximova SN, Marelli JP, Young A, Pishak S, Verica JA, Guiltinan MJ (2006) Over-expression of a cacao class I chitinase gene in Theobroma cacao L. enhances resistance against the pathogen, Colletotrichum gloeosporioides. Planta **224:** 740-749

McDowell JM, Cuzick A, Can C, Beynon J, Dangl JL, Holub EB (2000) Downy mildew (Peronospora parasitica) resistance genes in Arabidopsis vary in functional requirements for NDR1, EDS1, NPR1 and salicylic acid accumulation. Plant J **22:** 523-529

McKim SM, Stenvik GE, Butenko MA, Kristiansen W, Cho SK, Hepworth SR, Aalen RB, Haughn GW (2008) The BLADE-ON-PETIOLE genes are essential for abscission zone formation in Arabidopsis. Development **135:** 1537-1546

McNeil CL (2006) Chocolate in Mesoamerica : a cultural history of cacao. University Press of Florida, Gainesville

Meinhardt LW, Rincones J, Bailey BA, Aime MC, Griffith GW, Zhang D, Pereira GA (2008) Moniliophthora perniciosa, the causal agent of witches' broom disease of cacao: what's new from this old foe? Mol Plant Pathol **9:** 577-588

Melnick RL, Zidack NK, Bailey BA, Maximova SN, Guiltinan M, Backman PA (2008) Bacterial endophytes: *Bacillus* spp. from annual crops as potential biological control agents of black pod rot of cacao. Biological Control **46:** 46-56

Mitsuhara I, Ugaki M, Hirochika H, Ohshima M, Murakami T, Gotoh Y, Katayose Y, Nakamura S, Honkura R, Nishimiya S, Ueno K, Mochizuki A, Tanimoto H, Tsugawa H, Otsuki Y, Ohashi Y (1996) Efficient promoter cassettes for enhanced expression of foreign genes in dicotyledonous and monocotyledonous plants. Plant Cell Physiol **37:** 49-59

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

Mondego JM, Carazzolle MF, Costa GG, Formighieri EF, Parizzi LP, Rincones J, Cotomacci C, Carraro DM, Cunha AF, Carrer H, Vidal RO, Estrela RC, Garcia O, Thomazella DP, de Oliveira BV, Pires AB, Rio MC, Araujo MR, de Moraes MH, Castro LA, Gramacho KP, Goncalves MS, Neto JP, Neto AG, Barbosa LV, Guiltinan MJ, Bailey BA, Meinhardt LW, Cascardo JC, Pereira GA (2008) A genome survey of Moniliophthora perniciosa gives new insights into Witches' Broom Disease of cacao. BMC Genomics **9:** 548

Motamayor JC, Risterucci AM, Lopez PA, Ortiz CF, Moreno A, Lanaud C (2002) Cacao domestication I: the origin of the cacao cultivated by the Mayas. Heredity **89:** 380-386

Mou Z, Fan W, Dong X (2003) Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. Cell **113:** 935-944

Mukhtar MS, Nishimura MT, Dangl J (2009) NPR1 in plant defense: It's not over 'til It's turned over. Cell **137:** 804-806

Mur LA, Kenton P, Lloyd AJ, Ougham H, Prats E (2008) The hypersensitive response; the centenary is upon us but how much do we know? J Exp Bot **59:** 501-520

Nawrath C, Metraux JP (1999) Salicylic acid induction-deficient mutants of Arabidopsis express PR-2 and PR-5 and accumulate high levels of camalexin after pathogen inoculation. Plant Cell **11:** 1393-1404

Ndamukong I, Abdallat AA, 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

Norberg M, Holmlund M, Nilsson O (2005) The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. Development **132:** 2203-2213

Penninckx IA, Thomma BP, Buchala A, Metraux JP, Broekaert WF (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

Phillips-Mora W, Wilkinson MJ (2007) Frosty pod of cacao: a disease with a limited geographic range but unlimited potential for damage. Phytopathology **97:** 1644-1647

Pieterse CM, Van Loon LC (2004) NPR1: the spider in the web of induced resistance signaling pathways. Curr Opin Plant Biol **7:** 456-464

Pieterse CM, van Wees SC, Hoffland E, van Pelt JA, van Loon LC (1996) Systemic resistance in Arabidopsis induced by biocontrol bacteria is independent of salicylic acid accumulation and pathogenesis-related gene expression. Plant Cell **8:** 1225-1237

Ploetz RC (2007) Cacao diseases: important threats to chocolate production worldwide. Phytopathology **97:** 1634-1639

Poelman EH, Broekgaarden C, Van Loon JJ, Dicke M (2008) Early season herbivore differentially affects plant defence responses to subsequently colonizing herbivores and their abundance in the field. Mol Ecol **17:** 3352-3365

Pogany M, von Rad U, Grun S, Dongo A, Pintye A, Simoneau P, Bahnweg G, Kiss L, Barna B, Durner J (2009) Dual roles of reactive oxygen species and NADPH oxidase RBOHD in an Arabidopsis-Alternaria pathosystem. Plant Physiol **151:** 1459-1475

Polesani M, Bortesi L, Ferrarini A, Zamboni A, Fasoli M, Zadra C, Lovato A, Pezzotti M, Delledonne M, Polverari A (2010) General and species-specific transcriptional responses to downy mildew infection in a susceptible (Vitis vinifera) and a resistant (V. riparia) grapevine species. BMC Genomics **11:** 117

Purdy LH, Schmidt RA (1996) STATUS OF CACAO WITCHES' BROOM: biology, epidemiology, and management. Annu Rev Phytopathol **34:** 573-594

Queiroz VT, Guimaraes CT, Anhert D, Schuster I, Daher RT, Pereira MG, Miranda VRM, Loguercio LL, Barros EG, Moreira MA (2003) Identification of a major QTL in cocoa (Theobroma cacao L.) associated with resistance to witches' broom disease. Plant Breeding **122:** 268-272

Racki LR, Yorgey P, Dewdney J, Ausubel FM (2003) NPR2: A Novel Arabidopsis Defense Response Gene. Poster, Plant Biology 2003, Honolulu, Hawaii USA

Rao MV, Paliyath G, Ormrod DP, Murr DP, Watkins CB (1997) Influence of salicylic acid on H2O2 production, oxidative stress, and H2O2-metabolizing enzymes. Salicylic acid-mediated oxidative damage requires H2O2. Plant Physiol **115:** 137-149

Rayapuram C, Baldwin IT (2007) Increased SA in NPR1-silenced plants antagonizes JA and JA-dependent direct and indirect defenses in herbivore-attacked Nicotiana attenuata in nature. Plant J **52:** 700-715

Rochon A, Boyle P, Wignes T, Fobert PR, Despres C (2006) The coactivator function of Arabidopsis NPR1 requires the core of its BTB/POZ domain and the oxidation of C-terminal cysteines. Plant Cell **18:** 3670-3685

Rombauts S, Dehais P, Van Montagu M, Rouze P (1999) PlantCARE, a plant cis-acting regulatory element database. Nucleic Acids Res **27:** 295-296

Ryals J, Weymann K, Lawton K, Friedrich L, Ellis D, Steiner HY, Johnson J, Delaney TP, Jesse T, Vos P, Uknes S (1997) The Arabidopsis NIM1 protein shows homology to the mammalian transcription factor inhibitor I kappa B. Plant Cell **9:** 425-439

Sakamoto K, Tada Y, Yokozeki Y, Akagi H, Hayashi N, Fujimura T, Ichikawa N (1999) Chemical induction of disease resistance in rice is correlated with the expression of a gene encoding a nucleotide binding site and leucine-rich repeats. Plant Mol Biol **40:** 847-855

Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W, Amores FM, Motamayor JC (2007) Development of a marker assisted selection program for cacao. Phytopathology **97:** 1664-1669

Scott RJ, Spielman M, Dickinson HG (2004) Stamen structure and function. Plant Cell **16 Suppl:** S46-60

Shah J (2003) The salicylic acid loop in plant defense. Curr Opin Plant Biol **6:** 365-371

Shah J, Tsui F, Klessig DF (1997) Characterization of a salicylic acid-insensitive mutant (sai1) of Arabidopsis thaliana, identified in a selective screen utilizing the SA-inducible expression of the tms2 gene. Mol Plant Microbe Interact **10:** 69-78

Shimono M, Sugano S, Nakayama A, Jiang CJ, Ono K, Toki S, Takatsuji H (2007) Rice WRKY45 plays a crucial role in benzothiadiazole-inducible blast resistance. Plant Cell **19:** 2064-2076

Smyth DR, Bowman JL, Meyerowitz EM (1990) Early flower development in Arabidopsis. Plant Cell **2:** 755-767

Spoel SH, Dong X (2008) Making sense of hormone crosstalk during plant immune responses. Cell Host Microbe **3:** 348-351

Spoel SH, Johnson JS, Dong X (2007) Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. Proc Natl Acad Sci U S A **104:** 18842-18847

Spoel SH, Koornneef A, Claessens SM, Korzelius JP, Van Pelt JA, Mueller MJ, Buchala AJ, Metraux JP, Brown R, Kazan K, Van Loon LC, Dong X, Pieterse CM (2003) NPR1 modulates cross-talk between salicylate- and jasmonate-dependent defense pathways through a novel function in the cytosol. Plant Cell **15:** 760-770

Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P, Dong X (2009) Proteasome-mediated turnover of the transcription coactivator NPR1 plays dual roles in regulating plant immunity. Cell **137:** 860-872

Sticher L, Mauch-Mani B, Metraux JP (1997) Systemic acquired resistance. Annu Rev Phytopathol **35:** 235-270

Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG (2005) Sequence and structural analysis of BTB domain proteins. Genome Biol **6:** R82

Stout MJ, Thaler JS, Thomma BP (2006) Plant-mediated interactions between pathogenic microorganisms and herbivorous arthropods. Annu Rev Entomol **51:** 663-689

Swanson JD, Carlson JE, Guiltinan MJ (2008) Comparative Flower Development in Theobroma Cacao Based on Temporal Morphological Indicators. International Journal of Plant Sciences **169:** 1187-1199

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol **24:** 1596-1599

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:** 4673-4680

Tiburcio RA, Costa GG, Carazzolle MF, Mondego JM, Schuster SC, Carlson JE, Guiltinan MJ, Bailey BA, Mieczkowski P, Meinhardt LW, Pereira GA (2010) Genes acquired by horizontal transfer are potentially involved in the evolution of phytopathogenicity in Moniliophthora perniciosa and Moniliophthora roreri, two of the major pathogens of cacao. J Mol Evol **70:** 85-97

Torres MA, Dangl JL (2005) Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. Curr Opin Plant Biol **8:** 397-403

Uppalapati SR, Ishiga Y, Wangdi T, Kunkel BN, Anand A, Mysore KS, Bender CL (2007) The phytotoxin coronatine contributes to pathogen fitness and is required for suppression of salicylic acid accumulation in tomato inoculated with Pseudomonas syringae pv. tomato DC3000. Mol Plant Microbe Interact **20:** 955-965

Uquillas C, Letelier I, Blanco F, Jordana X, Holuigue L (2004) NPR1-independent activation of immediate early salicylic acid-responsive genes in Arabidopsis. Mol Plant Microbe Interact **17:** 34-42

Valdes-Lopez O, Hernandez G (2008) Transcriptional regulation and signaling in phosphorus starvation: what about legumes? J Integr Plant Biol **50:** 1213-1222

Van der Ent S, Verhagen BW, Van Doorn R, Bakker D, Verlaan MG, Pel MJ, Joosten RG, Proveniers MC, Van Loon LC, Ton J, Pieterse CM (2008) MYB72 is required in early signaling steps of rhizobacteria-induced systemic resistance in Arabidopsis. Plant Physiol **146:** 1293-1304

van Hulten M, Pelser M, van Loon LC, Pieterse CM, Ton J (2006) Costs and benefits of priming for defense in Arabidopsis. Proc Natl Acad Sci U S A **103:** 5602-5607

van Wees SC, Glazebrook J (2003) Loss of non-host resistance of Arabidopsis NahG to Pseudomonas syringae pv. phaseolicola is due to degradation products of salicylic acid. Plant J **33:** 733-742

Verica JA, Maximova SN, Strem MD, Carlson JE, Bailey BA, Guiltinan MJ (2004) Isolation of ESTs from cacao (*Theobroma cacao L.*) leaves treated with inducers of the defense response. Plant Cell Rep **23:** 404-413

Vlot AC, Klessig DF, Park SW (2008) Systemic acquired resistance: the elusive signal(s). Curr Opin Plant Biol **11:** 436-442

Voelckel C, Borevitz JO, Kramer EM, Hodges SA (2010) Within and between whorls: comparative transcriptional profiling of Aquilegia and Arabidopsis. PLoS One **5:** e9735

von Arnim AG, Deng XW (1994) Light inactivation of Arabidopsis photomorphogenic repressor COP1 involves a cell-specific regulation of its nucleocytoplasmic partitioning. Cell **79:** 1035-1045

Walter M, Chaban C, Schutze K, Batistic O, Weckermann K, Nake C, Blazevic D, Grefen C, Schumacher K, Oecking C, Harter K, Kudla J (2004) Visualization of protein interactions in living plant cells using bimolecular fluorescence complementation. Plant J **40:** 428-438

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 Pathog **2:** e123

Watling R (2002) Tropical mycology. CABI, Wallingford

Wood G, Lass R (1985) Cocoa. Longman Scientific & Technical Press, New York

Wood GAR, Lass RA (1985) Cocoa, Ed 4th. Longman, London

Yang C, Bolotin E, Jiang T, Sladek FM, Martinez E (2007) Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. Gene **389:** 52-65

Yang P, Chen C, Wang Z, Fan B, Chen Z (1999) A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. The plant journal **18:** 141 - 149

Yang P, Chen C, Wang Z, Fan B, Chen Z (1999) A pathogen- and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. Plant J **18:** 141 - 149

Yang Y, Qi M, Mei C (2004) Endogenous salicylic acid protects rice plants from oxidative damage caused by aging as well as biotic and abiotic stress. Plant J **40:** 909-919

Yu D, Chen C, Chen Z (2001) Evidence for an important role of WRKY DNA binding proteins in the regulation of NPR1 gene expression. Plant Cell **13:** 1527-1540

Yuan Y, Zhong S, Li Q, Zhu Z, Lou Y, Wang L, Wang J, Wang M, Yang D, He Z (2007) Functional analysis of rice NPR1-like genes reveals that OsNPR1/NH1 is the rice orthologue conferring disease resistance with enhanced herbivore susceptibility. Plant Biotechnol J **5:** 313-324

Zhang HZ, Cai XZ (2005) Nonexpressor of pathogenesis-related genes 1 (NPR1): a key node of plant disease resistance signalling network. Sheng Wu Gong Cheng Xue Bao **21:** 511-515

Zhang S, Klessig DF (2001) MAPK cascades in plant defense signaling. Trends Plant Sci **6:** 520-527

Zhang X, Chen S, Mou Z (2010) Nuclear localization of NPR1 is required for regulation of salicylate tolerance, *isochorismate synthase 1* expression and salicylate accumulation in *Arabidopsis*. J Plant Physiol **167:** 144-148

Zhang Y, Cheng YT, Qu N, Zhao Q, Bi D, Li X (2006) Negative regulation of defense responses in Arabidopsis by two NPR1 paralogs. Plant J **48:** 647-656

Zhang Y, Fan W, Kinkema M, Li X, Dong X (1999) Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the *PR-1* gene. Proc Natl Acad Sci U S A **96:** 6523-6528

Zhang Y, McCormick S (2007) A distinct mechanism regulating a pollen-specific guanine nucleotide exchange factor for the small GTPase Rop in Arabidopsis thaliana. Proc Natl Acad Sci U S A **104:** 18830-18835

Zhang Y, Tessaro MJ, 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 X, Cheng C, Gao Q, Liu J, Guo X (2008) Molecular cloning and characterization of *GhNPR1*, a gene implicated in pathogen responses from cotton (*Gossypium hirsutum L.*). Biosci Rep **28:** 7-14

Zhao X, Huang J, Yu H, Wang L, Xie W (2010) Genomic survey, characterization and expression profile analysis of the peptide transporter family in rice (Oryza sativa L.). BMC Plant Biol **10:** 92

Zhou JM, Trifa Y, Silva H, Pontier D, Lam E, Shah J, Klessig DF (2000) NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the *PR-1* gene required for induction by salicylic acid. Mol Plant Microbe Interact **13:** 191-202

Zimmermann P, Hennig L, Gruissem W (2005) Gene-expression analysis and network discovery using Genevestigator. Trends Plant Sci **10:** 407-409

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

APPENDIX: SEQUENCES OF THE THEOBROMA CACAO GENES CLONED IN THIS THESIS

>gi|297186123|gb|HM117159.1| Theobroma cacao cultivar Scavina6 Non-expressor of PR1 (NPR1) gene, complete cds

ATGGATAACAGAAATGGCTTTTCGGATTCAAACGAAATCAGCAACAACAGCAGCACGTGCTGCA
TCGCGGCAGCAACCAACAGCGAAACGCTCGCTTCATCAGAACCGTTGAACACTCCCGACATAGCA
GCTCTTCAAATTCTTTCCAGAAACCTCGAGTCCGTTTTCGAATCAACGGACTCCGATTCCTTATA
CTCCGATGCCAAAATTGGGCTTTCCTCGGGCCGTGAAGTCCCCGTCCACCGCTGCATTTTATCGG
CGAGGAGTTCCGTTTTCAAGACCGTGTTTTCGGGGCTAAAAGATAGAGGAGCTAAGTTCGAGCT
GAAAGAGTTGGCTAGGGACTATGAGATTGGCTATGATTCTCTTGTTGCGGTTCTGGCTTACTTGT
ATAGTGGAAGAGTGAGATCGTTGCCGAGAGGCGTTTGCGTTTGCGTTGACGATGATTGCTCCCAC
TTGGCTTGCAGGCCGGCTGTTGATTTCGTGGCTGAAGTTTTATACGCAGCTTTTACTTTTCAGGT
CTCTGAATTGATTTCCCTTTATCAGAGGCATCTTTTAGACATTATCGACAAGGTTGAAATGGATG
ATATATTGGTGGTTCTTTATGTTGCAAATATGTGCGGTAATACTTGTGAAAGACTGCTGGCAAA
GTGTATAGAGACTCTTGTAAAATCTGATGTTGATATTGTAACACTTGACAAGGCCTTGCCTTATC
ACATTGTGAAACAAATCATGGATTCCCGTCTGGAACTTGGTTTGGACAAGCCTGAGAACACAGG
TTTTCCTGATAAACATGTGAAGAGGATACATCGTGCTTTGGATTCGGATGATGTTGAATTAGCG
AGAATGCTACTGAAAGAGGGTCATACTAATTTAGATGAGGCAAGTGCACTTCACTATGCCGTGG
CATACTGTGATGCGAAGACCACAACTGAATTACTTGACCTTGGACTTGCTGATGTTAACCGTAGA
AACTCAAGGGGATACACTGTGTTGCATGTTGCGGCAATGAGGAAAGAGCCTAAGATTATAGTAT
CTCTTTTAACCAAAGGTGCCCGGCCATCTGATCTCACCCTAGATGGTAGGAAAGCTTTTCAGATC
TCAAAGCGACTCACCAGGGCTGCAGATTACTATATGTCTACTGAGGAAGGAAAGGCTTCTCCAAA
GGACCGTTTGTGCGTTGAGATACTGGAGCAGGCAGAAAGAAGAGATCCATTGCTTGGAGAAGCT
TCTCTTTCTCTTGCCATGGCTGGTGATGATCTACGGATGAAATTATTGTATCTTGAAAATAGAGT
TGGATTGGCAAAACTTCTTTTCCCCATGGAAGCAAAAGTTGCGATGGACATTGCTAAAGTGGAT
GGAACATCTGAGTTCACATTAGCCAGCATCAATTCCAACAAATTAAATGATGCCCAAAGAACAA
CTGTGGACTTGAATGAGGCACCTTTTAGAATTCAGGAGGAGCATCTAAATAGGCTGAAAGCACT
TTCTAGAACAGTGGAGCTCGGCAAACGCTTTTTCCCTCGTTGCTCAGAAGTGCTGAACAAGATCA
TGGACGCTGACGACTTATCTCAGCTAGCATGCGGAGGAAATGACACTCCAGAGGAGCGACTTGTT
AAAAAGCAAAGGTACGTGGAACTCCAAGATGTACTGAGCAAGGCATTCAATGAGGATAAAGTAG
AGTTTGACAGATCAACCATCTCATCCTCTTCTTCATCAAAGTCAATAGGGGTGAGCAGACCTAAT
GGTAAGCTAACTGGTAGTGGTAGGGGCGGTTAG

>gi|297186123|gb|HM117159.1| Theobroma cacao cultivar Scavina6 Non-expressor of PR1 (NPR1) gene, genomic DNA

ACTTTCTTTCCTCTCAGTAATCCATGGCGGTTTCTCTCAATTTCCTTTTCCTTTAGTTCCACAACT
CAAAGCCCCCTTTTCCATCGCAGAATCCTCCGTCTCTCTGAACCGTTTCTTCTTCAGCTACCAACA
GAAATGCCTTTGAGTCTCAAATAGTTTTCCTTTTTTTTTTCCTTATGGATAACAGAAATGGCTTT
TCGGATTCAAACGAAATCAGCAACAACAGCAGCACGTGCTGCATCGCGGCAGCAACCAACAGCGA
AACGCTCGCTTCATCAGAACCGTTGAACACTCCCGACATAGCAGCTCTTCAAATTCTTTCCAGAA

ACCTCGAGTCCGTTTTCGAATCAACGGACTCCGATTCCTTATACTCCGATGCCAAAATTGGGCTT
TCCTCGGGCCGTGAAGTCCCCGTCCACCGCTGCATTTTATCGGCGAGGAGTTCCGTTTTCAAGAC
CGTGTTTTCGGGGCTAAAAGATAGAGGAGCTAAGTTCGAGCTGAAAGAGTTGGCTAGGGACTAT
GAGATTGGCTATGATTCTCTTGTTGCGGTTCTGGCTTACTTGTATAGTGGAAGAGTGAGATCGT
TGCCGAGAGGCGTTTGCGTTTGCGTTGACGATGATTGCTCCCACTTGGCTTGCAGGCCGGCTGTT
GATTTCGTGGCTGAAGTTTTATACGCAGCTTTTACTTTTCAGGTCTCTGAATTGATTTCCCTTTA
TCAGGTCAGTAACTACTCATGTTTTTAGCCTTTCTCAATCATGTTTAAGTCTTAATTATATAGTA
ATACTACGAATTTGTAAATTTGTGAGGAATTATATAGTTTGGGGAAAAAAAAGTTATGATGTTA
TTTTCTTTATACATTCCAGTAATTTTTGTCAAAACTAATCCCTATAACAATTTGCAAGTTTTGAT
TTTCAGGGTTCGATTCATATTTAGTTTTGCTATTTGCTTTTTCCTTTAAAAAGCAAAAGAGTAT
GTTTTGTGGTGAACAACAGAATGTATGCTTTAAGTAATGTTTGAATGAATGTTATATTTGAGTT
GTTTGCTTATTTTAAGATTTTGCTTTCTTTTTTTTTTTAATTTGTGCTCTGATTTTGATTCCAGA
GGCATCTTTTAGACATTATCGACAAGGTTGAAATGGATGATATATTGGTGGTTCTTTATGTTGC
AAATATGTGCGGTAATACTTGTGAAAGACTGCTGGCAAAGTGTATAGAGACTCTTGTAAAATCT
GATGTTGATATTGTAACACTTGACAAGGCCTTGCCTTATCACATTGTGAAACAAATCATGGATTC
CCGTCTGGAACTTGGTTTGGACAAGCCTGAGAACACAGGTTTTCCTGATAAACATGTGAAGAGG
ATACATCGTGCTTTGGATTCGGATGATGTTGAATTAGCGAGAATGCTACTGAAAGAGGGTCATA
CTAATTTAGATGAGGCAAGTGCACTTCACTATGCCGTGGCATACTGTGATGCGAAGACCACAACT
GAATTACTTGACCTTGGACTTGCTGATGTTAACCGTAGAAACTCAAGGGGATACACTGTGTTGC
ATGTTGCGGCAATGAGGAAAGAGCCTAAGATTATAGTATCTCTTTTAACCAAAGGTGCCCGGCC
ATCTGATCTCACCCTAGATGGTAGGAAAGCTTTTCAGATCTCAAAGCGACTCACCAGGGCTGCAG
ATTACTATATGTCTACTGAGGAAGGAAAGGCTTCTCCAAAGGACCGTTTGTGCGTTGAGATACT
GGAGCAGGCAGAAAGAAGAGATCCATTGCTTGGAGAAGCTTCTCTTTCTCTTGCCATGGCTGGTG
ATGATCTACGGATGAAATTATTGTATCTTGAAAATAGAGGTAATTTTCAGATCATATTCAAATC
CTTTTCCTTTTCATGCTAGCTGCTTTATCTTGGATGGCCCTTCTCAGTTACTTTCTTCAAATGGT
CAGGTTTTTAGCCATTAGAATTCCCATTTCTTTATAGTTTCAACCTATGATTTTGCAGTAAGGAG
CCCTTTGTTTTTTCCTTTGAATCTCCAATATGTGTTTTTACCGGTGAGTAGTCATGCAGGTGATG
CTGTGGGCTTCCATCAGTCATATTTACTTGTCCAAGAAAATAGTTTTATACAGACCAAAAGTCGG
TAGAGAATAGCAACGTATTGCAATTTGTGAACCAGAATGGAAAATAACAACTTCGACGGAGTGA
ACAATGTGCTTTTGCTTTGCAAGGAAATGCTGACCACACCTATTATCATAGGAAATGTTAGACCA
ATGCAAAGGGAATTAAGGCAATTACCTGACTATTATAATAGTACTATATACTTAACACAGTGAA
ATTATTGCACACCAAAATTCTCTTTTTCTCTTGCTCTTCTAAATCAAAGTATGAAACTCCTCTCT
TCTCTCTCTCGACTACTGCAAGTTTGTCTATGCTTGATTTATTCATTCGTTCTTTGGTTTAGGTC
ATTAGTTGAAAATCTGAATGCCACATAGATGTGGTTTTTCCTTAATCACTGATCGCATTAAACA
GAGAAGAATAACAATAACTAGAACAAGAACAAACCGAGTTATTTCTCTCCGAACACATGGGGTT
GGCAATGTAGACCGCATTGTCTAACTGTGCCCTATTTATAACCCTGTCTTGAATGGTTTCAGTAA
ACGTGAAAATGAGATTTTTTTTTTTATATAGTATTAGAGAAGAATTAAAAGAAAATTAATGATA
GCATAGTTCCACTACTCTTATAGCATGTTTGGTTCGCAGAATGGAATGGAATAGAATAGAATAA
TTATTTTGTGGAAATAGAATAAGGAAGGAATAGAATAGAATGACTATTATGTAGTTTGGTTTAT
GGAATGGAATAGGGTAGGAATTGTTATTATAACTTATTTCTAATGTTTGAGTGGTAGGAATAAT
TTTGAAAAAATTAAATTAAAAAATAAATGATAAAATGACAAAATTACCTTTTATTATAATTTAG

TAAACTAGACAAAATTACTCATTATTATAATTTAATAATATTTTTCTTTTTTAATCTTTTAATT
TTATTTTCATTAGAAATTTAAAAATTTATTATTAAATTATTATTATGATTTTTTTCCAATTTCT
TTTTATTTTTTCTTTATTTGTATTTTCATTTTTTAATCTTTTAATTTCTATTTTCATTAAAAATT
TAAACATTAATAGTTTATTATTAAATTATTAATATAATTTTTTAGTTTAATTTCTTTTTACTTT
TTTCTTTATTTTCATTTTATATTTGTAGGTCGTAATTATTTTCAATAAATAAAGATATTTCTGT
CCAATAAATTTTTGACTCTATTCCATAGTCATGGAATAGTTATTATCATGGGGAGGTGGGGACT
GGTCAAGCCAAGATTTTGAGGTGTAGCCATTCTTCCTCACCTCAAAAACCAACCAAACAAGAGAA
TAAAAATGAGCAAGGAATGAGGGAGGAATGGCTTAGCTATTCCCCGTACAAAACTTGCCATTAA
TTTTCTTTTCTTTTGAATTGTAAATATTAGCTTGGTATGTTTTTGCCTTATTTTCTTCCACAATA
TCAGCATAGAGTTTCTGATCTTTCTTTATTGGATAGATAGGTAGTTTGGTTGTTGAAAAGAGAT
ATCTCTCTGTCTCTGGTGAATATGTCAATTGGTGGAAACAACAAAATTAAAAGAAAGCCTAGGC
AAATTTATGCTAGGATATCTTTCTGTTGAAAAATGTATTGTGCTAAGTCAGCGTTTCATCTCTG
GACTCAAAACCATTAACACCTTCTACTTCGTTATGAATTGTTAATGATTTAAGCCAGAGGTTCAC
TGGCCATATCATAGAGTCAGAGGATTGGGGGTTTTAAGTTTGTCCTTCTTCTCTACTGATGCGGT
CCCAAGTTATTTGCAACCCTGGAAGCATCTCTACCTATATCTAATTTTGAACTCTCCATCATATT
TCTTATTGTTGTTGAAAAGTGATGGTTATGTTATCGATTTCTTACTTACCTTACATTTTGCAGTT
GGATTGGCAAAACTTCTTTTCCCCATGGAAGCAAAAGTTGCGATGGACATTGCTAAAGTGGATG
GAACATCTGAGTTCACATTAGCCAGCATCAATTCCAACAAATTAAATGATGCCCAAAGAACAAC
TGTGGACTTGAATGAGGCACCTTTTAGAATTCAGGAGGAGCATCTAAATAGGCTGAAAGCACTT
TCTAGAACAGGTACCGACTGATGCTTTACATCAATATTTTCTCTTTTCCTGTTAATGCCCCAATT
CATTTTCTTGGGTTTCTCCTTGCCAAAGAATCACATGATGAACCAGTGTAATGTCCTGGTAAGCT
ATTTCAGGAGTTGTAGATTTTGAGCGTCAGACTAGCATATTGACATTAATTTGGAAAAATTTGA
ATAAACATCATGTTGATGAGTAACTGCATTTACTCTAAACCCCTCTCTCTGTGGCAGTGGAGCTC
GGCAAACGCTTTTTCCCTCGTTGCTCAGAAGTGCTGAACAAGATCATGGACGCTGACGACTTATC
TCAGCTAGCATGCGGAGGAAATGACACTCCAGAGGAGCGACTTGTTAAAAAGCAAAGGTACGTG
GAACTCCAAGATGTACTGAGCAAGGCATTCAATGAGGATAAAGTAGAGTTTGACAGATCAACCA
TCTCATCCTCTTCTTCATCAAAGTCAATAGGGGTGAGCAGACCTAATGGTAAGCTAACTGGTAGT
GGTAGGGGCGGTTAGACCATAGGCCTAATGCAATCTTTTGTAGTTGTATCATAATAATTCCGTT
TCTTTTTGACCTTTGGTTATTTTTTGGGTTTTTGCTTTGTATACTATAAATTAGAACATCATTGA
TCACAGGAAGATGTGGACAAAAGATACCGTTAATTTGCTTT

> Theobroma cacao cultivar criollo *NPR1* gene (TcNPR1), complete cds

ATGGATAACAGAAATGGCTTTTCGGATTCAAACGAAATCAGCAACAACAGCAGCACGTGCTGCA
TCGCGGCAGCAACCAACAGCGAAACGCTCGCTTCATCAGAACCGTTGAACACTCCCGCCATAGCA
GCTCTTCAAATTCTTTCCAGAAACCTCGAGTCCGTTTTCGAATCAACGGACTCCGATTCCTTATA
CTCCGATGCCAAAATTGGGCTTTCCTCGGGCCGTGAAGTCCCCGTCCACCGCTGCATTTTATCGG
CGAGGAGTTCCGTTTTCAAGACCGTGTTTTCGGGGCTAAAAGATAGAGGAGCTAAGTTCGAGCT
GAAAGAGTTGGCTAGGGACTATGAGATTGGCTATGATTCTCTTGTTGCGGTTCTGGCTTACTTGT
ATAGTGGAAGAGTGAGATCGTTGCCGAGAGGCGTTTGCGTTTGCGTTGACGATGATTGCTCCCAC
TTGGCTTGCAGGCCGGCTGTTGATTTTGTGGCTGAAGTTTTATACGCAGCTTTTACTTTTCAGGT
CTCTGAATTGATTTCCCTTTATCAGAGGCATCTTTTAGACATTATCGACAAGGTTGAAATGGATG

ATATATTGGTGGTTCTTTATGTTGCAAATATGTGCGGTAATACTTGTGAAAGATTGCTGGCAAA
GTGTATAGAGACTCTTGTAAAATCTGATGTTGATATTGTAACACTTGACAAGGCCTTGCCTTATC
ACATTGTGAAACAAATCATGGATTCCCGTCTGGAACTTGGTTTGGACAAGCCTGAGAACACAGG
TTTTCCTGATAAACATGTGAAGAGGATACATCGTGCTTTGGATTCGGATGACGTTGAATTAGTG
AGAATGCTACTGAAAGAGGGTCATACTAATTTAGATGAGGCAAGTGCACTTCACTATGCCGTGG
CATACTGTGATGCGAAGACCACAACTGAATTACTTGACCTTGGACTTGCTGATGTTAACCGTAGA
AACTCAAGGGGATACACTGTGTTGCATGTTGCGGCAATGAGGAAAGAGCCTAAGATTATAGTAT
CTCTTTTAACCAAAGGTGCCCGGCCATCTGATCTCACCCTAGATGGTAGGAAAGCTTTTCAGATC
TCAAAGCGACTCACCAGGGCTGCAGATTACTATATGTCTACTGAGGAAGGAAAGGCTTCTCCAAA
GGACCGTTTGTGCGTTGAGATACTGGAGCAGGCAGAAAGAAGAGATCCATTGCTTGGAGAAGCT
TCTCTTTCTCTTGCCATGGCTGGTGATGATCTACGGATGAAATTATTGTATCTTGAAAATAGAGT
TGGATTGGCAAAACTTCTTTTCCCCATGGAAGCAAAAGTTGCGATGGACATTGCTAAAGTGGAT
GGAACATCTGAGTTCACATTAGCCAGCATCAATTCCAACAAATTAAATGATGCCCAAAGAACAA
CTGTGGACTTGAATGAGGCACCTTTTAGAATTCAGGAGGAGCATCTAAATAGGCTGAAAGCACT
TTCTAGAACAGTGGAGCTCGGCAAACGCTTTTTCCCTCGTTGCTCAGAAGTGCTGAACAAGATCA
TGGACGCTGACGACTTATCTCAGCTAGCATGCGGAGGAAATGACACTCCAGAGGAGCGACTTGTT
AAAAAGCAAAGGTACGTGGAACTCCAAGATGTACTGAGCAAGGCATTCAATGAGGATAAAGTAG
AGTTTGACAGATCAACTATCTCATCTTCTTCTTCATCAAAGTCAATAGGGGTGAGCAGACCTAAT
GGTAAGCTAACTGGTAGTGGTAGGGGCGGTTAG

> Theobroma cacao cultivar Scavina6 *NPR1-like3* gene (TcNPR3), complete cds

ATGGCGTATTTATCTGAGCCATCGTCTTCTTTGAGCTTCAGTTCATCTTCACATCTATCAAATGG
CTCAATTACTCACAATATACCCTCCTTTACTGTTCCTGAAACTGGGGCCAGTCTTGAAGTTATAA
GTTTGACCAAGCTTAGCACCAGCTTGGAGCAACTCGTTAATGACAACGGTCCTGATTTTAGTGAT
GCTGATATAGTTGTTGAAGATGTTCCTGTTGGTGTTCATAGATGTATTTTAGCTGTTAGGAGTA
AGTTTTTTAATGAGTTATTCAAGAAAGGAAATGGGTCTTGTGAGAAAGAAGGAAAACCAAGTTA
TAACATGAGTGAGTTGTTACCTTATGGCAAGATTGGACTTGAAGCTTTTCGGATTCTCTTGCATT
ATTTGTACACTGGGAAGCTGAGGCCTTCTCCCATGGAGGTTTCAACATGTGTTGATAATGTTTGT
GCTCATGATGCTTGTCGACCCGCCATCAATTTTGCCGTGGAGTTGATGTATGCATCATCGATATT
TCAAATACCGGAGCTGGTTTCACTTTTTCAGCGACGTCTCCTTAATTTTGTTGAGAAGGCTCTTG
TAGAAGATATTATCACAATCCTTGTGGTTGCTTTCCATTGTCAATGCAGTCAGCTTGTTTCTCAA
TGTGTTGATAGAGTGGCACGGTCAGATCTTGATAGCATCTCTATTGAGAAAGAGCTTCCTTATG
AAGTTGCGGAGAGTATTCGATTGCTTCGCCGCAAGTCTCCACCTGATGGTGAAGATAATGAGGCA
GTGGTTGACCCTTTGCGAGAAAAAAGAATCCGGAGAATACATAAAGCATTAGACTCTGACGATG
TTGAACTTGTTAAACTACTTTTAACTGAGTCTGACATAACTTTGGATGATGCTGCTGCCCTCCAT
TATGCTGCAGCATATTGTGATCCCAAGGTTGTCTCTGAGGTTCTTGGCCTGCGTCTGGCTGATGT
CAATCTGCGAAATTCTCGTGGTTACACAGTTCTTCACATCGCTGCAATGCGTAAAGAACCATCAG
TGATAATGTCTCTTCTGGCAAAAGGGGCCTCTGCTTCAGAGTTGACAGTGGATGGACGAAGTGC

AGTTAACATCTGCCAGAGGTTGACAAGACCAAAGGATTATCATGCAAAAACAGAGCAAGGAAAG
GAAACAAATAAAGACAGGATATGCATTGATGTTCTAGAAAGGGAAATGAGGAGGAATCCAATGG
CCGGGGATGTGTCTGTTACTTCCCATACGTTGGCTGATGATCTACACATGAGACTTCTGTATCTA
GAGAACAGAGTGGCTTTTGCAAGGTTACTTTTCCCTAGTGAAGCAAAACTTGCTATGGACATAG
CACATGCTGAGACAACCTCTGAGTTTGCTGGTCTTTGTGCATCAAAAGGTTCAAATGGAAACTTA
AGGCAGGTTGACTTGAATGAGACACCCATTATGCAGAAGAAAAGACTTCTTGCAAGGATGGAGG
CCCTCATGAAAACAGTGGAGATGGGTCGGCGCTATTTCCCTCATTGCTCAGAAGTTCTAGATAAG
TTCATGGAGGATGACCTTCCTGATTTATTTTACCTTGAGAAGGGCTCTTCAGAGGAGCAGAAAA
TAAAGAGATCACGTTTCAGGGAGCTGAAGGATGATGTTCAGAAGGCATTTAGTAAGGACAAGGC
TGAGTTTAACCGCACTGGTTTGTCTTCATCCTCATCATCATCATCTTTAAAAGATGGTGGCCCTT
ACAAGCTCAGAAAATTGTGA

> Theobroma cacao cultivar criollo *NPR1-like3* gene (TcNPR3), complete cds

ATGGCGTATTTATCTGAGCCATCGTCTTCTTTGAGCTTCAGTTCATCTTCACATCTATCAAATGG
CTCAATTGCTCACAATATACCCTCCTTTACTGTTCCTGAAACTGGGGCCAGTCTTGAAGTTATAA
GTTTGACCAAGCTTAGCACCAGCTTGGAGCAACTCGTTAATGACAACGGTCCTGATTTTAGTGAT
GCTGATATAGTTGTTGAAGATGTTCCTGTTGGTGTTCATAGATGTATTTTAGCTGTTAGGAGTA
AGTTTTTTAATGAGTTATTCAAGAAAGGAAATGGGTCTTGTGAGAAAGAAGGAAAACCAAGTTA
TAACATGAGTGAGTTGTTACCTTATGGCAAGATTGGACTTGAAGCTTTTCGGATTCTCTTGCATT
ATTTGTACACTGGGAAGCTGAGGCCTTCTCCCATGGAGGCTTCAACATGTGTTGATAATGTTTGT
GCTCATGGTGCTTGTCGACCCGCCATCAATTTTGCCGTGGAGTTGATGTATGCATCATCGATATT
TCAAATACCGGAGCTGGTTTCACTTTTTCAGCGACGTCTCCTTAATTTTGTTGAGAAGGCTCTTG
TAGAAGATATTATCACAATCCTTGTGGTTGCTTTCCATTGTCAATGCAGTCAGCTTGTTTCTCAA
TGTGTTGATAGAGTGGCACGGTCAGATCTTGATAGCATCTCTATTGAGAAAGAGCTTCCTTATG
AAGTTGCGGAGAGTATTCGATTGCTTCGCCGCAAGTCTCCACCTGATGGTGAAGATAATGAGGCA
GTGGTTGACCCTTTGCGAGAAAAAAGAATCCGGAGAATACATAAAGCATTAGACTCTGACGATG
TTGAACTTGTTAAACTACTTTTAACTGAGTCTGACATAACTTTGGATGATGCTGCTGCCCTCCAT
TATGCTGCAGCATATTGTGATCCCAAGGTTGTCTCTGAGGTTCTTGGCCTGCGTCTGGCTGATGT
CAATCTGCGAAATTCTCATGGTTACACAGTTCTTCACATCGCTGCAATGCGTAAAGAACCATCAG
TGATAATGTCTCTTCTGGCAAAAGGGGCCTCTGCTTCAGAGTTGACAGTGGATGGACGAAGTGC
AGTTAACATCTGCCAGAGGTTGACAAGACCAAAGGATTATCATGCAAAAACAAAGCAAGGAAAG
GAAACAAATAAAGACAGGATATGCATTGATGTTCTAGAAAGGGAAATGAGGAGGAATCCAATGG
CCGGGGATGTGTCTGTTACTTCCCATACGTTGGCTGATGATCTACACATGAGACTTCTGTATCTA
GAGAACAGAGTGGCTCTTGCAAGGTTACTTTTCCCTAGTGAAGCAAAACTTGCTATGGACATAG
CACATGCTGAGACAACCTCTGAGTTTGCTGGTCTTTGTGCATCAAAAGGTTCAAATGGAAACTTA
AGGCAGGTTGACTTGAATGAGACACCCATTATGCAGAAGAAAAGACTTCTTGCAAGGATGGAGG
CCCTCATGAAAACAGTGGAGATGGGTCGGCGCTATTTCCCTCATTGCTCAGAAGTTCTAGATAAG
TTCATGGAGGATGACCTTCCTGATTTATTTTACCTTGAGAAGGGCTCTTCAGAGGAGCAGAAAA
TAAAGAGATCACGTTTCAGGGAGCTGAAGGATGATGTTCAGAAGGCATTTAGTAAGGACAAGGC
TGAGTTTAACCGCACTGGTTTGTCTTCATCCTCATCATCATCATCTTTAAAAGATGGTGGCCCTT
ACAAGCTCAGAAAATTGTGA

> Theobroma cacao cultivar criollo *NPR1-like4* gene (TcNPR4), complete cds

ATGGAATATGGGAATGAAATTTCATCACCATTAAGCTTTGCCCCATCCTCTTACTTATCGGATGG
GTCTGGTGGTCACCTTATAGAAGCTGCCACAAGTACGGAAGACCCTGGGGCCAATCTTGAAATTC
TGAGCCTGAGCAGGCTCAGCTGCAGTCTTGAAAAGTTGTTGGTCGATCAGGAATATGATTACAG
CGATGCTGAGATAGTCGTGGAGGGCAACATTGCTGTGGGGGTGAATAGGTGCATATTGGCTGCG
CGGAGCCAGTACTTCCATGAGCTTTTCAGGAAAGGGAGAGATGACAATTCCATGAATAAGGAAG
GCAAACCACAGTATCTGATGTCAGAGCTAGTGCCTCATGGTCGGGTGGGGTATGAAGCCTTCAAA
GTTTTCTTGAACTATTTATATACTGGAAAGGTTAAGCCGTCACCTAGGGAAGTCTCAACCTGTAT
GGACGATGCTTGTGCCCATGATGCCTGTGGTCCTCTGATTACTTATGCTTTGGAGTTGATGTATG
CTTCTGCCACATTTCAGATGAAAGAGCTGGTTTTGCTTGTTCAGCGCCATCTTTTGAACTCTGTC
GGGAAGGCACTTGTGGAAGATGTAATCCCAATTCTTGTGGCCGCTTTTCACTATAAACTTAACCC
GCTTCTATATGACTGCATCCAAAGAGTGGCAACATCGGGTCTCGATGATGCATGTTTAGAGAAA
GAACTTCCTGGGGAAGTTTATCACGAGATCAGATCAGTCCGTCTCAAATCTCATAAGCAAGAGGC
TGAGCCAGGTGCAGTGGAGCTAGACCCAATGCTTGAAAAGAGAATCAGGAGAATTCACAAGGCA
TTGGACTCTGACGATGTTGAACTGCTCAGGCTTCTTCTGACGGAATCCAATGTCACATTAGATGA
TGCATATGCACTCCACTATGCTGCTGCCTACTGTGATCCTAAGGTTGTTAATGAGGTGCTCAGTC
TTGGCTCGGCTAATGTCAATCTTAGGAACCCCCGAGGGTACACGGTTCTTCAAGTCGCTGCAAGA
CGTAAGGAACCATCAGTGCTGGTGGCCCTGCTGAATAAGGGAGCCTACGTGGCAGAGACTACACC
TGATGGCCAAACTGCTGTCGCGATTTGCAGGAGACTAACTCGGCCCAAGGATTATAATGAGAAC
AAAAAGCAGGGGGAGCTATCCAATAAAGATCGTCTGTGCATTGATGTCCTTGAGAGAGAGATGC
GGAATTCAGTAGAATCTTATAATCAGGGAGTTTCATCACAAGTGATGCATGATGATTTACACAT
GAAACTGGATTACTATGAAAATCGAGTCTCATTTGCGCGGTTGTTGTTTCCTGCTGAAGCCAAGC
TGGCTATGGAAATCGCAGATGCAGATTCAAACATAAGGGAAGTAAATTTGAATGAATCACCTAG
CATACAGACCAAAAGGCTTCAGTTGAGACTGCAAACCCTGCTTAGAACAGTGGAAACTGGGAGG
CGCTACTTTCCACTTTGTTCAGACGTGCTTGACAAGTTTCTGGTGGATGATATGTCTGATCCTTC
CTTAGTTGAGGAAGGAAGTTCAGAAGAGCAGAGACTGAAGAAAAGGAGATTCACAGAACTGAAA
GAGGAGCTGCAACAGGCATTTTACAAGGACATAGAACAAAAGAAAAGGTCGACTTTGTCCCCAT
CATGCTCAGCCTCCTCATCTACTGCTAAAGAGGGTGCTGCTGCTGCGCCGCGTCACAAAACTAGA
AGGAAGTGA

> Theobroma cacao cultivar criollo *BLADE-ON-PETIOLE2* gene (TcBOP2), complete cds

ATGAGTAGCTTCGAGGAGTCCTTACGATCTCTCTCTCTGGACTACCTCAATCTTCTCATCAATGG
CCAGGCTTTCTCCGATGTAACTTTCAGCGTAGAGGGTCGCTTGGTACATGCCCATAGGTGTATTT
TAGCTGCAAGGAGTCTCTTCTTCAGGAAATTCTTCTGCGGACCTGATCCTCCATCTGGGTTAGAC
CCGGTTGGTTCCAGGATGAACCCAGCATCAGCAGCAGCTGCAGGGTCTAGGCCTGGTGTTATACC
AGTGAACTCAGTGGGATACGAGGTGTTCTTGCTGCTGTTACAGTTCTTGTATAGTGGCCAAGTCT
CTATTGTGCCTCAAAAGCATGAGCCCAGGCCTAATTGTTCTGAAAGAGCGTGTTGGCATACGCAT
TGCACCTCTGCCGTTGATCTTGCTCTCGAGACTCTTGCTGCTGCTAGATACTTTGGTGTTGAACA
GCTTGCATTGCTCACCCAGAAACAATTGGCCAGCATGGTCGAGAAGGCCTCAATTGAAGATGTTA
TGAAAGTACTCATAGCTTCAAGAAAGCAAGACATGCACCAACTTTGGACTACTTGTTCCCACCTA

GTGGCGAAATCAGGCCTACCACCAGAAGTTCTCGCCAAACATCTTCCCATCGATGTGGTGGCCAA
AATTGAAGAGTTACGTCTCAAGTCATCCCTTGCTCGGCGCCCCCTAATCACTCATCACCACCACC
ACCATGATCTCACCTCGACCGCTGATCTTGAGGACCAGAAAATTCGGAGGATGAGAAGGGCTTTA
GACTCCTCAGATGTTGAGCTAGTCAAACTCATGGTTATGGGAGAGGGTCTGAATCTTGATGAAG
CATTGGCTTTACACTATGCTGTCGAGAACTGCAGCCGGGAAGTGGTTAAGACCTTGCTAGAGCTT
GGTGCAGCCGATGTTAACTACCGGGCCGGGCCTGCAGGCAAGACCCCGCTTCACATTGCAGCGGA
AATGGTATCACCAGATATGGTTGCAGTCCTTTTAGACCACCACGCGGACCCTAATGTAAGAACCG
TTGATGGGGTGACTCCTTTGGACATTCTTCGGACTCTAACCTCTGATTTCTTGTTCAAAGGTGCC
GTGCCGGGGCTGACTCACATTGAACCAAATAAGCTTAGGCTATGTCTTGAGCTTGTCCAATCAGC
AGCTTTGGTTATTTCGCGTGAAGAAGGGAGCGCAAATGCTCCAACTTCAACTGCAATTTACCCAC
CAATGAGTGATGAACATAATAGTAGCAGCAGTGGAAGCAATCTTGCTACCCTTAATTTGGATTC
AAGGTTGGTTTATCTCAATCTTGGTGCTACAGGTTCAACTCAGATGGGATCGAGAATGGAGGGA
GATGATGATAGCAGCCACAACAGCCAAAGAGAAGCCATGAACCGACATGACCCAACAATGTACC
ATCACTCTCATGACTTCTAG

VITA

Zi Shi

Education:

Doctor of Philosophy

The Pennsylvania State University 8/2004-12/2010

Graduate Degree Program Integrative Biosciences (IBIOS)

Bachelor of Science

China Agricultural University, Biological Sciences 9/2000- 7/2004

Experience and Certificate:

Teaching Assistants for five semesters in the Department of Biochemistry and Molecular Biology, Department of Biology and Department of Horticulture, The Pennsylvania State University

The Graduate School Teaching Certificate, The Pennsylvania State University

Publications:

Shi Z., Maximova S., Liu Y., Verica J. and Guiltinan M. Functional Analysis of the *Theobroma cacao NPR1* Gene in *Arabidopsis. BMC Plant Biology* 2010, **10**:248, In Press

Shi Z., Maximova S., Liu Y., Verica J. and Guiltinan M. NPR1-like protein 3 (NPR3) is a negative regulator of the transcriptional defense response during early flower development in *Arabidopsis* (Manuscript submitted)

Liu Y., **Shi Z.**, Maximova S., Payne M. and Guiltinan M. Proanthocyanidin synthesis in *Theobroma cacao*: Characterization of Genes Encoding Anthocyanidin Synthase, Anthocyanidin Reductase, and Leucoanthocyanidin Reductase (Manuscript submitted)

Yandeau-Nelson M., Laurens L., **Shi Z.**, Xia H., Smith A. and Guiltinan M. Starch Branching Enzyme IIa is required for proper diurnal cycling of starch in leaves of *Zea mays* (Manuscript submitted)

*Argout X, *Salse J, *Aury J, *Guiltinan M, Droc G, Gouzy J, Allegre M, Chaparro C, Legavre T, Maximova S, Abrouk M, Murat F, Fouet O, Poulain J, Ruiz M, Roguet Y, Rodier-Goud M, Barbosa-Neto J, Sabot F, Kudrna D, Ammiraju J, Schuster S, Carlson J, Sallet E, Schiex T, Dievart A, Kramer M, Gelley L, **Shi Z**, Bérard A, Viot C, Boccara M, Risterucci A, Guignon V, Sabau X, Axtell M, Ma Z, Zhang Y, Brown S, Bourge M, Golser W, Song X, Clement D, Rivallan R, Tahi M, Akaza J, Pitollat B, Gramacho K, D'Hont A, Brunel D, Infante D, Kebe I, Costet P, Wing R, McCombie W, Guiderdoni E, Quetier F, Panaud O, Wincker P, Sidibe-Bocs S, Lanaud C (2010) (*These authors contributed equally to this work.) Deciphering the genome structure and paleohistory of *Theobroma cacao*. Nature Genetics In Press