Balance Between Plant Growth and Defense: Transcriptional and Translational Control of Plant Immune System

by

Wei Wang

Department of Biology
Duke University

Date: ______________________

Approved:

___________________________
Xinnian Dong, Supervisor

___________________________
Philip Benfey

___________________________
Meng Chen

___________________________
Zhen-Ming Pei

___________________________
Joshua Socolar

___________________________
Tso-Pang Yao

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology in the Graduate School of Duke University

2012
ABSTRACT

Balance Between Plant Growth and Defense: Transcriptional and Translational Control of Plant Immune System

by

Wei Wang

Department of Biology
Duke University

Date:_______________________
Approved:

___________________________
Xinnian Dong, Supervisor

___________________________
Philip Benfey

___________________________
Meng Chen

___________________________
Zhen-Ming Pei

___________________________
Joshua Socolar

___________________________
Tso-Pang Yao

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biology in the Graduate School of Duke University

2012
Abstract

The activation and maintenance of plant immune responses require a significant amount of energy because they are accompanied by massive transcriptional reprogramming. Spurious activation of plant defense machinery can lead to autoimmune diseases and growth inhibition. So it is important for plants to tightly regulate the immune system to ensure the balance between growth and defense. However, neither the molecular mechanisms nor the design principles of how plants reach this balance are understood.

In this dissertation work, I showed how transcriptional and translational control of plant immune system can help avoid the constant immune surveillance and elicit a proper level of defense responses when necessary. These fine tunings of the immune system ensure the balance between growth and defense.

My research on the transcriptional regulation of plant defense responses led to the surprising discovery that even without pathogen, plant can 'anticipate' potential infection according to a circadian schedule under conditions that favor the initiation of infection. Functional analysis of 22 novel immune components unveiled their transient expression at dawn, when the infection is most likely to happen. This pulse expression pattern was shown to be regulated by the central circadian oscillator, CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) since these 22 genes are no longer induced in the cca1
mutant. Moreover, the temporal control of the transcription level of these 22 immune genes by CCA1 also fine tunes their expression pattern according to the perceptions of different pathogenic signals. At the basal defense level, the expression of these genes can be transiently induced upon perceptions of critical infection stages of the pathogen. When an elevated level of defense response is needed, the high expression levels of these genes are maintained to confer a stronger immunity against pathogen. Since this stronger form of defense may also cause the suicidal death of the plant cells, the interplay between the circadian clock and defense allows a better decision on the proper level of the immunity to minimize the sacrificial death. The circadian clock is also known to regulate the growth-related cellular functions extensively. So the circadian clock can help to balance the energy used in growth and defense through transcriptional regulation on both sides.

Besides the integrated control by the circadian clock, the translational control on a key transcription factor involved in the growth-to-defense transition can also maintain the balance between growth and defense. TBF1 is a major transcription factor that can initiate the growth-to-defense transition through transcriptional repression of growth-associated cellular functions and induction of defense-related machinery. Bioinformatics studies identified 2 upstream open reading frames (uORFs) encoding multiple phenylalanine at 5' of the translation initiation codon of TBF1. Under normal conditions, these 2 uORFs can repress the translation of TBF1 to prevent accidental activation.
However, pathogen infection may cause rapid and transient depletion of phenylalanine, a well-known precursor for cell wall components and the plant defense-related hormone salicylic acid (SA). This depletion signal can be reflected by the increase of uncharged tRNA\textsuperscript{Phe}, which subsequently leads to the phosphorylation of eIF2\(\alpha\) and the release of uORFs’ repression on TBF1. These findings provided the molecular details of how uORF-based translational control can couple transcriptional reprogramming with metabolic status to coordinately trigger the growth-to-defense transition.

In summary, my dissertation work has identified previously unrecognized regulatory mechanisms by which plant immune responses are balanced with growth. These new findings will further investigations into these novel interfaces between plants and pathogens. Future studies will definitely further improve our understandings of the plant-microbe interactions.
Contents

Abstract .................................................................................................................................................iv

List of Tables ........................................................................................................................................xi

List of Figures .......................................................................................................................................xii

List of Abbreviations ..........................................................................................................................xv

Acknowledgements ............................................................................................................................xxii

1.  Introduction to Plant Immune System and Regulatory Mechanisms...........................................1

   1.1. Introduction to plant immune system .........................................................................................1

      1.1.1. Overview ................................................................................................................................1

      1.1.1. Basal defense ........................................................................................................................4

      1.1.2. Effector-triggered immunity (ETI) .........................................................................................6

      1.1.3. Systemic acquired resistance (SAR) .....................................................................................10

   1.2. Introduction to plant circadian clock ........................................................................................14

      1.2.1. Overview ................................................................................................................................14

      1.2.2. Molecular architecture of central circadian clock .................................................................17

      1.2.3. Inputs and outputs of the circadian clock .............................................................................20

   1.3. Introduction to translational control ........................................................................................23

      1.3.1. Overview ................................................................................................................................23

      1.3.2. Global regulation via control on ternary complex formation ............................................26

      1.3.3. Gene-specific regulation through uORF ..............................................................................27

   1.4. Dissertation outline ....................................................................................................................28
2. Timing of Plant Immune Responses by a Central Circadian Regulator .......... 30
   2.1. Introduction ........................................................................................................... 30
   2.2. Results .................................................................................................................... 38
      2.2.1. Phenotypic clustering of immune gene mutants upon Emwa infection highlights the significance of PCD ................................................................. 38
      2.2.2. Mutants defective in PCD are also compromised in basal defense .......... 40
      2.2.3. Bioinformatics analysis on the immune genes suggests potential regulation by the circadian clock ................................................................. 41
      2.2.4. Circadian clock gene mutants are defective in ETI ........................................ 42
      2.2.5. Validation of RNA annealing selection ligation-sequencing (RASL-seq) .... 42
      2.2.6. RASL-seq analysis unveils ‘anticipation’ model ............................................ 44
      2.2.7. Diurnal difference in ETI, enhanced PCD and basal defense in CCA1 overexpression plants validate the ‘anticipation’ model ........................................ 47
   2.3. Discussion ............................................................................................................... 68
   2.4. Methods and Materials ........................................................................................ 71
      2.4.1. Arabidopsis and Hpa growth condition ............................................................ 71
      2.4.2. Phenotyping mutants upon Hpa infection ....................................................... 71
      2.4.3. Phenotypic clustering ....................................................................................... 72
      2.4.4. Bioinformatics analysis ..................................................................................... 72
      2.4.5. RASL-seq ........................................................................................................ 73
      2.4.6. RASL-seq data analysis ................................................................................... 73
      2.4.7. RNA extraction and quantitative PCR (qPCR) analysis ................................. 75
      2.4.8. Bioluminescence detection .............................................................................. 75
3.5. Acknowledgements

4. Summary and Perspectives

4.1. Reconcile between two anticipation system

4.2. Future research on translational control of plant immune system

References

Biography
List of Tables

Table 2-1: Phenotypic scores and ranks of gene mutants.................................................................36

Table 2-2: Bioinformatics evidences of potential circadian clock control on the new defense genes..................................................................................................................................................54

Table 3-1: Weight matrix of TL1 element derived from the promoter sequences of BTH-induced ER-resident genes .................................................................................................................................................................89

Table 3-2: Genome-wide distribution of TL1 element ........................................................................91
List of Figures

Figure 2-1: Characterization of mutants in RPP4-mediated resistance against Emwa ..... 34
Figure 2-2: Phenotypic clusters with multiple mutants alleles .................................... 49
Figure 2-3: Phenotypic clusters using data from 3 biological replicates ......................... 50
Figure 2-4: Principal component analysis of phenotypic scores ..................................... 51
Figure 2-5: Illustration of functional divergence of Group 1 and 2 mutants ..................... 52
Figure 2-6: Some of the RPP4-mediated resistance mutants are also compromised in basal defense .................................................................................................................. 53
Figure 2-7: Involvement of the circadian clock in plant immune system ................. 55
Figure 2-8: Different responsiveness of CCA1 and LHY to Emwa identified through RASL-seq under diurnal conditions ................................................................. 56
Figure 2-9: Different responsiveness of CCA1 and LHY to Emwa identified using CCA1:LUC and LHY:LUC lines under free-run conditions .................................................. 57
Figure 2-10: Confirmation of the unexpected expression pattern of CCA1 under control conditions using CCA1:LUC line .................................................................................. 58
Figure 2-11: Determination of the cluster number used in NMF ................................ 59
Figure 2-12: Membership distance for Cluster 1 and Cluster 2 genes determined by their expression patterns ............................................................................................................ 60
Figure 2-13: Convergence of NMF ................................................................................ 61
Figure 2-14: Time-course expression of NMF Cluster 1 genes .................................... 62
Figure 2-15: Time-course expression of the NMF Cluster 2 genes together with CCA1 and RPP4 as references ................................................................................................. 63
Figure 2-16: ‘Anticipation’ model illustrating the circadian clock control on plant immune system ........................................................................................................................... 64
Figure 2-17: SPP count after Emwa infection at dawn or dusk ...........................................65

Figure 2-18: Enhanced disease resistance to Hpa Noco observed in the CCA1\textsubscript{OE} transgenic line based on sporangiospore count at 7 dpi .........................................................66

Figure 2-19: Validation of the positive regulation of CCA1 on host cell death upon Emwa infection .................................................................67

Figure 3-1: Sequence logo of TL1 ..................................................................................90

Figure 3-2: Gene ontology analysis of TL1-containing genes ...........................................92

Figure 3-3: Venn diagrams of BTH-, elf26- and flg22-regulated genes .........................94

Figure 3-4: Venn diagram of enriched GO terms of TL1-containing genes ..................95

Figure 3-5: Enriched GO terms from Figure 3-4 ............................................................96

Figure 3-6: Venn diagram shows the numbers of TBF1-dependent SA-repressed (SA down), SA-induced (SA up), elf18-induced (elf18 up), and elf18-repressed (elf18 down) genes (P<0.05) .................................................................................................97

Figure 3-7: Heat maps of TBF1-regulated genes in total numbers (top), degrees of TBF1 dependency (middle), and numbers of TL1 elements in the gene promoters (bottom) in response to SA (a) and elf18 (b) treatment .........................................................98

Figure 3-8: TBF1 controls growth-to-defense transition during SAR and MTI ..........100

Figure 3-9: Schematic representation of uORF1 and uORF2 and exon I of TBF1 ..........101

Figure 3-10: Characterization of the uORFs of TBF1 ....................................................102

Figure 3-11: Ribosome profiles and translational status of TBF1 upon infection ..........104

Figure 3-12: The UBQ5 transcript association with polysomes is not altered by pathogen infection .........................................................................................................................106

Figure 3-13: uORF-mediated translation derepression mechanism ...............................107

Figure 3-14: Phenylalanine but not aspartate starvation can alleviate the translation inhibition effect of uORFs .............................................................................................................108
Figure 3-15: A model illustrating the translational control mechanism on TBF1 .......... 111
List of Abbreviations

AU p-value  approximately unbiased p-value
AvrPto  avirulent gene of Pseudomonas syringae pv. Tomato
AzA  azelaic acid
AZI1  AZELAIC ACID INDUCED 1
B.C.  Before Christ
BAK1  BRASSINOSTEROID RECEPTOR 1-ASSOCIATED KINASE
BIK1  BOTRYTIS INDUCED KINASE 1
BiP2  LUMENAL BINDING PROTEIN
BSMT1  BENZOIC ACID/SALICYLIC ACID CARBOXYL METHYLTRANSFERASE 1
CAB2  CHLOROPHYLL A/B-BINDING PROTEIN 2
CC  coiled-coil
CCA1  CIRCADIAN CLOCK ASSOCIATED 1
CCA1_{OE}  CCA1-overexpression line
CCD  charge coupled device
CHE  CCA1 HIKING EXPEDITION
ChIP-seq  chromatin immunoprecipitation-sequencing
Col-0  Columbia-0 ecotype of Arabidopsis
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT1</td>
<td>COMPROMISED RECOGNITION OF TURNIP CRINKLE VIRUS 1</td>
</tr>
<tr>
<td>DA</td>
<td>dehydroabietinal</td>
</tr>
<tr>
<td>DAMP</td>
<td>damage-associated molecular pattern</td>
</tr>
<tr>
<td>DHFR</td>
<td>dihydrofolate reductase</td>
</tr>
<tr>
<td>DIH</td>
<td>Discrete HR</td>
</tr>
<tr>
<td>DIR1</td>
<td>DEFECTIVE IN INDUCED RESISTANCE 1</td>
</tr>
<tr>
<td>dpi</td>
<td>days post inoculation</td>
</tr>
<tr>
<td>EDS5</td>
<td>ENHANCED DISEASE SUSCEPTIBILITY 5</td>
</tr>
<tr>
<td>EE</td>
<td>evening element</td>
</tr>
<tr>
<td>eIF</td>
<td>eukaryotic initiation factor</td>
</tr>
<tr>
<td>eIF2α-P</td>
<td>phosphorylated eIF2 at its α subunit</td>
</tr>
<tr>
<td>elf26</td>
<td>the first 26 amino acids of a MAMP signal, elongation factor thermo unstable</td>
</tr>
<tr>
<td>ER</td>
<td>endoplasmic reticulum</td>
</tr>
<tr>
<td>ETI</td>
<td>effector-triggered immunity</td>
</tr>
<tr>
<td>ETS</td>
<td>effector-triggered susceptibility</td>
</tr>
<tr>
<td>EXH</td>
<td>expanding HR</td>
</tr>
<tr>
<td>FHI</td>
<td>free hyphal intermediate</td>
</tr>
<tr>
<td>flg22</td>
<td>synthetic 22 amino acids from flagellin</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>FLS2</td>
<td>FLAGELLIN-SENSING 2</td>
</tr>
<tr>
<td>FRH</td>
<td>free hypha</td>
</tr>
<tr>
<td>G3P</td>
<td>Glycerol-3-phosphate</td>
</tr>
<tr>
<td>GCN4</td>
<td>GENERAL CONTROL NONDEREPRESSIBLE 4</td>
</tr>
<tr>
<td>GEF</td>
<td>guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>GI</td>
<td>GIGANTEA</td>
</tr>
<tr>
<td>GO</td>
<td>gene ontology</td>
</tr>
<tr>
<td>GOE</td>
<td>Great Oxidation Event</td>
</tr>
<tr>
<td>GUS</td>
<td>β-glucuronidase</td>
</tr>
<tr>
<td>Hpa</td>
<td><em>Hyaloperonospora arabidopsidis</em></td>
</tr>
<tr>
<td>HR</td>
<td>hypersensitive response</td>
</tr>
<tr>
<td>HSF</td>
<td>heat-shock factor</td>
</tr>
<tr>
<td>ICS1</td>
<td>isochorismate synthase 1</td>
</tr>
<tr>
<td>JA</td>
<td>jasmonic acid</td>
</tr>
<tr>
<td>LHY</td>
<td>LATE ELONGATED HYPOCOTYL</td>
</tr>
<tr>
<td>LRR</td>
<td>leucine-rich repeat</td>
</tr>
<tr>
<td>LRR-RK</td>
<td>LRR receptor kinase</td>
</tr>
<tr>
<td>LTB</td>
<td>lactophenol trypan blue</td>
</tr>
<tr>
<td>LUC</td>
<td>luciferase</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MAMP</td>
<td>microbe-associated molecular pattern</td>
</tr>
<tr>
<td>MAPK</td>
<td>mitogen-associated protein kinase</td>
</tr>
<tr>
<td>MeSA</td>
<td>methyl SA</td>
</tr>
<tr>
<td>Met-tRNA$_{\text{Met}}$</td>
<td>methionyl-initiator tRNA</td>
</tr>
<tr>
<td>MTI</td>
<td>MAMP-triggered immunity</td>
</tr>
<tr>
<td>MTX</td>
<td>methotrexate</td>
</tr>
<tr>
<td>NB</td>
<td>nucleotide-binding</td>
</tr>
<tr>
<td>NDR1</td>
<td>NON-RACE SPECIFIC DISEASE RESISTANCE-1</td>
</tr>
<tr>
<td>NF-$\kappa$B</td>
<td>NUCLEAR FACTOR KAPPA-LIGHT-CHAIN-ENHANCER OF ACTIVATED B CELLS</td>
</tr>
<tr>
<td>NIMIN</td>
<td>NON-INDUCIBLE IMMUNITY 1-INTERACTING</td>
</tr>
<tr>
<td>nm</td>
<td>nanometer</td>
</tr>
<tr>
<td>NMF</td>
<td>nonnegative matrix factorization</td>
</tr>
<tr>
<td>NPR1</td>
<td>NONEXPRESSOR OF PATHOGENESIS-RELATED PROTEINS 1</td>
</tr>
<tr>
<td>OOS</td>
<td>oospores</td>
</tr>
<tr>
<td>PABP</td>
<td>poly(A)-binding protein</td>
</tr>
<tr>
<td>PAD4</td>
<td>PHYTOALEXIN DEFICIENT 4</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS3</td>
<td>AVRPPHB SUSCEPTIBLE 3</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>PC</td>
<td>principal component</td>
</tr>
<tr>
<td>PCA</td>
<td>principal component analysis</td>
</tr>
<tr>
<td>PCD</td>
<td>programmed cell death</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PIC</td>
<td>preinitiation complex</td>
</tr>
<tr>
<td>Pip</td>
<td>pipecolic acid</td>
</tr>
<tr>
<td>PR</td>
<td>pathogenesis-related</td>
</tr>
<tr>
<td>PRR</td>
<td>pattern recognition receptor</td>
</tr>
<tr>
<td>PRR7</td>
<td>PSEUDO- RESPONSE REGULATOR 7</td>
</tr>
<tr>
<td>PRR9</td>
<td>PSEUDO- RESPONSE REGULATOR 9</td>
</tr>
<tr>
<td>PRX</td>
<td>PEROXIREDOXIN</td>
</tr>
<tr>
<td><em>Psm</em></td>
<td><em>Pseudomonas syringae</em> pv. <em>maculicola</em></td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative PCR</td>
</tr>
<tr>
<td>R</td>
<td>resistance</td>
</tr>
<tr>
<td>RACE</td>
<td>rapid amplification of cDNA end</td>
</tr>
<tr>
<td>RASL-seq</td>
<td>RNA annealing selection ligation-sequencing</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td>RPP4</td>
<td>RECOGNITION OF <em>PERONOSPORA PARASITICA</em> 4</td>
</tr>
<tr>
<td>SAG</td>
<td>SA glucoside</td>
</tr>
</tbody>
</table>
SAGT  SA GLUCOSYLTRANSFERASE
SAR  systemic acquired resistance
SA  salicylic acid
SFD1  SUPPRESSOR OF FATTY ACID DESATURASE DEFICIENCY 1
SPP  sporangiophore
TBF1  TL1-BINDING FACTOR 1
TF  transcription factor
TIR  Toll, interleukin-1 receptor, resistance protein
TL1  TRANSLOCON 1
TMV  tobacco mosaic virus
TOC1  TIMING OF CAB2 EXPRESSION 1
TRN  trailing necrosis
TTFL  transcriptional-translational feedback loop
TTSS  type III secretion system
UBQ5  ubiquitin 5
uORF  upstream open reading frame
UTR  untranslated region
WT  Wild-type
Y1H  yeast one-hybrid

xx
<table>
<thead>
<tr>
<th>Zeitgeber</th>
<th>time giver</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZIP-2</td>
<td>bZIP transcription factor-2</td>
</tr>
<tr>
<td>ZTL</td>
<td>ZEITLUPE</td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to take this opportunity to express my heartfelt appreciation to all the people who devoted their time and wisdom in helping me shape this dissertation.

First of all, I would like to thank my advisor, Dr. Xinnian Dong. Thank you for your continuous support both in science and in my life. Your science career is an example of how persistent pursuit of perfection can lead to the achievement of excellence. I really enjoyed the time we worked together on all of our papers and grant proposals. I also appreciate your understanding and great support when I tried to bring more statistics to my biological research. Of course, your help with my family will always be remembered.

I would like to thank Jinyoung Barnaby for your guidance when I just joined the lab and also your great collaboration in my research. Some of your work is described in this dissertation as an introduction to my study on the circadian clock. I want to thank Karolina Pajerowska-Mukhtar for engaging me in your study on TBF1, which led to a new direction of my research. Some of your findings served as the basis for my discoveries of translational regulatory mechanism of TBF1. I also would like to thank Yasuomi Tada. How amazing it is that you have co-authored all the papers I have had so far in the Dong lab! Your rich experience in biology, your insightful ideas and your great passion for research are all so memorable. Of course, I am also thankful to all the Dong lab members. You have made up a great team and a warm family.
Surely, science is not a monologue, reaches beyond a single lab, university and even country. Several great scientists have also contributed to my research and this dissertation. Dr. Xiang-Dong Fu and Dr. Hairi Li at UCSD, Dr. Mahmut Tör at University of Worcester all participated in the work described in Chapter 2. Dr. Christopher Nicchitta at Duke and Dr. Tao Pan at University of Chicago provided important suggestions and guidance in the studies presented in Chapter 3. I am grateful to your collaboration.

Lookers-on see most of the game. So advices from beyond the lab are crucial. I would like to attribute credits to my committee members as well as the advisors and postdocs during my rotation. It is my honor to have Drs. Philip Benfey, Meng Chen, Zhen-Ming Pei, Joshua Socolar, Tso-Pang Yao on my dissertation committee. It is your critical inputs that helped push my research forward and enlighten me from an independent perspective. Also many thanks to Drs. Zhen-Ming Pei, Taiping Sun, Rui Ye and Zhongling Zhang for your guidance during my rotations in the Pei lab and the Sun lab.

One cannot make bricks without straw. Funding is the straw for scientific research. I would like to thank the funding support to me from the NSF2010 and HHMI grants awarded to Xinnian. In addition, the fellowship from the Department of Biology also supported me through my first year at Duke and allowed me to attend the conferences out of town.
My parents have always been supportive. Your encouragement made me strong when I am alone in a foreign country. Your daily words and deeds have shaped me. I miss you all the time.

Lastly, I want to thank my wife, Mian Zhou. How lucky I am to have someone who can understand me, work with me and share the rest of the life together! And how wonderful it is that I can collaborate with my wife on science! Part of your work is introduced in the future direction part of this dissertation. And I surely look forward to our own bright future together.
1. Introduction to Plant Immune System and Regulatory Mechanisms

Plant immune system, the circadian clock and translational control mechanism have all been studied extensively. This chapter is not intended to cover all the findings but rather focus on the major discoveries relevant to the topic of this dissertation.

1.1. Introduction to plant immune system

1.1.1. Overview

It is important for animals and plants to fend against invasions of harmful microbes. Animals and plants have multiple defense layers.

Physical, chemical and biological barriers comprise the first non-specific defense line. Plants use waxy cuticle (Martin, 1964) while animals have skin (Kupper and Fuhlbrigge, 2004) to physically shield against the invasion attempts of pathogens. Antimicrobial compounds secreted by both plants (Dixon, 2001) and animals (Zasloff, 2002) may kill the pathogen directly. In addition, commensal flora can compete for the nutrients with pathogenic microbes, thus indirectly protecting plants (Jacobsen, 2006) and animals (Macpherson and Harris, 2004).

Successful pathogens can breach this first barrier and make a direct contact with the host cell membrane where the second defense line is mounted. Plants and animals use pattern recognition receptors (PRRs) to perceive conserved microbe-associated molecular patterns (MAMPs) (Boller and Felix, 2009). Detection of MAMPs by PRRs can
trigger MAMP-triggered immunity (MTI). Since the MAMPs are usually conserved over a wide range of pathogens, MTI is a broad-spectrum immunity.

MAMPs are essential components to pathogens for a free living state. Therefore pathogens cannot afford to lose MAMPs in order to evade MTI. Instead, more successful pathogens can deliver effectors into plant (Jones and Dangl, 2006; Mukhtar et al., 2011) and animal (Navarro et al., 2005) cells to suppress MTI. The suppression of MTI by effectors leads to effector-triggered susceptibility (ETS). The repertoires of the effectors secreted by different pathogens are distinct. Since most effectors only participate in the invasion of host cells and therefore they are fully dispensable when the microbe is living in a nonpathogenic lifestyle. So, the effectors are under significant diversifying selection pressure which leads to the explosive expansion of effectors. As a result, plants and animals need to recognize many more effectors than MAMPs.

To cope with this challenge, higher animals (jawed vertebrates) and plants have evolved different strategies. The adaptive immune system of jawed vertebrates takes advantage of the explosive proliferation feature of permutations and recombinations using limited sets of gene modules. Through somatic recombination, jawed vertebrates can generate a collection of structurally similar receptors that encompass almost infinite antigen recognition specificities. Therefore higher animals give tit for tat, using infinite receptors to fight against infinite number of antigens. The tactics employed by plants are strategically different from higher animals. First of all, plants also have dramatically
expanded the immune receptor repository of both PRRs (Lehti-Shiu et al., 2009) and intracellular immune receptors known as resistance (R) proteins (Meyers et al., 2002). Secondly, besides recognition by direct binding with effectors, R proteins can also detect the presence of effectors by monitoring their actions. Most effectors are evolved to modify the key components of MTI. Therefore, through surveillance on these relatively few key components, limited R proteins are needed to protect against a large array of effectors which target similar MTI components (Jones and Dangl, 2006; Mukhtar et al., 2011). Hence, these two strategies used by plants complement their lack of adaptive immune system. Both the direct and indirect recognitions of effectors can activate R proteins to trigger an elevated level of innate immunity in plants, called effector-triggered immunity (ETI).

Besides localized immune response, animals and plants can also prevent pathogen from spreading through elevation of systemic defense level. Higher animals have dedicated immune cells circulating to ensure a systemic surveillance. However, constrained in cell wall, plant cells are immobile. Systemic resistance in plants is established through the perception of mobile systemic immune signals sent out by the local cells undergoing MTI or ETI.

All these defense lines are not isolated. Their intimate interactions have built up a robust immune network. Take plant immunity for example, damage-associated molecular patterns (DAMPs) generated by the pathogenic perturbation of the physical
barriers can activate MTI (Henry et al., 2012) while the antimicrobial compounds produced during MTI and ETI can help to strengthen the chemical barriers. The connections between MTI and ETI are exemplified by the substantial overlap of transcriptionally induced genes during MTI and ETI (Navarro et al., 2004). Moreover, MTI, ETI and SAR are closely linked by plant hormone, salicylic acid (SA) (Jones and Dangl, 2006). As a network, these different immune responses should not be studied alone. Rather, a system view needs to be emphasized. Therefore I will present our current understandings of plant immunities based on the recent findings and highlight their interconnections which may be of a broader interest or pertinent to the discoveries described in this dissertation.

1.1.1. Basal defense

Basal defense can be defined as the disease resistance activated by virulent pathogens on susceptible hosts (Jones and Dangl, 2006). Therefore, physical/chemical/biological barriers and MTI are the major components of basal defense. While physical/chemical/biological barriers are passive defense lines, MTI is an active form of immunity which is triggered by the perception of MAMPs by PRRs. MAMPs’ functions are vital to the microbes. Their structures are highly conserved across a wide range of microorganisms. And they are normally absent in the host. These features make MAMPs perfect targets for immune-recognition. For example, plants can detect flagellin, the building block of flagellum of Gram-negative bacteria; chitin,
major component of fungal cell wall; glucan from oomycete; xylanase from ascomycete, etc (Boller and Felix, 2009). Although the significance of these MAMPs to microbial life subjects them to purifying selection pressure, recent discoveries that flagellin evolves faster than expected (Cai et al., 2011), suggest diversifying selection pressure specifically on the immunogenic epitopes to evade the recognition by plant PRRs despite the overall conservation of MAMPs. Nevertheless, most plants can recognize flagellin or even the synthetic epitope, flg22, which includes the 22 amino acids from the conserved N-terminus (Boller and Felix, 2009). Actually, the perception of flagellin by FLAGELLIN-SENSING 2 (FLS2) is among the best characterized MAMP/PRR pairs.

Flagellin was first identified as a MAMP by chance during an effort to study the effect of a harpin preparation on tomato cells (Felix et al., 1999). A forward genetic screen based on the growth inhibition effect of flg22 on Arabidopsis seedling led to the isolation of the first PRR, FLS2 (Gomez-Gomez and Boller, 2000). FLS2 encodes a leucine-rich repeat (LRR) receptor kinase (LRR-RK). The protein domain structure of FLS2 is quite modular in that the extracellular LRR domain binds flagellin; the intracellular Serine/Threonine kinase domain transduces the signal. Upon binding, FLS2 dimerizes with BRASSINOSTEROID RECEPTOR 1-ASSOCIATED KINASE (BAK1) within 2 minutes in a flagellin-dependent manner (Chinchilla et al., 2007). Then BAK1 trans-phosphorylates BOTRYTIS INDUCED KINASE 1 (BIK1), which in turn phosphorylates FLS2 and BAK1 (Lu et al., 2010; Zhang et al., 2010). BAK1 and BIK1
appear to be master regulators of MTI, as they interact with multiple PRRs (Monaghan and Zipfel, 2012). These responses are accompanied by other early signaling and responses including ion flux, oxidative burst, activation of mitogen-associated protein kinase (MAPK) cascade and receptor endocytosis (Boller and Felix, 2009). Within 30 minutes, significant transcriptional reprogramming is induced (Zipfel et al., 2004).

MTI responses triggered by different MAMPs are largely similar. They converge to BAK1 and BIK1, share similar early responses and induce almost identical downstream genes (Boller and Felix, 2009; Monaghan and Zipfel, 2012). Establishment of a converged immunity has an obvious advantage in integrating multiple input signals to determine the overall defense strength. However, this type of structure makes the converging points vulnerable to targeted attacks. In fact, BIK1 is a direct target of effectors from multiple pathogens (Feng et al., 2012; Zhang et al., 2010). Therefore a more robust immune network structure is needed.

1.1.2. Effector-triggered immunity (ETI)

As mentioned above, the convergence of MTIs triggered by different MAMPs exposes vulnerable points of the structure. More adapted pathogens can suppress MTI by modifying these weak points using effector proteins. Bacterial phytopathogens can directly inject 20-30 different effectors into host cells through a needle-like structure, called type III secretion system (TTSS) (Cunnac et al., 2009). There can be multiple ways for effectors to inactivate the same host targets and there are only limited numbers of
vulnerable points in MTI. So it is possible that different effectors may modify the same immune components involved in MTI in different manners. In fact, we have seen ample examples that multiple effectors target the same host component but modify it in different ways, all resulting in suppression of MTI (Dodds and Rathjen, 2010). For example, both *Pseudomonas syringae* effectors AvrPto and AvrPtoB target FLS2-BAK1 complex. While AvrPto functions as a kinase inhibitor (Shan et al., 2008; Xiang et al., 2008; Xing et al., 2007), AvrPtoB directs them for degradation (Gimenez-Ibanez et al., 2009; Rosebrock et al., 2007).

Plants will not await their doom. They use the intracellular R proteins to detect corresponding effectors and lead to ETI. Some R proteins can recognize the effectors through direct physical binding while others act in an indirect way (Jones and Dangl, 2006). As for the direct recognition, it is conceivable that diversifying selection pressure on effectors and corresponding R proteins will drive the fast evolution of both. This arm race has generated sequence polymorphism between alleles in the pathogen and plant populations. However, if plants solely rely on this ‘gene-for-gene’ interaction mode, they have to encode a huge number of R proteins in the genome which will significantly lower their global fitness (Tian et al., 2003). In fact, besides direct physical interaction, plants also focus on the limited set of effector targets through indirect recognition. The ‘guard’ hypothesis was proposed to summarize the underlying mechanism of the indirect recognition conceptually (Jones and Dangl, 2006). It suggested that R proteins
monitor the integrity of effector targets. Once these targets are modified by effectors, the corresponding R proteins can recognize the modified targets. Accordingly, a bowtie network structure can be expected in which a limited number of MTI components are targeted by many more effectors while also monitored by multiple R proteins. A recent genome-wide yeast two-hybrid screen using plant immune proteins and effectors from two divergent pathogens identified such a bowtie network structure (Mukhtar et al., 2011). Multiple effectors converge to a few plant immune protein hubs that are highly interconnected. And multiple R proteins also interact with these hubs. This discovery provided a genome-wide evidence to support the ‘guard’ model and demonstrated how a limited number of R proteins can detect a much larger pool of effectors. Although this bowtie structure appears similar to the converged MTI signaling pathway, the dedicated immune surveillance of R proteins on these hubs makes it less vulnerable.

While our knowledge of recognition has dramatically expanded, the elucidation of how effector recognition leads to R protein activation still remains as a great challenge. Conceptually, it is believed that R proteins are kept in an inactive form either by an accessory inhibitor protein or through inhibitory intramolecular interactions. Recognition of effectors or modified targets releases the accessory protein or triggers a conformational change of R proteins into an active form. However, biochemical studies on R proteins and their complexes are still challenging. We also know little about the signaling pathways downstream of R protein activation. Nevertheless, similar
physiological responses as MTI were observed including rapid ion flux, oxidative burst, activation of MAPK cascade and transcriptional reprogramming (Dodds and Rathjen, 2010).

Often, ETI is accompanied by localized cell death, called the hypersensitive response (HR). It features cytoplasmic shrinkage, chromatin condensation, mitochondrial swelling, vacuolization and chloroplast collapse (Mur et al., 2008). Although ETI can be uncoupled from HR in several cases (Maekawa et al., 2011), it is still widely used as an indicator for R protein activation. In accordance with the ‘zigzag’ model (Jones and Dangl, 2006) which described the elevated arm race between effectors and R proteins, the altruistic sacrifice of infected host cells during HR marks a higher defense level compared to MTI. MTI generally does not involve host cell death and it was accordingly believed that the immune responses during ETI occur quicker, last longer and appear more robust than those during MTI (Jones and Dangl, 2006; Tao et al., 2003; Tsuda and Katagiri, 2010). However, recent studies challenged this idea. First of all, MTI and ETI share common physiological responses upon activation. In addition, MTI may also cause HR (Bailey et al., 1990; Khatib et al., 2004; Ron and Avni, 2004; Wei et al., 1992) and SAR (Mishina and Zeier, 2007). Furthermore, besides strong MTI, ETI can also be relatively weak and slow (Thomma et al., 2011). Finally, transcriptional changes during MTI and ETI largely overlap (Tao et al., 2003). Taken together, the differences between MTI and ETI reflect a quantitative continuum rather than a qualitative
dichotomy, in which fine tuning is the key to determine the level of defense according to the specific danger signal perceived. The HR phenotype associated with ETI or sometimes MTI merely manifests how dangerous the detected pathogenic signal is, so that it is vitally necessary to deploy even a deadly level of defense at local cells to protect the rest of the plants.

1.1.3. Systemic acquired resistance (SAR)

SAR was first observed as a broad-spectrum long-lasting resistance in naïve tissues triggered by local tobacco mosaic virus (TMV) infection of tobacco leaves (Ross, 1961). It is an important form of plant immunity that systemic resistance against the spread of the local infection.

Plant hormone SA sits at the center of SAR as: (1) ETI can induce the biosynthesis of SA both locally and systemically (Malamy et al., 1990; Metraux et al., 1990); (2) mutants deficient in systemic SA accumulation upon ETI induction are compromised in SAR (Gaffney et al., 1993; Mishina and Zeier, 2006; Wildermuth et al., 2001); (3) SAR can be triggered by exogenous application of SA, or its biologically active analogs 2,6-dichloroisonicotinic acid (INA) and benzothiadiazole S-methyl ester (BTH) (Durrant and Dong, 2004). In Arabidopsis, SA is mainly synthesized via chorismate pathway in which ISOCHORISMATE SYNTHASE 1 (ICS1) plays a major role (Wildermuth et al., 2001). After generated in the chloroplast, SA can be converted to the inactive form, SA glucoside (SAG) by SA GLUCOSYLTRANSFERASE (SAGT) and stored in the vacuole
SA can also be converted to the volatile methyl SA (MeSA) by BENZOIC ACID/SALICYLIC ACID CARBOXYL METHYLTRANSFERASE 1 (BSMT1) (Chen et al., 2003).

Although SA is necessary and sufficient to SAR, grafting studies disprove its role as the mobile signal (Gaffney et al., 1993). Many years of research to identify the mobile signal of SAR through forward genetic screen and biochemical analysis of petiole exudates has isolated several candidate mobile signals including MeSA (Park et al., 2007), jasmonic acid (JA) (Truman et al., 2007), azelaic acid (AzA) (Jung et al., 2009), Glycerol-3-phosphate (G3P) (Chanda et al., 2011), dehydroabietinal (DA) (Chaturvedi et al., 2012) and pipecolic acid (Pip) (Dempsey and Klessig, 2012) as well as some novel components required for SAR such as DEFECTIVE IN INDUCED RESISTANCE 1 (DIR1) (Maldonado et al., 2002), AZELAIC ACID INDUCED 1 (AZI1) (Jung et al., 2009) and SUPPRESSOR OF FATTY ACID DESATURASE DEFICIENCY 1 (SFD1) (Nandi et al., 2004). However, their role as genuine mobile signals for SAR is still under debate. For example, the *Arabidopsis* SA methyltransferase mutant (*bsmt1*), which is compromised in converting SA to MeSA, does not have a defect in SAR, thus challenging the role of MeSA as a real mobile SAR signal. Similarly, failure to observe SAR defects in JA biosynthetic and response mutants also suggested against its role as a SAR signal (Attaran et al., 2009). Nevertheless, the discoveries of multiple SAR mobile signal candidates have improved our understanding of signaling pathways involved in SAR.
and indicate a more complex composition of systemic signals than previously expected, especially considering the interactions among these candidates (Dempsey and Klessig, 2012).

NONEXPRESSOR OF PATHOGENESIS-RELATED PROTEINS 1 (NPR1) is pivotal to SA signal transduction. Its importance is evident since isolation, as multiple genetic screens searching for SA insensitive mutants all converged to NPR1 (Cao et al., 1994; Cao et al., 1997; Delaney et al., 1995; Shah et al., 1997). Under resting conditions, NPR1 forms oligomer through disulfide bonds, which are redox-sensitive. Induction of SAR by exogenous application of SA triggers a biphasic redox fluctuation, which can be sensed by NPR1 (Mou et al., 2003). NPR1 responds to this redox change through alternating between oligomer and monomer. S-nitrosglutathione promotes oligomerization while thioredoxins facilitate the release of NPR1 monomer (Tada et al., 2008). SA induces the switch from NPR1 oligomer to monomer which is then translocated into nucleus (Kinkema et al., 2000) and interacts with transcription factors (TFs) TGA (Despres et al., 2003; Zhang et al., 1999; Zhou et al., 2000) and NON-INDUCIBLE IMMUNITY 1-INTERACTING (NIMIN) (Maier et al., 2011; Weigel et al., 2001; Weigel et al., 2005). As a transcription cofactor, NPR1 was shown to enhance the binding of TGA to the promotors of target genes (Despres et al., 2000; Johnson et al., 2003). After NPR1 fulfills its tasks, it is polyubiquitinylated by the Cullin 3 E3 ligase and degraded by the 26S proteasome to refresh the initiation complex (Spoel et al., 2009).
Recent identification of NPR1 paralogs, NPR3 and NPR4 as SA receptors, provided a more comprehensive picture of how the amount of NPR1 is finely tuned (Fu et al., 2012). Without infection, NPR1 is constitutively degraded by NPR4 to avoid spurious activation of defense. Challenged by pathogen, high level of SA accumulated at the local tissue, though disrupts NPR4-mediated degradation of NPR1, yet promotes NPR3-mediated degradation. Therefore the inhibitory effects of NPR1 on host cell death are relieved to allow the progression of HR and ETI at the infection sites. However, the intermediate level of SA induced in the systemic tissue disrupts NPR4-mediated degradation but is not high enough to initiate NPR3-mediated degradation. As a result, the high amount of NPR1 accumulated in these cells defines the spatial boundary for cell death and triggers SAR.

Activation of SAR involves a massive transcriptional reprogramming which accounts for 10% of the whole genome and 99% of these SAR-responsive genes are dependent on NPR1 (Wang et al., 2006). Members of these potential SAR executors include pathogenesis-related (PR) proteins which have antimicrobial activities and are widely used as a proxy for SAR (Sels et al., 2008); endoplasmic reticulum (ER)-resident proteins which are induced earlier than PR proteins and have been shown to accommodate the increased demand on protein folding and secretion of PR proteins (Wang et al., 2005); WRKY TFs which have versatile functions in regulating the plant defense (Pandey and Somssich, 2009). Bioinformatics analysis on the promoters of BTH-
induced ER-resident genes helps the identification of a novel cis-element called TRANSLOCON 1 (TL1). Further experiments established its role in the regulation of these genes (Wang et al., 2005). The finding of the TF TBF1 that regulates TL1 and further functional characterization of TBF1 are described in Chapter 3 of this dissertation. The translational control mechanism on TBF1 may be a general strategy adopted by immune system to regulate critical TFs, as a similar regulation was also observed for the immune TF ZIP-2 in nematode (Dunbar et al., 2012).

1.2. Introduction to plant circadian clock

1.2.1. Overview

The rotation of the earth on its own axis and revolution around the sun impart daily and annual rhythmicity. This highly dramatic yet recurrent environmental change drives the evolution of the circadian clock, a self-sustaining endogenous chronometer in virtually all life forms. Based on the recent observations that a universal oscillation of the oxidation status of PEROXIREDOXIN (PRX), a protein involved in the scavenging of toxic reactive oxygen species (ROS), exists in every kingdom of life (Edgar et al., 2012), the circadian clock has been postulated to emerge as a redox oscillator around 2.5 billion years ago. Independently, circadian clock may also appear as a transcriptional-translational feedback loop (TTFL) in cyanobacteria around the same time. The prevalence of photosynthetic bacteria during this time resulted in the rapid accumulation of toxic level of oxygen in the atmosphere, a historical event called the
Great Oxidation Event (GOE). The GOE triggered a global transition from anaerobic life to aerobic style. The two emerged forms of circadian clock helped to coordinate all aspects of life to the oxygen and ROS production/consumption cycles driven by the rotation of the earth (Loudon, 2012). This 2.5 billion years’ evolution witnessed the huge expansion of life forms as well as the better adaptation of all organisms to the oxygen level in the atmosphere. While oxygen has now largely relieved its impact on the subsistence of life, the circadian clock still serves to anticipate the daily and annual environmental fluctuations and couples all facets of physiology with these rhythms.

The first documentation of circadian rhythms dates back to 4th century B.C. when daily leaf movements of tamarind tree were noticed (Bretzl, 1903). Early studies of the plant circadian clock discovered four defining features of the circadian clock. The circadian clock is self-sustained (de Mairan, 1729), has a 24-hour period (de Candolle, 1832), can be entrained by environmental cues (de Candolle, 1832) and may buffer temperature fluctuations (Bünning, 1931). Despite these historical focuses on the plant circadian clock, the first central clock oscillator was isolated in fruitfly (Bargiello et al., 1984; Reddy et al., 1984). The gene, TIMING OF CAB2 EXPRESSION 1 (TOC1), which is responsible for the first identified clock mutant in plant (Millar et al., 1995) was not cloned until 2000 (Strayer et al., 2000). The blooming of the genome era and the birth of the next generation sequencing epoch have revolutionized our ways to study complex systems like the circadian clock. The sequencing of Arabidopsis genome is the key to the
identification of TOC1 (Strayer et al., 2000), while the application of next generation sequencing technology recently helped to reverse the verdict of TOC1’s positive effect on CCA1 and LATE ELONGATED HYPOCOTYL (LHY) (Huang et al., 2012). Moreover, with the help of genomics and next generation sequencing technologies, people have isolated many more new components in plant central oscillators (Nagel and Kay, 2012) and unveiled the widespread circadian control on all spectra of physiologies (Michael et al., 2008).

The input signals, a central oscillator network and the output responses constitute an early understanding of plant circadian clock as a linear signaling pathway. However, accumulating evidences revealed extensive and complex interactions among these three components. Inputs can directly affect outputs both dependent and independent of the central clock. The central clock can gate its own sensitivities to the input signals. Outputs can feedback to central clock as well as the inputs. These mutual interactions obscure the discrete identities of inputs, the central clock and outputs. Therefore, plant circadian clock should be viewed as a complex system, whose emergent global characteristics cannot be generalized by studies on individual parts (Harmer, 2009). Instead, a system biology perspective needs to be considered.

So I will start the review on plant circadian system with our most recent understandings of the molecular architecture of central oscillators and emphasize the paradigmatic integration of theoretical modeling and genetic analysis through which
some historic breakthroughs were achieved in dissecting this complex system. As the reciprocal interactions between the central clock and its inputs or outputs blur the boundary between inputs and outputs, I will describe these two components together and emphasize the mutual effects among the inputs, the central clock and the outputs.

1.2.2. Molecular architecture of central circadian clock

After the initial cloning of TOC1 (Strayer et al., 2000) and the ensued discoveries of the reciprocal regulations between TOC1 and CCA1/LHY (Alabadi et al., 2001), a single TTFL was proposed to explain the oscillatory behavior of the central circadian clock. In this model, TOC1 promotes the transcription of CCA1 and LHY while CCA1 and LHY in turn repress the transcription of TOC1, thus leading to the oscillations. Theoretical modeling suggests that this single TTFL can maintain the correct period in constant darkness and light, be entrained to 24 hours’ light/dark cycles and produce correct waveforms of TOC1, CCA1/LHY expression during the day. However, it fails to reproduce the waveforms during the night and shows hypersensitivity to the complete or even partial loss-of-functions of any component which is contrary to the genetic data (Locke et al., 2005a). An extended version of single TTFL was later proposed which incorporated the effects of light on TOC1, a hypothetical gene X connecting TOC1 and CCA1/LHY and elevated protein degradation rate of TOC1 at night. Although it improved the reproducibility of the waveforms of TOC1, this model still cannot correctly predict the short period phenotype of cca1 or lhy mutant (Locke et al., 2005b).
These limitations have largely been alleviated by the introduction of a second hypothetical gene Y which activates TOC1 and is suppressed by both TOC1 and CCA1/LHY (Locke et al., 2005b). This modification of the previous model has qualitatively changed the topology from a single negative feedback loop to an interlocked feedback loop network. Moreover, the simulations of the characteristic features of this hypothetical Y enabled identification of GIGANTEA (GI) as a potential candidate for Y. The identity of Y is still under investigation as GI may contribute to Y activity indirectly. Nevertheless, it is owing to the system perspective of mathematical modeling rather than intuitive reasoning that allowed quantitative predictions of dynamic features of Y and the experimentally confirmed role of GI in central clock (Kim et al., 2007; Martin-Tryon et al., 2007; Sawa et al., 2007). This interlocked loop structure was then quickly expanded by incorporating newly discovered components into a three TTFLs network (Locke et al., 2006). It also allowed a qualitative and quantitative view of how light and clock interact with each other to regulate flowering time (Salazar et al., 2009).

The search for the hypothetical gene X was not fruitful. A large-scale yeast one-hybrid screen using CCA1 promoter and a genome-wide collection of TFs identified CCA1 HIKING EXPEDITION (CHE) as a middleman between TOC1 and CCA1 (Pruneda-Paz et al., 2009). However, instead of a direct positive relay to CCA1, TOC1 exerts its positive effect on CCA1 through CHE in a double negative way. Moreover,
CHE has no effect on LHY, which warrants a continuing quest for the hypothetical X. The search for X was finally terminated in an unexpected way. Three back-to-back studies from genetics, system biology, biochemistry and genomics testified TOC1 as a genuine TF which represses instead of activates CCA1 and LHY (Gendron et al., 2012; Huang et al., 2012; Pokhilko et al., 2012). More interestingly, theoretical simulation of this new architecture of the plant central clock suggested its underlying repressilator structure (Pokhilko et al., 2012), which is a famous gene regulatory circuit once artificially built in bacteria to create robust oscillations (Elowitz and Leibler, 2000). These findings lay the foundation for the current molecular framework of the plant central circadian clock.

These stepwise improvements in understanding of the overall blueprint of the plant central clock have gradually changed our fundamental knowledge of the design principles of a biological favorable oscillator. The isolations of additional central oscillators as well as revelation of the transcriptional, posttranscriptional and posttranslational regulations on the central clock components have greatly contributed to our realization of the high complexity needed for a robust internal chronometer and the multi-layer hierarchical controls on clock. However, the exact biological meanings of this complexity are yet to be understood (McClung, 2011). The increased number of components and their complex interactions with existing players in the central clock unequivocally points to a high complexity of this network in plants. Further considering
the fact that the unicellular green algae have a single feedback loop clock, one may think that the increased complexity in plants represents an evolutionary progress associated with more evolved organisms for a better flexibility and robustness compared to unicellular life. However, the observed robustness and sensitivity to environmental cues of the circadian clock in green algae challenged this instinct (Corellou et al., 2009; Morant et al., 2010; Thommen et al., 2010). Therefore, besides the continuous cloning of new circadian oscillators, discoveries of novel regulation mechanisms and expansions of the basic topology, we need to pay more attentions to the evolutionary design logic underpinning the structural topology and molecular composition.

1.2.3. Inputs and outputs of the circadian clock

Light, temperature, hormone levels, organic nitrogen intermediates, intracellular calcium and even hydration can be potent inputs to the central clock, while reciprocally, clock can modulate its sensitivity to these inputs through various pathways (McClung, 2011).

Light can control the transcription of core central oscillators including CCA1 (Wang and Tobin, 1998), LHY (Martinez-Garcia et al., 2000) and GI (Locke et al., 2005b). It can also regulate the mRNA stability of CCA1 (Yakir et al., 2007) and translation of LHY (Kim et al., 2003). Moreover, a blue light receptor ZEITLUPE (ZTL) also functions in the central clock to degrade TOC1 and be stabilized by GI. This dual role of ZTL shows the intimacy between light and clock (Kim et al., 2007). Although a less potent
Zeitgeber (time giver) compared to light, temperature is still a well-known input of the central clock. Temperature difference as small as 1 °C is enough to entrain Kalanchöe clock (Oltmanns, 1960). Although the limited knowledge of plant temperature sensor constrains the understandings of temperature entrainment, PSEUDO-RESPONSE REGULATOR 7 (PRR7) and PRR9 have already been shown to be essential in this process (Salome and McClung, 2005). Besides these two conventional Zeitgebers, various hormones play a role in regulating the clock (Hanano et al., 2006; Legnaioli et al., 2009). In addition, the discoveries described in Chapter 2 of this dissertation suggested that pathogen infection can also change the transcription level of CCA1 under diurnal and constant light conditions.

On the other side, the circadian clock regulates the entire life cycle of plants including the germination of seeds (Bryant, 1972; Zhong et al., 1998), hypocotyl elongation (Dowson-Day and Millar, 1999), transition from vegetative to reproductive development (Salazar et al., 2009), flower opening time (van Doorn and Van Meeteren, 2003), scent emission (Dudareva et al., 2003; Kolosova et al., 2001; Verdonk et al., 2003) and nectar secretion (Pesti, 1976) to attract pollinators and dormancy of woody plants during winter (Ramos et al., 2005). At the cellular level, the opening and closure of stomata follows a diurnal rhythm and its sensitivity to light is gated by the circadian clock (Gorton et al., 1993). More importantly, the circadian clock tightly controls the energy-producing process such as photosynthesis and the energy-consuming processes.
like nitrogen assimilation to ensure the energy balance (Harmer et al., 2000). In fact, the promoter of light-harvesting CHLOROPHYLL A/B-BINDING PROTEIN (CAB2) was historically used as a robust oscillatory reporter in genetic screens to identify central clock components. Genome-wide expression analysis (Covington et al., 2008) and enhancer trapping experiment (Michael and McClung, 2003) suggested that one-third of the Arabidopsis transcriptome display oscillations under free-run conditions after entrained by photocycles. Up to 90% of the Arabidopsis transcriptome cycles under certain conditions according to a more comprehensive survey of plants grown under thermocycles, photocycles or free-run conditions (Michael et al., 2008). It is evident that the circadian clock can regulate the abovementioned input signals through modulating the transcription of key players involved in the biosynthesis, metabolism and signaling pathways of hormones. The findings in Chapter 2 demonstrate the first example of how central oscillator CCA1 is engaged by plant immune system to time different immune responses according to the perceptions of distinct pathogenic signals. These results have also been supported by the recent chromatin immunoprecipitation-sequencing (ChIP-seq) data of TOC1 that genes associated with biotic stress are overrepresented in TOC1 target list (Huang et al., 2012).

Taken together, the reciprocal interactions among the central clock, inputs and outputs establish a densely connected web, in which any isolated view on individual
node or interaction cannot provide a holistic understanding of the emergent behaviors of the entire system.

1.3. **Introduction to translational control**

1.3.1. **Overview**

Transcriptional regulations on gene expression have attracted much attention over the decades. However, the ultimate biological functional units are mainly proteins. Therefore, in principle, the rate of translation rather than the rate of transcription should be more correlated with the abundance of protein. Indeed, a genome-wide measurement of relative contributions to the cellular protein levels confirmed the predominant role of translation rate (Schwanhausser et al., 2011). This highlights the direct and critical impact of translational control on cellular responses to environmental stimuli.

There are several advantages of translational control over transcriptional regulation. First of all, as a downstream process, translational control enables instantaneous responses with actual production of newly synthesized proteins. Second, enclosed in nucleus, transcriptional control lacks the spatial dimensions possessed by translational regulation. Therefore, localized protein production can only be achieved by translational control. There are a lot of different components involved in translation. Regulation on any of these targets can trigger translational control. Therefore translational control has many action modes and is very flexible. Moreover, the fine
tuning can be achieved when both translational and transcriptional regulations are deployed.

Translation involves three sequential steps including initiation, elongation and termination. During initiation, multiple processes result in the correct positioning of 80S ribosome at the start codon of the mRNA. Then the ribosome reads through the mRNA to synthesize the encoded polypeptide until it recognizes the stop codon, a trigger for termination. This is followed by the release of the newly synthesized polypeptide and dissociation of the subunits of ribosome. Translational regulation can take place in any one of these steps, but most regulations happen at the initiation step.

Despite a few exceptions, the paradigmatic and also the most common mode of initiation in eukaryotes is cap-dependent initiation of mRNAs with poly(A) tail. This is best understood in yeast (Lackner and Bahler, 2008). The first step of initiation is the assembly of ternary complex and 40S ribosomal subunit to form 43S preinitiation complex (PIC). The ternary complex is composed of eukaryotic initiation factor 2 (eIF2), methionyl-initiator tRNA (Met-tRNA\textsuperscript{Met}) and GTP while eIF2 itself is a hetero-trimer consisting of α, β and γ subunits. With the help of eIF1 and eIF1A, ternary complex is recruited to 40S subunit. Additional factors including eIF3 and eIF5 also reside with PIC. At the meantime, the poly(A) tail and the 5′ cap structure of the mRNA to be translated are bound by poly(A)-binding proteins (PABPs) and eIF4F respectively. The binding between PABPs and eIF4F circularizes mRNA, which has been considered as a structural
basis for 3’ untranslated region (UTR) binding proteins to regulate initiation (Gebauer and Hentze, 2004). With concerted contributions from eIF4F, eIF3 and PABP, PIC is attached to the mRNA ready for scanning. Then PIC scans along the mRNA until it reaches a start codon. Upon recognition of a start codon by the anticodon of Met-tRNA$_{\text{Met}}$ together with the help of eIF1, eIF2-GTP is hydrolyzed by eIF5 and becomes the inactive eIF2-GDP. As the tasks have been fulfilled, most initiation factors dissociate from PIC. Finally, 60S ribosomal subunit is recruited by eIF5B-GTP to form 80S ribosome, during which eIF5B-GTP is hydrolyzed to provide the energy. At this point, the ribosome acquires its ability to synthesize the encoded peptide from the mRNA. With the plethora of components involved in initiation, it is not difficult to imagine a large number of regulation targets. Initiation factors, mRNA itself and the composition of ribosome have all been shown to be translational regulation targets. These regulations can be either global or gene-specific. Since initiation factors are usually universal, the regulations on eIFs generally have a broad impact. On the other hand, the cis-regulatory elements like uORF and hairpin structure in transcript can endow the mRNA with gene- or even transcript-specific regulation. However, these regulations are not always isolated. In the sections below, I will use the example of the translational control on common genes and GENERAL CONTROL NONDEREPRESSIBLE 4 (GCN4) upon amino acid starvation to illustrate a global
regulation via control on ternary complex formation, a gene-specific regulation through uORF and their coordinated function to overcome this abiotic stress.

1.3.2. Global regulation via control on ternary complex formation

Amino acids are building blocks for all proteins. Therefore they are fundamental components of translation. When challenged by amino acid starvation, rational responses should include translational repression of nonessential proteins at the moment to save the limited reservoirs of amino acids and the induction of amino acid biosynthesis pathway to overcome the shortage. Indeed, global translation is repressed through the modulation of ternary complex formation.

As it was discussed in the overview, after initiation, GTP in the ternary complex is hydrolyzed, which generates eIF2-GDP. Compared to eIF2-GTP, the binding affinity between eIF2-GDP and Met-tRNA\textsubscript{Met} is reduced by tenfold. To revive it, a guanine nucleotide exchange factor (GEF), eIF2B is required. However, phosphorylated eIF2 at its α subunit (eIF2α-P) can serve as a strong competitor to eIF2B, impeding the regeneration of active eIF2. But how is eIF2α-P generated (Hinnebusch, 2005)? The lack of amino acid will inevitably cause an increase in uncharged tRNA which does not carry amino acid. This increase enhances tRNA binding affinity to the protein kinase GCN2 and activates the kinase activity of GCN2. GCN2 then phosphorylates eIF2 to generate eIF2α-P. The regeneration inhibition of the active eIF2 decreases the abundance of ternary complex, therefore repressing global translation initiation.
1.3.3. Gene-specific regulation through uORF

Repression of global translation alone is not an effective solution to cope with amino acid starvation. An orchestrated induction of amino acid biosynthesis pathways has to occur.

First of all, to assure the coordination, the genes involved in amino acid biosynthesis should be co-induced. In fact, GCN4, the master transcriptional regulator, can promote the transcription of more than 30 such genes, spanning 12 different pathways in response to amino acid starvation (Hinnebusch, 2005).

The question is how GCN4 is induced when the global translation is repressed. The secret lies in its 5’ UTR which contains four uORFs (Hinnebusch, 2005). Under normal conditions, PIC usually initiates at uORF1 and continues scanning to reinitiate at uORF2, uORF3 or uORF4. However, uORF3 and uORF4 can function as effective blocks to prevent reinitiation at GCN4. During amino acid starvation, the ternary complex becomes limited. The recruitment of ternary complex becomes significantly slower. Therefore due to the short spacing between uORF1 and uORF4, while the small subunit of ribosome scans this spacing, it is less likely to recruit the ternary complex. However, the spacing between uORF1 and GCN4 is long enough so that when the small subunit reaches the start codon of GCN4, the ternary complex can be recruited in time. In this way, ribosome simply ignores the start codons of uORF2, uORF3 and uORF4 en route to GCN4.
Executed in a concerted manner, the global and gene-specific translational regulations save the starved cells by decrease in expenditure and increase in synthesis at the same time.

1.4. Dissertation outline

Plant immune system is powerful. However, constant activation of immunity requires a huge amount of energy and is also hazardous to the plant itself. Therefore it is important to modulate the strength and timing of defense responses to ensure the right level of immunity at the right time in the right place.

Chapter 2 of this dissertation focuses on the circadian clock regulation of a set of novel immune regulators which participate in both basal and ETI. The interactions between the circadian clock and the immune system can dictate distinct expression dynamics of these immune genes when MAMPs or effectors are perceived. Moreover, even without the presence of pathogens, these genes can be regulated in a circadian manner to establish a proper level of defense at the most vulnerable time of the day. Considering the dominant regulatory role of the circadian clock in energy-production related pathways, circadian control on defense echoes its cardinal function to balance energy utilization. So in the case of defense, circadian is engaged to efficiently equilibrate energy consumptions in growth and defense.

Little is known about how SAR is controlled to coordinate growth and defense against the imminent pathogen infections. Chapter 3 aims to demonstrate that
translational control on the key regulator of SAR, TBF1 which is also a molecular switch for growth-to-defense transition, can tightly couple the metabolic status of the cell with the induction of immunity.

Finally, Chapter 4 highlights some of the novel research directions enlightened by the discoveries in Chapter 2 and 3.
2. Timing of Plant Immune Responses by a Central Circadian Regulator

Modified from:


2.1. Introduction

As described Chapter 1, ETI is the principle immune mechanism against phytopathogens, which is under strong selection pressure. The interaction between obligate biotrophic pathogens and their host plants provides an important model system to study ETI, owing to their intimately linked life cycles.

Hyaloperonospora arabidopsidis (Hpa, formerly known as Peronospora parasitica), an obligate biotrophic oomycete which can cause downy mildew disease in Arabidopsis, is one of the few pathogens that infect Arabidopsis in nature. Though the observation of their interactions has been documented since 1865 (Slusarenko and Schlaich, 2003), the model system for plant immune research was not established until 1990 (Koch and Slusarenko, 1990). The clearly defined infection steps of Hpa allow better dissection of the corresponding resistance mechanisms blocking these steps. During the asexual life cycle of Hpa, the sporangiospores germinate on the leaves. The infection pegs then penetrates epidermal cells. The hyphae grow between cells and then form a feeding
structure called the haustoria by penetrating cell wall and invaginating cell membrane to rob nutrients from the plant cells.

*Hpa* is likely an ancient pathogen to *Arabidopsis* because of the tremendous phenotypic, genotypic and geographic diversity of the interactions between the different ecotypes of *Arabidopsis* and the isolates of *Hpa* (Slusarenko and Schlaich, 2003). Accordingly, different isolates of *Hpa* are named after the location where they were found and a susceptible *Arabidopsis* accession, both of which are abbreviated to two letters. For instance, one of the isolates used in this study, Emwa was found in East Malling = Em and is virulent on the Wassilewskija ecotype of *Arabidopsis*, thus wa. Based on this rule, strain Noco is virulent on Columbia-0 (Col-0).

The compatibility between different isolates of *Hpa* and ecotypes of *Arabidopsis* follows the gene-for-gene hypothesis (Flor, 1971). The perception of effectors from *Hpa* by R proteins in *Arabidopsis* will lead to ETI and render the *Hpa* isolate harboring this effector avirulent to this ecotype, while lack of perception results in ETS and virulence. For example, Emwa is avirulent on Col-0 ecotype due to the presence of the R protein, **RECOGNITION OF PERONOSPORA PARASITICA** 4 (RPP4) in Col-0.

Despite the successful cloning of multiple RPP genes, the downstream pathway remains largely unknown. Even less is clear about the regulatory system that controls the downstream defense pathway. To identify these missing parts, a comprehensive reverse genetic screen was previously initiated in our lab. Wild-type plants (WT) and
*rpp4* mutant were subjected to the challenge from Emwa. A time course whole genome expression profiling was performed at 0, 0.5, 2 and 4 days post inoculation (dpi). A total of 106 genes were identified to be differentially expressed in WT and *rpp4* at 2 dpi. These genes were induced earlier than the previously reported immune regulators including ENHANCED DISEASE SUSCEPTIBILITY 5 (*EDS5*), PHYTOALEXIN DEFICIENT 4 (*PAD4*), AVRPPHB SUSCEPTIBLE 3 (*PBS3*), ICS1, NON-RACE SPECIFIC DISEASE RESISTANCE-1 (*NDR1*) and *EDS1* (van der Biezen et al., 2002), which were known to function downstream of R gene activation. T-DNA insertion mutants of these 106 genes were assayed with respect to their susceptibility to Emwa. Among them, 22 gene mutants displayed enhanced susceptibility compared to wild type based on sporangiophore (SPP) growth and other disease symptoms (for example, chlorosis) by microscopic inspection. For most of the 22 genes, at least two homozygous mutant T-DNA alleles were tested.

To identify specific resistance defects in each mutant, the infected plants were stained with lactophenol trypan blue (LTB) 7 dpi and scored for the occurrence of the seven phenotypes represented in Figure 2-1a. Since the sampling was done at 7 dpi, we were able to capture extreme as well as intermediate phenotypes of the entire infection cycle. Discrete HR (DIH) represents RPP4-mediated HR cell death, whereas SPP (an asexual reproductive structure that emerges on the leaf surface) indicates completion of the pathogen life cycle. Besides these two extreme phenotypes, oospores (OOS) are the
sexual spores formed *in planta*; trailing necrosis (TRN) depicts the presence of host cell necrosis along the path of hyphal growth; ‘free hypha’ (FRH) stands for the hyphal growth without host cell necrosis; while free hyphal intermediate (FHI) suggests disrupted hyphal growth. We also examined the presence of expanding HR (EXH), which depicts larger clusters of dead host cells surrounding infection sites than DIH. As shown in Figure 2-1b, *rpp4* mutant had the highest percentage of leaves with SPP, confirming that its resistance to Emwa is completely compromised. WT had the highest score of DIH, indicating its resistance against Emwa. The phenotype scores are also presented numerically in Table 2-1 and the mutants are ranked on the basis of their SPP scores.
Figure 2-1: Characterization of mutants in RPP4-mediated resistance against Emwa

(This is a collaborative study result with Dr. Jinyoung Barnaby)

a, The seven phenotypes scored in WT, rpp4, and 22 mutants 7 dpi with Emwa.

SPP, Sporangioophore; TRN, Trailing necrosis; OOS, Oospore; FRH, Free hypha;
FHI, Free hyphal intermediate; EXH, Expanding HR; DIH, Discrete HR. 

Bar graphs showing distributions of phenotype (PT) scores (% of occurrence in 40 leaves/genotype) of the mutants (black bars) compared with WT (blue bar) and rpp4 (red bar). Mutants are ranked from the highest to the lowest with no correspondence from panel to panel. The error bars, 95% confidence intervals based on binomial distribution.
Table 2-1: Phenotypic scores and ranks of gene mutants

(This is a collaborative study result with Dr. Jinyoung Barnaby)

<table>
<thead>
<tr>
<th>No.</th>
<th>Description</th>
<th>SPP</th>
<th>TRN</th>
<th>OOS</th>
<th>FRH</th>
<th>FHI</th>
<th>EXH</th>
<th>DIH</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPP4</td>
<td>RPP4</td>
<td>98*</td>
<td>55*</td>
<td>73*</td>
<td>100*</td>
<td>25*</td>
<td>8*</td>
<td>3*</td>
</tr>
<tr>
<td>1</td>
<td>LRR-kinase</td>
<td>78*</td>
<td>10*</td>
<td>88*</td>
<td>88*</td>
<td>93*</td>
<td>20*</td>
<td>5*</td>
</tr>
<tr>
<td>2</td>
<td>Calcineurin-like phosphoesterase</td>
<td>78*</td>
<td>15*</td>
<td>100*</td>
<td>100*</td>
<td>65</td>
<td>13*</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>ADR1-like1</td>
<td>75*</td>
<td>0</td>
<td>80*</td>
<td>95*</td>
<td>85*</td>
<td>5*</td>
<td>3*</td>
</tr>
<tr>
<td>4</td>
<td>Aspartic protease 38</td>
<td>68*</td>
<td>13*</td>
<td>95*</td>
<td>90*</td>
<td>18*</td>
<td>5*</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>UDP-glucosyl transferase 72E1</td>
<td>65*</td>
<td>18*</td>
<td>95*</td>
<td>90*</td>
<td>93*</td>
<td>8*</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Ankyrin</td>
<td>65*</td>
<td>13*</td>
<td>95*</td>
<td>90*</td>
<td>83*</td>
<td>28*</td>
<td>10*</td>
</tr>
<tr>
<td>7</td>
<td>Disease resistance protein</td>
<td>65*</td>
<td>15*</td>
<td>90*</td>
<td>80*</td>
<td>90*</td>
<td>45*</td>
<td>43*</td>
</tr>
<tr>
<td>8</td>
<td>Protein kinase family protein</td>
<td>65*</td>
<td>0</td>
<td>95*</td>
<td>73*</td>
<td>95*</td>
<td>15*</td>
<td>25*</td>
</tr>
<tr>
<td>9</td>
<td>ABC transporter family protein</td>
<td>58*</td>
<td>0</td>
<td>93*</td>
<td>93*</td>
<td>95*</td>
<td>8*</td>
<td>0*</td>
</tr>
<tr>
<td>10</td>
<td>ATPase-like domain containing protein</td>
<td>55*</td>
<td>10*</td>
<td>60*</td>
<td>88*</td>
<td>95*</td>
<td>18*</td>
<td>8*</td>
</tr>
<tr>
<td>11</td>
<td>Ankyrin repeat protein</td>
<td>55*</td>
<td>3*</td>
<td>85*</td>
<td>90*</td>
<td>45</td>
<td>43*</td>
<td>28*</td>
</tr>
<tr>
<td>12</td>
<td>Lipid class 3 family protein</td>
<td>55*</td>
<td>23*</td>
<td>100*</td>
<td>55*</td>
<td>100*</td>
<td>58</td>
<td>13*</td>
</tr>
<tr>
<td>13</td>
<td>Constitutive shade avoidance 1</td>
<td>43*</td>
<td>50*</td>
<td>75*</td>
<td>50*</td>
<td>75*</td>
<td>25*</td>
<td>25*</td>
</tr>
<tr>
<td>14</td>
<td>Cyclic nucleotide gated channel 3</td>
<td>40*</td>
<td>10*</td>
<td>83*</td>
<td>13</td>
<td>93*</td>
<td>85*</td>
<td>55</td>
</tr>
<tr>
<td>15</td>
<td>Lipid transfer family protein</td>
<td>38*</td>
<td>0</td>
<td>88*</td>
<td>43*</td>
<td>88*</td>
<td>60</td>
<td>8*</td>
</tr>
<tr>
<td>16</td>
<td>Tyrosine aminotransferase 3</td>
<td>28*</td>
<td>5*</td>
<td>95*</td>
<td>35*</td>
<td>93*</td>
<td>60</td>
<td>28*</td>
</tr>
<tr>
<td>17</td>
<td>Glycoside hydrolase</td>
<td>25*</td>
<td>5*</td>
<td>78*</td>
<td>35*</td>
<td>78*</td>
<td>40*</td>
<td>40*</td>
</tr>
<tr>
<td>18</td>
<td>Cold-responsive protein</td>
<td>25*</td>
<td>3*</td>
<td>95*</td>
<td>30*</td>
<td>85*</td>
<td>65</td>
<td>48</td>
</tr>
<tr>
<td>19</td>
<td>Aspartyl protease family protein</td>
<td>20*</td>
<td>5*</td>
<td>90*</td>
<td>38*</td>
<td>90*</td>
<td>70</td>
<td>50</td>
</tr>
<tr>
<td>20</td>
<td>BON-association protein 1</td>
<td>20*</td>
<td>8*</td>
<td>73*</td>
<td>30*</td>
<td>85*</td>
<td>78*</td>
<td>35*</td>
</tr>
<tr>
<td>21</td>
<td>LRR-kinase</td>
<td>8</td>
<td>0</td>
<td>60*</td>
<td>28*</td>
<td>58*</td>
<td>80*</td>
<td>50</td>
</tr>
<tr>
<td>22</td>
<td>Ser/Thr protein kinase family protein</td>
<td>8</td>
<td>0</td>
<td>68*</td>
<td>40*</td>
<td>73*</td>
<td>83*</td>
<td>38*</td>
</tr>
<tr>
<td>WT</td>
<td>Wild type</td>
<td>5</td>
<td>0</td>
<td>48</td>
<td>15</td>
<td>40</td>
<td>63</td>
<td>55</td>
</tr>
</tbody>
</table>
Phenotype scores are percentage in 40 leaves per genotype. SPP, sporangiophore; TRN, trailing necrosis; OOS, oospore; FRH, free hypha; FHI, free hyphal intermediate; EXH, expanding hypersensitive response; DIH, discrete hypersensitive response. * P<0.05.
2.2. Results

2.2.1. Phenotypic clustering of immune gene mutants upon Emwa infection highlights the significance of PCD

The statistical tests applied to the phenotypic scores in Table 2-1 demonstrated the significant difference between WT and mutants. Although this helped to confirm the observed enhanced disease susceptibility to Emwa during the reverse genetic screen in a quantitative way, it cannot integrate the information embedded in all these 7 phenotypes to unveil the common features of different gene mutants. The inherent high-dimension feature of this analysis requires effective and efficient data digestion. Data clustering is a widely used method to handle such high-dimension datasets. It aims at the identification of groups of subjects with similar features within the group and distinct between groups. Through this technique, we may learn the underlying structure of the data, classify the objects and compress or reduce the data dimension.

Here I applied Pvclust, a hierarchical clustering method, to the phenotypic scores (Suzuki and Shimodaira, 2006). A major advantage of this method over the standard hierarchical clustering lies in the implementation of approximately unbiased (AU) p-value to account for the uncertainty of clusters. The multiscale bootstrap resampling strategy used to generate AU p-value is superior to the ordinary bootstrap method in terms of bias.

Hierarchical clustering of the mutants using the phenotype scores of multiple mutant alleles (Figure 2-2) places them into two groups. The power of this phenotypic
clustering is suggested by both the high AU p-values and the fact that different mutant alleles of the same gene reside in the same big cluster. Additionally, similar groupings are also obtained with data from three biological replicates (Figure 2-3). The Group 1 mutants (red numbers in Figures 2-2 and 2-3) seem to be defective in R gene-mediated PCD (low DIH and EXH scores) with high disease symptoms (FRH and SPP). In contrast, the Group 2 mutants (blue numbers) appeared to be intact in PCD with high EXH and DIH scores and milder symptoms (low FRH and SPP scores).

To identify the main discriminators among the 7 phenotypes studied and determine the redundancy of them, a principal component analysis (PCA) was performed (Figure 2-4). The derived eigenvectors indicate that 63.8% of the phenotype variations could be accounted for by principal component 1 (PC1) with indicators of resistance, DIH and expanding hypersensitive response (EXH), as positive contributors and disease phenotypes, SPP and free hypha (FRH), as negative contributors. While this identifies these four phenotypes as the main discriminators, it also quantitatively suggests the contradictory relationship between host cell death and susceptibility symptoms. If PC2 is considered, which together with PC1 can explain over 75% of the data variance, six out of the seven phenotypes tested are important to separate the gene mutants.

Taken together, these observations suggest that RPP4 regulates at least two separate responses (Figure 2-5): Group 1 genes are required for R-mediated PCD, as
mutations in these genes led to low DIH and EXH scores and formation of FRH, TRN, and SPP. The Group 2 genes are probably involved in defense responses independent of PCD. Loss of these latter functions resulted in pathogen invasion even in the presence of PCD. One important conclusion from these data is that PCD is the predominant resistance response against Emwa because the Group 1 mutants were more susceptible, on the basis of the SPP scores (except Mutant 12), than the Group 2 mutants. This is supported by the eigenvector composition where PC1 (DIH and EXH) was the major contributor to the phenotypic variations (Figure 2-4). This finding is consistent with the fact that Emwa is an obligate biotrophic pathogen. Suicidal death of the host cell means the end of the pathogen life cycle. The functional diversity of the Group 1 genes indicates that RPP4-mediated PCD is orchestrated by changes in multiple biological processes, rather than a single triggering event.

2.2.2. Mutants defective in PCD are also compromised in basal defense

As it was mentioned in Chapter 1, the connection between the signal-specific R-mediated resistance and the broad-spectrum basal defense has been speculated previously. Transcriptional profiling suggests that temporal control of similar transcriptomes may distinguish R-mediated immunity from basal resistance. Identification of a large number of genes involved in RPP4-mediated resistance allowed a systematic test of this hypothesis. The mutants were thus subjected to infection by the virulent isolate, Noco, to which a cognate R gene is absent in Col-0, to test their potential
involvement in basal defense. 10 of the mutants showed enhanced disease susceptibility (Figure 2-6), while 8 out of these 10 gene mutants belong to Group 1. Therefore these genes do have dual roles in both basal defense and ETI.

2.2.3. Bioinformatics analysis on the immune genes suggests potential regulation by the circadian clock

The identification of common components involved in both pro- and anti-cell death immune responses raises the question of how activation of similar sets of genes causes PCD in RPP4-specific resistance against Emwa and nonspecific basal resistance against Noco without cell death. To understand the differential regulation of these immune mechanisms, I analyzed the promoter regions of these 22 genes. Using the Athena program (O’Connor et al., 2005), I found significant enrichment of only one cis-regulatory element, the evening element (EE), which is regulated both positively and negatively by the circadian regulator, CCA1 (Harmer et al., 2000; Harmer and Kay, 2005). Further examination showed that 14 of the 22 genes contain either evening element and/or the CCA1-binding site and/or have rhythmic expression patterns (Table 2-2) (Michael et al., 2008). Importantly, the promoter region of RPP4 also contains two evening elements and its expression shows a circadian rhythm. Collectively, these bioinformatics clues suggest the potential circadian clock regulations on these novel defense genes.
2.2.4. Circadian clock gene mutants are defective in ETI

To confirm the involvement of the circadian clock in defense experimentally, circadian clock gene mutants were inoculated with Emwa. The infection was carried out at dawn, the time when *Hpa* spores are normally disseminated in nature (Slusarenko and Schlaich, 2003). The *cca1* mutant (Salk_067780) and *ztl-4* (a mutant of *ZTL*) (Salome and McClung, 2005) showed compromised resistance whereas a CCA1-overexpression line (CCA1_{OE}) (Wang and Tobin, 1998) showed enhanced resistance (Figure 2-7). Surprisingly, *lhy* (Schaffer et al., 1998), the mutant of the CCA1 homolog, LHY responded as wild type. It is remarkable that while CCA1 and LHY function redundantly in conferring circadian rhythms, CCA1 appears to have evolved to link clock and defense functions.

2.2.5. Validation of RNA annealing selection ligation-sequencing (RASL-seq)

The confirmation of the involvement of the circadian clock in plant immune system through both bioinformatics analysis and genetic verification revealed a new interface between plant host and biotrophic pathogen. The dynamic control by the circadian clock on these defense genes requires a periodic sampling scheme to investigate this regulatory mechanism. To elucidate the combinatorial effects of CCA1, RPP4 and Emwa on the temporal expression pattern of these genes, a large scale expression profiling was performed in which WT, *cca1, rpp4* plants were sampled every 2 hours for 46 hours after infected by Emwa or control solution. In collaboration with Dr.
Xiang-dong Fu at UCSD, I adapted RNA annealing selection ligation-sequencing (RASL-seq) technology originally developed in the mammalian system (Li et al., 2012) to Arabidopsis. Briefly, I designed two oligos specific to the sequences flanking an intron of the gene of interest to distinguish the mRNA from the genomic DNA because only in the mRNA the intron is removed and the two primer targets become adjacent to each other. This allows annealing and ligation of the two oligos. Finally a barcode based polymerase chain reaction (PCR) and high throughput sequencing were performed to generate reads of genes from all the samples in a single run from which expression levels can be inferred by the occurrence of a defined gene sequence with a specific barcode corresponding to the designated sample.

It is crucial to validate the quality of RNA preparations and the RASL-seq. The rhythmic expression of LHY was not significantly altered in WT or rpp4 (Figure 2-8) provided an internal control for the rest of the genes tested. This result also indicates that RPP4-mediated defense does not disrupt the overall running of the clock, but rather engages CCA1. This finding was further confirmed using transgenic line expressing CCA1:LUC and LHY:LUC under constant light condition (Figure 2-9). Firefly luciferase (LUC) can decarboxylate luciferin in an ATP-dependent manner, which releases a photon at 560 nanometer (nm). The intensity of this light emission, a direct reflection of the transcription level (McClung, 2006), can be quantified by a sensitive charge coupled device (CCD) camera. Similar to the RASL-seq results, LHY:LUC expression remained
unchanged, whereas CCA1:LUC expression was significantly induced and even became arrhythmic upon Emwa challenge.

Finally, while the CCA1 expression pattern changes caused by pathogen infection or in the rpp4 mutant background was explainable, the altered expression profile of CCA1 under control conditions in WT was puzzling at first. Similar expression pattern of CCA1 under control condition was observed in the CCA1:LUC transgenic line showed that this result was not an artifact. I reasoned that since this control condition was set to optimize Hpa growth, which involved alternating of different humidity levels, the unexpected expression pattern of CCA1 might be caused by the humidity changes. I hypothesize that CCA1 may actually sense this environmental cue and transduce the signal to the downstream defense genes so that plants can ‘anticipate’ the potential infections under this favorable condition. This hypothesis will be further discussed and tested in the following sections.

2.2.6. RASL-seq analysis unveils ‘anticipation’ model

Just like the phenotypic scores, RASL-seq result is an even higher dimension dataset with 24 time points, 3 genotypes, 2 treatment conditions and more than two dozens of genes. Therefore, to enable a manageable inference, a data clustering is needed.

The high dimension of this dataset makes Pvclust computationally infeasible. Therefore I performed an alternative clustering analysis, nonnegative matrix factorization (NMF), which has been successfully applied to study high dimensional
expression profiles (Tamayo et al., 2007). Again, the expression patterns of these 22 defense genes can be divided into two clusters based on cophenetic correlation coefficient (Figure 2-11). The membership distance of each gene to its cluster is illustrated by the circle radii in Figure 2-12. Since the algorithm of NMF involves random sampling procedure, it is important to ensure the convergence of the algorithm through monitoring of the cost values associated with increasing numbers of runs (Figure 2-13).

These two clusters corresponded roughly to the two phenotypic groups determined by reverse genetic analysis (Table 2-2). Most of the Cluster 1 genes containing evening element in their promoters are involved in R-gene-mediated PCD and were therefore the focus of further concern (Figure 2-14). Anyhow, the expression patterns of the Cluster 2 genes are shown in Figure 2-15.

Consistent with the fact that evening element is enriched in the Cluster 1 gene promoters (Table 2-2), the weighted mean expression of these genes largely overlaps with the expression patterns of CCA1. In wild-type control (Col CK), Cluster 1 genes showed a rhythmic expression pattern with a single sharp peak every evening. In cca1 (cca1 CK), the expression peaks were greatly diminished, confirming that CCA1 is an activator of these defense genes.

Based on the detailed temporal expression patterns, I propose the following ‘anticipation’ model which conceptually describes the circadian clock control on defense
genes to time different immune responses according to the perception of different environmental cues and pathogenic signals (Figure 2-16).

The rhythmic expression of the defense genes in the absence of pathogen indicates that plants are programmed to ‘anticipate’ infection according to a circadian schedule. The CCA1-mediated pulse expression of the defense genes coincides with the time of *Hpa* sporulation which mainly occurs at night and the time of spore dissemination which takes place at dawn (Slusarenko and Schlaich, 2003).

Upon Emwa infection, Cluster 1 genes showed drastically different expression patterns in wild type (Col Emwa) and *rpp4* (*rpp4* Emwa) (Figure 2-14). Without RPP4, the expression of the defense genes peaked at the 6-, 16- and 24-h time points which coincided with the expected time of *Hpa* spore germination, formation of penetration hyphae and establishment of primary haustoria in mesophyll cells, respectively (Donofrio and Delaney, 2001). These passive responses to the key infection stages of *Hpa* are not effective against the invasion.

In the presence of RPP4, the 6-h expression peak was diminished (Figure 2-14). The perception of the pathogen effector by RPP4 led to the gradual and sustained expression of defense genes. We propose that the prolonged expression of these defense genes, which are normally pulse-expressed at dawn, results in PCD of the infected host cells and pathogen resistance. CCA1 has a role in controlling this prolong expression.
since the duration of this induction becomes shorter in CCA1 although not completely abolished.

2.2.7. Diurnal difference in ETI, enhanced PCD and basal defense in CCA1 overexpression plants validate the ‘anticipation’ model

According to the ‘anticipation’ hypothesis, the pulse expression pattern of the defense genes is intended to prepare the plants for the potential attack of *Hpa* during dawn. We shall expect to see enhanced disease susceptibility if the plants are infected at dusk, an unexpected time of invasion. Indeed, this is what I have recorded (Figure 2-17). I performed Emwa infection not only at the normal ‘dawn’ infection time but also at ‘dusk’. I found that if the plants were inoculated at dusk, significantly higher levels of susceptibility were observed in both WT and *rpp4*. This indicates that both ETI and basal resistance are controlled in a diurnal manner so that the defense level is higher at dawn. It is interesting to notice that RPP4 is also involved in this regulation since the loss of RPP4 reduced the difference of susceptibility between dawn and dusk. CCA1 clearly plays a predominant role in conferring resistance at dawn because in *cca1*, more *Hpa* Emwa growth was observed compared to WT. Moreover, the diurnal difference of susceptibility was completed abolished in *cca1* because CCA1 and the CCA1-regulated defense genes are not expressed at all times.

Besides the anticipation under normal conditions, the model (Figure 2-16) also proposed that CCA1 positively controls the induction of the defense genes during basal resistance. This is supported by the fact that CCA1OE was more resistant to Noco than
WT, confirming its role in the positive regulation of basal defense (Figure 2-18).

However, how CCA1 perceives the critical infection stages of Emwa which leads to this specific timing of defense gene expression will require future research.

Finally, to validate the involvement of CCA1 in regulating Cluster 1 defense genes during RPP4-mediated PCD, I checked the cell death phenotypes of cca1 mutant as well as CCA1oe (Figure 2-19). In accordance with its positive role, DIH score (a measure of PCD) of cca1 was dramatically lower than WT while the overexpression line had statistically significantly more cell death occurrences than WT.

Taken together, these experiments together with the RASL-seq data support the ‘anticipation’ hypothesis in which central circadian oscillator CCA1 is engaged by plant immune system to time preventive, basal or elevated defense responses according to the perceptions of environmental cues or pathogenic signals.
Figure 2-2: Phenotypic clusters with multiple mutants alleles

Mutants were clustered on the basis of their phenotype scores. Second allele, ‘A’.

Group 1, red; Group 2, blue. au, approximately unbiased P-values (0–100%, the higher the number the more significant).
Figure 2-3: Phenotypic clusters using data from 3 biological replicates

Hierarchical clustering performed using the mean phenotype scores of the three biological replicates with bootstrap (100,000 times) showed that mutants were divided into 2 groups similar to Figure 2-2 (Mutants were colored according to Figure 2-2). Distance was measured by the standard correlation coefficient (average linkage; scale 0-1)
Figure 2-4: Principal component analysis of phenotypic scores

Eigenvectors derived from PCA. The percentage of phenotypic variations captured by each PC is shown.
Figure 2-5: Illustration of functional divergence of Group 1 and 2 mutants

A diagram showing that the Group 1 mutants are defective in RPP4-mediated PCD, whereas the Group 2 mutants are compromised in defense responses with intact PCD.
Figure 2-6: Some of the RPP4-mediated resistance mutants are also compromised in basal defense

Enhanced disease susceptibility to Noco based on sporangiospore count 7 dpi (n=3). Error bars represent standard error of mean. Mutants were colored according to Figure 2-2. * P<0.05, ** P<0.01, *** P<0.001.
**Table 2-2: Bioinformatics evidences of potential circadian clock control on the new defense genes**

<table>
<thead>
<tr>
<th>NMF cluster</th>
<th>EE</th>
<th>CCA1</th>
<th>Circadian Corr.</th>
<th>SPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AAAATATCT</td>
<td>AAA/CAATCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(+)-140</td>
<td>(+)-195</td>
<td>0.86</td>
<td>RPP4</td>
</tr>
<tr>
<td></td>
<td>(+)-270</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-)-58</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>(+)-380</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>(-)-1724</td>
<td></td>
<td>0.85</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(+)-88</td>
<td>(-)-11</td>
<td>0.95</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>(-)-593</td>
<td></td>
<td>0.88</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>(+)-1471</td>
<td>(1)-1447</td>
<td>0.87</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>(+)-412</td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>(+)-143</td>
<td>(-)-218</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>(-)-693</td>
<td></td>
<td>0.88</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>(-)-370</td>
<td>(+)-1982</td>
<td></td>
<td>12</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14</td>
</tr>
</tbody>
</table>

Evening element is enriched in the promoters of these new defense genes (P<e-5). NMF, nonnegative matrix factorization; CCA1, CCA1-binding sites; Circadian correl., circadian correlations. +, sense; -, antisense.
Figure 2-7: Involvement of the circadian clock in plant immune system

SPP count at 7 dpi by Emwa (n=3). Error bars represent standard error of mean. * P<0.05, *** P<0.001.
Figure 2-8: Different responsiveness of CCA1 and LHY to Emwa identified through RASL-seq under diurnal conditions

Expression profiles of CCA1 and LHY measured by RASL-seq. CK, control; Emwa, *Hpa* Emwa infected. White bars, day; black bars, night.
Figure 2-9: Different responsiveness of CCA1 and LHY to Emwa identified using *CCA1:LUC* and *LHY:LUC* lines under free-run conditions.

CK, control; Emwa, *Hpa* Emwa treated; ZT, Zeitgeber time. Error bars represent standard error of mean (n=18). This experiment has been repeated three times with similar results.
Figure 2-10: Confirmation of the unexpected expression pattern of CCA1 under control conditions using CCA1:LUC line

Catalytic activity of luciferase was measured in the protein extracts from the time course samples of CCA1:LUC experimented under the same conditions used for the RASL-seq. Two biological replicates were included. The white bars, the day; black bar, the night.
Figure 2-11: Determination of the cluster number used in NMF

A range of cluster numbers from 2 to 22 were implemented in NMF with 300 runs (10,000 iterations/run) and the corresponding cophenetic correlation coefficients were calculated. When cluster number equaled 2, the highest cophenetic correlation coefficient was achieved.
Figure 2-12: Membership distance for Cluster 1 and Cluster 2 genes
determined by their expression patterns

The two expression clusters overlap with the phenotypic grouping. For each
gene, its membership distance in the expression cluster is represented by the
radius of the circle. The calculation is detailed in Methods and Materials.
Figure 2-13: Convergence of NMF

300 runs (10,000 iterations/run) were performed using 2 clusters and the corresponding costs were calculated for each run. The NMF algorithm converged with fewer than 50 runs.
Figure 2-14: Time-course expression of NMF Cluster 1 genes

CI, confidence interval; CK, control; Emwa, *Hpa* Emwa infected. White bars, day; black bars, night.
Figure 2-15: Time-course expression of the NMF Cluster 2 genes together with CCA1 and RPP4 as references

The white bars, the day; black bar, the night. CK, control; Emwa, *Hpa* Emwa inoculated.
Figure 2-16: ‘Anticipation’ model illustrating the circadian clock control on plant immune system

This model shows circadian regulation of the defense genes in anticipation of infection under normal conditions, in basal and R-gene-mediated resistance. The blocked arrows represent defense against infection.
Figure 2-17: SPP count after Emwa infection at dawn or dusk

WT, *cca1* and *rpp4* were infected with Emwa at dawn or dusk according to the photoperiod in the growth chamber. At 7 dpi, SPP numbers were counted (n=3).

Error bars represent standard error of mean. * P<0.05, *** P<0.001.
Figure 2-18: Enhanced disease resistance to *Hpa* Noco observed in the CCA1<sub>OE</sub> transgenic line based on sporangiospore count at 7 dpi

The *rpp5* mutant in which the R protein responsible for the resistance against Noco in Landsberg (Ler-0) ecotype background is mutated is used as a positive control. The wild-type Ler-0 is used as a negative control. Error bars represent standard deviation (n=3, One-way ANOVA and Bonferroni’s Multiple Comparison Test); ** P < 0.01, *** P < 0.001. This experiment has been repeated twice with similar results.
Figure 2-19: Validation of the positive regulation of CCA1 on host cell death upon Emwa infection

Occurrence of DIH was counted at 7 dpi by Emwa. Error bars represent 95% confidence interval based on binomial distribution. * P<0.05, *** P<0.001.
2.3. Discussion

The circadian modulation of plant immune system has been postulated since the documentations of altered survival rate of clock mutants upon pathogen infection in fruitfly (Lee and Edery, 2008; Shirasu-Hiza et al., 2007). However, except some speculations on the possible links between the circadian clock and the regulation of plant defense system, the connection between the circadian lock and plant immunity has not been established firmly (Roden and Ingle, 2009). In this Chapter and the associated paper, I identified the circadian control on plant immunity serendipitously (Wang et al., 2011).

Initially, we aimed to find the signaling components of ETI. The identification of 22 novel defense genes expressed earlier than the previously reported immune regulators did appear to be a successful start. However, their diverse ontologies set a bottleneck to the downstream studies. Nevertheless, the power of this functional reverse genetic screen has been substantiated by the recent elucidation of the critical functions of some of these 22 genes. For example, gene #10, ATPase-like domain containing protein was identified as COMPROMISED RECOGNITION OF TURNIP CRINKLE VIRUS 1 (CRT1) which interacts with diverse R proteins and is involved in the defense response against turnip crinkle virus (Kang et al., 2008). Also more recently, gene #10 was again found to be a key regulator of gene silencing (Moissiard et al., 2012), which is in
accordance with the accumulating awareness of the intimacy between epigenetic regulation and ETI (Spoel and Dong, 2012).

Every gene in the list has a story of its own that is related to well-known aspects of plant defense or recently discovered novel interfaces between plant and pathogen. However, instead of investigating the biological functions of these novel immune regulators, their involvement in ETI and basal defense, two ostensibly contradictory immune responses with respect to host cell fate, prompt me to investigate the regulatory mechanisms controlling their expression patterns. A combination of bioinformatics, genetics and genomics approaches helped to suggest and support the ‘anticipation’ hypothesis.

While this study identified a novel interface between plant and pathogen, more questions have been raised. First of all, is this circadian regulation specific to the oomycete tested? Second, is it specific to microbes or can be extended to other invaders like insects? Third, besides basal defense and ETI, is SAR also controlled by the circadian clock? Fourth, while this study focused mainly on the host side, what about the circadian clock of the pathogen? From a similar logic, while the circadian clock can regulate immunity, what about the effect of defense responses on circadian? Of course, besides these questions from a macroscopic level, the underlying molecular mechanisms are yet to be elucidated.
The studies on plant circadian clock and immunity have developed rapidly. Large collections of genetic tools from both sides over the decades can and have already facilitated the novel discoveries linking both fields. Demonstration of the temporal variations of Arabidopsis against bacterial pathogen Pseudomonas syringae extended the circadian regulation of plant defense beyond oomycete (Bhardwaj et al., 2011). At the same time, since the experiment was conducted under constant light conditions, it also eliminated the effects from light, thus confirming the direct role of the circadian clock. Later, Arabidopsis was shown to synchronize its defense with the circadian feeding behavior of insects, which is also controlled by the circadian clock (Goodspeed et al., 2012). This then extended the clock control of plant defense beyond microbes. More importantly, this study reported the diurnal oscillations of key defense-related hormones JA and SA in opposite phases. Considering the antagonism between JA and SA which are involved in the immunities against necrotrophic and biotrophic pathogens respectively, this circadian regulation perfectly coordinates the functions of these two hormones. One step further, the molecular bridge connecting the circadian clock and JA signaling was soon found to be the negative regulation of a circadian component, TIME FOR COFFEE (TIC) on a key TF of JA signaling, MYC2 (Shin et al., 2012). While still in its infancy, the research on the interactions between the circadian clock and plant immune system are growing rapidly.
'Nothing in biology makes sense except in light of evolution’. While the history of the arm race between plants and pathogens is a vivid depiction of coevolution by itself (Jones and Dangl, 2006), the discoveries of a universal transcriptional-translational feedback loop (TTFL)-independent circadian clock in all kingdoms have just dramatically changed our fundamental understandings of the evolutionary trajectories of the circadian clock (Edgar et al., 2012). Hence, the interactions between TTFL-independent circadian clock and plant immunity are yet to be explored, which will ultimately unfold the history of their interactions from an evolutionary perspective.

2.4. Methods and Materials

2.4.1. Arabidopsis and Hpa growth condition

Arabidopsis seedlings were grown for 10 days at 16–18°C, 12-h day length, 80-100% relative humidity before Hpa infection through spray of a spore suspension (5 × 10⁵ spores/ml in distilled H₂O) at dawn according to the photoperiod of the plant growth chamber. Hpa Emwa and Hpa Noco were subcultured and inocula prepared using methods modified from previous reports (Holub et al., 1994).

2.4.2. Phenotyping mutants upon Hpa infection

Seven days after inoculation with Hpa Emwa infection, phenotypes were scored following LTB staining (Bowling et al., 1997). Leaves were vacuum-infiltrated twice in a solution of phenol, lactic acid, glycerol and water (1:1:1:1) plus 2.5 mg/ml trypan blue. The tubes containing the samples were placed in a boiling water bath for 2 minutes and
allowed to cool for overnight. The leaves were destained in the chloral hydrate solution and then treated with 70% glycerol. Whole leaves were analyzed and photographed with a MZ8 stereo microscope (Leica) and a PM-C35 camera (Olympus). Detailed examination of *Hpa* structures was conducted with an Olympus BX60F compound microscope and differential interference contrast optics. To measure *Hpa* Noco infection, infected leaves were collected in 1ml water, and sporangiospores were counted.

### 2.4.3. Phenotypic clustering

Hierarchical clustering was performed using the phenotypic scores containing multiple mutant alleles or mean phenotype scores of the three biological replicates using Pvclust with bootstrap (100,000 times). The clustering of mutants and phenotypes were performed sequentially.

### 2.4.4. Bioinformatics analysis

The statistical significance of over-represented TF binding elements was calculated using a hypergeometric probability model. The following equation was used to provide the P-values:

\[
P = 1 - \sum_{x=0}^{k-1} \frac{m \cdot \binom{n-m}{n-x}}{\binom{N}{n}}
\]

\(N\) is the total number of promoters in the genome, \(n\) is the number of promoters in the genome containing the specified TF binding element, \(m\) is the size of the selected set of promoters, and \(x\) is the number of promoters with the specified element in the selected set. Because multiple hypotheses were tested in the analysis, the Bonferroni
correction was used. The genome-wide occurrences of these elements in the promoters are used as controls.

The circadian correlation coefficients were fetched from the time course microarray under circadian conditions through Diurnal database (Mockler et al., 2007). The significance cutoff was set to 0.8.

2.4.5. RASL-seq

The growth conditions (12/12 light/dark cycle, 16–18°C, 80–100% humidity), which were optimized for Hpa infection, were different from those used in traditional circadian studies. Samples were collected every 2 h after inoculation and the remaining plants were kept to ensure successful pathogen inoculation. Total RNA for each sample (1 mg) was used for RASL-seq. Primer (gene-specific with flanking 5' or 3' universal sequences) annealing to mRNA and ligation were designed and carried out accordingly (Yeakley et al., 2002). Barcoded primers were then added to each sample to convert the ligated products to individual libraries, which were pooled from all samples and subjected to multiplex sequencing using Solexa GAI (Illumina).

2.4.6. RASL-seq data analysis

The readings from RASL-seq were assumed Poisson distribution. Only those samples with mean readings significantly above zero (Pr (mean=0)<0.01) were considered for further analysis. The reading for each sample was first divided by the
corresponding reading of control, ubiquitin 5 (UBQ5; AT3G62250), and then standardized. The resulting matrix was used for clustering analysis.

NMF algorithm was used to cluster the genes. The number of the clusters was determined by comparing the cophenetic correlation coefficient for a range of cluster numbers (from 2 to 22). The cophenetic correlation coefficient is a measurement of how faithfully the result of NMF clustering preserves the pairwise distances between the original data points. As shown in Figure 2-11, two clusters generated the highest cophenetic correlation coefficient, which means two clusters can reflect the original data more faithfully than more clusters. Divergence was used as the update rule and cost measurement. Minimum of the data was subtracted from the data matrix to ensure that there were no negative numbers in the matrix. Because the NMF algorithm iteratively updates the decomposition of the data matrix, 300 runs with 10,000 iterations/run were performed to reach the convergence (Figure 2-13). The membership indicators from NMF clustering were used as weights to calculate the weighted mean expression pattern shown in Figures 2-14, 2-15. The weights were also used to determine the radii of circles in Figure 2-12. Smaller radius indicates a higher membership of the gene to the corresponding cluster.
2.4.7. RNA extraction and quantitative PCR (qPCR) analysis

RNA extraction was performed as described previously (Cao et al., 1994). cDNA synthesis (SuperScript III, Invitrogen) and qPCR (SYBR Green PCR kit, Qiagen) were performed according to the manufacturer’s protocols.

2.4.8. Bioluminescence detection

Protein was extracted and bioluminescence intensity measured using the Luciferase Assay System (Promega) according to the manufacturer’s manual. A Victor3 (PerkinElmer) multilabel reader was used to detect the bioluminescence. Substrate (100 ml) was added using an automatic injector. After 3s shaking, 2s delay, the signal was captured for 20s. Log₁₀ transformation was performed to the raw signals to ensure the normal distribution of the data. After subtraction of the blank, the data were normalized according to the total protein concentrations determined by the Bradford method (Bio-Rad). The resulting data were then standardized.

2.4.9. Free-run assay

Seeds were sterilized in 2% Plant Preservative Mixture (PPM, Plant Cell Technology) in the dark at 4°C for 4 days before plating on Murashige and Skoog (MS) plate (3% sucrose, 1.5% agar) and grown in a 12/12 h light/dark growth chamber for 9 days. At the dawn and the dusk of the ninth day, 2.5mM luciferin in 0.05% Triton-X 100 was sprayed onto the seedlings. At the dawn of the tenth day, the seedlings were treated
by distilled H$_2$O or $Hpa$ Emwa before being placed in a constant light chemiluminescence box. The bioluminescence signals were captured by CCD camera.

2.5. Acknowledgements

I would like to thank my collaborators, Drs. Jinyoung Barnaby, Yasuomi Tada, Hairi Li, Mahmut Tör, Daniela Caldelari, Xiang-Dong Fu for their help. I am grateful to Dr. Tobin for providing CCA1OE transgenic line. I appreciate Dr. McClung for providing CCA1:LUC, LHY:LUC, ztl-4 lines. This work was supported by a grant from NSF (MCB-0519898) to Dr. Xinnian Dong and a grant (HG004659) to Dr. Xiang-Dong Fu.
3. Translational Control on a Key Transcription Factor of SAR

Modified from:


3.1. Introduction

In Chapter 2, I described that circadian regulation on defense genes’ expression pattern can help ‘anticipate’ the forthcoming of pathogen invasion under favorable conditions for infection. Besides this recently discovered anticipation mechanism, SAR has been known for long to be the preventative immune response. As an emergent stress response of naïve cells, induction of SAR involves massive transcriptional reprogramming to prioritize defense over growth through the diversion of the energy from growth-associated cellular functions to synthesis of defense-related machinery. As it was introduced in Chapter 1, besides the up-regulation of executors like PR proteins, ER-resident proteins are induced even earlier than PR proteins to meet with the increased demand of protein folding and secretion. TGA TFs were shown to be responsible for the induction of PR proteins. However they do not control ER-resident genes. Therefore the identification of TF that regulates ER-resident genes during SAR is the key to our understanding of the transcriptional reprogramming mechanism.
Moreover, ER-resident genes have also been shown to participate in MTI (Nekrasov et al., 2009; Saijo et al., 2009). This further emphasized the necessity to isolate this TF for a better understanding of both SAR and MTI.

A similar approach like the one in Chapter 2 was used to identify the cis-element enriched in the promoter regions of ER-resident genes that have been induced during SAR (Wang et al., 2005). Instead of a known element like the evening element, a novel consensus was found which was named TL1. While identification of a novel element indicates a potential expansion of the types of TFs involved in SAR, it also became an obstacle to the subsequent isolation of the corresponding TF, TBF1.

Conventional solutions usually involve unbiased whole genome screens. Rather than a laborious and blind screening, the isolation of TBF1 is an exemplary case in which the bioinformatics study played a predominant role (Pajerowska-Mukhtar et al., 2012). First of all, the TL1 sequence was used to search for the potential TF in TFSEARCH database including information from plants as well as arthropods, vertebrates and yeast (Heinemeyer et al., 1998). The result suggested that heat-shock factors (HSFs) may bind TL1. There are 21 HSF-like genes in Arabidopsis. While it is already quite manageable to test all these 21 candidates through yeast one-hybrid (Y1H) assay, the discovery power of bioinformatics surely did not cease at this point. Exploration of the publically available genome-wide expression profiles narrowed the shortlist down to only one gene, HSF4. HSF4 (also known as HsfB1), was strongly induced by BTH as well as
virulent and avirulent *Pseudomonas syringae* pv. *maculicola* (*Psm*) ES4326 bacteria. In addition, *Arabidopsis* HSF4 and its tomato homolog do not functionally complement the yeast *hsf1* mutant and its overexpression has little effect on heat-shock protein expression or thermotolerance. This evidence implies that HSF4 may be the TBF1 regulating ER-resident genes rather than involved in thermotolerance. To test this hypothesis, the promoter of LUMENAL BINDING PROTEIN 2 (BiP2) was selected as a reporter. BiP2 is induced during SAR in an NPR1-dependent manner. Y1H, electrophoretic mobility shift and ChIP assays were performed to verify the regulation of HSF4 on BiP2 *in vitro* and *in vivo*. Besides the confirmation of the binding ability, genetic data also support HSF4 as TBF1. Loss of HSF4 leads to the reduction of PR protein secretion, enhanced susceptibility to *Psm* and compromised SAR. Taken together both biochemical and genetic evidences demonstrated that HSF4 is the *TL1*-binding TF, TBF1.

The isolation of TBF1 allows further characterization of how it functions upon induction of immune system and how the regulation of its own expression is connected with pathogen infection. These are the main focuses of the studies below.

### 3.2. Results

#### 3.2.1. Genome-wide distribution of *TL1* element

Despite the identification of TBF1, the derivation of *TL1* from BTH-induced ER-resident genes already allows us to infer its potential functions beyond regulating this
gene category. Based on the aligned \textit{TL1} sequences of the BTH-induced ER-resident genes (Wang et al., 2005), I calculated the weight matrix of \textit{TL1} (Table 3-1) and generated graphical representation of \textit{TL1} which illustrates the degeneracy at each nucleotide (Figure 3-1). To initiate a genome-wide identification of genes with \textit{TL1} in their promoter sequences, 1000-bp upstream sequences with cutoff at the adjacent gene were fetched from the \textit{Arabidopsis} Information Resource website (http://www.arabidopsis.org). To control the level of degeneracy, the total weight of the hit was restricted to be more than 664, which allows a total of 54 different variants.

44.78\% genes of the whole genome have at least one \textit{TL1} in their promoter regions (Table 3-2). Gene ontology (GO) analysis revealed a diverse cellular functions potentially regulated by TBF1 through \textit{TL1} element, including various metabolic processes, flower development, responses to external and endogenous stimuli, signal transduction, transportation. (Figure 3-2). It is interesting that these proteins are preferentially located at Golgi apparatus, nucleus and membrane, especially plasma membrane, despite the fact that \textit{TL1} was initially deduced from the promoters of ER-resident genes. This implies that the enrichment of ER-resident genes may be specific to SAR induction.

Therefore, to explore the distribution of \textit{TL1} in the promoters of genes with a defense context, I retrieved the list of BTH-, flg22- or elf26- (the first 26 amino acids of a MAMP signal, elongation factor thermo unstable) regulated genes from public
expression profile database (detailed in Methods and Materials). As expected, TL1 is not significantly enriched in BTH-regulated genes, since it was derived from BTH-regulated and ER-resident genes (Table 3-2). However, TL1 is over-represented in MAMP-regulated genes, suggesting its involvement in MTI. While the percentages of the overlap among BTH-, elf26- and flg22-regulated genes remain largely unchanged when the presence of TL1 in their promoters are considered (Figure 3-3), a Venn diagram comparison of the GO analyses on these TL1-containing genes revealed notable enrichment of GO categories (Figures 3-4, 3-5).

The GO categories including response to biotic stimulus and biosynthetic processes are enriched in genes regulated by BTH, elf26 and flg22, which is consistent with the context of these treatments. On the other hand, the GO terms such as flower development, cell differentiation and location in Golgi apparatus are clearly not major players in SAR or MTI, as none of them are enriched in BTH-, elf26- or flg22-regulated genes. The GO terms central to the Venn diagram which are shared by all five genes lists, include response to abiotic, endogenous and external stimuli, signal transduction, cellular protein modification process. This is in agreement with our understanding of plant immune system, as the innate immunity involves an extensive crosstalk with other cellular functions. Therefore it is not surprising to see even the enrichment of response to abiotic stress. Besides, these GO categories shared by different treatments, it is also important to notice the GO terms specific to a single treatment. There are 14 GO terms
enriched in genes regulated by BTH only, including ER-resident genes, photosynthesis, generation of precursor metabolites and energy. While the enrichment of ER-resident genes recapitulates the original derivation of TL1, the over-representation of energy-related genes implies TL1’s potential role in coordinating the energy balance between growth and defense. On the other hand, no GO terms are specific to a single MAMP treatment. The GO categories over-represented in elf26-regulated genes are identical to those enriched in flg22-responsive genes, which is in accordance with the convergence of signaling pathways of MTIs triggered by different MAMPs.

3.2.2. TBF1 triggers transcriptional reprogramming to control the growth-to-defense transition during MTI and SAR

The bioinformatics analysis of TL1 element suggested its role in regulating both MTI and SAR. The identification of its corresponding TF, TBF1, allows a genome-wide test of the dependencies of MTI and SAR on TL1 using tbf1 mutant. The WT and tbf1 plants treated with SA of elf18 were used for microarray hybridization. 1269 and 1792 TBF1-dependent genes were found to be differentially regulated by SA and elf18, respectively (P<0.05, fold change>2). However, only about 8% of these genes were regulated by both signals, indicating that TBF1 regulates different output genes in SAR and MTI (Figure 3-6). This is largely consistent with the similar analysis of TL1-containing genes (Figure 3-3b). Moreover, the total numbers of significantly induced and repressed genes (the top heatmaps in Figures 3-7a, 3-7b), the degrees of TBF1 dependency (the middle heatmaps), and the numbers of TL1 elements present in the
gene promoters (the bottom heatmaps) indicate that TBF1 plays a greater role in SA- and elf18-mediated transcription repression than in induction. This finding is in agreement with the previous work indicating that class B-HSFs mainly act as repressors of target gene expression (Czarnecka-Verner et al., 2004; Prandl et al., 1998).

GO analysis identified a significantly enriched cluster of SA-induced secretory pathway genes (P<0.001; Figure 3-7a ‘membrane proteins’), again recapitulating the original derivation of TL1. In addition, I found several major functional categories significant at P<0.001, comprising genes known to encode defense-related proteins such as NPR1, a TGA-class TF, several WRKY family members, EDS5, metacaspase 2 and membrane-associated proteins like SNAP33. Intriguingly, a strong enrichment of ribosomal proteins was identified among the elf18-induced TBF1-dependent genes, suggesting considerable translation reprogramming following MAMP induction. Upon SAR induction, TBF1 represses genes encoding chloroplast proteins (Figure 3-7a), an effect that is known to be associated with SA. Interestingly, chloroplast function-related genes were even more profoundly repressed by elf18 (Figure 3-7b). These genes encode several structural and regulatory proteins of the chloroplast, e.g., a subunit of photosystem II, chloroplast o-succinylbenzoyl-CoA ligase, plastid ribosomal protein, a subunit of the chloroplast NAD(P)H dehydrogenase complex and components involved in thylakoid membrane biogenesis. Loss-of-function mutants of these genes display a variety of developmental defects such as decreased photosynthesis rate, reduced
chloroplast number, pale pigmentation, dwarfism and lethality (Leister and Schneider, 2003). Somewhat unexpectedly, the GO analysis indicated that elf18 treatment had a significant inhibitory effect on both abiotic stress and defense responses through TBF1 (Figure 3-7b). However, these repressed genes are involved in JA, ethylene and auxin biosynthesis or signaling pathways which are known to be suppressed during SA-mediated defense. The genes of these abovementioned GO terms are likely to be regulated by TBF1 directly since they all have TL1 in their promoters.

Taken together, the GO analysis indicates that activation of TBF1 during MTI and SAR can trigger a massive transcriptional reprogramming to divert the energy from growth-related cellular processes to the armory of immunity through repression of photosynthesis and induction of defense machinery. Consistently, loss of TBF1 in tbf1 mutant alleviated the inhibitory effects of MTI or SAR on growth, while tbf1 and WT have similar growth rates in the absence of elf8 or SA (Figure 3-8). Moreover, plants overexpressing TBF1 have severe developmental defects and are lethal, further confirming its role in repressing growth once activated. Therefore, both genomic and genetic data suggest TBF1 as a key molecular switch for the growth-to-defense transition.

3.2.3. Identification of uORFs in 5’ UTR of TBF1

As a pivotal TF involved in the growth-to-defense transition, TBF1 itself should be tightly regulated to prevent spurious activation and inhibition of growth while still maintain appropriate sensitivity to the induction of immunity.
The analysis of TBF1 mRNA through the 5’ and 3’ rapid amplification of cDNA ends (RACE) detected two upstream open reading frames (uORFs) 5’ of the TBF1 start codon (Figure 3-9). A comparative genomics approach aimed to identify conserved peptide uORFs in Arabidopsis and rice also identified the second uORF of TBF1 as a conserved uORF in these two plant species, which implies its potential regulatory functions (Hayden and Jorgensen, 2007).

### 3.2.4. uORFs’ repression on TBF1 translation can be relieved by pathogen infection

The presence of uORFs does not necessarily ensure they can affect the translation of the downstream genic ORF. Therefore, to test whether these two uORFs may affect the translation initiation of TBF1, a fusion between the 5’ UTR of TBF1 containing both uORFs, the first exon of TBF1 and the β-glucuronidase reporter gene (GUS) was constructed (abbreviated as uORF1-uORF2-GUS). Additional three constructs were generated with the start codon mutated (ATG to CTG) for uORF1 (uorf1-uORF2-GUS), uORF2 (uORF1-uorf2-GUS), or both uORFs (uorf1-uorf2-GUS). These reporter constructs were driven by the constitutive 35S promoter to allow detection of only translational differences.

For a quick check, we first transiently expressed these constructs in Nicotiana benthamiana leaves and quantified the GUS activities (Figure 3-10a). 1.5- and 3.5-fold increases in GUS activities were detected in uorf1-uORF2-GUS and uORF1-uorf2-GUS when uORF1-uORF2-GUS as used as a control. Mutating both uORFs in uorf1-uorf2-GUS
resulted in a 3.5-fold elevation in GUS activity over WT. Therefore, both uORFs have inhibitory effects on TBF1 translation, while uORF2 is epistatic to uORF1.

The inhibition of TBF1 by these two uORFs under normal conditions and demonstrated a critical role of TBF1 during defense suggest a pathogen-mediated derepression mechanism. To verify this hypothesis, we quantified GUS activities of these constructs in transgenic Arabidopsis lines in response to Psm ES4326 carrying the avirulent effector, avrRpt2. The perception of this avirulent bacterial strain, which can induce MTI, ETI, and SAR, caused a rapid induction of GUS activity in the uORF1-uORF2-GUS transgenic lines (Figure 3-10b). Interestingly, this increase was not observed in the uorf1-uorf2-GUS transgenic lines. So it is clear from these results that Psm ES4326/avrRpt2 could rapidly alleviate the inhibitory effects of the uORFs on translation of the downstream gene.

To further confirm this conclusion with respect to the endogenous TBF1, a ribosome profiling was performed to monitor the translation status of endogenous TBF1 upon pathogen infection (Figure 3-11). TBF1 transcript increased drastically in the polysomal fractions of the gradient within 30 min of Psm ES4326/avrRpt2 inoculation, consistent with the results from the studies using the GUS reporter. Moreover, this polysomal association appeared to be dynamic as the TBF1 transcript decreased 1 hour post inoculation. These translational control events seem to be specific to TBF1, as the
polysomal association of the housekeeping gene UBQ5 transcript was not perturbed (Figure 3-12).

3.2.5. uORF may serve as a metabolic sensor of phenylalanine metabolism

The demonstration of this pathogen-induced uORF-mediated translational control on TBF1 brought us additional question. How does pathogen derepress TBF1?

With the rich knowledge obtained through the studies on the translational control on GCN4, we may have the following working model of how pathogen could derepress TBF1. Pathogen invasion may induce amino acid starvation stress, which subsequently results in the increase of uncharged tRNA, activation of the kinase activity of GCN2, phosphorylation of eIF2α by GCN2 and translation reinitiation at the start codon of TBF1. This is actually what we have observed. Inoculation of Psm ES4326/avrRpt2 caused rapid induction of uncharged phenylalanine tRNA (tRNA\textsubscript{Phe}) with 30 minutes, while the phosphorylation of eIF2α was also detected around the same time (Figure 3-13). Intriguingly, the induction of uncharged tRNA seems to be specific to tRNA\textsubscript{Phe}, since similar effects were not observed when aspartic acid tRNA was tested.

It is notable that both uORFs of TBF1 are enriched in aromatic amino acids, especially in phenylalanine (Figure 3-9). The average frequency of aromatic amino acids for species sequenced up till now is 7.63-7.86%, while this goes up to 27% in uORF1 and 19% in uORF2. Considering the fact that aromatic amino acids are precursors for a variety of plant metabolites including SA, it is tempting to consider these uORFs as
metabolic sensors to detect the increase consumption of phenylalanine potentially caused by pathogen invasion and then release their repression on TBF1 translation. Although technical limitations hindered the direct measurement of this pathogen-induced aromatic amino acid content changes, the detection of tRNA\(^{Phe}\) induction partly support this idea. Moreover, the deprivation of phenylalanine rather than aspartic acid can help release the repression effect of these uORFs on the translation of downstream genic ORF in a dihydrofolate reductase (DHFR)-based yeast reporter system (Figure 3-14), further supporting the metabolic sensor function of these uORFs.
Table 3-1: Weight matrix of TL1 element derived from the promoter sequences of BTH-induced ER-resident genes

<table>
<thead>
<tr>
<th>Position</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0</td>
<td>88</td>
<td>85</td>
<td>0</td>
<td>91</td>
<td>91</td>
<td>12</td>
<td>62</td>
<td>65</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>38</td>
<td>26</td>
</tr>
<tr>
<td>G</td>
<td>100</td>
<td>12</td>
<td>3</td>
<td>100</td>
<td>6</td>
<td>3</td>
<td>82</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>T</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>
Figure 3-1: Sequence logo of *TL1*

This logo was generated according to the promoter sequences of BTH-induced ER-resident genes (Wang et al., 2005) using WebLogo (Crooks et al., 2004). The overall height of each stack represents the sequence conservation at that position (measured in bits). The height of symbols within the stack indicates the relative frequency of the corresponding nucleic acid at that position.
Table 3-2: Genome-wide distribution of TL1 element

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total</th>
<th>Number of genes with TL1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BTH 24h</td>
<td>2182 (6.51%)</td>
<td>993 (45.51%, p=0.26)</td>
</tr>
<tr>
<td>elf26 1h</td>
<td>891 (2.66%)</td>
<td>450 (50.51%, p=2.0e-4)</td>
</tr>
<tr>
<td>flg22 0.5h</td>
<td>750 (2.24%)</td>
<td>371 (49.47%, p=3.7e-3)</td>
</tr>
<tr>
<td>BTH∪elf26∪flg22</td>
<td>2711 (8.09%)</td>
<td>1261 (46.51 %, p=2.6e-2)</td>
</tr>
<tr>
<td>Whole Genome</td>
<td>33518</td>
<td>15008 (44.78%)</td>
</tr>
</tbody>
</table>

Genes that are regulated by BTH, elf26 and flg22 were fetched from corresponding microarrays. The percentages of these genes among the whole genome were included in the parentheses. The percentages of genes with TL1 in their promoters were calculated based on the corresponding total. The p-values were calculated using hypergeometric distribution. BTH 24h, 24 hours after BTH treatment; elf26 1h, 1 hour after elf26 treatment; flg22 0.5h, 0.5 hour after flg22 treatment; BTH∪elf26∪flg22, genes regulated by any of the three treatments.
Figure 3-2: Gene ontology analysis of TL1-containing genes

Enriched GO terms of all the genes with TL1 element in their promoters are represented by different colors of nodes. The color scale is based on the adjusted
P-values from $e^{-2}$ to $e^{-7}$. The nodes in white are not enriched GO terms but are only shown as the direct parents of the nodes that are enriched. The unidirectional edges linking different nodes indicate the parent-child relationship. The size of nodes is proportional to the number of genes included in the corresponding GO term.
Figure 3-3: Venn diagrams of BTH-, elf26- and flg22-regulated genes

a, Venn diagrams of BTH-, elf26- and flg22-regulated genes from public expression profiles. b, Venn diagrams of BTH-, elf26- and flg22-regulated TL1-containing genes. BTH 24h, 24 hours after BTH treatment; elf26 1h, 1 hour after elf26 treatment; flg22 0.5h, 0.5 hour after flg22 treatment.
Figure 3-4: Venn diagram of enriched GO terms of TL1-containing genes

BTH 24h, 24 hours after BTH treatment; elf26 1h, 1 hour after elf26 treatment;
flg22 0.5h, 0.5 hour after flg22 treatment; BTH∩elf26∩flg22, enriched GO terms of
genes regulated by all three treatments; genome-wide, enriched GO terms of all
TL1-containing genes.
Figure 3-5: Enriched GO terms from Figure 3-4

Enriched GO terms from different overlap regions of Figure 3-4 are shown in the same color scheme. The unidirectional edges linking different nodes indicate the parent-child relationship.
Figure 3-6: Venn diagram shows the numbers of TBF1-dependent SA-repressed (SA down), SA-induced (SA up), elf18-induced (elf18 up), and elf18-repressed (elf18 down) genes (P<0.05)
Figure 3-7: Heat maps of TBF1-regulated genes in total numbers (top), degrees of TBF1 dependency (middle), and numbers of TL1 elements in the gene promoters (bottom) in response to SA (a) and elf18 (b) treatment. Top ranked GO terms were determined using Database for Annotation, Visualization and Integrated Discovery (DAVID). Scale indicates the log-
transformed P-values of down-(blue) and up-(yellow) regulated genes (top),
yellow lines indicate TBF1 dependency (middle), and yellow lines correspond to
the numbers of TL1 elements in the gene promoters (bottom).
Figure 3-8: TBF1 controls growth-to-defense transition during SAR and MTI (This data was provided by Dr. Karolina M. Pajerowska-Mukhtar)

Fresh weight of ten seedlings grown for 10 days on plates with MS growth media (ctrl), or MS supplemented with increasing concentrations of SA or 10 μM elf18. Error bars represent standard deviation of three replicates. This experiment was repeated three times with similar results. Statistical analysis was performed using Student’s t test, * P<0.05, ** P<0.01, *** P<0.001.
Figure 3-9: Schematic representation of uORF1 and uORF2 and exon I of *TBF1*

The phenylalanines (F) in uORF1 and uORF2 are highlighted in red, and the stop codons are shown as asterisks. ‘+1’ represents the translational start of TBF1 and -451, -265, and -217 represent the upstream positions of the 5’ end of the transcript, the start codon for uORF1, and the start codon for uORF2, respectively.
Figure 3-10: Characterization of the uORFs of *TBF1*

(This is a collaborative study result with Dr. Karolina M. Pajerowska-Mukhtar)

a, Quantification of GUS activity in *Nicotiana benthamiana* leaves transiently expressing *uORF1-uORF2-GUS* (WT), *uorf1-uORF2-GUS*, *uORF1-uorf2-GUS*, and *uorf1-uorf2-GUS*. This experiment has been repeated three times with similar results. b, Quantification of translational inhibitory effect exerted by uORFs in transgenic T3 plants expressing *uORF1-uORF2-GUS* (two independent transformants 6-1 and 9-4) or *uorf1-uorf2-GUS* (two independent transformants 7-
3 and 8-3) at various time points after inoculation with *Psm* ES4326/avrRpt2 (OD$_{600\text{nm}} = 0.02$). Error bars represent standard deviation from three different replicates. Experiment was repeated at least three times with similar results.
Figure 3-11: Ribosome profiles and translational status of TBF1 upon infection

a, Ribosome profiles of WT plants at 0, 0.5 and 1 hour after inoculation with Psm ES4236/avrRpt2 (OD_{600nm} = 0.02). The fractions containing monosome and polysome were annotated based on the absorbance at 254 nm (A_{254nm}). b, The TBF1 transcript abundance in corresponding fractions of ribosome profile is normalized against Alien Alert® control transcript and expressed in arbitrary units (AU). Error bars represent standard error. This experiment was repeated
using two biological replicates (each with three technical replicates) with similar results.
Figure 3-12: The *UBQ5* transcript association with polysomes is not altered by pathogen infection

*UBQ5* expression in samples obtained from WT plants at 0, 0.5 and 1 hour after inoculation with *Psm* ES4326/avrRpt2 (OD<sub>600nm</sub> = 0.02). The *UBQ5* transcript abundance normalized against Alien Alert® control transcript is expressed in arbitrary units (AU). Error bars represent standard error. This experiment was repeated using two biological replicates (each with three technical replicates) with similar results.
Figure 3-13: uORF-mediated translation derepression mechanism

(Figure 3-13b is provided by Dr. Yasuomi Tada)

**a**, tRNA analysis of WT plants at various time points after inoculation with \( Psm \) ES4326/avrRpt2 \((\text{OD}_{600\text{nm}} = 0.02)\). Northern blot using probes against \( \text{tRNA}^{\text{Phe}} \) or \( \text{tRNA}^{\text{Asp}} \) was performed to detect charged and uncharged \( \text{tRNA}^{\text{Phe}} \) or \( \text{tRNA}^{\text{Asp}} \). This experiment was repeated using three biological replicates with similar results.  

**b**, Phosphorylated form of eIF2α was detected using a phospho-specific antibody in the total protein extracts from leaves of 3-week-old WT plants collected at various time points after inoculation with \( Psm \) ES4326/avrRpt2 \((\text{OD}_{600\text{nm}} = 0.02)\). Ponceau S stain was used to determine equal loading.
Figure 3-14: Phenylalanine but not aspartate starvation can alleviate the translation inhibition effect of uORFs

(This result is provided by Dr. Karolina M. Pajerowska-Mukhtar)

The effects of phenylalanine and aspartate starvation on the translational inhibitory function of uORFs were measured by growth of the yeast strain \textit{aro7} (\textit{phe}, \textit{tyr}) transformed with the \textit{uORF1-uORF2-DHFR} or \textit{DHFR} reporter in medium containing methotrexate, an inhibitor of the endogenous DHFR. Optical densities for cultures containing two different concentrations of phenylalanine (Phe; 15 and 75 mg/L) as well as cultures lacking Asp and supplemented with tobramycin (TOB), an inhibitor of yeast tRNA\textsubscript{Asp} aspartylation, were recorded over the course of 32 hours. Error bars represent standard deviation from nine technical replicates derived from three independent experiments.
3.3. Discussion

The genome-wide studies on TL1 distribution implies its importance in regulating various cellular functions, while a similar analysis focused on its role in defense suggests that TL1 may control the growth-to-defense transition. The discovery of TBF1 allows a systematic verification of this hypothesis. TBF1’s repression effect on growth-related biological functions and induction effect on defense-related machinery were revealed by whole genome expression profiling experiment on tbf1 mutant. At the genetic level, loss of TBF1 can negate the growth inhibition effect of SA and MAMP signal. Based on these results, we can reach the conclusion that TBF1 is a major molecular switch to control the growth-to-defense transition. Due to its vital impact of the subsistence of plants, a fine tuning of TBF1 abundance is critical. A uORF-mediated translational repression on TBF1 is deployed to ensure its inactive translation under normal conditions (Figure 3-15, left panel). Upon pathogen attack, plants may experience an aromatic amino acids starvation caused by infection. This emergent starvation will trigger the stepwise amino acid starvations responses including induction of uncharged tRNA, activation of eIF2α kinase, phosphorylation of eIF2α and derepression of TBF1 (Figure 3-15, right panel). Although the detailed molecular mechanism underlying this uORF-mediated translation derepression has yet to be elucidated, the significance of this strategy is backed by the observed sequence conservation of the uORF also upstream of the TBF1 homolog gene in rice (Hayden and
Jorgensen, 2007). It is likely that the regulation on the TBF1 homolog is also operated in the similar way.

However, due to the technical hindrances of identifying uORFs experimentally, our understanding of uORFs is still limited. Hence, the importance of uORFs has been greatly underestimated. In fact, the number of uORFs in plants was estimated to be 60%, while quite a few of them are even conserved at the uORF and genic ORF sequence level in both monocots and dicots (Hayden and Jorgensen, 2007). Besides the sequence-dependent uORFs which may encode bioactive peptide, sequence-independent uORFs can also have a great impact on the translation of the downstream genic ORF like the case of GCN4. A similar situation may also exist in plants, since multiple stress triggers including amino acid starvation, purine deprivation, wounding, cold shock, SA, JA and ethylene can lead to the phosphorylation of eIF2α in a GCN2-dependent way (Lageix et al., 2008). Therefore it will be interesting to identify all the uORFs and explore their regulatory functions under these stress conditions beyond the computational speculation. The emerging approaches to empower this genome-wide study such as ribosome footprinting technique will be introduced and some perspectives will be briefly envisaged in Chapter 4.
Figure 3-15: A model illustrating the translational control mechanism on TBF1

TBF1 is regulated through rapid increases in uncharged and charged tRNA\textsubscript{Phe}, phosphorylation of eIF2\textalpha, and ribosomal read-through of uORFs, leading to the growth-to-defense transition.
3.4. Methods and Materials

3.4.1. Genome-wide search of TL1 element

To perform a genome-wide search for the TL1, 1000-bp upstream sequences with cutoff at the adjacent gene were fetched from the Arabidopsis Information Resource website (http://www.arabidopsis.org). Degeneracy of the TL1 element was evaluated as Table 3-1. To control the level of degeneracy, the total weight of the hit was restricted to be more than 664. The exact and degenerate TL1 motifs were searched for using the scan_for_matches software (http://iubio.bio.indiana.edu/soft/molbio/pattern/scan_for_matches.readme). To explore the distribution of TL1 in defense-related context, I examined the promoters of the genes regulated by BTH (available at http://affy.arabidopsis.info/narrays/experimentbrowse.pl, experiment ID: NASCARRAYS-392) (Wang et al., 2006) and by the MAMP signals flg22 and elf26 (available at http://www.ebi.ac.uk/arrayexpress/, experiment ID: E-MEXP-547) (Zipfel et al., 2006). The Venn diagrams were produced using Venny (http://bioinfo.genotoul.fr/index.php?id=116). The Cytoscape plugin, BiNGO was used for GO analysis (Maere et al., 2005).

3.4.2. Microarray

Arabidopsis plants (Col-0 and tbf1 mutant) were grown on soil (Metro Mix 360) at 22°C under a 16/8 hr light/dark cycle for 3 weeks and treated with 1 mM SA for 6 hours (spray) or 10 μM elf18 for 2 hours (infiltration into leaves). Mock treatments with water...
were included for both spray and infiltration. The RNA, extracted with TRIzol (Ambion) and labeled with MessageAmp Premier RNA Amplification Kit (Ambion), was hybridized with GeneChip *Arabidopsis* ATH1 Genome Array (Affymetrix, Santa Clara, CA) and subsequently washed and scanned at the Duke Microarray Facility. Experiments were repeated three times using independently grown and treated plants. The resulting data were normalized using Gene-Spring GX Software (RMA algorithm; Agilent). Two-way ANOVA with Benjamini-Hochberg multiple comparison correction was used to identify TBF1-dependent genes (i.e., with significant interaction between genotypes and treatments, P<0.05). The SA- and elf18-responsive genes (fold change>2) were found through unpaired Student’s t-test with Benjamini-Hochberg multiple comparison correction (P<0.05). The Venn diagram was adapted from Venny. To generate the heatmaps of SA- or elf18-induced and repressed genes, -log10 P-values of induced genes and log10 P-values of repressed genes from Student’s t-test were used. Therefore, higher positive values represent greater induction and lower negative values indicate greater repression. For TBF1 dependence, -log10 P-values from two-way ANOVA were used.

### 3.4.3. RACE-PCR

RACE-PCR analyses were performed as described in manufacturer’s protocol (SMART™ RACE cDNA Amplification Kit, Clontech, Mountain View, CA, USA).
3.4.4. uORF amino acid enrichment calculations

Aromatic amino acids contents in both uORFs were determined by comparing the frequencies of phenylalanine, tyrosine and tryptophan in uORFs (uORF1 - 27%, uORF2 - 19%) to the average frequency of aromatic amino acids reported for species sequenced so far (7.63-7.86%) (ExPASy proteomics server http://expasy.org/sprot/relnotes/relstat.html) (Brooks et al., 2002).

3.4.5. Quantitative β-galactosidase assay

The assay was modified from a previously published protocol (Zhang and Bremer, 1996). In brief, 0.1 ml of yeast transformant extract was added to 0.9 ml of Z buffer and warmed to 28°C. Reactions were initiated upon addition of 0.2 ml of o-Nitrophenyl-β-D-galactopyranosidase substrate (4 mg/ml) in Z buffer and terminated with 0.5 ml of 1 M Na₂CO₃. Reactions were terminated within the linear range of the assay (OD₄₂₀nm < 1.0). β-galactosidase activity in yeast supernatants was normalized to the protein concentrations. Data are averages of three dilutions of the extracts.

3.4.6. Ribosome profiling

Before extraction, a spike-in control was added into the pulverized leaf tissue at a concentration of 10⁷ copies of Alien qRT–PCR Inhibitor Alert (Agilent Technologies, USA) per mg of fresh weight. 500 mg pulverized leaf tissue was hydrated on ice for 10 minutes with occasional vortexing in 3 ml of extraction buffer containing 0.2 M Tris (pH = 9.0), 0.2 M KCl, 0.025 M EGTA, 0.035 M MgCl₂, 1% (w/v) Brij-35, 1% (v/v) Triton X-100,
1% (v/v) Igepal CA 630, 1% (v/v) Tween 20, 1% (w/v) sodium deoxycholate, 1% (v/v) polyoxyethylene 10 tridecyl ether, 5 mM dithiothreitol, 1mM phenylmethylsulfonyl fluoride, 50 μg/mL cycloheximide, and 50 μg/mL chloramphenicol. Hydrated tissue was centrifuged at 16,000 g for 15 minutes. The supernatant was then separated in a 10 mL continuous (15-60% w/v) sucrose gradient containing 400 mM potassium acetate, 25 mM potassium HEPES (pH = 7.2), 15 mM magnesium acetate, and 200 μM cycloheximide by ultracentrifugation at 35,000 rpm using SW 41Ti rotor (Beckman Coulter, Germany) for 10 hours at 4°C. The gradients were fractionated into 36 fractions of about 330 μl each using automated Density Gradient Fractionation System (Teledyne Isco Inc., USA) with a simultaneous A$_{254nm}$ trace. Total RNA was extracted from the fractions containing ribosomes using TRIzol reagent (Invitrogen) according to the instructions provided by the manufacturer. mRNA was further precipitated using 2 M LiCl overnight. cDNA was prepared and qPCR analyses performed as described above.

### 3.4.7. tRNA analysis

Total RNA was extracted using TRIzol reagent (Invitrogen) according to the instructions provided by the manufacturer. Total RNA was then dissolved in 0.1 M sodium acetate (pH 5.0). mRNA was precipitated using 2 M LiCl overnight. 2 volumes of isopropanol were added to the supernatant to precipitate tRNA. After washing with 100% ethanol, tRNA was dissolved in 0.1 M sodium acetate. 1μg tRNA was separated by acid urea polyacrylamide gel electrophoresis (PAGE) and transferred to NEF 976
GeneScreen Plus Hybridization Transfer Membrane (PerkinElmer) according to (Kohrer and RajBhandary, 2008). Specific tRNA species were detected by hybridization using digoxigenin labeled DNA probe according to the manufacturer’s instruction (DIG High Prime DNA labeling and detection starter kit II, Roche Applied Science). The signal was visualized using CCD camera.

Probes sequences:

**tRNA-Phe:**

`AGCGTGCGATCGAACAACGCACCTTCAGATCTTCAGTCTGCTCTCCCAACTGAGCTA`

**tRNA-Asp:**

`GCCGGGATCGAACCCCGGTCAACCGTCGACAGGCGGAAATACTTACCACTATACTAC`

### 3.4.8. Western blotting

The anti-phospho eIF2α Western blotting experiment was performed as described previously (Lageix et al., 2008), using leaf tissue infected with *Psm* ES4326/avrRpt2 (OD$_{600nm}$ = 0.02) over the indicated time periods. The protein extraction was carried out in the presence of a phosphatase inhibitor PhosSTOP (Roche), Protease Inhibitor Cocktail (Sigma Aldrich) and proteasome inhibitor MG-115 (Sigma Aldrich). The primary antibody was the phospho-specific α-p-eIF2α (pS51; Epitomics, Burlingame,
CA, 1090-1; 1:1000 dilution, overnight at 4°C) and the secondary antibody was goat anti-rabbit (Bio-Rad, 1:4000 dilution, 1 hour, RT).

3.4.9. Yeast growth assay using the DHFR reporter

Since a phenylalanine-deficient *Arabidopsis* mutant has not been identified to date, we used a yeast chorismate mutase deletion strain, *aro7*, which is auxotrophic for phenylalanine and tyrosine (Ball et al., 1986). The DHFR reporter gene carried by pTB3 plasmid was engineered to make an unstable enzyme (Tucker and Fields, 2001) and to contain L22F/F31S mutations that confer resistance to methotrexate (MTX) (ErcikanAbali et al., 1996a; ErcikanAbali et al., 1996b). The *uORF1-uORF2* of TBF1 was fused to the coding region of the DHFR reporter and subsequently integrated into the genome of yeast strain BY4742 through homologous recombination. Equal amounts of yeast culture grown in liquid media (SD-Leu) were inoculated into SD-Leu-Phe double drop-out media supplemented with 15 mg/L or 75 mg/L phenylalanine. Alternatively, the yeast cultures were grown in Phe-rich, Asp-deficient media supplemented with 15 mM tobramycin (TOB) (Sigma, St. Louis, MI, USA), a known inhibitor of yeast tRNA<sub>Asp</sub> aspartylation. MTX was added to all cultures at the final concentration of 80 μM to inhibit the endogenous DHFR activity. Yeast growth, which was dependent on the expression of the recombinant DHFR reporter in the presence of MTX, was measured using optical density (OD<sub>600nm</sub>) during a 32-hour time course.
3.5. Acknowledgements

I would like to thank my collaborators, Drs. Karolina M. Pajerowska-Mukhtar, Yasuomi Tada, Chandra L. Tucker for their help. I am grateful to Dr. Christopher Nicchitta for assistance in the ribosome profiling experiment, Dr. Tao Pan for input in the tRNA study. This work was supported by a grant from NSF (MCB-0519898) to Dr. Xinnian Dong and Grants-in-Aid for Scientific Research (No. 23120520) from the Ministry of Education, Culture, Sports, Science and Technology of Japan to Dr. Yasuomi Tada.
4. Summary and Perspectives

4.1. Reconcile between two anticipation system

Time is an important element of plant immune system. While MTI is generally weaker than ETI, if plants are given a sufficient lead time, MTI can be potent enough to fend off the virulent pathogens (Katagiri and Tsuda, 2010). This unveils the underpinning reason why plant circadian regulator has been engaged by immune system to time different defense responses. The anticipation power of circadian clock supplements the immune system with this critical lead time to effectively mount a proper level of defense at the appropriate time. Originated from the similar strategy, SAR also intends to provide naïve cells with this lead time to prepare for the possible future infection.

However, the co-existence of the circadian clock and SAR in the systemic tissue raises more questions. How are these two anticipation systems reconciled to reach a consistent and coordinated decision? How do they communicate with each other at the molecular level? Moreover, the newly discovered redox clock further complicated the situation. How do the redox clock and TTFL clocks talk to each other? What is the relationship between the redox clock and SAR?

Besides these new challenges, the identification of redox clock actually provided a potential link between the circadian clock and SAR. As it was introduced in Chapter 1, 99% of SAR-responsive genes are dependent on NPR1, which is also a co-receptor for SA.
Induction of SAR in systemic cells triggers conformational change of NPR1, which is subsequently translocated into nucleus. In the nucleus, NPR1 interacts with TGA and initiates a massive transcriptional reprograming. The conformational transition of NPR1 between oligomer and monomer is responsive to the redox oscillations induced by SA. Therefore, I propose that the redox-sensitive conformational transition and the subsequent nuclear translocation of NPR1 may transduce the redox oscillation signal from the redox clock to the TTFL clock. Indeed, TGA binding sites are found in the promoter regions of \textit{TOC1} and \textit{CHE}. On the other hand, CHE binding sites are also identified in the promoter of \textit{ICS1}, a key player in the biosynthesis of SA. Therefore we have a closed loop in which the redox clock may modulate the expression of TOC1 and CHE through the redox-sensitive NPR1 nuclear translocation while CHE then regulates cellular redox clock via its control on \textit{ICS1}. Since NPR1 is tightly controlled through conformation transition, nuclear translocalization, phosphorylation and degradation, the integration of NPR1 in the communication between the redox and TTFL clocks may ensure a robust yet sensitive circadian system.

While this model can help to explain how clock may regulate SAR through the control on SA synthesis, it also postulates a potential impact of SA on clock in an NPR1-dependent way. Indeed, this is what people have noticed empirically. SA can suppress the rhythmic leaf movement of WT plants, an easily observed circadian output physiology. Meanwhile, this rhythmic movement still persists in \textit{npr1} mutant after SA
treatment. Apart from these anecdotal phenomena, more formal experiments are needed to investigate this novel interface between the circadian clock and plant immune system.

4.2. Future research on translational control of plant immune system

The uORF-mediated translational control on TBFI described in Chapter 3 was echoed by a recent study showing the similar regulatory mechanism on a key TF of defense, ZIP-2 in Caenorhabditis elegans (Dunbar et al., 2012). The inhibitory effect of its conserved uORF can also be alleviated upon infection. Therefore, this regulatory strategy seems to be conserved in both plants and animals. Besides this uORF-mediated posttranscriptional regulation, two additional ways to modulate at this level have emerged and attracted much attention these years. One is the cytoplasmic processing bodies (P-bodies) which may serve as a site of mRNA degradation or temporary storage. The second one is the regulation by small RNAs. The discovery of P-bodies in plant cells is of particular interest, as its temporary storage function may also contribute to SAR. The mRNA of defense machinery in the systemic cells may have been prepacked in these P-bodies to allow for an instantaneous initiation of defense upon infection. It would be important to determine whether the key TFs of SAR are also in these P-bodies to facilitate the rapid induction of downstream genes.

In order to investigate these mechanisms at the genome level, a global probing of the mRNAs in these structures are needed. However, while whole genome transcription profiling has become a routine methodology to study the transcriptional regulation
systematically, a similar holistic approach to dissect translational control is still quite a challenge since the techniques used to study translation at the single gene level cannot be easily adapted to the genome-wide scale study. As a result, this layer of regulation has largely been studied at single-gene level, greatly impeding our in-depth understanding. With the emergence of next generation sequencing technologies, this bottleneck has recently been removed via ribosome footprinting approach in yeast (Ingolia et al., 2009) and subsequently in the mammalian system (Ingolia et al., 2011). By sequencing of ribosome-protected mRNA from monosome and polysome, one can detect the translational status of every gene at the same time. One caveat from the study results using this approach is the fact that many uORFs use noncanonical start codons, which are completely missed by computational detection of uORFs based on the sequence. Therefore, experimental search for the uORFs is less prone to false negative. However, up till now, similar method has not been adapted to study the translational control on plant immune system. It is foreseeable that the application of this method in the studies of plant immunity will reveal more uORFs and allows a systematic understanding of their biological functions in the near future.
References


Czarnecka-Verner, E., Pan, S., Salem, T., and Gurley, W.B. (2004). Plant class B HSFs inhibit transcription and exhibit affinity for TFIIB and TBP. Plant Mol Biol 56, 57-75.


Ron, M., and Avni, A. (2004). The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. The Plant cell 16, 1604-1615.


Biography

On June 27, 1985, Wei Wang was born in Shanghai, China. My early interests during high school were chemistry and physics, while I started to learn more about biology and pharmacy in Fudan University in Shanghai. When I was a sophomore, the training in experimental pharmacology which involves animal sacrifice made me uncomfortable. At the meantime, my research proposal to investigate the mechanism of aluminum tolerance of wild rice got funded by Chun-Tsung Scholarship, a fellowship established by a Nobel Prize laureate to support undergraduate research. This opportunity brought me into a plant research lab, where I began my research in plant molecular biology. My subsequent studies on the phosphate transporter of rice brought about my first research article, while my work on the aluminum tolerance was also awarded the excellent thesis by Chun-Tsung Scholarship.

My research experience prompted me to pursue a graduate study in plant biology after graduation from the college. I flew from China to America, where I started my graduate student life in Duke University. I spent my first year rotating in the labs of Drs. Zhenming Pei and Taiping Sun, learning the calcium imaging technique and genetic characterization. My third rotation study with Dr. Xinnian Dong on plant immune system attracted my interest in this interaction between two living organisms. So I joined her lab and started my Ph.D. candidacy. During my research, I have encountered multiple statistics-related questions, which alerted me to the importance of
the adequate knowledge of statistics to ensure a success in this genomics era. With the support from Xinnian, I took all the major courses in statistics and became adept at it. This has greatly contributed to the subsequent publication of my first paper in the Dong lab.

During my stay in the Dong lab, I have attended multiple conferences, workshops and summer school training. I participated in the 20th and 21st NCBC Plant Molecular Biology Retreat in North Carolina. I also attended XIV International Congress on Molecular Plant-Microbe Interactions in Quebec. Every year I joined the trip to Boston for a joint group meeting with members from the labs of Drs. Frederick Ausubel, Xinnian Dong and Shauna Somerville. The two-week summer school at MIT presented me with a systems view on biology which is quite distinct from a statistical perspective. All these activities fulfilled my desire for all kinds of knowledge and have broadened my horizons.

Now it is a time for me to set up a new journey to explore the greater unknowns and experience a different facet of the world.

Publications:

Yan, S.P., Wang, W., Durrant, W., Song, J.Q., Dong, X.N. Synergism between DNA damage response and immunity in plants (In preparation)


*, co-first author

Honors/Awards/Fellowships:

2010 Department Travel Award

2006-2007 People’s Scholarship, First Award

2005-2006 People’s Scholarship, Second Award

2004-2005 People’s Scholarship, Second Award

2005 Chun-Tsung Scholarship

2003-2004 People’s Scholarship, Third Award