The molecular interactions of soybean circadian clock with abiotic stresses and soybean cyst nematode (Heterodera glycines)

Lijun Cao

Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/etd

Part of the Genetics Commons

Recommended Citation
Cao, Lijun, "The molecular interactions of soybean circadian clock with abiotic stresses and soybean cyst nematode (Heterodera glycines)" (2019). Graduate Theses and Dissertations. 17415.
https://lib.dr.iastate.edu/etd/17415

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
The molecular interactions of soybean circadian clock with abiotic stresses and soybean cyst nematode (*Heterodera glycines*)

by

Lijun Cao

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics and Genomics

Program of Study Committee:
Wei Wang, Co-major Professor
Steven A. Whitham, Co-major Professor
Thomas J. Baum
Bing Yang
Peng Liu

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2019

Copyright © Lijun Cao, 2019. All rights reserved.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMENCLATURE</td>
<td>iv</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER 1. GENERAL INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Abstract</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Introduction to the Circadian Clock</td>
<td>1</td>
</tr>
<tr>
<td>1.2.1 Overview of the Circadian Clock</td>
<td>1</td>
</tr>
<tr>
<td>1.2.2 The Origin of the Circadian Clock</td>
<td>2</td>
</tr>
<tr>
<td>1.3 The Plant Circadian Clock</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1 The Central Oscillator of the Plant Circadian Clock</td>
<td>4</td>
</tr>
<tr>
<td>1.3.2 Inputs of the Plant Circadian Clock</td>
<td>6</td>
</tr>
<tr>
<td>1.3.3 Outputs of the Plant Circadian Clock</td>
<td>8</td>
</tr>
<tr>
<td>1.4 The Plant Immune System</td>
<td>9</td>
</tr>
<tr>
<td>1.4.1 Microbe-Associated Molecular Pattern (MAMP)-Triggered Immunity (MTI)</td>
<td>9</td>
</tr>
<tr>
<td>1.4.2 Effector-Triggered Immunity (ETI)</td>
<td>10</td>
</tr>
<tr>
<td>1.4.3 Systemic Acquired Resistance (SAR)</td>
<td>12</td>
</tr>
<tr>
<td>1.5 The Plant Circadian Clock Regulates Defense against Abiotic Stresses</td>
<td>13</td>
</tr>
<tr>
<td>1.6 The Plant Circadian Clock Regulates Defense against Biotic Stresses</td>
<td>16</td>
</tr>
<tr>
<td>1.7 References</td>
<td>18</td>
</tr>
<tr>
<td>CHAPTER 2. COMPREHENSIVE MAPPING OF ABIOTIC STRESS INPUTS TO THE SOYBEAN CIRCADIAN CLOCK</td>
<td>31</td>
</tr>
<tr>
<td>2.1 Abstract</td>
<td>31</td>
</tr>
<tr>
<td>2.2 Introduction</td>
<td>32</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>35</td>
</tr>
<tr>
<td>2.3.1 Identification of Time-Indicating Genes</td>
<td>35</td>
</tr>
<tr>
<td>2.3.2 Application of Molecular Timetable Method</td>
<td>37</td>
</tr>
<tr>
<td>2.3.3 Validation of Global Rhythm Changes and Identification of Specific Targets</td>
<td>39</td>
</tr>
<tr>
<td>2.3.4 Assessment of the Effects of Abiotic Stresses on Circadian Leaf Movement</td>
<td>43</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>43</td>
</tr>
<tr>
<td>2.5 Materials and Methods</td>
<td>48</td>
</tr>
<tr>
<td>2.6 Acknowledgements</td>
<td>52</td>
</tr>
<tr>
<td>2.7 Author Contributions</td>
<td>53</td>
</tr>
<tr>
<td>2.8 References</td>
<td>53</td>
</tr>
<tr>
<td>2.9 Figures</td>
<td>58</td>
</tr>
<tr>
<td>CHAPTER 3. BIDIRECTIONAL INTERACTION BETWEEN SOYBEAN ROOT CIRCADIAN CLOCK AND SOYBEAN CYST NEMATODE (HETERODERA GLYCINES)</td>
<td>73</td>
</tr>
<tr>
<td>3.1 Abstract</td>
<td>73</td>
</tr>
<tr>
<td>3.2 Introduction</td>
<td>74</td>
</tr>
<tr>
<td>3.3 Results and Discussions</td>
<td>76</td>
</tr>
</tbody>
</table>
3.3.1 SCN Perturbed Global Circadian Rhythm in Syncytia .................................................. 76
3.3.2 Identification of Circadian Genes Targeted by SCN using RASL-seq .................. 77
3.3.3 SCN Perturbed the Expression Rhythms of Two Circadian Genes in Soybean Hairy Roots .......................................................................................................................... 78
3.3.4 Soybean Seedlings Are More Resistant to SCN at Dawn than at Dusk ............. 80
3.3.5 Soybean Clock Mutant Hairy Root Lines Showed Disrupted SCN Infection Rhythmicity ....................................................................................................................................... 81
3.3.6 Identification of Genes in the Interaction between Soybean Root Circadian Clock and SCN using Time-course RNA sequencing (RNA-Seq) .................................................... 82
3.4 Materials and Methods .................................................................................................. 84
3.5 Author Contributions ..................................................................................................... 87
3.6 References ...................................................................................................................... 87
3.7 Figures ............................................................................................................................ 93

CHAPTER 4. GENERAL CONCLUSIONS AND FUTURE PLANS ................................................. 110
4.1 Mapping of Abiotic Stress Inputs to Soybean Circadian Clock ........................... 110
4.2 Future Direction for Impacts of Abiotic Stresses on Soybean Circadian Clock ...... 111
4.3 The Interaction of Soybean Root Circadian Clock and Soybean Cyst Nematode (SCN) ....................................................................................................................................... 111
4.4 Future Plan for the Interaction of Soybean Root Circadian Clock and SCN ......... 112
4.5 References ...................................................................................................................... 114
4.6 Figures ............................................................................................................................ 115
### NOMENCLATURE

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA</td>
<td>Abscisic Acid</td>
</tr>
<tr>
<td>ABAR/CHLH/GUN5</td>
<td>ABA-Related/H Subunit of the Magnesium-Protoporphyrin IX Chelatase/Genomes Uncoupled 5</td>
</tr>
<tr>
<td>AZI1</td>
<td>Azelaic Acid Induced 1</td>
</tr>
<tr>
<td>BAK1</td>
<td>Brassinosteroid Receptor 1-Associated Kinase 1</td>
</tr>
<tr>
<td>bHLH</td>
<td>Basic Helix-Loop-Helix</td>
</tr>
<tr>
<td>CAB2</td>
<td>Chlorophyll A/B-Binding Protein 2</td>
</tr>
<tr>
<td>CAT2</td>
<td>Catalase 2</td>
</tr>
<tr>
<td>CBF1</td>
<td>C-Repeat Binding Factor 1</td>
</tr>
<tr>
<td>CBS</td>
<td>CCA1 Binding Site</td>
</tr>
<tr>
<td>CC</td>
<td>Coiled-Coil</td>
</tr>
<tr>
<td>CCA1</td>
<td>Circadian Clock Associated 1</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge Coupled Device</td>
</tr>
<tr>
<td>CEBiP</td>
<td>Chitin Elicitor Binding Protein</td>
</tr>
<tr>
<td>CHE</td>
<td>CCA1 Hiking Expedition</td>
</tr>
<tr>
<td>CO</td>
<td>CONSTANS</td>
</tr>
<tr>
<td>COP1</td>
<td>Constitutive Photomorphogenic 1</td>
</tr>
<tr>
<td>COR</td>
<td>Cold Response</td>
</tr>
<tr>
<td>COS</td>
<td>Cosine Function Combined with a Linear Trend</td>
</tr>
<tr>
<td>CRT/DRE</td>
<td>Repeat/Dehydration Response Element</td>
</tr>
<tr>
<td>CT</td>
<td>Circadian Time</td>
</tr>
<tr>
<td>CV</td>
<td>Coefficient of Variation</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>DEX</td>
<td>Dexamethasone</td>
</tr>
<tr>
<td>DIR1</td>
<td>Defective in Induced Resistance 1</td>
</tr>
<tr>
<td>EC</td>
<td>Evening Complex</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>Elongation Factor Tu</td>
</tr>
<tr>
<td>ELF</td>
<td>Early Flowering</td>
</tr>
<tr>
<td>EPR1</td>
<td>Early Phytochrome Responsive 1</td>
</tr>
<tr>
<td>ETI</td>
<td>Effector-Triggered Immunity</td>
</tr>
<tr>
<td>ETS</td>
<td>Effector-Triggered Susceptibility</td>
</tr>
<tr>
<td>FAR1</td>
<td>Far-Red Impaired Response 1</td>
</tr>
<tr>
<td>FDR</td>
<td>False Discovery Rate</td>
</tr>
<tr>
<td>FFT-NLLS</td>
<td>Fast Fourier Transform-Non-Linear Least Squares</td>
</tr>
<tr>
<td>FHY3</td>
<td>Far-Red Elongated Hypocotyl 3</td>
</tr>
<tr>
<td>FKF1</td>
<td>Flavin-Binding Kelch Repeat F Box 1</td>
</tr>
<tr>
<td>flg22</td>
<td>the 22 Conserved Amino Acids of Flagellin Peptide</td>
</tr>
<tr>
<td>FLS2</td>
<td>Flagellin-Sensitive 2</td>
</tr>
<tr>
<td>FT</td>
<td>Flowering Locus T</td>
</tr>
<tr>
<td>G3P</td>
<td>Glycerol-3-Phosphate</td>
</tr>
<tr>
<td>GAs</td>
<td>Gibberellins</td>
</tr>
<tr>
<td>GI</td>
<td>GIGANTEA</td>
</tr>
<tr>
<td>GRP7</td>
<td>Glycine-Rich RNA Binding Protein 7</td>
</tr>
<tr>
<td>HDAC6</td>
<td>Histone Deacetylase 6</td>
</tr>
<tr>
<td>HY5</td>
<td>Elongated Hypocotyl 5</td>
</tr>
<tr>
<td>ICS1</td>
<td>Isochorismate Synthase 1</td>
</tr>
</tbody>
</table>
ISR  Induced Systemic Resistance
JA   Jasmonic Acid
LDL1/2 Lysine-Specific Demethylase 1 (LSD1)-Like Histone Demethylases 1/2
LHY  Late Elongated Hypocotyl
LNK  Night Light-Inducible and Clock-Regulated
LRR  Leucine-Rich Repeat
LS   Lomb-Scargle Periodogram
LUX  LUX ARRHYTHMO
LWD  Light-Regulated WD
MAMPs Microbe-Associated Molecular Patterns
MAPK Mitogen-Activated Protein Kinase
MESA Maximum Entropy Spectral Analysis
MeSA Methyl Salicylate
MTI  MAMP-Triggered Immunity
PAS  Per-ARNT-Sim
PBA1 Proteasome Subunit Beta Type-6
PIF3 Phytochrome Interacting Factor 3
PR   Pathogenesis-Related
PRRs Pseudo-Response Regulators
PRR  Pattern Recognition Receptors
RASL-seq RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIN4</td>
<td>RPM1 Interacting Protein 4</td>
</tr>
<tr>
<td>RIPK</td>
<td>RPM1-Induced Protein Kinase</td>
</tr>
<tr>
<td>RLKs</td>
<td>Receptor-Like Kinases</td>
</tr>
<tr>
<td>RLPs</td>
<td>Receptor-Like Proteins</td>
</tr>
<tr>
<td>RNA-Seq</td>
<td>RNA Sequencing</td>
</tr>
<tr>
<td>RNS</td>
<td>Reactive Nitrogen Species</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>RVE8</td>
<td>Reveille 8</td>
</tr>
<tr>
<td>SA</td>
<td>Salicylic Acid</td>
</tr>
<tr>
<td>SAR</td>
<td>Systemic Acquired Resistance</td>
</tr>
<tr>
<td>SCN</td>
<td>Soybean Cyst Nematode</td>
</tr>
<tr>
<td>SOS2</td>
<td>Salt Overly Sensitive 2</td>
</tr>
<tr>
<td>TCP</td>
<td>Teosinte Branched 1/Cycloidea/PCF</td>
</tr>
<tr>
<td>TIC</td>
<td>Time for Coffee</td>
</tr>
<tr>
<td>TIR</td>
<td>Toll/Interleukin 1 Receptor</td>
</tr>
<tr>
<td>TOC1</td>
<td>Timing of Cab2 Expression 1</td>
</tr>
<tr>
<td>TPL</td>
<td>TOPELESS</td>
</tr>
<tr>
<td>VIGS</td>
<td>Virus-Induced Gene Silencing</td>
</tr>
<tr>
<td>ZT</td>
<td>Zeitgeber Time</td>
</tr>
<tr>
<td>ZTL</td>
<td>ZEITLUPE</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

I would like to take this opportunity to express my sincere appreciation to the people I’ve met these years. Thank you, to each and every person that has brought me here, made my time at Iowa more than I can imagine, and left me lifelong friendships that I will forever cherish.

First of all, I am very grateful to my advisor Dr. Wei Wang for the instruction and support throughout my research. Wei guided me step by step throughout the work, from sampling to data collecting and analysis, and supported me in job applications.

Thanks to my advisor Dr. Steven A. Whitham and committee member, Dr. Thomas J. Baum. They helped find funding resources and support me in completing graduate study at ISU. Their continued supports in research, job applications, and life are greatly appreciated.

Dr. Bing Yang is one of the kindest people I’ve ever met. I knew Bing from the Ph.D. application interview in 2015. Then, I was honored to study at ISU and rotated in his lab. After that, he became one of my committee members, giving me suggestions in conducting research, taking courses, and supporting me in the job applications. His help and kindness will always be remembered.

I would like to thank Dr. Peng Liu. I took the RNA-Seq course instructed by her. Peng is a very conscientious and patient professor. And I am very grateful by all the efforts and help she made to support me through the Ph.D. study.

I would like to thank my family, my wife, my son, my parents, my sister, my brother-in-law, my parents-in-law and my sister-in-law. Thanks to my wife, Huimin Li, for being with me through these ups and downs, my son, Jinghua Cao, for bringing so much happiness into my life.
Thanks to my friends, Changtian Chen, Zihao Zheng, Yang Yang, Yingzhou Du, Qiang Liu, Qian Dong, Jian Cheng, Ruirui Xu, etc. They are always around to help whenever I have troubles. A friend in need is a friend indeed. They really make a difference in my life.

Thanks to my colleagues, the department faculties and staffs, and all the others. Every single one of you has made a positive impact on my life and has changed me for the better.
Plants, like many other living organisms, evolved circadian clock to increase fitness by synchronizing physiological processes with environmental oscillations. In one hand, the circadian clock helps plants prepare for almost every growth routine, such as germination, flowering, and photosynthesis. On the other hand, the circadian clock regulates plant defense against biotic and abiotic stresses, and these stresses feedback to modulate the clock function. However, most of these findings so far made were from studies of *Arabidopsis*. Basically, very little is known about soybean (*Glycine max*) circadian clock and its interactions with stresses.

In Chapter 2, to unveil the impacts of various abiotic stresses on soybean circadian clock, we provided a two-module pipeline consisting of the Molecular Timetable module and the RASL-seq (RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing) module. Using the Molecular Timetable method, we identified 3,695 time-indicating genes in soybean seedlings. Then, using these time-indicating genes, we re-analyzed the publicly available soybean transcriptomes related to abiotic stresses. We found that mild drought caused little changes to the global circadian rhythm while severe drought induced a significant phase shift. And, a 30-minute heat shock was sufficient to cause a phase advancement. However, prolonged heat had a rather mild effect on the global circadian rhythm. These findings suggest that the global circadian rhythm responses to drought and heat are duration-dependent.

In addition to drought and heat, prolonged exposure to iron deficiency caused phase delay of the leaf circadian clock in two near-isogenic lines, Clark and IsoClark. However, in roots, phase of the global circadian clock was advanced in Clark while it was delayed in IsoClark. Alkaline stress also caused organ-specific changes to the global circadian rhythm in the wild soybean,
Glycine soja. These findings suggest organ-specific circadian rhythm responses to both iron deficiency and alkaline stresses.

Then, 49 soybean circadian clock genes were surveyed using RASL-seq. It enabled us to identify the clock components targeted by each stress. Lastly, to test whether the stress-induced transcriptional changes of soybean clock genes could lead to changes in clock-controlled physiological outputs, we assessed the effect of abiotic stresses on circadian leaf movement. We found that perturbation in individual or subset of clock components might not necessarily jeopardize clock outputs.

In Chapter 3, I used the two-module pipeline to study the impact of soybean cyst nematode (SCN) on soybean root circadian clock. First, I used the time-indicating genes to re-analyze the publicly available soybean transcriptome data conducted on both SCN feeding and surrounding sites. I found that SCN caused significant changes to the global circadian rhythm in the feeding site. However, circadian rhythm disturbance was not found in the surrounding site. Then, using RASL-seq, I identified that the expression rhythms of 12 circadian genes were modified by SCN. 5 showed significant phase shift and 8 showed significant period changes (1 showed both phase and period changes). Next, I used a live bioluminescence imaging system to confirm the results in soybean hairy roots. The vector carrying a luciferase gene driven by the promoter of the identified circadian gene (promoter-luciferase) was constructed and transformed into soybean hairy roots. I found that only one gene Early Phytchrome Responsive 1 (EPR1) promoter-luciferase activity showed robust circadian rhythm under prolonged constant condition. Then, I performed SCN infection assays and found an obvious phase shift to the EPR1 promoter-luciferase activity after SCN infection. In addition, I also found SCN caused a significant change to the average activity
of the CCA1 promoter-luciferase. Overall, I found that the promoter-luciferase activity rhythms of 2 circadian genes were modified by SCN in soybean hairy roots.

On the other hand, I also studied how the root circadian clock regulates soybean defense against SCN. I found that soybean seedlings were more resistant to SCN at dawn than at dusk. Then, I generated the soybean clock mutant hairy root lines by overexpressing one CCA1 gene (OE) and found a disrupted SCN infection rhythm in OE lines. In addition, OE lines are more resistant to SCN than the control lines.

Next, using a time-course RNA sequencing experiment, I identified key circadian rhythmically expressed and non-circadian-rhythm genes in the interaction between soybean root circadian clock and SCN. Roles of these genes still need further investigations.

These results enable us to uncover novel inputs and outputs to the soybean circadian clock. And, these unique responses in soybean demonstrate the necessities to directly study crop circadian clocks, and our discovery may serve as a broadly applicable procedure to facilitate these researches.
CHAPTER 1. GENERAL INTRODUCTION

1.1 Abstract

As sessile organisms, plants cannot escape from unfavorable environmental challenges or dangers, such as abiotic and biotic stresses. However, they are able to defend themselves with various strategies. Plant circadian clock has emerged as a cardinal signaling hub integrating diverse immune signals to gate plant defenses and coordinate metabolic growth needs. And stresses also regulate the circadian clock function. These bidirectional interactions constitute a complicated regulatory network. This introduction is focused on the review of circadian clock origin, plant circadian clock network, and plant immune system.

1.2 Introduction to the Circadian Clock

1.2.1 Overview of the Circadian Clock

The Earth's rotation around its axis results in robust rhythmic environmental changes—day and night. Most living organisms have evolved circadian clock to use this predictable information and synchronize their internal activities with this daily cycle. Living organisms not only respond to these environmental cues, but they anticipate them and adjust their biological behaviors accordingly (Karapetyan et al., 2018). The circadian rhythm distinguishes itself from the other biological rhythms by three characteristics: firstly, it is endogenously generated and self-sustaining, which means, circadian components oscillate even under constant environmental conditions, typically constant light (or dark) and constant temperature; secondly, the rhythm is temperature-compensated, for example, the period (~24h) of the circadian rhythm is relatively unchanged over a range of ambient temperatures (McClung, 2006); thirdly, the rhythm is entrainable, the circadian rhythm can be reset by external stimuli, such as temperature and light. These external stimuli are called Zeitgebers, which means “time giver”. For example, with the entrainment of local time cues,
people experiencing jet lag from traveling across time zones adjust their circadian rhythm to sync with the local time.

The circadian rhythm is regulated by the central oscillator of the circadian clock, which consists of complicated transcriptional and translational feedback networks. The input pathways, the central oscillator, and the output pathways are the three essential parts of the circadian clock. Input signals are external and internal cues that entrain the circadian clock while the output pathways are responses governed by the clock, and the central oscillator plays a pivotal role in integrating input stimuli, generating the circadian rhythm, and coordinating output responses.

1.2.2 The Origin of the Circadian Clock

The circadian clock controls almost all physiological processes, including sleep-wake cycles, the carbohydrate, lipid and protein metabolism, the regulation of growth hormones, stress responses, and immune pathways (Fu et al., 2002; Fu et al., 2005; Matsuo et al., 2003; Pregueiro et al., 2006; Turek et al., 2005). However, the origin of the circadian clock is still unclear.

The circadian clock is unique in a specific organismal system, but commons are shared throughout different systems (Schwartz et al., 2017). Genetic screens have revealed many circadian genes in multiple systems, including but not limited to human, mice, *Drosophila melanogaster* (fruit fly), *Neurospora*, plants, and cyanobacteria (Rosbash et al., 2009). These findings identified lots of clock components with similar function. For example, mammals and *Drosophila* share many clock orthologs (Hamilton et al., 2008; Kay et al., 1997; Young et al., 2001; Yu et al., 2006). The *Drosophila* proteins, including the Per-ARNT-Sim (PAS) domain-containing transcription factor heterodimer Clock-Cycle, have an ortholog in mammals: CLK-BMAL1 (Rosbash et al., 2009). The negative regulator “period” (PER) in *Drosophila* has orthologs in mammals: PER1 and PER2 (Rosbash et al., 2009). These findings indicate there might be a similar clock mechanism existed in a common ancestor of insects and mammals, probably
from more than 500 million years ago, before the separation of insects and mammals (Rosbash et al., 2009). The fact that plant circadian components sharing low sequence similarities with their relatives in mammals may indicate an independent origin (Rosbash et al., 2009). In the filamentous fungus Neurospora crassa, the clock mechanism is very similar, but non-orthologous, to those of mammals, flies, and plants (Brunner et al., 2006). Furthermore, bacteria circadian clock compositions appear completely unrelated to that in mammals, plants, and fungi (Nakajima et al., 2005). Kai C protein phosphorylation cycle is assumed to be a basic timing process of the circadian clock in bacteria. The self-sustainable oscillation of Kai C phosphorylation can be constructed in vitro by incubating Kai A, B, C, and ATP. Based on these facts, some scientists hypothesized there might be multiple origins of the circadian clock (Rosbash et al., 2009).

In addition, antioxidant proteins called peroxiredoxins accept or give up oxygen molecules in a circadian rhythmic way. These oxidation-reduction cycles were found in a wide variety of organisms, including human blood cells which do not have nuclei (O’Neill et al., 2011), fruit fly, marine algae Ostreococcus tauri, plant Arabidopsis thaliana, a fungus Neurospora crassa, and a cyanobacteria Synechococcus elongatus (Edgar et al., 2012). Peroxiredoxins themselves are not clock geared. The redox homeostatic mechanisms were supposed to have an intimate co-evolution with the cellular circadian clock, probably after the Great Oxidation Event ~2.5 billion years ago (Edgar et al., 2012).

The clock proteins of animals, plants, and cyanobacteria are very different, but scientists found that the catalytic domains of some kinases that set the speed of the clock (Chiu et al., 2011) are conserved throughout different systems (Brott et al., 1998; Causton et al., 2015; Chiu et al., 2011). Findings suggest that kinases are important for establishing the circadian clock rhythm. Some scientists even suggest that kinases may originally form a simple clock, similar to
cyanobacteria’s Kai A, B, C system. With this simple clock, organisms add more gears to form the current clocks, or there may be no common ancestor of the circadian clock, and cells may simply fall into rhythmic patterns driven by biochemical reactions (Brott et al., 1998; Causton et al., 2015).

1.3 The Plant Circadian Clock

1.3.1 The Central Oscillator of the Plant Circadian Clock

Most of the plant circadian clock components are transcription factors. The current understanding of the molecular architecture of plant circadian clock is primarily from studies in the model organism, Arabidopsis thaliana. In Arabidopsis, the central oscillator of the circadian clock is composed of a large number of transcription factors arranged in many feedback loops (McClung, 2019).

The initial loop consists of an evening-phased transcription factor Timing of Cab2 Expression 1 (TOC1) and two partially redundant morning-phased transcription factors, Circadian Clock Associated 1 (CCA1) and Late Elongated Hypocotyl (LHY). TOC1 is the first circadian component identified in Arabidopsis. Strayer et al. cloned TOC1 gene from a forward genetic screen based on a shortened circadian rhythm mutant toc1 (Strayer et al., 2000). TOC1 negatively regulates CCA1 and LHY expression. Conversely, CCA1 and LHY bind to TOC1 promoter at a CCA1 Binding Site (CBS) called the Evening Element and repress TOC1’s expression. Overexpression of TOC1, CCA1 or LHY results in circadian arrhythmicity in plants (Huang et al., 2012; Makino et al., 2002; Más et al., 2003a; Michael et al., 2002; Schaffer et al., 1998; Wang et al., 1998). In addition, CCA1 and LHY are also negatively regulated by themselves. CCA1, LHY, and TOC1 constitute a core negative feedback loop in the central oscillator.

This initial loop is further connected with many other loops to create the complex architecture of the plant circadian clock (Alabadi et al., 2001). At dawn, CCA1 and LHY have the highest expression level and repress the expression of Pseudo-Response Regulators (PRRs,
including PRR5, 7, and 9), GIGANTEA (GI), Evening Complex (EC) [including Early Flowering 3 and 4 (ELF3 and ELF4), and LUX ARRHYTHMO (LUX)] (Nagel et al., 2015). PRRs and TOC1 are sequentially expressed and repress the expression of CCA1 and LHY (Adams et al., 2015; Kamioka et al., 2016). Light-Regulated WD 1 (LWD1) and LWD2 are recruited to DNA by Teosinte Branched 1/Cycloidea/PCF 20 (TCP20) and TCP22, and act as co-activators to promote the expression of CCA1, PRR7, PRR9, and TOC1 (Wu et al., 2008; Wu et al., 2016). In the afternoon, transcriptional activator Reveille 8 (RVE8) (and probably RVE4 and RVE6) recruits co-activator Night Light-Inducible and Clock-Regulated 1 (LNK1) and LNK2 to increase the transcription of PRR5, PRR9, TOC1, GI, LUX, and ELF4 (Farinas et al., 2011; Hsu et al., 2013; Rawat et al., 2011). ELF4 could also be activated by Far-Red Elongated Hypocotyl 3 (FHY3), Far-Red Impaired Response 1 (FAR1), and Elongated Hypocotyl 5 (HY5) (Li et al., 2011; McClung, 2019). In the evening, TOC1 expression peaks and acts as a repressor to repress almost all the circadian components, such as GI, LUX, and ELF4 (Gendron et al., 2012; Matsushika et al., 2000). The EC shows the highest expression late at night, and it represses the expression of TOC1, GI, and ZTL, and PRRs, which indirectly promotes the expression of CCA1 and LHY (Greenham et al., 2015; McClung, 2019; Yang et al., 2018). While CCA1 and LHY negatively regulate the expression of TOC1 by directly binding to its promoter, a recent study indicates that CCA1 and LHY also regulate the expression of TOC1 by histone modifications. CCA1 and LHY accumulate highest in the morning, and they recruit LDL1/2-HDAC6 [Lysine-Specific Demethylase 1 (LSD1)-Like Histone Demethylases 1/2 (LDL1/2) and Histone Deacetylase 6 (HDAC6)] histone modification complex to repress target gene including TOC1 by histone demethylation and deacetylation (Hung et al., 2018). In the evening, CCA1 and LHY have low expression level, while
TOC1 expresses high because LDL1/2 and HDAC6 are released from the promoter of TOC1 (Hung et al., 2018).

The protein degradation and stabilization also play important roles in circadian clock function. The GI and ZTL form a heterodimer, thereby stabilizing each other in the cytosol (Kim et al., 2013). ZTL further mediates proteasomal degradation of TOC1. The degradation is interrupted by the interaction of ZTL with PRR3 (hindering ZTL access) and PRR5 (helping TOC1 enter into the nucleus) (Baudry et al., 2010; Más et al., 2003b). Additionally, ZTL targets TOC1 and PRR5 for ubiquitination and promotes proteasomal degradation (Wang et al., 2010). Furthermore, ELF3 physically interacts with Constitutive Photomorphogenic 1 (COP1) and GI, which promotes GI’s degradation (Yu et al., 2009).

1.3.2 Inputs of the Plant Circadian Clock

Environmental stimuli, such as light, temperature, and humidity, entrain the circadian clock, and biotic and abiotic stresses modulate the clock functions (Lu et al., 2017).

Light and temperature are the two major clock input signals. The light-mediated entrainment is achieved at transcriptional, translational, and post-translational levels (Ito et al., 2012; Rugnone et al., 2013; Staiger et al., 2011). The circadian clock also regulates its own sensitivity to the light input (Fankhauser et al., 2002).

Light regulates the degradation of PRRs, TOC1, and GI (Farré et al., 2007; Ito et al., 2007; Kiba et al., 2007). In addition, it was reported that PRR5, PRR9, and TOC1 regulate the circadian period in a light-dependent manner (Eriksson et al., 2003; Más et al., 2003a). In plants, red and far-red lights are perceived by phytochromes while UV-A and blue lights are perceived by cryptochromes. The phytochrome exists in an inactive form known as the Pr form. Red light converts phytochrome into an active form known as the far-red light absorbing (Pfr) form (Quail, 2002). It is proposed that the Pfr form accelerates the clock while the Pr form slows down the
There are more than 10 photoreceptors showing circadian oscillation (Edwards et al., 2015). Excitation of these photoreceptors results in speeding up of the circadian clock rhythm (Devlin et al., 2000; Herrero et al., 2012; Somers et al., 1998a; Somers et al., 1998b). The mechanism of how photoreceptors modulate the clock is not well understood. But it is known that 3 out of 5 identified Arabidopsis phytochromes, PHYB, PHYC, and PHYD, physically interact with ELF3 (Kolmos et al., 2011; Liu et al., 2001). A recent study showed that the phyAphyBphyCphyDphyE quintuple mutant had a longer period under a relatively high fluence of a continuous red-light condition and a shorter period under a relatively low fluence rate condition (< 10 µmol/m²/s⁻¹) compared with the wild type. ELF4 was found to be a potential target of PHYA signaling pathway. Phytochrome Interacting Factor 3 (PIF3) dimer binds directly to the promoters of CCA1 and LHY, and Pfr interacts with the promoter-bound PIF3 (Martínez-García et al., 2000). ZTL was also found to be involved in the light regulation of the circadian clock. ZTL was reported to be a blue-light photoreceptor (Kim et al., 2007). Under blue light, ZTL is stabilized by GI and consequently negatively regulates CCA1, LHY, and TOC1 (Kim et al., 2007).

Compared with the studies of light-mediated entrainment, the mechanism of how temperature entrains the circadian clock is largely unknown. Although daily temperature oscillation entrains the circadian clock, circadian rhythm maintains robust and unchanged through a broad range of physiological temperatures. PRR7 and PRR9 were shown to be essential for this process. In the prr7prr9 double mutant, plants show overcompensated due to hyperactivation of CCA1 and LHY (Salomé et al., 2010). In addition, PHYB signaling pathway is also involved in the temperature compensation effect, and it was found recently that PHYB is also a temperature sensor in plants (Legris et al., 2016). Temperature regulates the expression of CCA1, LHY, TOC1, and GI. When temperature increases (from 17 to 27 °C), amplitude and peak expression level of TOC1 and
GI increase whereas they decrease for LHY. When temperature decreases (from 17 to 12 °C), amplitude and peak expression level of CCA1 and LHY increase (Gould et al., 2006). On the other hand, temperature regulates alternative splicing of CCA1, LHY, PRR3, PRR5, PRR7, PRR9, and TOC1 (Filichkin et al., 2012; James et al., 2012; Kwon et al., 2014). For example, CCA1 has at least two alternative splicing forms, CCA1α and CCA1β. CCA1α dominates under low temperature while CCA1β dominates under high temperature. And CCA1α was shown to be essential for cold tolerance (Seo et al., 2012).

In addition to these external stimuli, internal metabolites also entrain the plant circadian clock. Sugar production was shown to be an input of the circadian clock, and it modulates plant circadian clock by regulating the transcription of PRR7 (Haydon et al., 2013).

1.3.3 Outputs of the Plant Circadian Clock

On the other hand, the circadian clock controls diverse developmental processes of plant life. Many genes are under the control of the circadian clock. Over 40 % of all protein-coding genes have circadian rhythmic transcription in mammals (Zhang et al., 2014). Approximately 30 % of Arabidopsis expressed genes in seedlings are under clock control (Covington et al., 2008). Recently, the central clock oscillators, including TOC1, PRR5, PRR7, RVE8, and EC, were found to directly control the transcripts of many genes involved in various biological processes, including germination, hypocotyl elongation, growth, reproduction, and pollination (Hsu et al., 2013; Huang et al., 2012; Liu et al., 2013; Nakamichi et al., 2012). Two Basic Helix-Loop-Helix (bHLH) transcription factors, PIF4 and PIF5, are crucial for determining the hypocotyl elongation rate in seedlings. Circadian components, such as EC, are required for the proper expression of PIF4 and PIF5 (Nusinow et al., 2011). In addition, the circadian clock regulates flowering initiation of Arabidopsis in two ways. First, GI and Flavin-Binding Kelch Repeat F Box 1 (FKF1) induce the transcription of Flowering Locus T (FT), and FT is important for flower formation in Arabidopsis
(Lu et al., 2012; Nusinow et al., 2011). Second, CONSTANS (CO) induces the transcription of FT in the leaf (Huang et al., 2005; Wigge et al., 2005), and the transcription of CO is under circadian regulation. Transcription of CO has a circadian oscillation with peak expression at the end of the day under long day (10 hours light/6 hours day extension/8 hours dark) and at night under short day (10 hours light/14 hours dark) (Suárez-López et al., 2001).

Taken together, the reciprocal regulation of the circadian clock with input and output pathways establishes a complicated network in which the inputs and outputs are connected by the central circadian oscillators. Given the complexity of the circadian regulation, further efforts are important in understanding the full map of the plant circadian clock.

1.4 The Plant Immune System

Unlike vertebrates, plants do not have circulatory systems and mobile immune cells (the vertebrate adaptive immune system relies on the almost infinite diversity of antigen receptors produced by the immune cell in the circulatory system) to protect them against bacterial, fungal, and viral pathogens (Spoel et al., 2012). However, plants are able to defend themselves through elaborate strategies.

1.4.1 Microbe-Associated Molecular Pattern (MAMP)-Triggered Immunity (MTI)

The plant cell walls, waxes, trichomes (leaf hairs), antimicrobial chemicals provide the first layer of physical and chemical obstacles against phytopathogens (Malinovsky et al., 2014). If these obstacles are overcome, pathogens are still confronted by defense responses of plant cells. The first action of plant cell defenses is pathogen recognition. MAMPs, such as bacterial flagellin and Elongation Factor Tu (EF-Tu), fungal chitin and xylanase, and oomycete heptagulcan, can be recognized by Pattern Recognition Receptors (PRR) and activate MTI in plants (Boller et al., 2009). PRR are membrane-bound Receptor-Like Kinases (RLKs) or Receptor-Like Proteins (RLPs). There are many RLKs in Arabidopsis, and some of them are involved in plant immunity
while others have key roles in plant development, symbiosis, and self-incompatibility (Boller et al., 2009; Spoel et al., 2012). The well-studied example of MAMP-PRR interaction is the *Pseudomonas syringae* (*P. syringae*) flg22 (the 22 Conserved Amino Acids of Flagellin Peptide) and the plant flg22 receptor Flagellin-Sensitive 2 (FLS2) (Gómez-Gómez et al., 2000; Robatzek et al., 2006). FLS2 locates in the plasma membrane and undergoes ligand-induced endocytosis when binding directly to flg22 (Robatzek et al., 2006). After binding, FLS2 interacts with several co-receptors, including Brassinosteroid Receptor 1-Associated Kinase 1 (BAK1) and a Leucine-Rich Repeat (LRR) kinase, to activate MTI signaling pathway (Chinchilla et al., 2007; Heese et al., 2007). The early MTI signaling responses occur within several minutes. These include ion fluxes, the oxidative burst, and the activation of a downstream Mitogen-Activated Protein Kinase (MAPK) cascade. The signal transduction further induces transcriptional changes and the production of antimicrobial compounds, such as Pathogenesis-Related (PR) proteins and phytoalexins (Yi et al., 2014). As a late response, MTI triggers callose deposition to reinforce the cell wall (Boller et al., 2009). FLS2 has three main domains, the LRR domain for ligand binding, a transmembrane domain, and a serine/threonine kinase domain. Some pathogen effectors, including *P. syringae* pv. *tomato* AvrPto1, were proven to help virulence by suppressing immune proteins (Zipfel et al., 2008), and result in Effector-Triggered Susceptibility (ETS) (Jones et al., 2006). To cope with these effectors, plants have intracellular immune receptors, the Resistance (R) proteins. These R proteins can recognize certain effector proteins and activate ETI
(also called, R-gene-mediated resistance). This immune response is usually characterized by hypersensitive response associated with rapid cell death at the infection site and the production of antimicrobial molecules in the surrounding cells, such as the chitinase and β-1,3-glucanase (Wu et al., 2014). R proteins typically consist of a variable amino terminus [could be either a Coiled-Coil (CC) domain or a Toll/Interleukin 1 Receptor (TIR)], a nucleotide-binding domain, and an LRR domain (NB-LRR). A mechanism of R protein activation has been proposed that: in the absence of the ligand, the three domains of R protein interact intracellularly, inhibit nucleotide exchange and hydrolysis of the central nucleotide-binding domain, thereby inactivating the R protein. Following binding to the ligand, the intracellular interaction is released, which results in nucleotide exchange and hydrolysis of the nucleotide-binding domain. This possible conformational change activates R proteins and induces the downstream signaling which leads to ETI (Axtell et al., 2003a; Axtell et al., 2003b; Kim et al., 2005). The recognition of pathogen effectors by R proteins is described by several proposed models. One of them is the “Guard Hypothesis” model, stating that pathogen effectors modify or perturb the pathogen-targeted plant self-proteins, resulting in activation of R proteins. This is exemplified in the guarded cellular target RPM1 Interacting Protein 4 (RIN4) by R proteins RPM1 and RPS2 in Arabidopsis (Mackey et al., 2002; Kim et al., 2005). RIN4 physically interacts with RPM1 and RPS2 and maintains these R proteins in an inactive state. Following infection by pathogens, such as P. syringae, which injects effector proteins AvrB and AvrRpm1 into the plant cell, RIN4 is phosphorylated by RPM1-Induced Protein Kinase (RIPK) (Liu et al., 2011). Phosphorylation of RIN4 activates RPM1. Another P. syringae effector AvrRpt2 induces direct cleavage of RIN4, thereby activating R protein RPS2 (Chung et al., 2011; Liu et al., 2011). The activation of R protein by pathogen effectors often leads to programmed cell death in the infected sites (Spoel et al., 2012). The
mechanism of R protein-mediated programmed cell death is not fully understood. However, in the defense response of *Arabidopsis* against bacterial pathogen *P. syringae pv. tomato*, Proteasome Subunit Beta Type-6 (PBA1) induces caspase-3-like protein activity. This activity is required for membrane fusion of the central vacuole and plasma membrane, which results in the production of anti-bacterial factors, programmed cell death, and promoting signaling from the vacuole (Hatsugai et al., 2009). Overall, ETI is generally a stronger immune response of plants compared with MTI (Spoel et al., 2012).

### 1.4.3 Systemic Acquired Resistance (SAR)

Unlike the circulatory system in animals, plants transport the mobile immune signals, possibly Methyl Salicylate (MeSA), azelaic acid, Glycerol-3-Phosphate (G3P), and lipid-transfer proteins [such as Defective in Induced Resistance 1, (DIR1) and Azelaic Acid Induced 1, (AZI1)], from the infected sites to systemic tissues to establish SAR (Chanda et al., 2011; Chaturvedi et al., 2012; Maldonado et al., 2002; Návarová et al., 2012; Park et al., 2007; Spoel et al., 2012). SAR is typically characterized by the accumulation of SA and the induction of PR proteins. The initial response of SAR is accompanied by the accumulation of SA. SA methyltransferase converts SA to MeSA, and SA methyltransferase was shown to be required in plant tissues that generate immune signals. MeSA is reversely converted to SA by MeSA esterase. MeSA esterase is required for signal perception in systemic tissues (Spoel et al., 2012). Exogenous application of MeSA induces systemic immunity in tobacco plants (Park et al., 2007). However, it is unknown whether the MeSA-induced SAR is dependent on SA methyltransferase or not. In addition, *Arabidopsis* mutant defective in the biosynthesis of G3P fails in activating SAR, and exogenous application of G3P can rescue SAR in distal tissues. An unknown cofactor from SAR-induced vascular exudates enhances G3P function (Chanda et al., 2011). In addition to SA and G3P, azelaic acid identified from pathogen infection-induced plant vascular exudates is also a potential mobile immune signal
The lipid-transfer proteins DIR1 and AZI1 are involved in the function and translocation of the mobile immune signals (Maldonada et al., 2002). Accumulation of these mobile immune signals in uninfected plant cells induces the secretion of PR proteins along with antimicrobial chemicals, histone methylation, and other chromatin modifications, which enables plants to confer a broad-spectrum of pathogen resistance (Spoel et al., 2012).

1.5 The Plant Circadian Clock Regulates Defense against Abiotic Stresses

Extreme environmental conditions, like extreme temperature, water, and salt, cause widespread plant damages throughout the world. Crop losses caused by water deficit were estimated to be $130 billion in Europe from 1977 to 2007 (Europea et al., 2007). The drought affected 80% of the crop-growing regions in the USA in 2012 (Hoerling et al., 2013), and it was predicted to be more frequent due to global warming (Cook et al., 2015). It is urgent to understand the mechanisms of how plants are affected by these stresses and how they react to them, thus developing crop varieties with improved tolerance to abiotic stresses. Abiotic stresses affect plant cellular homeostasis via a variety of mechanisms (Jajic et al., 2015). They are often associated with the generation of Reactive Oxygen Species (ROS), which may result in significant damage to the cell structure (Banerjee et al., 2007). For example, cold causes a phase transition from lamellar to hexagonal II in the plasma membrane resulting from apposition of the plasma membrane and internal endomembrane. During cold stress, large endocytic vesicles were observed in the protoplast, and they resulted in protoplast bursting during thawing or rehydration (Dowgert et al., 1984; Uemura et al., 1995; Webb et al., 1994). Heat also causes damage to the plasma membrane by increasing its fluidity and leading to leakage (Allakhverdiev et al., 2008). In addition, salinity induces stomatal closure and inhibits leaf expansion in the short term. In the long term (days to weeks), salinity induces the build-up of ions in the shoot to toxic concentration. These changes cause premature senescence or even death of plants (Negrão et al., 2017).
Transgenic improvement of resistance to abiotic stresses in crops has been proven to be a very powerful method. For example, Monsanto’s GM drought-tolerant corn “DroughtGard” was developed by overexpressing a cold-shock protein CspB originated from *Bacillus subtilis* in maize (Eisenstein et al., 2013). But genetic modification by overexpression has limitations. Stress-tolerant lines are usually associated with growth arrests and detrimental pleiotropic effects. A specific transgenic line is usually tolerant of only one type of stresses. (Hussain et al., 2011). Thus, revealing mechanisms that can naturally enhance a wide range of stress tolerances would be ideal approaches in improving crop performance. The circadian clock possibly represents one such mechanism.

Various stresses, including drought, cold, salinity, and heat, have been implicated in the modulation of the circadian clock. Synchronization of the endogenous metabolic pathway with these stresses by circadian clock is crucial for plant fitness and adaptation. In addition, the importance of controlling the timing of hormone function is exemplified, because almost all the stresses-associated responses are under the control of plant hormones (Mizuno et al., 2008; Kazan et al., 2013). Consistently, many genes involved in the hormone syntheses and responses are also regulated by the circadian clock (Covington et al., 2008; Dodd et al., 2005; Michael et al., 2003).

It has been long known that the circadian clock influences plant cold responses through the C-Repeat Binding Factor 1 (CBF1), CBF2, and CBF3 transcription factors (Eriksson et al., 2011). A conserved response hub to plant cold stress is the induction of CBFs. These proteins bind to the Repeat/Dehydration Response Element (CRT/DRE) in the promoters of the *Cold Response (COR)* genes. The circadian clock regulates *CBFs* transcriptional level on a seasonal and daily basis. CCA1, TOC1, PRR5, and PRR7 bind to the promoters of some or all these *CBFs* (Eriksson et al.,
CCA1 positively regulates CBFs and promotes cold-tolerance while PRR5 and PRR7 negatively regulate CBFs and inhibit cold-tolerance (Nakamichi et al., 2016).

In Arabidopsis, most of the heat-inducible genes express during the day, and heat-repressed genes express at night (Grundy et al., 2015). Mutations in CCA1 and LHY lead to disrupted sensitivity to heat (Kant et al., 2008). Insoluble aggregates induced by heat stress are toxic to cells. Recent findings showed that circadian clock component ZTL targeted proteins for proteasomal degradation, which removes the heat-induced protein aggregates, thus enhancing heat-tolerance of plants (Gil et al., 2017).

The drought triggers multiple signal pathways and leads to protection against water loss. In Arabidopsis, drought responses include both ABA-dependent and ABA-independent pathways (Gil et al., 2017). ABA increases at the onset of drought stress (Achuo et al., 2006). Transcriptome analysis showed that many of the ABA-regulated genes are under the control of TOC1. Reciprocally, TOC1 expression is induced by ABA. TOC1 binds to the promoter of ABA-regulated gene, ABA-Related/H Subunit of the Magnesium-Proto-Porphyrin IX Chelatase/Genomes Uncoupled 5 (ABAR/CHLH/GUN5) and represses its expression. ABAR/CHLH/GUN5 induces TOC1 expression around midday (Grundy et al., 2015; Huang et al., 2012; Seo et al., 2012). A toc1 mutant is defective in stomata closure and displays altered drought resistance (Legnaioli et al., 2009). PRR7 and Time for Coffee (TIC) were also shown to be involved in the ABA-dependent and ABA-independent drought resistance pathways (Liu et al., 2013; Sanchez-Villarreal et al., 2013).

Plants response to salt stress largely overlaps with cold and drought stresses and depends on the ABA pathway (Achuo et al., 2006). CCA1, LHY, and TOC1 also regulate plant salt stress response through the CBF transcription factors (Eriksson et al., 2011; Kant et al., 2008). These
regulations rely on GI. Normally, GI interacts with Salt Overly Sensitive 2 (SOS2), which results in the inactivation of SOS1. SOS1 play an important role in sodium ion transportation. Under the high salt condition, GI is degraded by proteasomal degradation, which releases SOS2 and activates SOS1. This way, the sodium ions are then exported outside of the cell, and the salt stress is alleviated (Kim et al., 2013).

1.6 The Plant Circadian Clock Regulates Defense against Biotic Stresses

Plant circadian clock functions as a strategic planner in the immune system (Karapetyan et al., 2018). The circadian clock-regulated stomatal closure is among the earliest plant immune responses. Stomata are the entry holes for some pathogens, and consequently, plants are more resistant to pathogens when stomatal apertures are smaller (Korneli et al., 2014; Zhang et al., 2013). Right after the initiation of MTI, circadian components CCA1 and LHY regulate stomatal closure through Glycine-Rich RNA Binding Protein 7 (GRP7) (Zhang et al., 2013). GRP7 binds to mRNAs of FLS2 and EFR and regulates plant immune responses (Nicaise et al., 2013). A clock component, TIC, was also reported to be involved in stomatal defense (Korneli et al., 2014).

Furthermore, R-gene-mediated host resistance is also under the circadian clock control. The circadian clock is involved in the coordination of R-gene-mediated programmed cell death and basal defense. For example, the R protein Recognition of Peronospora parasitica 4 (RPP4) regulates genes involved in both the basal defense and programmed cell death. Clock component CCA1 activates RPP4 and facilitates the proper expression of the target defense genes involved in programmed cell death (Seo et al., 2015). The well-known clock-regulated R-gene-mediated host resistance is demonstrated by Wang et al. (Wang et al., 2011). The obligate biotrophic oomycete pathogen Hyaloperonospora arabidopsidis (Hpa) causes downy mildew disease on Arabidopsis leaves. Due to the presence of the R gene RPP4 (Wang et al., 2011), Arabidopsis Columbia (Col-0) accession is resistant to the avirulent Hpa isolate Em wa1. The rpp4 mutant is defective in R-
gene-mediated defenses. 22 defense genes were identified through transcriptome analysis and mutant screen. Significant enrichment of the ‘evening element’ (evening element is regulated both positively and negatively by circadian component CCA1) is found in the promoter regions of these genes. Additionally, 14 out of these 22 genes have rhythmic expression patterns. This circadian-regulated defense allows plants to anticipate pathogen infection time and promotes defense response at that time (Wang et al., 2011).

In addition, the biosynthesis and signaling pathway of plant defense hormones SA and JA are both under circadian clock regulation. SA and JA are key plant defense hormones against pathogens (Ingle et al., 2015; Smith et al., 2009). The expression of *Isochorismate Synthase 1* (*ICS1*), a key enzyme involved in the biosynthesis of SA, is under the control of clock component CCA1 Hiking Expedition (CHE) (Pruneda-Paz et al., 2009; Zheng et al., 2015). Both *ICS1* expression and SA concentration peak at midnight (Zheng et al., 2015). A circadian component, TIC, promotes the degradation of the basic-helix-loop-helix transcription factor MYC2 (Kazan et al., 2013). In *Arabidopsis*, MYC2 is a master regulator of many aspects of the JA signaling pathway, and its function is induced by pathogen effectors in roots and required for Induced Systemic Resistance (ISR) immune responses (Shin et al., 2012). MYC2 also connects the JA with signaling pathways of other phytohormones, like SA, Auxin, Abscisic Acid (ABA), and Gibberellins (GAs) (Erb et al., 2012).

Studies on the interaction of *Arabidopsis* and *P. syringae* showed that the host susceptibility is dependent on the time of infection. *Arabidopsis* is more susceptible to *P. syringae* during the light period than the dark period. Under constant light, spray infection of wild type *Arabidopsis* with *P. syringae pv. tomato* DC3000 showed that the greatest susceptibility happened when infected in the subjective morning. It was proposed that the susceptibility differences are due
to the clock-regulated SA pathway and stomata open-and-close status (Sharma et al., 2015; Zhang et al., 2013).

Plant circadian clock also gates defense responses indirectly through controlling ROS and Reactive Nitrogen Species (RNS) homeostasis. ROS and RNS play important roles in MTI to strengthen the cell wall and prevent pathogens from spreading to the other parts of plants (Karapetyan et al., 2008). The core circadian component CCA1 regulates the expression of Catalase 2 (CAT2) and controls ROS and RNS concentrations in plants (Lai et al., 2012).

Overall, plant circadian clock plays a central role in both plant growth and immune system, and it is able to balance the energy needs for vital activities of growth and defense (Karapetyan et al., 2018).

1.7 References


CHAPTER 2. COMPREHENSIVE MAPPING OF ABIOTIC STRESS INPUTS TO THE SOYBEAN CIRCADIAN CLOCK

Meina Li\textsuperscript{a,d,1}, Lijun Cao\textsuperscript{d,1}, Musoki Mwimba\textsuperscript{e}, Yan Zhou\textsuperscript{f}, Ling Li\textsuperscript{h}, Mian Zhou\textsuperscript{d,i}, Patrick S Schnable\textsuperscript{f}, Jamie A. O'Rourke\textsuperscript{f,g}, Xinnian Dong\textsuperscript{e,2}, Wei Wang\textsuperscript{b,c,d,2}

\textsuperscript{a}School of Life Sciences, Guangzhou University, Guangzhou, 510006, China; \\
\textsuperscript{b}School of Life Sciences, Peking University, Beijing, China; \\
\textsuperscript{c}Peking University - Tsinghua University Joint Center for Life Sciences, Beijing, China; \\
\textsuperscript{d}Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, 50011, USA; \\
\textsuperscript{e}Howard Hughes Medical Institute-Gordon and Betty Moore Foundation, Department of Biology, Duke University, Durham, North Carolina 27708, USA; \\
\textsuperscript{f}Department of Agronomy, Iowa State University, Ames, IA 50011, USA; \\
\textsuperscript{g}United States Department of Agriculture, Agricultural Research Service, Corn Insects and Crop Genetics Research Unit, Ames, IA 50011, USA; \\
\textsuperscript{h}Department of Biological Sciences, Mississippi State University, Starkville, MS 39762 \\
\textsuperscript{i}School of Life Sciences, Capital Normal University, Beijing, China. \\
\textsuperscript{1}M.L. and L.C. contributed equally to this work; \textsuperscript{2}To whom correspondence may be addressed.

E-mail: oneway1985@pku.edu.cn or xdong@duke.edu

Modified from a manuscript under review in Proceedings of the National Academy of Sciences of the United States of America

2.1 Abstract

Plants evolved the circadian clock to increase fitness by synchronizing physiological processes with environmental oscillations. Crop fitness was artificially selected through domestication and breeding, and the circadian clock was identified by both natural and artificial
selections as a key to improved fitness. Despite progress in Arabidopsis, our understanding of the crop circadian clock is still limited, impeding the rational improvement of crop circadian clock for enhanced fitness. To unveil the interactions between crop circadian clock and various environmental cues, we provide a comprehensive mapping of abiotic stress inputs to the soybean circadian clock through a two-module discovery pipeline. Using the Molecular Timetable method, we computationally surveyed publicly available abiotic stress-related soybean transcriptomes to identify stresses that have dramatic impacts on the global rhythm. These predictions were then experimentally confirmed using a multiplexed RNA sequencing technology. Specific clock components modulated by each stress were further identified. This comprehensive mapping uncovered novel inputs to the plant circadian clock, which included heat and alkaline stresses. Moreover, short-term iron deficiency had opposite impacts on the clocks of soybean and Arabidopsis by targeting different clock components. Comparisons between soybeans with different iron uptake efficiency suggest that phase modulation may be a potential mechanism to alleviate iron deficiency symptoms in soybean. These unique responses in soybean demonstrate the necessities to directly study crop circadian clocks, and our discovery pipeline may serve as a broadly applicable tool to facilitate these explorations.

2.2 Introduction

Plants, like many other living organisms, have an internal timekeeper called the circadian clock that synchronizes physiological processes to allow anticipation of recurring seasonal and daily environmental changes and coordinate metabolism accordingly. The current understanding of the molecular architecture of the plant circadian clock is primarily from studies performed in the model organism, Arabidopsis thaliana. In Arabidopsis, the clock system is composed of three parts: a central oscillator driven by interlocked transcriptional-translational feedback loops that generate the endogenous circadian rhythms; input pathways that integrate environmental cues to
the oscillator function; and output pathways that control diverse physiological processes, including growth, flowering time, and stress responses (Hsu and Harmer, 2014 and Greenham and McClung, 2015). While light is considered the dominant input to the plant circadian clock, cumulating studies have demonstrated that a wide spectrum of stress signals also have feedback regulation on the circadian clock in Arabidopsis (Greenham and McClung, 2015). These emerging discoveries suggest that the plant circadian clock may function as a central signaling hub integrating diverse stress signals to balance the energy needs for stress tolerance and growth.

In contrast to Arabidopsis circadian clock studies, little is known about the impacts of environmental stresses on the circadian clock in most crop plants despite of the fact that circadian clock genes have been repeatedly identified to associate with key agronomic traits (Bendix et al., 2015; Kim et al., 2012; Preuss et al., 2012). Artificial selection has even resulted in the deceleration of the circadian clock in cultivated tomatoes during domestication to allow the better adaptation to the long summer days they encountered as they were moved away from the equator (Muller et al., 2016). Our limited understanding of the environmental impacts on crop circadian clocks represents a key knowledge gap impeding further germplasm improvement through more targeted selection on specific circadian traits. To overcome these obstacles, we developed a reliable and cost-effective two-module discovery pipeline in soybean for comprehensive mapping of abiotic stress inputs to the circadian clock.

Soybean (Glycine max L. Merrill) is the primary source of the world’s supply of vegetable protein and oil. Soybean demand has grown rapidly due to its wide range of applications in food, feed, and industrial products. However, soybean production is threatened by a number of abiotic stresses such as drought (Rodrigues et al., 2015), soil salinity (Ge et al., 2010; Ge et al., 2011) and nutrient deficiency (O'Rourke et al., 2009; Severin et al., 2010). Whether these environmental
stresses perturb the soybean circadian clock and how they may interface with specific components of the central oscillator remain largely unknown, even though the clock has already been suggested to influence soybean yield. Modulation of central clock components through overexpression of a B-box protein involved in light signaling has led to soybean grain yield increases year after year in multiple transgenic events in multi-location field trials (Preuss et al., 2012). Natural variation of soybean ELF3, a key component of the central oscillator network, has been shown to improve soybean’s adaptation to tropics (Lu et al., 2017). Therefore, deciphering the interacting network between environmental stress signals and the soybean circadian clock may help enhance the abiotic stress tolerance of soybean.

Our two-module discovery pipeline consists of the Molecular Timetable module, a computational method and the RASL-seq (RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing) module, a large-scale targeted sequencing technology. The Molecular Timetable method is a genome-scale global rhythm analysis method, first developed in a mouse study (Ueda et al., 2004) and later adapted for analysis of Arabidopsis (Kerwin et al., 2011) and tomato (Higashi et al., 2016). This method relies on the average expression profiles of time-indicating genes to assess the global rhythm based on a single-time-point genome-wide expression profile rather than a time-course experiment. Therefore, it can be readily applied to existing genome-wide expression profiles. Using this computational method, we surveyed publicly available soybean transcriptomes related to abiotic stresses including 19 datasets with 306 microarrays and RNA-Seq samples (Belamkar et al., 2014; Carvalho et al., 2014; Duressa et al., 2010a; Duressa et al., 2010b; Ge et al., 2010; Ge et al., 2011; Ha et al., 2015; Kidokoro et al., 2015; Le et al., 2012; O’Rourke et al., 2009; Prince et al., 2015; Rodrigues et al., 2015; Severin et al., 2010; Tamang et al., 2014; Tripathi et al., 2015; Wei et al., 2015; Weston et al., 2011;
Whaley et al., 2015) and identified specific abiotic stresses that dramatically perturbed the global rhythm. To enable affordable experimental validation of these computational predictions, RASL-seq was adapted to soybean. RASL-seq is a probe-based targeted sequencing method that can accommodate profiling of hundreds of genes in hundreds of samples in a single HiSeq 2000/2500 sequencing lane (Li et al., 2012), making large-scale expression profiling cost-effective. By applying RASL-seq to over 300 time-course experiment samples with various abiotic stress treatments, we confirmed the computational predictions and further identified specific circadian clock components targeted by each abiotic stress. Through the integration of the Molecular Timetable method and the RASL-seq technology, we discovered that the soybean circadian clock responds to short-term iron deficiency in a significantly different manner from that in Arabidopsis. We also discovered, for the first time, that alkaline stress has a pervasive impact on the circadian clock components in plants.

2.3 Results

2.3.1 Identification of Time-Indicating Genes

Central to the Molecular Timetable method is the availability of the time-indicating genes derived from genome-wide time-course circadian expression profiles. To identify the time-indicating genes in soybean, the transcriptome of soybean cultivar Williams 82 was analyzed by RNA-Seq in samples subjected to constant light conditions during a two-day time-course experiment (Figure 2.1 A). Briefly, soybean seedlings were grown in soil under the diurnal condition (16 h Light/8 h Dark) for 9 days. At the end of the 9th day, they were exposed to constant light conditions, and unifoliolate leaves were collected starting at ZT0 (Zeitgeber Time) after 1-day acclimation in constant light. The samples were harvested every 4 hours for 2 days and subjected to 100 bp pair-end Illumina sequencing. The average clean reads were 57.6 million per sample with the highest reads being 65.3 million, providing a high-coverage expression profile.
The sampling was performed under the constant light condition to eliminate the effect of photoperiod and reveal the function of the endogenous circadian clock.

To select the time-indicating genes, the expression profiles of each gene were fit against 1,440 cosine curves (1 curve/min) to assess their rhythmicity. Coefficient of Variation (CV) of the time-course expression profile of each gene was used to select for oscillatory genes with high relative amplitude to ensure the robustness against noise (See Materials and Methods for details). In total, 3,695 genes were selected as time-indicating genes (Figure 2.1 B). Therefore, their phases reflect the time of the day. Based on the calculated phase, these 3,695 genes were further assigned to 24 phase groups with the peak expression time between CT0 (Circadian Time) and CT23. Hereafter, the 24 time-indicating groups will be termed CT groups.

To assess the performance of these time-indicating genes, we first applied the Molecular Timetable method to the time-course RNA-Seq dataset itself. Through nonlinear regression, we derived Phase24, a period-corrected estimation of circadian sampling time, for each RNA-Seq sample and fit it to the actual sampling time (Figure 2.1 C). The estimated linear relationship is statistically indistinguishable from the diagonal line, clearly showing that the Molecular Timetable method using these time-indicating genes can provide a good reflection of the actual sampling time. However, Phase24 does not simply equal the sampling time. Instead, it is a reflection of the actual body time of the sample, which is the endogenous time of the sample set by the circadian clock (Ueda et al., 2004). In fact, the Molecular Timetable method has captured the subtle body time changes that occurred during the two-day exposure to constant light conditions (Figure 2.1 D). While the global rhythm of samples collected at ZT0 and ZT24 appeared visually indistinguishable, asynchrony started to emerge over time (Figure 2.1 D). It became evident that the global rhythm of samples collected at ZT20 and ZT44 had significant differences.
To further validate the performance of these time-indicating genes, we tested them on a well-designed time-course profiling experiment related to mild drought stress (Figure 2.2 A and 2 B) performed by Rodrigues, et al. (Rodrigues et al., 2015). Although the experimental conditions, including cultivar, photoperiod, light intensity, temperature, and humidity, were quite different from our circadian time-course RNA-Seq experiment (Table S2.1), the estimated Phase24 still closely matched the sampling time of the control samples (Figure 2.2 B). Taken together, the time-indicating genes that we identified can be used reliably in the Molecular Timetable method for an accurate reflection of the global rhythm of each individual biological sample based on its genome-wide expression profile.

2.3.2 Application of Molecular Timetable Method

To identify specific abiotic stresses that may perturb the global rhythm in soybean, we first assembled the publicly available soybean transcriptomes related to abiotic stresses which included 19 datasets with 306 RNA-Seq and microarray samples generated from different platforms (Table S2.1). In order to enhance comparability, whenever available, the raw data were re-summarized and standardized. Then the expression matrices of the time-indicating genes were used, through nonlinear regression, to estimate Phase24, the key oscillation parameter.

Although the Molecular Timetable method can detect subtle changes due to the large degree of freedom used in the statistical test, here we focus on more dramatic perturbations. First, we investigated the datasets with soil-grown plants, the most similar growth conditions to our circadian time-course profiling experiment. Mild drought (without irrigation for 3 days) caused negligible changes to the global rhythm (Figure 2.2 A and 2 B) (Rodrigues et al., 2015). However, more severe drought (without irrigation for 6 days) induced a drastic phase shift (Figure 2.2 C) (Le et al., 2012). In nature, drought stress can be accompanied by heat stress. However, different from mild drought, even a 30-minute heat shock was strong enough to cause a phase advancement
of more than 8 h (Figure 2.2 D) (Weston et al., 2011). Surprisingly, prolonged heat stress had a rather milder perturbation of the global rhythm (Figure S2.1) (Kidokoro et al., 2015). This may be due to the temperature compensation capacity of soybean circadian clock after adaptation to prolonged heat stress. These results suggest that the responses of soybean global rhythm to stresses like drought and heat are duration-dependent.

Next, we applied the Molecular Timetable to datasets with hydroponically-grown plants in which both leaf and root transcriptomes have been obtained. Prolonged exposure to iron deficiency caused phase delay of leaf circadian clock in the two near-isogenic lines Clark (iron-efficient) and IsoClark (iron-inefficient) (Figure 2.3 A and B) (O'Rourke et al., 2009). The phase lagged in a similar manner in IsoClark compared with Clark. In roots, however, the phase changes were opposite in Clark and IsoClark under iron deficiency condition: it was advanced in Clark and delayed in IsoClark. These perturbations were largely recovered in IsoClark but not in Clark when iron was replenished (Figure 2.3 C and D) (Severin et al., 2010). These results imply that the responses to iron deficiency may be organ-specific, probably due to their differences in functions. Roots are responsible for iron assimilation while leaves are where most of the iron is stored and utilized. Besides the organ-specific responses in Clark, it is also important to note iron deficiency causes phase-delay in both the leaf and root in IsoClark. Considering the difference in iron utilization efficiency of these two near-isogenic lines, these differences suggest a strong connection between the phase change and iron utilization efficiency.

Besides iron deficiency, alkaline stress also causes organ-specific changes in the wild soybean, *Glycine soja* (Figure 2.3 E, F, and Figure S2.2) (Ge et al., 2010; Ge et al., 2011). Different from the iron deficiency datasets, the alkaline-treated samples were not accompanied by corresponding control samples at each time point. Therefore, a direct comparison was not possible
for the identification of potential changes in phase. To solve this problem, we performed a regression analysis and used model selection to look for phase changes. If no phase shift was induced by alkaline treatment, the estimated Phase24 in leaves should follow a linear relationship with the actual sampling time. However, F-test indicated that Phase24 follows a quadratic relationship with the sampling time ($p < 0.0001$), and the coefficient of the quadratic is significantly lower than zero ($p < 0.0001$) (Figure 2.3 E and F) (Ge et al., 2011). This means alkaline treatment causes statistically significant phase advances of the global rhythm in leaves. In roots, however, the quadratic curve fits no better than the linear relationship (Figure S2.2) (Ge et al., 2010). Therefore, alkaline does not cause a significant phase shift in roots.

Finally, we extended the Molecular Timetable method to all the abiotic stress-related soybean transcriptomes and summarized the corresponding Phase24. Except for a few experiments without appropriate controls from which we could not draw conclusions, most abiotic stresses perturb the phase to some extent. This suggests that the phase of soybean circadian clock is sensitive to environmental cues, probably allowing rapid response to the ever-changing environment by adjusting the gene expression accordingly.

**2.3.3 Validation of Global Rhythm Changes and Identification of Specific Targets**

The Molecular Timetable method has offered an efficient and sensitive way to detect phase perturbations triggered by diverse stress signals. However, one major issue is the interpretation of the observed stress-induced phase shift. Such a phase shift may reflect a true phenotype if the corresponding control and treated samples were taken at a similar time of a day. Yet, sampling in some published experiments did not take this into consideration. Obvious phase shifts were detected even among biological replicates in some datasets. Therefore, it is imperative to validate these computational predictions experimentally.
We performed two-day time-course circadian experiments with four stresses including mild drought, heat shock, short-term iron deficiency, and short-term alkaline stress. For soil-grown experiments, we sampled unifoliolate leaves to minimize the effects of developmental stages on transcription. For hydroponically-grown plants, we collected the first trifoliolate leaf and root tissues for the iron-deficiency experiment, and only the first trifoliolate leaf for alkaline stress. With three biological replicates and sharing common control samples, these experiments generated 315 samples (Table S2.2). Since we are only interested in examining the effects of these stresses on the circadian clock components and clock output genes, it is not cost-effective to perform genome-wide expression profiling on this large number of samples using either microarray or RNA-Seq. Instead, we adapted RASL-seq, a multiplexed targeted sequencing method (Wang et al., 2011).

Through bioinformatic analysis, we first identified 49 soybean orthologs of 16 key Arabidopsis circadian clock components. For each of these genes, three pairs of gene-specific primers were first tested during preliminary analysis, and the best pair was chosen for RASL-seq based on its efficiency (Li et al., 2012).

Due to the large number of soybean clock gene orthologs surveyed, we applied a very high stringency level on the robustness of the circadian oscillation ($p < 10^{-10}$) to focus on the genes with robust circadian oscillation and statistically significant changes caused by stress treatments (FDR < 0.05) (Figure 2.4 and Figure 2.5). Only the genes with phase shift over 4 hours and the period between 22-26 hours in control samples were shown in the main figures. Consistent with the prediction from the Molecular Timetable analysis, the short-term mild drought did not cause any statistically significant changes (Figure S2.3). Heat shock caused both phase advances and period changes of 1 FKF1 ortholog, 1 CHE ortholog and 1 PRR7 ortholog under high stringency level
(Figure 2.4 A and Figure 2.5 A). The phases of three PRR genes were advanced more than 5 hours including 1 PRR5 ortholog, 1 PRR7 ortholog, and 1 PRR9 ortholog after heat shock treatment. These observed phase changes are in agreement with what has been observed through the Molecular Timetable analysis (Figure 2.2 D).

Short-term iron deficiency mainly caused phase changes in leaves rather than roots in the iron-inefficient line, IsoClark. It resulted in phase advances of evening genes including 8 orthologs of 6 key evening components (Figure 2.4 B). Although short-term iron deficiency only caused phase advance of 1 CCA1/LHY ortholog in roots; but under the low stringency condition, 4 morning-phased genes were altered in period (Figure S2.3). Two CCA1/LHY orthologs and 1 RVE6 ortholog had periods less than 18 hours under the control condition in the iron-inefficient IsoClark. This suggested that iron deficiency shortened the circadian period in soybean and that morning-phased genes may play a role in the regulation of iron assimilation in roots. This organ-specific response was unlikely due to the less synchronized circadian clock in the root, since only 10 out of the 49 clock genes did not have robust circadian oscillation in roots while 6 failed the oscillation test in leaves under the low stringency (Figure S2.3). Therefore, consistent with the implication from the Molecular Timetable analysis, soybean circadian clock’s response to iron deficiency is organ-specific. Besides organ specificity, different treatment durations also elicited different circadian clock responses. Compared with long-term iron deficiency (Figure 2.3 A to D, IsoClark), short-term treatment induced more perturbations on the leaf circadian clock than on the root clock. Furthermore, the type of changes was also the opposite, with long-term iron deficiency causing phase delay of the leaf circadian clock while short-term stress led to phase advance.

Surprisingly, short-term alkaline stress has a profound impact on the soybean circadian clock in leaves when grown hydroponically. Alkaline stress triggered a phase shift in 8 out of 16
different key clock components under the high stringency condition (Figure 2.4 C), and 25 out of all the 55 soybean clock gene and output gene orthologs tested under the low stringency condition (Figure S2.3). Interestingly, consistent with the Molecular Timetable analysis, all these phases were advanced, suggesting a coherent perturbation to the whole circadian clock network. Despite this prevalent influence on phase, alkaline stress also disturbed period, lengthening the period of 2 morning-phased genes including 1 CCA1/LHY ortholog and 1 RVE6 ortholog, 1 output gene ortholog CAT2 under high stringency condition and shortening the period of 1 LUX ortholog (Figure 2.5 B).

To enable a more holistic view of the whole RASL-seq analysis results, we applied a lower stringency condition (oscillation $p < 0.05$, FDR < 0.05) and summarized all the phase and period changes together in a heatmap (Figure S2.3). The four stresses except for drought which did not cause any changes on the circadian phase, all produced phase advances of different genes with different degrees. The period was not altered as dramatically as the phase. Among the four stresses, heat shock has the most prevalent impact on period. It caused arrhythmicity of more than half of the soybean circadian clock orthologs tested, suggesting that soybean circadian clock is highly sensitive to heat shock. Different orthologs of the same clock components had largely similar responses to the same stress. For example, the short-term alkaline stress and the iron deficiency caused the phase advance of two TOC1 orthologs, Glyma04g33110 and Glyma06g21120. Heat shock and short-term alkaline stress lengthened the period of two ELF3 orthologs, Glyma07g01601 and Glyma08g21115. Nevertheless, different CCA1/LHY orthologs do respond to iron deficiency in roots differently in terms of their period. This summarized heatmap also allowed us to identify more functional orthologs (Figure S2.3). Among the six CCA1/LHY
orthologs, five of them specifically respond to short-term alkaline stress including the two with lower sequence similarities.

2.3.4 Assessment of the Effects of Abiotic Stresses on Circadian Leaf Movement

To test whether the stress-induced transcriptional changes of the soybean clock genes can lead to changes in other clock-controlled physiological outputs, we chose to measure circadian leaf movement as it has been widely used as a proxy to study plant circadian clock. We performed the two-day time-course experiments with alkaline stress, mild drought, heat shock, and short-term iron deficiency using the imaging setup, specific soybean cultivar or species and plant growth conditions provided in Materials and Methods and in Table S2.2. The circadian parameters including Phase24 and period were derived through nonlinear regression. We found that the wild soybean, *Glycine soja*, showed very robust circadian leaf movement and alkaline stress caused a phase advancement (Figure 2.6) which is consistent with the results from the Molecular Timetable analysis (Figure 2.3 F) and RASL-seq analysis (Figure 2.4 C). Interestingly, while mild drought caused negligible perturbations to the clock gene expression according to the Molecular Timetable analysis (Figure 2.2 B) and RASL-seq analysis (Figure S2.3), it resulted in significant changes of the Phase24 and period of leaf movement (Figure S2.4 A). On the contrary, heat shock and short-term iron deficiency did not cause significant changes in leaf movement (Figure S2.4 B and C) despite their impacts on the clock gene expression (Figure S2.3). These data are consistent with facts that different plant physiological processes are controlled by distinct subsets of clock components (Greenham and McClung, 2015). Perturbation in individual or subsets of clock components may not necessarily jeopardize every clock behavior (Matsuzaki et al., 2015).

2.4 Discussion

In developing our two-module discovery pipeline to assess rhythmicity changes, sensitivity and accurateness are prerequisites. To this end, we have tested different algorithms with different
design principles, assumptions, and complexity. An online rhythmicity analysis system called BioDare (Zielinski et al., 2014) was developed with 6 period analysis algorithms including Fast Fourier Transform-Non-Linear Least Squares (FFT-NLLS) (Plautz et al., 1997), mFourfit (Edwards et al., 2010), Maximum Entropy Spectral Analysis (MESA) (Burg, 1972), Enright Periodogram (EPR) (Enright, 1965, and Sokolove and Bushell, 1978), Lomb-Scargle Periodogram (LS) (Lomb, 1976) and Spectrum resampling (SR) (Costa et al., 2013). In this study, we also devised a nonlinear regression strategy based on a Cosine Function Combined with a Linear Trend (COS) (See Materials and Methods for details). To compare the performance of these 7 methods, the time-course RNA-Seq data of 55 clock candidate genes were used as the test dataset (Dataset S8). The smallest variation of estimated periods was achieved by LS, followed by COS and then FFT-NLLS (Figure S2.5 A). However, the LS algorithm has a poor false negative rate. The expression profiles of genes (e.g. Glyma03g42221 and Glyma04g33110) which obviously showed excellent oscillation were considered as arrhythmic by LS (Figure S2.5 B). FFT-NLLS also suffered from the same problem with examples like Glyma03g42221, Glyma08g05130, and Glyma11g14490 (Figure S2.5 B). In contrast, the COS method developed in our study achieved a sound assessment of rhythmicity without high false negative rate. Therefore, COS was used in our analysis pipeline to examine the influences that abiotic stresses have on the soybean circadian clock.

Our previous knowledge of the abiotic stress effects on the soybean circadian clock is limited. Marcolino-Gomes et al. observed amplitude perturbation by mild and severe drought stresses on representative clock genes in a drought-sensitive soybean cultivar under the diurnal condition (Marcolino-Gomes et al., 2014). However, these observed changes were confounded by both the endogenous circadian clock and the exogenous periodic light condition changes, making
it hard to distinguish which one was the major contributor to this phenotype. Our RASL-seq analysis suggests that mild drought had a negligible effect on the phase and period of the soybean circadian clock gene expression (Figure S2.3). Therefore, it is likely that the effects recorded in Marcolino-Gomes, et al. were mainly due to the exogenous periodic light changes instead of the endogenous circadian clock (Marcolino-Gomes et al., 2014).

The effect of iron deficiency on the circadian clock was investigated extensively in *Arabidopsis* (Chen et al., 2013; Hong et al., 2013; Salome et al., 2013; Wagner et al., 2017). Short-term iron deficiency causes period lengthening under circadian conditions (Chen et al., 2013) and phase delay under diurnal conditions (Salome et al., 2013). Despite the molecular details uncovered, the biological significance of these phenotypes remains unclear (Salome et al., 2013). The observed soybean circadian clock responses in our analysis have provided some clues. As a key element for photosynthesis and redox regulation, iron homeostasis is tightly controlled through the competition between the iron uptake and iron storage pathways. Many genes involved in iron homeostasis are regulated by the circadian clock. It has been suggested that the iron uptake pathway is induced during the day to facilitate photosynthetic processes (Darbani et al., 2013). To maintain homeostasis, and avoid iron toxicity, excess iron is stored in ferritin proteins. At night, iron uptake is repressed and stored iron is released (Darbani et al., 2013). RASL-seq analysis suggested that short-term iron deficiency mainly targeted evening-phased circadian components for phase advances (Figure 2.4 B). Long-term iron deficiency caused phase delay in both the iron-efficient wild-type cultivar, Clark, and the inefficient Clark mutant, IsoClark in leaves (Figure 2.3 C). The phase delay may prolong the iron uptake period to compensate for the iron deficient conditions. Although utilizing the same strategy as Clark, IsoClark displays lengthened periods including 1 *LUX* ortholog in leaves and 3 morning-phased clock gene orthologs in roots under
short-term iron deficiency probably due to its iron inefficient feature (Figure S2.3). This is similar to the Arabidopsis mutants defective in iron homeostasis which also have longer periods even when grown under iron sufficient conditions (Hong et al., 2013). Future experiments comparing Clark and IsoClark will provide more mechanistic insights into the biological function of the circadian clock in the regulations of the short-term and long-term iron deficiency.

Alkaline stress is also one of the major abiotic stresses limiting crop productivity. However, to our knowledge, its impact on the circadian clock has never been explored in any plant species. Yet this connection was captured by our computational analysis (Figure 2.3 E and F) and its pervasive modulations on the phase of clock components have been experimentally revealed (Figure 2.4 C, 2.5 B, 2.6, and Figure S2.3). Interestingly, salinity stress which usually co-exists with alkaline stress also causes phenomenal perturbations in phase (Figure S2.6). Since alkaline stress reduces the solubility of many minerals, including iron (Darbani et al., 2013), its pervasive impacts on the phase echo the effect of iron deficiency. Whether alkaline stress and iron deficiency share similar signaling components to modulate the circadian clock and the biological significance of these modulations remain to be determined in future studies.

One interesting general observation made in our study is that only a subset of soybean circadian clock components was perturbed by different abiotic stresses (Figure S2.3). This is consistent with results from other studies. In Arabidopsis, comparisons of the genome-wide expression profiles of samples under the thermo-cycle condition and constant conditions also found that the phases of only a subset of clock-related genes were shifted by temperature (Michael et al., 2008). In rice, while temperature and solar radiation had strong impacts on the expression of certain clock-related genes, the internal time of the plants was not affected by weather, photocycle or developmental age in the field (Matsuzaki et al., 2015). These observations raised
one interesting question: why perturbations on a subset of clock-related genes do not transmit to other clock genes or even the circadian clock system as a whole. One possible explanation is that the interlocked multi-loop structure and nonlinearity of the circadian clock system have enhanced the robustness against the endogenous and exogenous perturbations on a limited number of clock components (Saithong et al., 2010). Based on our bioinformatic analysis, the soybean genome has a lot more copies of clock-related genes than *Arabidopsis*. It is likely that the regulatory network of soybean circadian clock system may have higher complexity and is potentially more robust than *Arabidopsis* circadian clock. In-depth mechanistic and systems studies are required for a better understanding of soybean circadian clock.

Similar levels of complexity were observed in the effects of abiotic stresses on the clock outputs. Using leaf movement as a proxy for the clock behavior, we found a phase advancement in leaf movement that is consistent with the gene expression analysis in response to alkaline stress, but not with mild drought, heat shock or short-term iron deficiency. We hypothesize that the differences between the rhythmic leaf movement and clock gene expression in response to abiotic stresses are probably due to the distinct molecular mechanisms underlying the rhythmic leaf movement of *Arabidopsis* and soybean. Even though molecular details responsible for *Arabidopsis* leaf movement have not been fully resolved, circadian leaf movement is conferred by the differential growth of the adaxial and abaxial sides of the petiole (Polko et al., 2012; Rauf et al., 2013). However, leguminous plants like soybean have a specialized leaf motor organ, pulvinus, to regulate reversible changes in leaf position. Opposing volume changes resulting from ion and water fluxes in opposite parts of the pulvinus empower the reversible leaf movement without dependency on the growth (Moran, 2007). Isolated bean pulvinus protoplasts could maintain its rhythmic swelling under continuous light for more than 8 days (Mayer and Fischer, 1994).
Therefore, while the circadian regulation on growth is likely to be the main contributor to the rhythmic leaf movement of *Arabidopsis*, the circadian regulation on pumps, carriers, ion and water channels in pulvinus is probably the key driving force for that of soybean (Mayer and Fischer, 1994 and Moran, 2007). Alkaline stress is likely to affect the ion flux through its direct impact on pH and mild drought may significantly influence the water flux, thus perturbing the rhythmic swelling of pulvini. According to the current literature, it is unclear whether heat shock or short-term iron deficiency may affect the swelling of pulvini. It is possible that the clock components involved in response to heat shock and short-term iron deficiency may not regulate pulvinus swelling. An alternative explanation for the differences in circadian clock gene expression changes and rhythmic leaf movement changes may be attributed to tissue-specificity of the circadian clock (Endo, 2016). Since different tissues were studied for the leaf movement assay (pulvin of first trifoliolate) and RASL-seq analysis (blades of unifoliolate or trifoliolate), it is possible that the circadian clock in these tissues was not tightly coupled.

Taken together, our discovery pipeline has helped reveal novel insights into the environment-circadian clock interplay in soybean that cannot be extrapolated from what has been learned in *Arabidopsis*. The comprehensive high-resolution datasets generated are also valuable assets for quantitative modeling of the crop circadian clock system, which is paramount to the systemic understanding of this complex network. Importantly, this discovery pipeline is not limited to application in soybean. It can be readily applied to any species with genome information and publicly available transcriptome datasets.

### 2.5 Materials and Methods

Plant materials, growth conditions, and treatments. Seeds of soybean Williams 82, Clark, IsoClark and *Glycine soja* were used in this study. For the RNA-Seq analysis, seedlings of Williams 82 were grown in soil (Sun Gro Horticulture, Vancouver, BC, Canada) under long-day
conditions (16 h light/8 h dark) for 9 days in growth chamber with the following conditions: 28 °C, 50 % relative humidity and the light intensity of 100 μmol/m²/s. On the 10th day, the light was switched to constant conditions. Seedlings were sampled every 4 hours for 44 hours starting at ZT0 on the 11th day. For each time point, three biological replicates were obtained, and each sample was pooled from the unifoliolate leaves of three individual seedlings. The plant materials, growth conditions, and treatments related to the time course experiment for RASL-seq are summarized in Table S2.2.

RNA-Seq and data analysis. Total RNA was extracted using RNeasy Plant Mini Kits (Qiagen). The RNA quality was checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). The transcriptome sequencing was performed using the Illumina HiSeq platform (Illumina, Inc., San Diego, CA, USA) and conducted at BGI Americas (bgi-international.com/us). The data processing is performed mainly according to Li et al. (Li et al., 2015). Briefly, for each sample, RNA-Seq data was collected from three lanes, and the raw reads were trimmed and cleaned separately. The cleaned reads are combined and aligned to soybean reference genome Glycine_max_Wm2.a1.v1.0 using the spliced alignment identification tool TopHat (V2.1.0). Glycine_max_Wm2.a1.v1.0 was used for better comparison since most of the transcriptome data used for the Molecular Timetable analysis were originally mapped to this version of the reference genome. The aligned reads were counted using HTseq-count (Version: 0.6.0). Only genes with non-zero FPKM values in all 36 samples were included for further analysis. The RNA-Seq data generated in this study were deposited in NCBI’s Gene Expression Omnibus (GSE94228).

Identification of time-indicating genes. To identify time-indicating genes that have robust circadian oscillation with high relative amplitude, the FPKM values of each gene were analyzed
through two filters. First, the expression profile of each gene was fit against 1,440 test cosine curves. These curves all have unit amplitude and a 24-hour period, but they had different phases (0-24 h) measured at 1 min increments. The correlation coefficient (R) and the phase of the best-fitting curve were assigned to each gene. The genes with $R \geq 0.9$ were selected for further filtering.

Second, the Coefficient of Variation (CV) was used as a proxy for relative amplitude and calculated for each gene using FPKM values. A total of 3,695 genes with $CV \geq 0.3$ and $R \geq 0.9$ were selected as time-indicating genes. The phases of these timing-indicating genes were rounded to an integer ranging from 0 to 23. The time-indicating genes with the same rounded phase were considered as one Circadian Time (CT) group, resulting in 24 CT groups.

Assembly of soybean abiotic stress-related transcriptomes and normalization. Abiotic stress-related soybean transcriptome datasets were assembled from diverse online sources including NCBI’s Gene Expression Omnibus, EMBL-EBI’s ArrayExpress, PLEXdb. Whenever available, raw data from single-color microarray datasets were re-normalized using RMA algorithm with baselines of each array adjusted to the median. The two channels from two-color microarray datasets were individually normalized using the quantile normalization method and median adjusted. RPKM and FPKM values were generated and used for single-end and pair-end RNA-Seq datasets, respectively. Whenever the raw datasets are not available in the submitted files, processed expression matrices supplied by the original authors were directly used. The original experimental designs of these transcriptomes were listed in Table S2.1.

Application of Molecular Timetable method. The expression profiles of time-indicating genes were extracted from the re-normalized transcriptome datasets. Standardization was first applied to each time-indicating gene across different samples within one experiment so that each time-indicating gene has the same average expression level and standard deviation.
Standardization was then applied to each sample within one experiment so that each sample has the same average expression level and standard deviation. Three sufficient statistics including the number of genes, the mean and standard deviation of standardized expression levels of time-indicating genes were calculated for each of the 24 CT groups and used for the estimation of oscillation parameters for each sample. To estimate the circadian rhythm parameters, the three sufficient statistics for the 24 CT groups of each sample were fit to the following equation in GraphPad Prism 6:

\[
Expression \ level = Amplitude \times \cos \left( \frac{2\pi}{Period} \times CT \ group - \frac{2\pi}{24} \times Phase24 \right) + aX + b
\]

The period was constrained to be more than 12 h but less than 36 h. Samples with non-converging fitting were considered as arrhythmic. The resulting best-fitted value, standard error, and degree of freedom for each parameter were used for statistical tests. To improve comparability and discount the confounding effects of period differences on the comparisons of phases (peak expression time of a day), the phase was scaled to a 24 h period, yielding Phase24.

Identification of soybean circadian clock and output genes. To identify orthologs of the *Arabidopsis* circadian clock and output genes in the soybean genome, the amino acid sequences (The *Arabidopsis* Information Resource; http://www.arabidopsis.org) of the corresponding *Arabidopsis* proteins were used as queries in BLAST searches (tBLASTn) against the *G. max* genome v1.0 using the Phytozome database (http://phytozome.net/soybean). The most similar representative sequences were selected on the basis of whether they had alignment e-values equal to 0 or using the cutoff \( E \) value \( \leq 10^{-12} \).

RASL-seq. Total RNA was extracted using Trizol reagent (Invitrogen, CA, USA) according to the manufacturer’s protocol. The RASL-seq procedure was performed as described (Li et al., 2012). The primers are listed in Dataset S2.3. The sequencing data was de-multiplexed
and aligned with the gene-specific primers. The resulting count matrix was first blanked using the reads from the average reads of blank wells and then standardized. Nonlinear regression was applied to derive Phase24 and period. The best fit value, standard error, degree of freedom of Phase24, and period were used for statistical tests with multiple comparison correction to obtain False Discovery Rate (FDR). The predicted values from nonlinear regression were fit against the standardized expression profile of each gene in each experimental condition through linear regression. The corresponding $p$ values from associated ANOVA tests on these linear regressions were used as measurements of the circadian oscillation robustness.

Leaf movement assay. Leaf movement was measured as described previously with minor modifications (Kim et al., 2008). Images under constant light were recorded hourly for 2 days using 4K Ultra HD Camera (GoPro Hero 5). Leaf movement was assessed by measuring the distance from tip to tip of the symmetrical leaf pair of first trifoliolate using Fiji/ImageJ (NIH). Data normalized by Z score transformation were used in the analysis of circadian rhythm. Results represent mean ± standard error. Four biological replicates were obtained. The plant materials, growth conditions, and treatments were used the same as the RASL-seq experiments.

2.6 Acknowledgements

We thank G. Tylka and S. Whitham for sharing soybean seeds and P. Benfey for enlightening discussions on the project. We thank BGI for processed sequencing data. This work was supported by the startup funds from Peking University and Peking University-Tsinghua University Joint Center for Life Sciences and USDA National Institute of Food and Agriculture, Hatch project 3808 to W.W., and grants from NIH (R35GM118036-04) and Howard Hughes Medical Institute-Gordon and Betty Moore Foundation (through Grant GBMF3032) to X.D., and
from the United States Department of Agriculture Agricultural Research Service (USDA-ARS) project 3625-21220-005-00D to J.O.

2.7 Author Contributions

W.W, M.L, and L.C designed the research project. M.L and L.C performed the experiments. W.W, M.L, and L.C analyzed the data and drafted the manuscript.

2.8 References


2.9 Figures

Figure 2.1. Identification of time-indicating genes and validation of Molecular Timetable method in soybean. (A) Sampling scheme for circadian time-course RNA-Seq of soybean unifoliolate. The digits before “ZT” mark the sampling time-points. White box, day. Black box, night. Grey box, subjective night under the constant light condition. ZT, Zeitgeber Time. (B) Heatmap of standardized expression levels of 3,695 time-indicating genes. Three biological replicates were shown as adjacent columns within each sampling time. Genes were sorted based on their peak expression time and organized as rows. (C) Molecular Timetable method can estimate actual sampling time with high precision. Estimated Phase24 was plotted against actual sampling time. Linear regression was performed and y = x line was plotted as a reference. (D) Normalized expression levels of time-indicating genes binned into 24 Circadian Time (CT) group based on their phases. Samples from the same time of a day from day 1 and day 2 were plotted as blue and red respectively. Mean and standard error were plotted.
Figure 2.2. The drought degree correlates with perturbations on global rhythm while heat shock has a dramatic impact on soil-grown soybean seedlings. (A) Normalized expression levels of time-indicating genes in control and mild drought treated samples. Treatments were performed at ZT0 and samples were harvested at the indicated time. Mean and standard error were plotted. ZT, Zeitgeber Time. CT, Circadian Time. (B) Estimated Phase24 has a good linear relationship with reported sampling time. Linear regression was performed using estimated Phase24 and reported sampling time. Adjusted $R^2 = 0.9999$ for both control and drought-treated samples. $y = x$ was plotted as a reference. (C) Severe drought stress perturbs global rhythm and causes dramatic phase shift. (D) Heat shock stress (30 min heat treatment at 42 °C) perturbed global rhythm dramatically. Mean and standard error were plotted. ****, $p < 0.0001$ (Student’s t-test). (A) and (B) were derived from GSE69469. (C) was derived from GSE40627. (D) was derived from GSE26198.
Figure 2.3. Circadian phases are altered in hydroponically-grown soybean seedlings under long-term iron deficiency and alkaline stress. (A, B) Long-term iron deficiency causes different rhythm changes in soybean leaves from cultivars with different iron utilization efficiency. Normalized expression levels of time-indicating genes were plotted in (A). Estimated Phase24 was shown in (B). Clark, iron-efficient. IsoClark, iron-inefficient. Fe sufficient, seedlings grown with 100 μM Fe(NO₃)₃. Fe limited, seedlings grown with 50 μM Fe(NO₃)₃. ****, p < 0.0001 (Student’s t-test with Holm-Sidak multiple comparison correction). (C, D) Long-term iron deficiency causes different rhythm changes in soybean roots from cultivars with different iron utilization efficiency. Normalized expression levels of time-indicating genes were plotted in (C). Estimated Phase24 was shown in (D). DAP, days after planting. Fe sufficient, seedlings grown with 100 μM Fe(NO₃)₃. Fe limited, seedlings grown with 50 μM Fe(NO₃)₃. Fe sustained, seedlings grown with sufficient iron for 14 days. Fe recovered, seedlings grown with limited iron for 12 days and then with sufficient iron for another 2 days. ****, p < 0.0001 (Student’s t-test with Holm-Sidak multiple comparison correction). (E, F) Alkaline stress causes phase advance in leaves. Normalized expression levels of time-indicating genes were shown in (E). h, hour. Estimated Phase24 were plotted against reported sampling time in (F). A quadratic curve (blue solid line) fits the data better than a straight line (black dash line), p < 0.0001 (Exact F-test). The coefficient for the quadratic is significantly smaller than 0, suggesting phase advancement induced by alkaline stress. p < 0.0001 (Exact F-test). Mean and standard error were plotted. CT, Circadian Time. (A) and (B) were derived from GSE10730. (C) and (D) were derived from GSE22227. (E) and (F) were derived from GSE20323.
Figure 2.4. Abiotic stresses induced phase shift of the global circadian rhythm. Heat shock (A), short-term iron deficiency (B) and short-term alkaline stress (C) induced phase advance of specific circadian components in soybean leaves. The angular coordinates represent Phase24. The $-\log_{10}$ transformed oscillation $p$ values represent the robustness of the oscillation and were plotted as radial coordinates. The circle represents control samples and the triangle represents treated samples. The size of the symbols is proportional to the Standard Error of Mean (SEM) of Phase24 as illustrated in the legend. The arrows from control to treated samples represent the direction of phase shift. Student’s $t$-test with Benjamini and Hochberg multiple comparison correction was used to compare control and treated samples and derive False Discovery Rate (FDR). To highlight statistically significant changes and apply a high stringency level on oscillation robustness, only genes with oscillation $p < 10^{-10}$ and FDR < 0.05 were plotted.
Figure 2.5. Abiotic stresses induced period changes to the global circadian rhythm. Heat shock (A) and short-term alkaline stress (B) changes to period of specific circadian components in soybean leaves. Student's *t*-test with Benjamini and Hochberg multiple comparison correction was used to compare control and treated samples and derive False Discovery Rate (FDR). To highlight statistically significant changes and apply a high stringency level on oscillation robustness, only genes with oscillation $p < 10^{-10}$ and FDR < 0.05 were plotted. Mean and standard error were plotted.
Figure 2.6. The wild soybean has robust circadian leaf movement and alkaline stress changes the global rhythm dramatically. (A) Leaf movement of soybean under alkaline stress. The 7-day-old seedlings of *Glycine soja* were transferred from germination paper to hydroponic unit system and grown for 10 days under the ambient light condition in the greenhouse. The seedlings were kept in constant light on the 9th day. On the 10th day, half of the seedlings were treated with 50 mmol/L NaHCO$_3$ (pH 8.5) at ZT0 and then the movement of the first unifoliolate leaf was recorded hourly. Leaf movement data represent mean ± standard error. (B) Alkaline stress-induced phase advance of circadian leaf movement. (C) The period of circadian leaf movement was lengthened after alkaline stress. (D) Alkaline stress increased the relative amplitude of circadian leaf movement. The circadian rhythm parameters including Phase24, period and relative amplitude were derived through nonlinear regression. The white and gray regions in the trace plot indicate subjective light and dark periods, respectively. **, *p* < 0.01. ****, *p* < 0.0001 (Student's *t*-test).
Figure S2.1. Prolonged heat stress (5 hours’ treatment) causes mild phase delay. (A) Normalized expression levels of time-indicating genes. CT, Circadian Time. (B) Prolonged heat stress causes a mild but significant phase delay. ****, *p* < 0.0001 (Student's *t*-test). Mean and standard error were plotted. (A) and (B) were derived from E-MTAB-2852.
Figure S2.2. Short-term alkaline stress does not cause phase advance in roots. (A) Normalized expression levels of time-indicating genes. CT, Circadian Time. h, hour. (B) Estimated Phase24 were plotted against reported sampling time. A quadratic curve (black dash line) does not fit the data better than a straight line (blue solid line), $p > 0.05$ (Exact F-test). Mean and standard error were plotted. (A) and (B) were derived from GSE17883.
Figure S2.3. Summary of soybean circadian clock responses to abiotic stresses under low statistical stringency level. Student's *t*-test with Benjamini and Hochberg multiple comparison correction was used to compare control and treated samples and derive False Discovery Rate (FDR). When control value is more than treated value, \(-\log_{10}(\text{FDR})\) was plotted (yellow). When control value is less than treated value, \(-\log_{10}(\text{FDR})\) was plotted (blue). To highlight only the statistically significant changes, FDRs that are more than 0.05 were masked and represented as light grey. Genes with insignificant oscillation \((p > 0.05)\) in either control or treated samples were masked and represented as grey. The total number of genes with statistically significant changes were summarized and listed for each stress. For each clock gene group, the locus names were arranged in the descending order based on their sequence similarity to their orthologs in *Arabidopsis*.
Figure S2.4. Assay of soybean circadian leaf movement in response to abiotic stresses under constant light conditions. (A) Leaf movement of soybean under mild drought stress. (B) Leaf movement of soybean in response to heat shock. (A) and (B) Seedlings of Williams 82 were grown in soil under 16 h L/8 h D for 14 days, released into constant light for 24 hours and then the movement of the first unifoliolate leaf was recorded hourly. (C) Leaf movement of soybean under short-term iron deficiency. The 7-day-old IsoClark seedlings were transferred from germination paper to hydroponic unit system and grown for 10 days in iron sufficient condition (100 µM Fe(NO₃)₃) under the ambient light condition in the greenhouse. The seedlings were kept in constant light starting at the 9th day. On the 10th day, the treated seedlings were switched to iron limited condition (50 µM Fe(NO₃)₃) at ZT0 and then the movement of the first unifoliolate leaf was recorded hourly. The circadian rhythm parameters including Phase24, period and relative amplitude were derived through nonlinear regression. The white and gray regions in the trace plot indicate subjective light and dark periods, respectively. *, p < 0.05. **, p < 0.01 (Student's t-test).
Figure S2.5. Comparison of methods for period analysis. (A) Period distributions of soybean orthologs of circadian clock components analyzed with different methods. The RNA-Seq result from this study was used for period analysis. The COS is the method used in this study to derive circadian parameters including Phase24, period and relative amplitude. The other six methods are available in BioDare 2 website (https://sourceforge.net/projects/biodare/). FFT-NLLS, Fast Fourier Transform-Non-Linear Least Squares. MESA, Maximum Entropy Spectral Analysis. EPR, Enright Periodogram. LS, Lomb-Scargle Periodogram. SR, Spectrum Resampling. (B) Four representative genes from analysis in (A). The period of circadian genes with obvious circadian rhythm was accurately estimated by the COS method but considered as arrhythmic by the LS Periodogram or FFT-NLLS method.
Figure S2.6. Salinity stress causes dramatic phase shift. (A) Normalized expression levels of time-indicating genes. CT, Circadian Time. (B) Salinity stress causes dramatic phase shift. ****, $p < 0.0001$ (Student's $t$-test). Mean and standard error were plotted. (A) and (B) were derived from GSE411.
Table S2.1. Experimental conditions of soybean transcriptome datasets under abiotic stress treatments.

<table>
<thead>
<tr>
<th>ID</th>
<th>Stress Condition</th>
<th>Genotype</th>
<th>Photoperiod</th>
<th>Light Intensity</th>
<th>Temperature</th>
<th>Humidity</th>
<th>Age</th>
<th>Tissue</th>
<th>Treatment</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Drought</td>
<td>Williams 82</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>25 °C</td>
<td>50 %</td>
<td>30 d</td>
<td>Leaf and root</td>
<td>0/0.5/1/2/3/5 h</td>
<td>GSE49537</td>
</tr>
<tr>
<td>2</td>
<td>Drought</td>
<td>BR16</td>
<td>14 h L/10 h D</td>
<td>500 μmol/m²/s</td>
<td>28 °C</td>
<td>80 %</td>
<td>V1</td>
<td>Leaf</td>
<td>0/4/8/12/16/20 h</td>
<td>GSE69469</td>
</tr>
<tr>
<td>3</td>
<td>Drought</td>
<td>Williams 82</td>
<td>12 h L/12 h D</td>
<td>80 μmol/m²/s</td>
<td>30 °C</td>
<td>50 %</td>
<td>R2</td>
<td>Trifoliate leaves</td>
<td>6 d</td>
<td>GSE40627</td>
</tr>
<tr>
<td>4</td>
<td>Drought</td>
<td>Williams 82/DT2008</td>
<td>10 h L/12 h D</td>
<td>150 μmol/m²/s</td>
<td>30 °C</td>
<td>50 %</td>
<td>14 d</td>
<td>Root</td>
<td>0/2/10 h</td>
<td>GSE65553</td>
</tr>
<tr>
<td>5</td>
<td>Drought</td>
<td>Conquesta transgenic lines</td>
<td>Ambient</td>
<td>NA</td>
<td>15-35 °C</td>
<td>65-85 %</td>
<td>V3/ V4</td>
<td>Third trifoliate leaves</td>
<td>25 d</td>
<td>GSE50408</td>
</tr>
<tr>
<td>6</td>
<td>Drought and water recovery</td>
<td>PI567690/Pan a</td>
<td>14 h L/10 h D</td>
<td>NA</td>
<td>28 °C</td>
<td>NA</td>
<td>V5</td>
<td>4th trifoliate leaf</td>
<td>15 d</td>
<td>GSE70310</td>
</tr>
<tr>
<td>7</td>
<td>Submergence and reoxygenation</td>
<td>Williams 82</td>
<td>Constant light</td>
<td>50 μmol/m²/s</td>
<td>NA</td>
<td>NA</td>
<td>4 d</td>
<td>All aerial tissue</td>
<td>1 d/5 d/ 5 d +1 d reoxygenation</td>
<td>GSE51710</td>
</tr>
<tr>
<td>8</td>
<td>Heat shock</td>
<td>Williams 82</td>
<td>14 h L/10 h D</td>
<td>350 μmol/m²/s</td>
<td>22 °C</td>
<td>60 %</td>
<td>NA</td>
<td>First expanded Leaf</td>
<td>22/33.25/40,75/43.8 °C, 30 min</td>
<td>GSE26198</td>
</tr>
<tr>
<td>9</td>
<td>Heat shock</td>
<td>Williams 82</td>
<td>12 h L/12 h D</td>
<td>200 ± 25 μmol/m²/s</td>
<td>25 °C</td>
<td>NA</td>
<td>21 d</td>
<td>Leaf, stem &amp; root</td>
<td>42 °C 30 min &amp; 5 h</td>
<td>E-MTAB-2852</td>
</tr>
<tr>
<td>10</td>
<td>Alkaline</td>
<td>G07256 (Glycine soja)</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>24 °C</td>
<td>60 %</td>
<td>21 d</td>
<td>Root</td>
<td>0/0.5/1/3/6/12/24 h</td>
<td>GSE17883</td>
</tr>
<tr>
<td>11</td>
<td>Alkaline</td>
<td>G07256 (Glycine soja)</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>24 °C</td>
<td>60 %</td>
<td>21 d</td>
<td>Leaf</td>
<td>0/0.5/1/3/6/12/24 h</td>
<td>GSE20323</td>
</tr>
<tr>
<td>12</td>
<td>Salinity</td>
<td>Williams 82</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>25 °C</td>
<td>NA</td>
<td>14 d</td>
<td>RNA pools from multiple time points</td>
<td>7 d</td>
<td>GSE41125</td>
</tr>
<tr>
<td>13</td>
<td>Salinity/ Dehydration</td>
<td>Williams 82</td>
<td>NA</td>
<td>NA</td>
<td>25 °C</td>
<td>60 %</td>
<td>V1</td>
<td>Root</td>
<td>0/1/6/12 h</td>
<td>GSE57252</td>
</tr>
<tr>
<td>14</td>
<td>Melatonin/ Salinity</td>
<td>SuiNong 28</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>30-35 °C</td>
<td>25 °C</td>
<td>NA</td>
<td>Leaf</td>
<td>7 d</td>
<td>GSE57960</td>
</tr>
<tr>
<td>15</td>
<td>Iron deficiency</td>
<td>Clark/Isoclark</td>
<td>Ambient</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>7 d</td>
<td>Root</td>
<td>10/14 d</td>
<td>GSE22227</td>
</tr>
<tr>
<td>16</td>
<td>Iron deficiency</td>
<td>Clark/Isoclark</td>
<td>Ambient</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>7 d</td>
<td>Second trifoliate leaves</td>
<td>14 d</td>
<td>GSE10730</td>
</tr>
<tr>
<td>17</td>
<td>Aluminum</td>
<td>PI416937/ Young</td>
<td>16 h L/8 h D</td>
<td>100 μmol/m²/s</td>
<td>28 °C</td>
<td>NA</td>
<td>3 d</td>
<td>Root</td>
<td>12/72 h</td>
<td>GSE18518</td>
</tr>
<tr>
<td>18</td>
<td>Aluminum</td>
<td>PI416937</td>
<td>16 h L/8 h D</td>
<td>NA</td>
<td>28 °C</td>
<td>NA</td>
<td>NA</td>
<td>Root</td>
<td>2/12/48/72 h</td>
<td>GSE18423</td>
</tr>
<tr>
<td>19</td>
<td>Ozone</td>
<td>PI548379/ PI438471</td>
<td>NA</td>
<td>254 ± 13 μmol/m²/s</td>
<td>34 ± 1 °C</td>
<td>61 ± 1%</td>
<td>25 d</td>
<td>5th trifoliate leaf</td>
<td>30-minute increments starting 5 hours after introduction of ozone, 4 time points</td>
<td>GSE59076</td>
</tr>
</tbody>
</table>
Table S2.2. Summary of RASL-seq experiment conditions.

<table>
<thead>
<tr>
<th>Abiotic stress</th>
<th>Drought</th>
<th>Heat</th>
<th>Fe deficiency Leaf</th>
<th>Fe deficiency Root</th>
<th>NaHCO$_3$ stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotypes</td>
<td>Williams 82</td>
<td>Williams 82</td>
<td>IsoClark</td>
<td>IsoClark</td>
<td>Glycine soja</td>
</tr>
<tr>
<td>Photoperiod</td>
<td>16 h L/8 h D</td>
<td>16 h L/8 h D</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
</tr>
<tr>
<td>Light intensity</td>
<td>100 μmol/m$^2$/s</td>
<td>100 μmol/m$^2$/s</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
</tr>
<tr>
<td>Temperature</td>
<td>28 °C</td>
<td>28 °C</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
</tr>
<tr>
<td>Humidity</td>
<td>50 %</td>
<td>50 %</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
<td>Ambient, June-July 2016. USDA green house, ISU</td>
</tr>
<tr>
<td>Age</td>
<td>14 d</td>
<td>14 d</td>
<td>17 d</td>
<td>17 d</td>
<td>17 d</td>
</tr>
<tr>
<td>Tissue</td>
<td>Unifoliolate</td>
<td>Unifoliolate</td>
<td>First Trifoliolate</td>
<td>Root</td>
<td>First Trifoliolate</td>
</tr>
<tr>
<td>Growth and treatment</td>
<td>Plants were grown in soil for 14 days and given same amount water every day. For the drought treated plants, watering was suspended on day 12th; the control plants were maintained same water every day. On 14th day, the light condition was switched to constant light. The samples were harvested every 4 hours for 44 hours starting at ZT0 on the 15th day.</td>
<td>Plants were grown in soil for 14 days. On 14th day, the light condition was switched to constant light. The samples were harvested every 4 hours for 44 hours starting at ZT0 on the 15th day.</td>
<td>The growth condition was mainly according to the conditions used in dataset GSE22227 (See the reference in Dataset S2) with some changes as below. One week after planting, plants were transferred to hydroponic unit system and grow for 10 days in iron efficient condition (100 μM Fe(NO$_3$)$_3$). (on 9th day, change to constant light). On 10th day, treated seedlings were switched to iron limited condition (50 μM Fe(NO$_3$)$_3$) at ZT 0 and started to collect samples every 4 hours for 44 hours.</td>
<td>The growth condition was mainly according to the conditions used in dataset GSE22227 (See the reference in Dataset S2) with some changes as below. One week after planting, plants were transferred to hydroponic unit system and grow for 10 days (on 9th day, change to constant light). On 10th day, half seedlings were treated with 50 mmol/L NaHCO$_3$ (pH 8.5) at ZT0 and started to collect samples every 4 hours for 44 hours.</td>
<td>Same as in Fe deficiency Leaf</td>
</tr>
<tr>
<td>Sampling notes</td>
<td>NA</td>
<td>The control samples are same as the control samples in the drought experiment.</td>
<td>The 0 h samples are shared between the control and treatment.</td>
<td>The 0 h samples are shared between the control and treatment.</td>
<td>The 0 h samples are shared between the control and treatment.</td>
</tr>
<tr>
<td>Replicates</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>

Growth and treatment:

- Plants were grown in soil for 14 days and given same amount water every day. For the drought treated plants, watering was suspended on day 12th; the control plants were maintained same water every day. On 14th day, the light condition was switched to constant light. The samples were harvested every 4 hours for 44 hours starting at ZT0 on the 15th day.
- Plants were grown in soil for 14 days. On 14th day, the light condition was switched to constant light. The samples were harvested every 4 hours for 44 hours starting at ZT0 on the 15th day.
- The growth condition was mainly according to the conditions used in dataset GSE22227 (See the reference in Dataset S2) with some changes as below. One week after planting, plants were transferred to hydroponic unit system and grow for 10 days in iron efficient condition (100 μM Fe(NO$_3$)$_3$). (on 9th day, change to constant light). On 10th day, treated seedlings were switched to iron limited condition (50 μM Fe(NO$_3$)$_3$) at ZT 0 and started to collect samples every 4 hours for 44 hours.
- The growth condition was mainly according to the conditions used in dataset GSE22227 (See the reference in Dataset S2) with some changes as below. One week after planting, plants were transferred to hydroponic unit system and grow for 10 days (on 9th day, change to constant light). On 10th day, half seedlings were treated with 50 mmol/L NaHCO$_3$ (pH 8.5) at ZT0 and started to collect samples every 4 hours for 44 hours.

12 time Points:

- 0/4/8/12/16/20/24/28/32/36/40/44 hour
- 0/4/8/12/16/20/24/28/32/36/40/44 hour
- 0/4/8/12/16/20/24/28/32/36/40/44 hour

Total samples:

- 72
- 36
- 69
- 69
- 69
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>SwissProt ortholog gene</th>
<th>Common Primer 1 (Primers specific to Index primer)</th>
<th>Common Primer 2 (Primers + gene specific primer)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9899</td>
<td>G-ATTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9889</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9869</td>
<td>G-CTGGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9849</td>
<td>G-CTGGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9629</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9519</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9499</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9389</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9279</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9169</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j9059</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8949</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8839</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8729</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8619</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8509</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8399</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8289</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8179</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j8069</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j7959</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j7849</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j7739</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
<tr>
<td>CCA1-HV</td>
<td>Gm07j7629</td>
<td>G-TGTTGATAGGCACGTCCTGTCAG</td>
<td>GCTGACGTCCTGTCAG</td>
</tr>
</tbody>
</table>
CHAPTER 3. BIDIRECTIONAL INTERACTION BETWEEN SOYBEAN ROOT CIRCADIAN CLOCK AND SOYBEAN CYST NEMATOIDE (HETERODERA GLYCINES)

Lijun Cao\textsuperscript{a}, Wei Wang\textsuperscript{a,b,*}

\textsuperscript{a}Department of Plant Pathology and Microbiology, Iowa State University, Ames, IA, 50011, USA;

\textsuperscript{b}School of Life Sciences, Peking University, Beijing, China;

\textsuperscript{*}Corresponding Author: Department of Plant Pathology and Microbiology, Iowa State University, 2213 Pammel Drive, Ames, IA 50011, USA.

Modified from a manuscript to be submitted to The Plant Cell

3.1 Abstract

Plant circadian clock has emerged as a cardinal signaling hub, and it integrates diverse immune signals to gate plant defense response against pathogens. However, almost all findings so far made are through studies of leaf circadian clock. Virtually nothing is known about the root circadian clock and its interaction with pathogens. To fill this knowledge gap, this research initiates the study of bidirectional interaction between soybean (\textit{Glycine max}) root circadian clock and soybean cyst nematode (SCN) (\textit{Heterodera glycines}). Through bioinformatics analysis using the time-indicating genes, I found SCN caused significant phase-shift to the global circadian rhythm in the infection site. Then, I identified that SCN caused circadian rhythm changes to 12 core circadian gene transcripts in soybean seedlings using RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing (RASL-seq). Using live bioluminescence imaging, I further found SCN caused circadian rhythm changes to 2 core circadian gene promoter-luciferase hairy root lines.
On the other hand, I also found that soybean root circadian clock regulated plant defense against SCN. Overexpression of a CCA1 gene in soybean hairy roots resulted in disrupted SCN infection rhythm.

To obtain a comprehensive understanding of this root circadian clock regulation, I performed a large-scale time-course transcriptional profiling and identified key circadian rhythmically expressed and non-circadian-rhythm genes in the interaction between soybean root circadian clock and SCN. These results enable us to uncover novel inputs and outputs to the root circadian clock.

### 3.2 Introduction

The circadian clock not only helps plants defense against various pathogens but also helps to anticipate when to defense. The previous studies on Arabidopsis discovered that even without the presence of pathogens, the circadian clock induced plant defense in the morning to “anticipate” the potential infection (Wang et al., 2011). Several other studies on various plant-pathogen systems have unveiled similar phenomena not only in Arabidopsis (Bhardwaj et al., 2011; Goodspeed et al., 2013; Windram et al., 2012; Zhang et al., 2013), but also in post-harvest vegetables and fruits (Goodspeed et al., 2013). Therefore, the circadian clock-regulated immune system appears to be a universal strategy adopted by various plants to fight against pathogens. Additionally, the plant immune hormone Salicylic Acid (SA) reinforces the circadian clock, and this reinforcement helps fine-tune the timing of Systemic Acquired Resistance (SAR) (Zhou et al., 2015). Jasmonic Acid (JA), another important hormone in plant defense against root-knot nematodes (Nahar et al., 2011), was also shown to interact with the plant circadian clock (Sanchez et al., 2016). While the regulation of plant defense by circadian clock has been recognized, accumulating studies also suggest that pathogens also regulate the plant circadian clock function. Bacterial and oomycete pathogens have been shown to perturb the plant circadian clock in Arabidopsis (Wang et al., 2011;
Zhang et al., 2013). Taken together, it is clear that the impacts between plant circadian clock and pathogens are bidirectional. The circadian clock regulates the temporal expression of defense genes against pathogens, while pathogens, in return, modulate the circadian clock function.

As an emerging research direction bridging plant circadian clock and pathogens, almost all discoveries made so far were focused on leaf circadian clock and leaf pathogens. Little is known about the interactions between the root circadian clock and root pathogens. This represents a key knowledge gap impeding further enhancement of crop resistance against root pathogens. To fill this gap, I initiated the study of the bidirectional interaction between soybean root circadian clock and soybean cyst nematode (SCN).

SCN is the most destructive pathogen of soybean, causing over $1 billion yield loss annually in the US alone (Allen et al., 2004; Koenning et al., 2010). An in-depth understanding of its interaction with soybean is essential to develop novel SCN-resistant varieties and mitigate economic loss. As an obligate biotrophic pathogen, SCN obtains nutrients exclusively from soybean roots through a feeding site called syncytium. To establish the syncytium, SCN injects effector proteins directly into the root, which induces dramatic physical and biochemical reprogramming (Fiorilli et al., 2012). This initial feeding cell then fuses with hundreds of adjacent cells to form an enlarged, multi-nucleated, and metabolically hyperactive feeding site (Allen et al., 2004; Gheysen et al., 2011; Mitchum et al., 2013; Williamson et al., 1996). Our limited understanding of how nematodes select and reprogram the initial feeding cells remains to be a big obstacle to combat this pathogen.

It is well-known that plant metabolism is extensively regulated by the circadian clock in anticipation of the recurring daily environmental changes to maximize fitness (Michael et al., 2008). The circadian clock was also shown to be epigenetically modulated for the reprogramming
of cell metabolism in plants showing hybrid vigor (Miller et al., 2015; Ni et al., 2009). It is possible that SCN may also modulate or even hijack the soybean root circadian clock and dictate soybean metabolism. Further dissection of the underlying signaling pathway may hold the key to understand how cyst nematodes reprogram plant cells.

While SCN perturbs the soybean root circadian clock, the soybean root circadian clock may also regulate plant defense against SCN. Similar interactions have been observed in Arabidopsis (Bhardwaj et al., 2011; Goodspeed et al., 2013; Windram et al., 2012; Zhang et al., 2013). However, the interactions between the root circadian clock and root pathogens can be more complicated than that in leaves. According to a recent study in Arabidopsis, the circadian clock in root cells is less synchronized compared with that in leaf cells (Takahashi et al., 2015). If a similar phenomenon also exists in soybean, the root circadian clock-regulated defense system may exhibit variations among different root cells in the same root. Indeed, I observed that SCN caused significant change to the global circadian rhythm only in the infection site. However, I also found that SCN modulated the expression of two key circadian genes in the whole root system. Therefore, it is likely SCN modulates clock function of the feeding cells to directly obtain nutrients. On the other hand, SCN also modulates the clock function of whole root cells to get a continued nutritional supply. These results have added to our understanding of the interactions between root circadian clock and pathogens.

3.3 Results and Discussions

3.3.1 SCN Perturbed Global Circadian Rhythm in Syncytia

Transcriptional profiling of soybean after SCN infection was conducted on both SCN infection site and the surrounding site using a laser capture microdissection coupled with high-density oligonucleotide microarray analysis (Ithal et al., 2007a, 2007b). The authors found that SCN induced significant transcriptional changes in both nematode infection site and the
surrounding site. In order to determine if SCN induces change to the soybean circadian clock, I re-analyzed these transcriptome data using the Molecular Timetable method (method presented in Chapter 2). Two days after infection, no significant circadian rhythm change was found in the uninfected soybean root cells (Figure 3.1). A significant change to circadian rhythm was only found in SCN infection site (Figure 3.2 A), and the phase shifted significantly from midday to early night after infection (Figure 3.2 D). SCN infection resulted in a slightly longer period (Figure 3.2 B) and lower amplitude (Figure 3.2 C) in the infection site compared with that in the mock treatment, but not statistically significant.

The results showed that SCN induced obvious change to the global circadian rhythm only in the infection site. However, there is another possible explanation due to the setup of these experiments. Samples used in these experiments were growing under diurnal condition: 16 h of light and 8 h of dark. The entraining effects from photoperiod may mask SCN’s impact on the circadian clock. Similar experiment under constant condition can help discount the photoperiod effects and reveal the impact on the circadian clock by SCN.

### 3.3.2 Identification of Circadian Genes Targeted by SCN using RASL-seq

To confirm SCN’s impact on soybean circadian rhythm and further identify specific clock components targeted by SCN, I performed two-day time-course RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing (RASL-seq) experiments. RASL-seq is cost-effective to simultaneously quantify expression profiles of several hundred genes under many different conditions (Li et al., 2015). Roots of soybean seedlings with SCN infection and mock infection (0.1 % agarose) were collected every 4 hours for 2 days (Figure 3.3 A). RNA was extracted and sent for sequencing. Through the statistical analysis, I identified the expression rhythms of 12 circadian genes were modified by SCN (Figure 3.3 B and C). Of these
genes, 5 showed significant phase changes (Figure 3.3 B) and 8 showed significant period changes (Figure 3.3 C).

Through time-course RASL-seq, I am able to pinpoint the circadian components targeted by SCN. However, the RNA samples were extracted from the whole roots, and I am still not sure if these changes are specific to the infection site or the surrounding site.

3.3.3 SCN Perturbed the Expression Rhythms of Two Circadian Genes in Soybean Hairy Roots

To figure out how SCN modifies soybean root circadian rhythm, I used a live circadian monitoring system. Basically, the vector harboring the luciferase gene driven by the promoter of the identified target gene was transformed into soybean hairy root, and the luciferase activity was monitored through a hyper-sensitive Charge Coupled Device (CCD) camera. This live imaging system allowed me to obtain images of luciferase activity in soybean hairy root with a high temporal and spatial resolution.

One important character of the circadian clock is that it is endogenously generated and self-sustainable, thus the expression of the circadian gene oscillates even under constant condition. I first checked the circadian rhythm robustness of the 12 genes identified by RASL-seq. Gene promoter-luciferase vectors were constructed, 11 out of 12 vectors were transformed successfully to soybean hairy roots (Figure S3.1). All 11 gene promoter-luciferase showed circadian oscillation under diurnal condition (Figure S3.1). However, only one gene, Glyma.16G217700, promoter-luciferase showed good circadian rhythm under constant condition (Figure S3.1). Therefore, I only focused on this gene promoter-luciferase line for later SCN infection assay.

The overall circadian rhythm change upon SCN infection was shown in Figure 3.4 A. I found a significant phase change after SCN infection under constant condition (Figure 3.4 B). I
also noticed the luciferase activity had a relatively weak circadian rhythm ($R^2 = 0.04814$) after SCN infection compared with mock infection, and it was arrhythmic 3 days after SCN infection.

These imaging data were collected from root tips, including SCN infection site and the surrounding site. This indicates that SCN perturbation of soybean root circadian clock is not restricted to the feeding sites, and it also influences the whole root circadian rhythm through targeting Glyma.16G217700. Glyma.16G217700 encodes an Early Phytochrome Responsive 1/Reveille 7 (EPR1/RVE7)-like protein. In Arabidopsis, overexpression of EPR1 results in promoting far-red light-induced cotyledon opening and delaying flowering (Kuno et al., 2003). The transcription of EPR1 in wild-type Arabidopsis exhibits circadian rhythm similar to that of CCA1 and LHY under constant light condition. EPR1 suppresses its own expression, suggesting that this protein is part of a regulatory feedback loop. Constitutive expression of CCA1 and LHY results in the loss of EPR1 rhythmicity, whereas overexpressing EPR1 has no effect on the central oscillator (Kuno et al., 2003). EPR1 is proposed to be a component of slave oscillators contributing to the refinement of circadian output pathways, ultimately mediating the correct oscillatory behavior of target genes (Kuno et al., 2003). However, different plants might have different circadian regulation pathways (Hsu et al., 2014). Glyma.16G217700 may have different roles in soybean as EPR1 in Arabidopsis. The function of Glyma.16G217700 is still largely unclear.

To investigate the evolutionary relationship of EPR1 among different plant species, phylogenetic analysis was performed using amino acid sequences according to their similarities by MEGA 7.0. Phylogenetic analysis revealed that Glyma.16G217700 was closest related to Reveille 7-like in Cajanus cajan (Figure 3.4 C). The lack of similarity between Glyma.16G217700 and Arabidopsis EPR1/RVE7 also implies functional difference.
In addition to the *EPR1*, I also monitored the activity of *CCA1* (Figure S3.2 A) and *TOC1* promoter-luciferase activities (most *TOC1* promoter-luciferase activities showed weak circadian rhythmic or arrhythmic, and the data are not shown). A significant change to the average expression of *CCA1* promoter-luciferase activity was observed (Figure S3.2 B). The change to *CCA1* average expression was not observed in RASL-seq, because the data was normalized to its average expression, and I only analyzed phase and period changes in RASL-seq.

### 3.3.4 Soybean Seedlings Are More Resistant to SCN at Dawn than at Dusk

The circadian-regulated defense response against a leaf obligate biotrophic pathogen, *Hyaloperonospora arabidopsidis*, was described in *Arabidopsis* (Wang et al., 2011). The intimate relationship and long co-evolution history make the host and the pathogen familiar with each other’s metabolic rhythm. It is not surprising that the plant circadian clock may evolve to regulate the defense genes to anticipate potential infection on a daily basis. Cyst nematodes are obligate biotrophic parasites. Likewise, it is very likely that the root circadian clock may also regulate defense against cyst nematodes. Understanding this plant-pathogen relationship will allow the modulation of plant defenses to fight against cyst nematodes in the future.

To directly test the involvement of the circadian clock in soybean defense against SCN, I performed time-course SCN infection assay using soybean seedlings (Figure 3.5). Seedlings were inoculated with hatched SCN at two time-points, ZT0 and ZT16, representing dawn and dusk, respectively. A significantly more adult female SCNs were observed 30 days after infected at dusk than that at dawn under diurnal condition (Figure 3.5 A).

The diurnal condition is more like a natural growth condition for soybean. However, photoperiod difference is a confounding factor, which makes it difficult to tell if the susceptibility difference is due to the circadian clock or photoperiod. Therefore, I performed similar experiments under constant condition. Soybean seedlings were transferred to constant light condition one day
before SCN inoculation. Similar results were observed, and more infected female nematodes were found when infected at dusk than that infected at dawn (Figure 3.5 B).

These results indicate that soybean in the morning is more resistant to SCN than at dusk. The observed SCN susceptibility difference can be explained by circadian clocks of soybean, SCN or both. Although little is known about the circadian clock of SCN, the soil-dwelling nematode Caenorhabditis elegans (C. elegans) displays circadian rhythms in swimming behavior (Saigusa et al., 2002) and responses to osmotic stress (Kippert et al., 2002). And a circadian homologous gene LIN-42 was shown to regulate the behavior of C. elegans (Van Wynsberghe et al., 2016).

To find out whether the susceptibility difference is due to SCN clock or not, I performed time-course nematode hatch experiment and measured the number of hatched nematodes (Figure 3.6 A) and size (Figure 3.6 B). No significant differences were observed among the three different hatch times (Figure 3.6). These results suggest that SCN circadian clock may not be the major determinant. However, further studies are needed to confirm the conclusion.

3.3.5 Soybean Clock Mutant Hairy Root Lines Showed Disrupted SCN Infection Rhythmicity

To further differentiate the contributing factors to susceptibility difference, I developed soybean clock mutant hairy root lines by overexpressing one of the key circadian gene CCA1 (Glyma.03G261800). Then, I performed time-course qPCR to check the gene expression level in both the control and overexpression (OE) lines (Figure 3.7 A and B). Soybean hairy roots were collected every 4 hours for 2 days. RNA was extracted and used for expression analysis. The OE lines showed significantly higher expression levels under both diurnal and constant conditions (Figure 3.7 A and B) than that of the control lines. The control lines had circadian rhythmic expression under both conditions, while weak circadian rhythm for OE lines was observed under
diurnal condition, and totally arrhythmic for OE lines was observed under constant condition (Figure 3.7 A and B).

A relatively slower growth rate of OE lines was observed (Figure 3.7 C and D), and significantly fewer branches were generated in the OE lines after growing for 9 days (Figure 3.7 E) compared with the control lines. These observations are consistent with those phenotypes observed in *Arabidopsis* by overexpressing *CCA1* gene (Mizoguchi et al., 2002; Wang et al., 1998).

To further determine if soybean root circadian clock is the major determinant of soybean susceptibility difference, I performed time-course SCN infection assay using both control and OE hairy root lines (Figure 3.8). Under both diurnal and constant conditions, the control lines showed good circadian oscillations of SCN infection rate (Figure 3.8 A and B) while circadian rhythms of infection for OE lines were disrupted (Figure 3.8 A and B).

These results confirm that the soybean root circadian clock is the main contributor to the susceptibility difference. I also noticed that OE lines were more resistant than the control lines, with fewer infected SCNs (Figure 3.9). This is especially obvious under constant condition (Figure 3.9 B).

### 3.3.6 Identification of Genes in the Interaction between Soybean Root Circadian Clock and SCN using Time-course RNA sequencing (RNA-Seq)

To further investigate the role of root circadian clock in the interaction between soybean and SCN, I performed a time-course RNA-Seq experiment using control and OE lines under constant condition. Control and OE hairy root lines were treated with both mock (0.1% agarose) and hatched SCNs at ZT0, and the treated root samples were collected every 4 hours for 2 days starting right after treatments.
A total of 144 mRNA samples were sequenced using Illumina High Throughput Sequencing (Hiseq-PE150), and the raw sequencing data were mapped to both soybean (*Glycine max* v2.0) (Schmutz et al., 2010) and SCN (*Heterodera glycines* isolate TN10) (Masonbrink et al., 2019) genomes.

To find out the circadian rhythmically expressed soybean genes, the time-course expression data were first subjected to DESeq normalization. The normalized data were then analyzed through nonlinear regression using GraphPad Prism 6. In total, 318 circadian rhythmic genes were identified in the control hairy root lines with the coefficient of determination $R^2 > 0.6$ (Figure 3.10 A). The overall expression profiles of the 318 genes were shown in heatmaps (Figure 3.10 A, B, C, and D). SCN infection (Figure 3.10 C and D) caused slightly change to the circadian rhythm compared with that in mock infection (Figure 3.10 A and B). However, compared with control lines (Figure 3.10 A and C), OE lines (Figure 3.10 B and D) showed obvious circadian rhythm disruption. Specifically, 2 arrhythmically expressed genes, 11 phase-shifted genes, 5 expression-decreased genes, and 4 other parameter-changed genes were identified in SCN infection hairy root lines compared with their relative circadian-rhythmically expressed genes in mock infection (Figure 3.10 E). For OE hairy root lines, 51 arrhythmically expressed genes, 12 phase-shifted genes, 22 amplitude-changed genes, 65 expression-increased genes, and 44 expression-decreased genes were identified compared with their relative circadian-rhythmically expressed genes in control lines (Figure 3.10 F).

In addition to the circadian rhythm genes, I identified 36,807 soybean genes without circadian rhythm (Figure S3.3), named as non-circadian-rhythm genes. Using DESeq2, I found that 27 genes were up-regulated (log2FoldChange > 1 and FDR < 0.05) while 18 genes were down-regulated (log2FoldChange < -1 and FDR < 0.05) in SCN infection hairy root lines compared with
the mock infection lines (Figure 3.11 A and B). 554 genes were up-regulated while 588 genes were down-regulated in OE hairy root lines compared with the control lines (Figure 3.11 C and D).

To identify the circadian rhythmically expressed SCN genes, the time-course transcriptome data were first subjected to DESeq normalization. Then, the time-course expression of each SCN gene was fit against 1,440 cosine curves. In total, 2,598 circadian rhythmic SCN genes were identified in control soybean hairy root lines with SCN infection (Figure 3.12 A and C) while 901 circadian rhythmic SCN genes were identified in OE soybean hairy root lines with SCN infection (Figure 3.12 B and C), and only 443 genes shared in both groups (Figure 3.12 C). These data clearly showed that CCA1 overexpression in soybean hairy roots also caused changes to the expression of infected-SCN genes, and less circadian rhythmically expressed SCN genes were identified in OE hairy root lines compared with control lines (Figure 3.12 C). More arrhythmically expressed SCN genes (Figure 3.12 C) and fewer infected SCNs (Figure 3.9) in OE hairy root lines indicate possibly decreased SCN fitness, as asynchronization of physiological processes with circadian clock decreases fitness in both animals (Husse et al., 2015) and plants (Mwimba et al., 2018).

3.4 Materials and Methods

Plant and nematode materials. Seeds of soybean Williams 82 were used throughout this study. Soybean cyst nematode (Heterodera glycines) Race 3 (HG3) was used and propagated on soybeans as described (Niblack et al., 1994). For nematode infection assay, SCN eggs were cultured for 3 days under constant 26 °C and dark conditions, and then the hatched SCNs were collected and used for infection.

RASL-seq: Seedlings of Williams 82 were grown in the soil (Sun Gro Horticulture, Vancouver, BC, Canada) under long-day diurnal conditions (16 h light/8 h dark) for 9 days in growth chambers with the following conditions: 28 °C, 50% relative humidity and 100 μmol/m²/s
light intensity. On the 10th day, the light was switched to constant condition (constant 100 μmol/m²/s light intensity). One day after, the seedlings were inoculated with 3,000 hatched SCNs or mock (same amount of water). 7 days after inoculation, root samples were collected every 4 hours for 2 days. For each time-point, three biological replicates were harvested. Detailed RASL-seq procedure is shown in Chapter 2.

Re-analysis of the available RNA-Seq data, non-linear regression assay using GraphPad was performed according to Chapter 2.

Soybean hairy root lines. Soybean gene promoters, luciferase gene, and terminator were PCR-amplified with Q5® High-Fidelity DNA Polymerase (NEB). For promoter-luciferase reporter lines, vector pTF101.1 (donated by Thomas Baum’s lab) was digested with restriction enzymes (NEB). Gibson Assembly was used for assembly of the vector with the cloned gene promoter, luciferase, and terminator (Gibson et al., 2009). For CCA1 overexpression, PCR-amplified CCA1 fragments and vector pG2RNAi2 (GenBank: KT954097) were digested with SwaI and BamHI (NEB) and ligated with DNA ligase (NEB). Transgenic hairy roots generation and nematode infection assay were performed as described (Liu et al., 2012). Soybean hairy roots were cultured under the diurnal condition, which is the cycling of 26 °C for 16 h and 20 °C for 8 h, and constant condition, which is constant 26 °C. All hairy roots were grown in the growth chamber with 50 % relative humidity and dark condition.

Live imaging system. Transgenic soybean hairy roots were cultured on sterilized 150 × 15 mm petri dishes. One day before imaging, roots were sprayed with 2.5 mM luciferin (Gold Biotechnology) in 0.02 % Triton X-100 (Sigma). A hypersensitive CCD iKon-L 936 camera with Nikon NIKKOR 50 mm f/1.2 Lens was used to capture the light generated. The data were analyzed using ImageJ. More than 4 biological replicates were used for each line of hairy roots.
Growth rate measurements. The length of the soybean hairy root was measured every 2 days, and the growth rate was calculated as newly generated root length relative to soybean hairy root length of two days before measurement. 8 biological replicates were used.

Quantitative real-time PCR. One-step qRT-PCR was performed as described (Noon et al., 2016). Three biological replicates were used.

Time-course RNA-Seq. Soybean hairy root samples were collected every 4 hours for 2 days. Total RNA was extracted using TRIzol (Invitrogen) RNA Isolation Protocol. The RNA quality was checked using an Agilent 2100 Bioanalyzer (Agilent Technologies, Pale Alto, CA, USA). Illumina High Throughput Sequencing (Hiseq-PE150) was used, and the raw sequencing data were mapped to both soybean and SCN genomes. After alignment, raw read counts were subjected to DESeq 2 (R package, Bioconductor-DESeq2) normalization. To find out the circadian rhythmically expressed soybean genes, the normalized data of each gene was fit to nonlinear regression in GraphPad Prism 6 according to Chapter 2. Genes with the coefficient of determination $R^2 > 0.6$ were identified as circadian rhythmically expressed genes, and the rest was non-circadian-rhythm genes. Circadian rhythm parameters, such as amplitude, phase, period, and average expression, were also calculated in GraphPad. To identify circadian rhythmically expressed genes in SCN, the raw read counts were first subjected to DeSeq2 normalization. For each time-point, the data were then normalized to the number of their relative infected SCNs (Figure S3.4). The time-course expression of each gene was subjected to Z-Score normalization and then fit against 1,440 test cosine curves (details can be seen from Chapter 2). The correlation coefficient (R) and the phase of the best-fitting curve were assigned to each gene. The genes with correlation coefficient (R) $\geq 0.5$ and the Coefficient of Variation (CV) $\geq 30$ were selected as circadian rhythmically expressed genes. For genes without circadian rhythm, raw reads counts
were analyzed using DESeq2. Differentially expressed genes were selected with parameters: log2FoldChange > 1 or < -1, and FDR < 0.05. Three biological replicates for each time-point were used.

Time-course infected SCN number measurement. 110 hatched SCNs were used for infection of both control and OE hairy root lines at ZT0. Right after infection, the root samples were collected every 4 hours for two days, the same as root samples used in RNA-Seq. The samples were further stained with acid fuchsin as described (Hussey et al., 1985). After staining, place samples on the inside of Petri-dish lids and embed in approximately 10 ml acidified glycerine (prepare freshly by adding 5-10 drops of 1 M HCl to glycerine) for each lid. Then slightly press down the Petri-dish bottom on the roots without trapping bubbles. A microscope was used to count the number of infected SCNs. For each time-point, more than three biological replicates were used.

Statistical analysis. Statistical analysis was performed using R languages (version 3.3.1) and GraphPad Prism 6. Heatmaps were generated using GraphPad Prism 6 and online software Heatmapper (http://www.heatmapper.ca). Venn diagrams were generated using jvenn online (http://jvenn.toulouse.inra.fr/app/example.html).

3.5 Author Contributions

W.W and L.C designed the research project. L.C performed the experiments, analyzed the data, and drafted the manuscript.

3.6 References


3.7 Figures

Figure 3.1. Soybean cyst nematode (SCN) does not perturb soybean circadian rhythm in root segments surrounding SCN infection sites. (A) Global circadian rhythm of root segments surrounding SCN infection sites in mock and SCN infection root samples. Changes to Period (B), Amplitude (C), and Phase (D) after SCN infection. The mean and the standard error of the mean are plotted. CT, Circadian Time. a.u., arbitrary unit.
Figure 3.2. SCN perturbs soybean circadian rhythm in syncytia. (A) Global circadian rhythm of syncytia in mock and SCN infection root samples. Changes to Period (B), Amplitude (C) and Phase (D) after SCN infection. The mean and the standard error of the mean are plotted. CT, Circadian Time. a.u., arbitrary unit. ****, \( p < 0.0001 \).
Figure 3.3. Changes to expression rhythms of circadian genes after SCN infection. (A) Sampling scheme for time-course RASL-seq. The digits before “ZT” mark the sampling time-points. White box, day. Black box, night. Grey box, subjective night. ZT, Zeitgeber Time. (B) SCN induces phase changes to the expression rhythms of circadian genes. The angular coordinates represent Phase24. The -log_{10} transformed oscillation p values represent the robustness of the oscillation and are plotted as radial coordinates. The circles represent mock infection samples and the triangle represents SCN infection samples. The size of the symbols is proportional to the standard error of the mean of phase as illustrated in the legend. The arrows point from mock to SCN infection samples represent the direction of phase shift. (C) SCN induces period changes to the expression rhythms of circadian genes. Student’s t-test with Benjamini and Hochberg multiple comparison is used to compare mock and SCN infection samples and derive False Discovery Rate (FDR). To highlight statistically significant changes and apply a high stringency level on oscillation robustness, only genes with oscillation p < 10^{-10} and FDR < 0.05 are plotted. The mean and the standard error of the mean are plotted.
Figure 3.4. Changes to *Glyma.16G217700* promoter-luciferase activity upon SCN infection and the phylogenetic analysis of Glyma.16G217700. (A) *Glyma.16G217700* promoter-luciferase activity in soybean hairy roots before and after mock (0.1% agarose, blue line) and SCN (3,000 J2 SCNs, red line) infection at ZT52 (black arrow). Red boxes under X-axis indicate 26 °C. Cyan boxes represent subjective 20 °C. (B) Phase changes before and after SCN infection. ****, \( p < 0.0001 \) (two-way ANOVA). The mean and the standard error of the mean are plotted. (C) The phylogenetic tree of Glyma.16G217700. The unrooted phylogenetic tree is constructed using MEGA 7.0. ZT, Zeitgeber Time. a.u., arbitrary unit.
Figure 3.5. Soybean seedings are more resistant to SCN at dawn than that at dusk. The number of female SCNs 30 days after infection performed at dawn and dusk under diurnal (A) and constant (B) conditions. The mean and the standard error of the mean are plotted. Unpaired Student's *t*-test is used. ****, *p* < 0.0001.
Figure 3.6. Time-course SCN hatch assay. (A) The number of hatched SCNs. (B) The size of hatched SCNs. All SCNs used in this experiment are hatched from eggs under 26 °C for 3 days. The mean and the standard error of the mean are plotted. Nonparametric Mann-Whitney test is used. ZT, Zeitgeber Time. a.u., arbitrary unit.
Figure 3.7. Expression of *Circadian Clock Associated 1 (CCA1)* and phenotypes of control and CCA1 overexpression (OE) soybean hairy roots. The relative expression level of *CCA1* under diurnal (A) and constant (B) conditions. Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. Blue boxes under X-axis indicate 20 °C. ZT, Zeitgeber Time. (C) Phenotypes of control and OE hairy roots at day 0 and day 9 of culturing. (D) Relative growth rates of control and OE. (n = 8 hairy root lines) (E) Branches of control and OE at day 0 and day 9. (n = 8 hairy root lines). The mean and the standard error of the mean are plotted. ****, p < 0.0005 (two-way ANOVA).
Figure 3.8. Time-course SCN infection assay. The number of adult female SCNs 30 days after infection under diurnal (A) and constant (B) conditions. Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. Blue boxes under X-axis indicate 20 °C. The mean and the standard error of the mean are plotted. ZT, Zeitgeber Time.
Figure 3.9. The number of infected female SCNs. The number of infected female SCNs under diurnal (A) and constant (B) conditions. Data is re-calculated from SCN infection assay in Figure 3.8. The mean and the standard error of the mean are plotted. Nonparametric test: Mann-Whitney test is used. ****, $p < 0.0001$. 
Figure 3.10. RNA-Seq analysis of circadian rhythmically expressed soybean genes. Heatmaps of time-course expression of the 318 circadian rhythmically expressed genes in control hairy root lines with mock infection (A), CCA1 overexpression hairy root lines with mock infection (B), control hairy root lines with SCN infection (C), and CCA1 overexpression hairy root lines with SCN infection (D). Data is subjected to Z-score normalization and arranged in groups by their phase time and placed in rows from ZT0 (up) to ZT23 (down). (E) The number of genes showing changes to the circadian rhythm in CCA1 overexpression hairy root lines compared with that in control hairy root lines. (F) The number of genes showing changes to circadian rhythm after SCN infection compared with that after mock infection. Arrhythmicity, arrhythmically expressed genes. Phase, phase-shifted genes. Amplitude, amplitude-changed genes. Expression Increased, average expression increased genes. Expression Decreased, average expression decreased genes. Others, genes showing changed expression trend. ZT, Zeitgeber Time. CKMock, control hairy root lines with mock infection. OEMock, CCA1 overexpression hairy root lines with mock infection. CKSCN, control hairy root lines with SCN infection. OESCN, CCA1 overexpression hairy root lines with SCN infection.
Figure 3.11. RNA-Seq data analysis of non-circadian-rhythm soybean genes. The number (A) and heatmap (B) of differentially expressed genes in CCA1 overexpression hairy root lines compared with that in control hairy root lines. The number (C) and heatmap (D) of differentially expressed genes in SCN infection hairy root lines compared with that in mock infection hairy root lines. Up, the number of up-regulated genes. Down, the number of down-regulated genes. The digit above every box represents the number of differentially expressed genes. Mock, mock infection. SCN, SCN infection. OE, Overexpression. Genes are arranged in rows in heatmaps by their genomic position, chromosome 1 (top) to chromosome 20 (bottom).
Figure 3.12. RNA-Seq analysis of circadian rhythmically expressed SCN genes. Heatmaps of time-course expression of the circadian rhythmically expressed SCN genes in control hairy root lines with SCN infection (A) and CCA1 overexpression hairy root lines with SCN infection (B). Data is first subjected to Z-score normalization and arranged in rows by their phase time, from ZT0 (up) to ZT23 (down). (C) Venn diagram of the number of circadian rhythmically expressed SCN genes in control (blue) and CCA1 overexpression (red) hairy root lines. ZT, Zeitgeber Time. CKSCN, control hairy root lines with SCN infection. OESCN, CCA1 overexpression hairy root lines with SCN infection.
Figure S3.1. Identification of genes with robust circadian rhythm. 11 transgenic soybean hairy root lines harboring circadian gene promoter-luciferase constructs are generated and imaged for 6 days (3 days of diurnal condition, 3 days of constant condition). Only one gene, Glyma.16G217700, showing robust circadian rhythm, which is indicated using the red curve. Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. Blue boxes under X-axis indicate 20 °C. ZT, Zeitgeber Time. a.u., arbitrary unit.
Figure S3.2. Changes to *Glyma.03G261800* promoter-luciferase activity upon SCN infection. (A) *Glyma.03G261800* promoter-luciferase activity in soybean hairy roots before and after mock (0.1 % agarose, blue line) and SCN (3,000 J2 SCNs, red line) infection at ZT52 (black arrow). Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. (B) Average expression changes before and after SCN infection. ****, p < 0.0001 (two-way ANOVA). The mean and the standard error of the mean are plotted. ZT, Zeitgeber Time. a.u., arbitrary unit.
Figure S3.3. The pie chart for all expressed soybean genes in RNA-Seq. 318 circadian rhythmically expressed genes are identified and indicated with red color. 36,807 genes are identified as non-circadian-rhythm genes and indicated with brown color. Circadian Rhythmic, circadian rhythmically expressed genes. Arrhythmic, non-circadian-rhythm genes.
Figure S3.4. The number of time-course infected SCNs of soybean hairy roots. Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. The mean and the standard error of the mean are plotted. ZT, Zeitgeber Time.
Table S3.1. Primers used in this study.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Name</th>
<th>Sequence</th>
<th>Usage</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glyma.03G261800</td>
<td>LCL2p_fwd</td>
<td>tggatagctggaagctggatgaATTATTTTTAAAAATAATTCTTAAGAAAA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LCL2p_rev</td>
<td>cgtctcatTTAGAAGACGTGCTGCTAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUC_fwd</td>
<td>aacccctctATGGAGGGCCTCCTGGCAAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>LUC_rev</td>
<td>gatttcggctTTACCGGCGATCTTCGCC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TEM_fwd</td>
<td>atctgagctggtcactgcattgcATTTTTTTCTTTTCTTTTTATG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TEM_rev</td>
<td>cgcgcgtactGTCCTCTAGACTGCTTTAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.03G261300</td>
<td>#9_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#9_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.01G162600</td>
<td>#10_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#10_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.03G177300</td>
<td>#11_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#11_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.16G217700</td>
<td>#12_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#12_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.14G210600</td>
<td>#17_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#17_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.14G210600</td>
<td>#14_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#14_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.10G230700</td>
<td>#15_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#15_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.10G285900</td>
<td>#16_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#16_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.16G032600</td>
<td>#17_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#17_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.17G062000</td>
<td>#18_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#18_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.06G196200</td>
<td>#19_Infusion_FWD</td>
<td>tgcgctgcctggaagctgacttagGGTTTAATTCTTCTTTGTTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#19_Infusion_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.03G261800</td>
<td>CCA1_2tss_Fwd</td>
<td>gtagccaccgcggcatTTAGAGAGGGCGCATCCT TTGGAGAAGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CCA1_2tss_Rev</td>
<td>tagttggtgcctcatTTTCACAAAGTTGCGGC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glyma.03G261800</td>
<td>CCA1_qPCR_FWD</td>
<td>AAGGGGAACGATGGACAGAGG</td>
<td></td>
<td>Overexpression</td>
</tr>
<tr>
<td></td>
<td>CCA1_qPCR_Rev</td>
<td>ACTTTTTTGCGTAGCTTCTG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.1 Mapping of Abiotic Stress Inputs to Soybean Circadian Clock

Plant circadian clock plays an important role in increasing fitness by synchronizing physiological processes with environmental oscillations (Greenham et al., 2015; Hsu et al., 2014). And studies in Arabidopsis demonstrate that various abiotic stresses influence circadian clock function (Sanchez et al., 2011). However, our understanding of the impacts of abiotic stresses on the soybean circadian clock is limited, which impedes the rational improvement of soybean circadian clock for enhanced fitness.

To unveil how abiotic stresses, such as drought, heat, iron deficiency, and alkaline stress, regulate soybean circadian clock, we devised a reliable and cost-effective pipeline and demonstrated its feasibility through mapping the abiotic stress “inputome” to soybean circadian clock. Specifically, we first built up the Molecular Timetable method in soybean and demonstrated its reliability. Then, using the time-indicating genes, we computationally surveyed the current available abiotic stress-related soybean transcriptome data and found that abiotic stresses caused profound changes to the global circadian rhythm. These results were further confirmed by RNA-Mediated Oligonucleotide Annealing, Selection, and Ligation with Next-Generation Sequencing (RASL-seq). In addition, we also identified specific circadian components modified by each stress. Specifically, we found that the interfaces between soybean circadian clock and abiotic stress signals were quite different from those in Arabidopsis. Furthermore, alkaline stress was identified as a circadian clock modulator for the first time in plants. Our findings highlight the greater complexity of soybean circadian clock than that in Arabidopsis. Our pipeline offers a broadly applicable and affordable tool for similar large-scale circadian clock studies in diverse species.
4.2 Future Direction for Impacts of Abiotic Stresses on Soybean Circadian Clock

Although circadian components targeted by each stress are identified, the mechanism of stress modulation of circadian genes is largely unknown. Future research should focus on revealing how abiotic stresses regulate soybean circadian genes.

We found that mild drought stress had little effect on the global circadian rhythm, and no statistically significant circadian rhythm change was observed for the expression of all 49 circadian clock genes surveyed. Based on these results, it is better to study the impacts of severe drought stress on the circadian clock.

The heat caused both phase and period changes to one Flavin-Binding Kelch Repeat F Box 1 (FKF1) ortholog, 1 CCA1 Hiking Expedition (CHE) ortholog, and 1 Pseudo-Response Regulator 7 (PRR7) ortholog in soybean, and many other PRR genes showed different parameter changes. Generation of soybean mutant lines would be important for further study. The knock-down and overexpression could be the choices. In addition, Virus-Induced Gene Silencing (VIGS) can be first used for the pretest.

Our results implied that the circadian responses to iron deficiency and alkaline stress are organ-specific. Soybean hairy root lines can be used to test this conclusion. Mutant soybean hairy root lines are much easier to generate than mutant soybean lines. So, for further study, it is better to focus on how these stresses modulate soybean root circadian clock. At the same time, researches on the interaction of the root clock and leaf clock might be very important to fully understand the mechanism.

4.3 The Interaction of Soybean Root Circadian Clock and Soybean Cyst Nematode (SCN)

The interaction between plant circadian clock and pathogens, like bacterial pathogens, oomycete pathogens, and insects, have been studied (Roden et al., 2009; Seo et al., 2015; Sharma et al., 2015; Spoel et al., 2014). However, almost all these discoveries were focused on leaf
circadian clock. Little is known what happens to the root circadian clock, and many microbes actually live near the roots in the soil. This represents a key knowledge gap impeding further enhancement of crop resistance against root pathogens. To fill this gap, I initiated the study of the interaction between soybean root circadian clock and SCN.

Using the Molecular Timetable method, I found SCN caused significant change to the global circadian rhythm in the feeding site. Then, I found that the expression rhythms of two circadian genes were modified by SCN using a time-course RASL-seq experiment as well as a living bioluminescence imaging assay.

To investigate how soybean root circadian clock regulates plant defense against SCN, I generated soybean clock mutant hairy root lines by overexpressing a *Circadian Clock Associated 1* (*CCA1*) gene. I observed a disrupted SCN infection rhythm of the mutant lines through a time-course infection experiment. In addition, the mutant lines are more resistant to SCN than the control lines. Then, using a time-course RNA-Seq experiment, I identified candidate regulators in the interaction of soybean root circadian clock and SCN.

**4.4 Future Plan for the Interaction of Soybean Root Circadian Clock and SCN**

I found that the expression rhythms of two circadian genes, one *CCA1* gene and one *Early Phytochrome Responsive 1* (*EPR1*) gene, were modified by SCN. SCN injects effector proteins into plant roots, and these effectors help SCN migrate in soybean roots, establish feeding site, and obtain nutrients (Gao et al., 2003; Gardner et al., 2018; Hamamouch et al., 2012; Hewezi et al., 2010; Juvale et al., 2018; Lee et al., 2010; Niblack et al., 2006; Noon et al., 2015). These effectors function in various ways, e.g., cell wall degrading enzymes help the nematode migrate; effector 10A06 plays an important role in suppressing plant defense system, specifically in the SA defense signaling pathway (Hewezi et al., 2010); CLE protein functions in cell-to-cell communications; and effector 19C07 helps auxin transportation (Hamamouch et al., 2012; Lee et al., 2010; Niblack
et al., 2006). There are more than 80 effector proteins identified in SCN so far (Gao et al., 2003; Gardner et al., 2018; Noon et al., 2015). Only a limited number of them have been functionally characterized (Juvale et al., 2018; Lee et al., 2018).

It is highly possible that these effector proteins also function to modulate soybean circadian clock. To figure out which effector regulates the circadian rhythmic expression of EPR1. I developed a hybrid Gateway vector carrying a Dexamethasone (DEX)-inducible SCN effector expression cassette and a luciferase reporter cassette driven by the promoter of EPR1 (Figure 4.1). 64 SCN effectors were cloned into this system, and 54 of the effector-vectors were transformed into soybean hairy roots for luciferase imaging. The effector-modulated EPR1 circadian rhythm change can be controlled by spraying of DEX and monitored by light intensity.

I also identified genes in the interaction of soybean root circadian clock and SCN using time-course RNA sequencing. These genes can be further verified using soybean mutant hairy roots. In addition, I also noticed that 2 genes in the Rhg1 (Resistance to Heterodera glycines 1) locus were up-regulated in CCA1 overexpression hairy root lines (not selected by our criterion) compared with that in control hairy root lines (Figure 4.2). One is Glyma.18G022400, encoding an amino acid transporter, and the other one is Glyma.18G022500, encoding an α-SNAP protein. In addition, a CCA1 binding site was found in the promoter (-360bp) of Glyma.18G022500. However, the circadian rhythmic expression of Glyma.18G022500 was not observed in soybean hairy roots (Figure 4.3 A). Interestingly, circadian rhythmic expression of Glyma.18G022500 was found in soybean seedlings (Figure 4.3 B). So, I proposed that in the root, CCA1 only regulates the expression levels of Glyma.18G022400 and Glyma.18G022500, while the expression rhythmicity of Glyma.18G022500 requires the aboveground soybean signals (Figure 4.4).
To further test this hypothesis, protein activities of Glyma.18G022400 and Glyma.18G022500 can be measured in both soybean hairy roots and soybean seedlings. Then, in vitro yeast one-hybrid and in vivo biotinylated pull-down associated with Mass Spectrometry can be used to figure out which proteins bind to the promoter of Glyma.18G022500. The candidate proteins are expected to have different activity patterns in soybean hairy roots and soybean seedlings.

4.5 References


### 4.6 Figures

Figure 4.1. The Dexamethasone (DEX)-inducible vector scheme. RB, the right border of T-DNA. LB, the left border of T-DNA. 35S, cauliflower mosaic virus 35S promoter. GVG, chimeric transcription factor containing GAL4 DNA-binding domain and VP16 transactivating domain, rat-GR (Glucocorticoid receptor), and HBD (hormone binding domain). E9T, the terminator of the pea ribulose bisphosphate carboxylase small subunit rbcS-E9. NosP, nopaline synthase promoter. HPT, hygromycin phosphotransferase gene. NosT, nopaline synthetase polyadenylation sequence terminator. Rbcs A3T, the terminator of pea rbc3-3A. ccdB cassette, gateway component. minimal 35S, CaMV 35S minimal promoter. 6xUAS, 6 x upstream activating sequence. TEM, terminator. LUC, luciferase coding sequence. Promoter inserted Early Phytochrome Responsive 1 (EPRI) promoter. The vector uses pTF101.1 as the backbone, and the Dex-inducible fragment is inserted into pTF101.1.
Figure 4.2. Expression of *Glyma.18G022400* and *Glyma.18G022500*. Expression of *Glyma.18G022400* (A) and *Glyma.18G022500* (B) in control (CK) and *CCA1* overexpression (OE) soybean hairy roots. The mean and the standard error of the mean are plotted. Unpaired Student's *t*-test is used. ****, *p* < 0.0005.
Figure 4.3. The time-course gene expression of *Glyma.18G022500*. Gene expression of *Glyma.18G022500* in soybean hairy roots (A) and soybean seedlings (B). Red boxes under X-axis indicate 26 °C. Cyan boxes under X-axis represent subjective 20 °C. The mean and the standard error of the mean are plotted. ZT, Zeitgeber Time. CKMock, control hairy root lines with mock infection. OEMock, overexpression hairy root lines with mock infection.
Figure 4.4. The proposed model. CCA1 along with other unknown factors regulates the expression of AAT and α-SNAP genes in the root. The circadian rhythmic expression of α-SNAP requires aboveground signals.