Control of Alternative Splicing by SICKLE/WARP2 is Required for Adaptation of the Plant Circadian Clock to Cool Temperatures

By Carine M. Marshall

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Plant Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Frank Harmon, Chair
Professor Sarah Hake
Professor Peter Quail
Professor Todd Dawson

Summer 2017

Abstract

Control of Alternative Splicing by SICKLE/WARP2 is Required for Adaptation of the Plant Circadian Clock to Cool Temperatures

by

Carine M. Marshall

Doctor of Philosophy in Plant Biology

University of California, Berkeley

Professor Frank G. Harmon, Chair

The circadian clock is a key mechanism for plants to anticipate and respond to daily and seasonal cycles of light and temperature. Cues of temperature and light cycles set the timing of circadian rhythms through the process of entrainment, yet the period length of rhythms remains constant because of temperature compensation. The molecular pathways that sense and respond to temperature are not well understood in plants. Recent studies in *Arabidopsis thaliana* suggest that alternative splicing of circadian clock transcripts is a fundamental mechanism that links ambient temperature perception with circadian clock rhythms. Splice variants of circadian clock transcripts are observed in response to temperature changes in the environment, but the source of splice variants and their effect on the circadian clock are not known. In addition, the core circadian clock genes *LATE ELONGATED HYPOCOTYL (LHY)* and *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* play an important role in modifying the circadian clock's response to temperature.

Here, we describe characterization of a new mutant allele of the Arabidopsis *SICKLE* (*SIC*) gene, *sic-3*, and the existing *sic-1* allele with respect to their effects on the Arabidopsis circadian clock. *sic* was first identified as, *warm acute response of PRR7 (warp2)*. *sic* has a unique temperature-specific defect in circadian clock function and in alternative splicing of circadian clock transcripts. *sic* exhibits low-amplitude or arrhythmic circadian rhythms under cool ambient temperature cycles, but not under light-dark entrainment. *sic* mutants also lengthen free running period in a manner consistent with impaired temperature compensation. *sic* mutant alleles accumulate splice variants of core circadian clock gene transcripts, including *LHY* and *CCA1*, which is exacerbated by cool temperatures. *sic* mutant alleles also modify transcription of circadian clock genes, particularly in cool temperatures. The *cca1 lhy* double mutant is epistatic to *sic*, indicating that control of *LHY* and *CCA1* splicing by SIC is needed for circadian clock function. Furthermore, the double and triple mutant combinations of *sic-3* with *cca1-1* and *lhy-20* indicate that CCA1 and LHY function redundantly to modify temperature compensation, and that this regulation is dependent on SIC in cool temperatures. Finally, expression of LHY protein is altered in *sic* mutants in a temperature-dependent manner.

SIC is a nuclear and cytoplasmic protein of unknown function, previously implicated in the control of alternative splicing, microRNA (miRNA) biogenesis, and stress responses (Zhan et al. 2012). SIC has homologous proline/serine rich proteins present only in angiosperms. The SIC protein interacts with spliceosome-associated proteins, including protein arginine methyltransferases PRMT4a and PRMT4b, as well as members of the NineTeen Complex (NTC), including the lariat debranching enzyme *DBR1*. The NTC is a multiprotein complex involved in transcription, transcription coupled DNA repair, and alternative splicing. This suggests that SIC maintains spliceosome efficiency at low temperatures through interaction with splicing factors, and this aspect of pre-RNA metabolism is critical for low temperature adaptation and for the circadian clock's response to temperature cues.

We propose the following model: alternative splicing converts circadian clock pre-mRNA transcripts to mature mRNA in a temperature-dependent manner, and the resulting splice variant pool determines the proper accumulation and timing of circadian clock proteins, especially LHY and CCA1. This form of post-transcriptional circadian clock regulation is crucial for temperature compensation and thermocycle entrainment in Arabidopsis. Where SIC fits into this model is unknown, but our data suggests two possibilities: 1) SIC modifies PRMT4a and PRMT4b activity, which then affects spliceosome activity, and/or 2) SIC modifies the NTC to regulate transcription, transcription coupled DNA repair, and alternative splicing. All of these processes are temperature-dependent, and it is clear that the activity of SIC is most important in cool temperatures. SIC's affect on *LHY* and *CCA1* alternative splicing appears to play a large role in regulating temperature compensation at cool temperatures. Discovery of the biochemical and molecular activities of SIC is certain to reveal important details about the functional link between alternative splicing and temperature responses within the Arabidopsis circadian clock, and holds the potential to provide new insights into the molecular processes involved in pre-mRNA processing.

This is dedicated to the sweet worldsong, singing itself in the dark.

“Where the trees are, the world is, and a sweet worldsong is singing itself in the dark.”
-Wendell Berry *(Remembering)*

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 4:

Acknowledgements

I would like to thank my advisor, Frank Harmon, for being a mentor, a teacher, an inspiration, and a friend. Thank you for believing in me since the first day, for giving me this beautiful thesis project, for the humorous and thoughtful discussions that made me excited for science, and, especially, thank you for treating me as an equal in mind and skill. I will never know a better advisor or work partner than you.

Thank you to the many PIs in Berkeley who mentored me over the years. Sarah Hake, Peter Quail, and Todd Dawson surely had many more important things to do, but they always gave wonderful feedback and support in my qualifying exam and thesis committees. Thank you to Patricia Zambryski for being my qualifying exam chair and PI mentor in my rotations. Thank you to Chelsea Specht for every step of her help in my education, including my undergraduate years, writing letters of recommendation for graduate school and fellowships, and for giving me the platform and advice to become a teacher.

I would also like to thank all members of the Harmon lab, from the members I never knew, such as Maritza Duarte, to the members I saw every day and who shaped my education. Claire Bendix was the welcoming leader when I first joined the lab, Virginia Tartaglio was the brilliant and talented undergraduate researcher who kick-started my thesis, and Dominica Rohozinski spent sleepless nights helping me collect tissue in timecourses; thank you all. In particular, thank you to Emma Kovak for being a dear friend and companion, intelligent lab mate, and for guaranteeing *warp2* is left in good hands. I would be remise to not thank all the members of the PGEC: Emilio Corona for always bringing a smile to our lab, China Lunde for bearing with my questions and helping me find last minute reagents, the many members of the USDA building that makes sure it runs smoothly, especially Rebecca the PGEC manager, Lil at the front desk, our janitor Jian Huang, and thank you to all the staff, PIs, postdocs, research assistants, and graduate students for help with protocols, finding reagents, and friendship.

So much of my work was accomplished by the many talented undergraduates who came through the Harmon lab. In particular, I would like to acknowledge Joyce Zhang, Jane Moon, Samantha Nguyen, and Virginia Tartaglio for their hard work.

Thank you to my wonderful cohort for the support, the friendship, and the learning. A particular thank you to Anna Josephson-Day, Kate Scheibel, and Becky Mackelprang for their dear friendship. Riva Bruenn was the leaf to my branch, and I could not have existed without her. Also, I would like to thank you Meagan Oldfather for helping me survive my first semester of teaching and studying for my qualifying exam.

Thank you to the PMB staff, past and present. A special thank you to Rocio Sanchez; no graduate student could survive without her, and I am forever indebted and thankful to

her help and smile. Also thank you to Dana Jantz for her very hard work and ability to make things happen.

Thank you to my family for everything. Without my parents I would not be a scientist, a farmer, or lover of plants. Thank you for all the love, support, and guiding over the years. Thank you to my brother for making me a big sister that had to earn the opportunity to be the person you look up to; eventually he will surpass me, and I could not be more pleased. Thank you to my family in France, California, and Australia for the support, even though they did not quite know what I was studying. Thank you to my wonderful in-law parents, sisters, and brothers for always being there for me. And thank you to all my friends, who always took the time to ask about my work and science, especially Anna Poteete, Shannon Burns, and Juli Steinbaum. Thank you to my housemates Rachna Ram, Sonia Fox, and Anya Fox for bringing joy to my life and making me not miss my home in Guinda.

Finally, thank you to my husband and best friend, Robert Hines, for everything. As you know, I could have accomplished all this without you, but never as well, or with as much joy. Thank you for challenging me intellectually, supporting me when I was low, teaching me that I am in school to learn, and for being the shade of my heart.

Curriculum Vitae - Carine M. Marshall

e-mail: carinemarshall@gmail.com

Education

University of California Berkeley: Plant and Microbial Biology Department **Aug 2012 – present**

Graduating Doctoral student in Plant Molecular Biology and Genetics.

University of California Berkeley: Plant and Microbial Biology Department **Aug 2005 – May 2009**

Bachelor's in Science for the study of Genetics and Plant Biology, College of Natural Resources.

Scientific Research & Professional Experience

UC Berkeley/USDA Plant Gene Expression Center: Berkeley, CA **Aug 2012 - Present**

- Doctoral graduate student researcher in the Plant and Microbial Biology Department, in Frank Harmon's lab. Investigation of the molecular pathway of temperature signaling to the *Arabidopsis thaliana* circadian clock through regulation of mRNA alternative splicing. Characterization of the novel mutant *sickle/warp2* using molecular biology and plant genetics to investigate the relationship between temperature and the plant circadian clock.

Sun Tracker Farm: Guinda, CA **Aug 2015 - Present**

- Co-owner/operator of Sun Tracker Farm, a 28-acre organic farm in the Capay Valley, specializing in annual fruits and vegetables such as melons, potatoes, tomatoes, eggplants, and leafy greens. Produce sold in the Napa Valley and East SF Bay area. In charge of all aspects of farm, including soil, water, and weed management, building infrastructure, accounting, marketing, harvesting, crop planning, and quality control of produce.

Golden State Bulb Growers: Moss Landing, CA **Apr 2010 - May 2011**

- Assistant manager and data analyst for large-scale field research on the growth and development of flowering bulbs/tubers (*Scilla peruviana*, *Zantedeschia aethiopica, Eucomis regia)* for GSBG, the largest international growers of flowering bulbs and tubers in the USA. Presented our research to the company board members and scientists, elucidating two years worth of data and its potential to drastically improve crop yields. Supervisor of bulb/tuber tissue culture production, quality control, and the training of new tissue culturists.

Combined Power LLC, formerly Advanced Lab Group: Santa Monica, CA **May - Oct 2009**

- Head Biologist of Combined Power LLC, formerly known as Advanced Lab Group, a company that designs Concentrated Solar Thermal systems. Previously, Advanced Lab Group investigated sustainable production of algae biodiesel in closed bioreactors. My role was to design and outfit the biology lab, as well as research and write original research protocols to test thermophilic algae and cyanobacterium viability for biofuels.

UC Berkeley: Berkeley, CA **Aug 2008 - May 2009**

- Undergraduate research assistant in the Algae biofuels lab of Anastasios Melis, investigating the efficiency of *Chlamydomonas reinhardtii* light-harvesting chlorophyll antenna. Quantified the Hydrogen and starch byproducts of *Rhodospirillum rubrum* as a potential source of bioenergy.

Suncrest Nurseries: Watsonville, CA **Intermittently throughout 2005-2009**

- Propagation worker in the horticulture department of Suncrest Nurseries, the largest retail nursery in native California ornamental plants. Propagated hundreds of ornamental plants for retail production.

Fellowships, Scholarships, Awards

American Dissertation Fellowship Alternate, American Association of American Women **Apr 2016**

NSF-GRFP Honorable Mention, NSF-Graduate Research Fellowships Program **Aug 2012**

Daniel Arnon Fellowship, National Academy of Sciences **Aug 2012**

Campus Outstanding Graduate Student Instructor, UC Berkeley GSI Teaching Center **Mar 2015**

Outstanding Graduate Student Instructor, Plant and Microbial Biology Department **Mar 2015**

First Place Poster Award, UC Berkeley Plan and Microbial Biology Departmental Retreat **Sep 2013**

Publications

- Ruijiao Xin, Ling Zhu, Patrice Salome, Estefania Mancini, **Carine M. Marshall**, Frank G. Harmon, Marcelo J. Yanovsky, Detlef Weigel, Enamul Huq, *SPLICING FACTOR FOR PHYTOCHROME SIGNALING promotes photomorphogenesis by regulating pre-mRNA splicing in Arabidopsis.* Proceedings in the National Academy of Science. **July 2017**.
- **Carine M. Marshall**, Virginia Tartaglio, Maritza Duarte, Frank G. Harmon, *The Arabidopsis* sickle *Mutant Exhibits Altered Circadian Clock Responses to Cool Temperatures and Temperature-Dependent Alternative Splicing.* The Plant Cell. **September 2016**; DOI:10.1105/tpc.16.00223
- Claire Bendix, **Carine M. Marshall**, Frank G. Harmon, *Circadian clock genes universally control key agricultural traits.* Molecular Plant. **2015** 8(8):1135-52; DOI: 10.1016/j.molp.2015.03.003

Professional Presentations

French Connection Plants Symposium: Stanford Carnegie Institute, CA**,** Invited Speaker **Jun 2016**

- "Unraveling WARP2: How alternative splicing regulates the Arabidopsis circadian clock's response to cool temperature"

Gordon Research Seminar: Gordon Conference in Holderness, NH, Invited Speaker **Jun 2016**

- "The Arabidopsis *warp2* mutant impairs circadian clock responses to temperature & temperature-dependent alternative splicing"

Department Seminar: Plant Gene Expression Center-USDA, Albany, CA, Speaker

- "How alternative splicing regulates the Arabidopsis circadian clock's response to cool temperature" **Oct 2016**
- "Working out the kinks in temperature sensing of the Arabidopsis circadian clock" **Mar 2016**
- "The search for a temperature signaling gene in the plant circadian clock" **Nov 2014**
- "The circadian clock mutant *warp2* exhibits defective temperature responses and alternative splicing" **Feb 2014**

UCSC Center for Agroecology and Sustainable Food Systems: UC Santa Cruz, Speaker **Sep 2015**

- "Plant molecular biology and its role in agriculture"

Plant Gene Expression Center Retreat: Marconi Center, CA, Poster Presentation **Dec 2015**

- "Unraveling WARP2: Working out the kinks in temperature sensing of the Arabidopsis circadian clock"

Diversity Break-out Session for Department Retreat: UC Berkeley, Moderator and organizer **Sep 2015**

- Question and group discussion on promoting diversity in the PMB department and STEM fields.

International Symposium on Plant Photobiology: Univ of Texas, Austin, TX, Poster **May 2015**

- "*WARP2* is necessary for temperature responses of the Arabidopsis circadian clock and alternative splicing"

UC Berkeley PMB Student/Postdoc Seminars: Berkeley, CA, Seminar Speaker

- "Working out the kinks in temperature sensing of the Arabidopsis circadian clock" **Dec 2015**
- "The circadian clock mutant *warp2* exhibits defective temperature responses and alternative splicing" **Feb 2014**

Plant and Microbial Biology Departmental Retreat: Berkeley, CA Poster Presentation **Sep 2013**

- "The search for a temperature signaling gene in the plant circadian clock"

Golden State Bulb Growers Board Meeting: Moss Landing, CA **Dec 2010**

- "Optimal conditions and timing for harvest, storage, planting, and growing of multiple cultivars of *Zantedeschia aethiopica* to decrease disease, and optimize sprouting and flowering uniformity"

Aquitaine France Chamber of Agriculture: Agen, France **Jun 2007**

- "The Pros and Cons of Transgenic Plants"

Teaching/Mentorship

Teaching Conference for Graduate Student Instructors (GSI),

Discipline-cluster workshop leader. **Aug 2016**

Undergraduate Research Mentor,

Mentor to seven undergraduate researchers (currently two). **May 2013 - Present**

Comparative Plant Morphology,

Laboratory and discussion section GSI. Aided in curriculum redesign. **Fall 2015**

Comparative Plant Morphology,

Laboratory and discussion section GSI. **Fall 2014**

Introduction to the Science of Living Organisms,

Laboratory GSI, and guest lecturer. **Spring 2014**

Leadership and Outreach Experience

Diversity Representative for the Plant and Microbial Biology Student Group **May 2015 - present**

- Graduate student representative to encourage diversity in STEM fields and the PMB department. Bring awareness and facilitate discussion, opportunities, and actions to create a more socially, racially, and gender diverse department. Student representative of the PMB Diversity Committee, in which university faculty target issues of racial and gender diversity in the PMB department.

Student Faculty Relations Committee for the UC Berkeley College of Natural Resources

Aug 2015 - present

- The committee considers and renders advice on matters concerning student activities, student problems, and student faculty relations. Maintains liaison with student organizations, student affairs officers, undergraduate advisers, and graduate advisers.

Graduate Peer Advisor for UC Berkeley Plant and Microbial Biology Department

Aug 2015 - present

- Graduate Peer Advisors create a comfortable community to address concerns of 1st and 2nd year PMB graduate students. We facilitate connections with senior PMB graduate students and provide access to campus and community resources.

Bay Area Scientists in Schools: Oakland, CA **Aug 2012 - present**

- Volunteer teacher in elementary schools throughout the East Bay to teach science experiments. Designed a lesson plan teaching elementary students how plants adapt different climates.

President of the UC Berkeley Plant Student Group **Aug 2013 – Aug 2014**

- President of the Graduate Plant Student Group on UC Berkeley campus. Organize outreach events with community and support for graduate student studies.

Laboratory and Research Skills

- **Plant Genetics & Biology:** recombinant genetics, plant crosses, Quantitative Trait Loci mapping, tissue culture, phenotypic scoring of different germplasms, seed sterilization and germination, *Agrobacterium*-mediated transformation, plant propagation and horticulture, greenhouse and field management, plant phylogenetics.

- **Molecular Biology:** DNA extraction (genomic and plasmid prep), RNA extraction, PCR and primer design, DNA sequencing and analysis, DNA restriction enzyme digest, DNA ligation, plasmid manipulation via construct design, DNA and protein electrophoresis, immunoblotting, quantitative real time PCR, reverse transcription PCR (RT-PCR), quantitative FAM RT-PCR, bioluminescence assays, circadian rhythm analysis with FFT-NLLS.
- **Project Management:** design and coordination of multi-year research projects while coordinating variable and control conditions, efficient and accurate data collection, statistical data analysis using programs such as R and Graphpad Prism, train and manage researcher teams for implementation of experimental conditions and data collection, ability to coordinate multiple field trials and research projects, coordination with field and greenhouse staff to run trials, time-management skills to ensure efficient use of time, team-worker.
- **Microbiology:** aseptic technique, transformation & cloning, preparation of competent bacterial cells.
- **Microscopy:** bright field, fluorescence microscopy.
- **Bioinformatics:** Basic skills in UNIX, R statistical computing, and Python. Experience in analysis of Next-Generation Mapping and RNAseq data.

Professional Skills & Attributes

- Native French and English speaker, fluent in Spanish. Dual citizen of the USA and France.
- Experience working in academic institutions and industry.
- Experience leading and managing small working groups with overall goals to advance larger project initiatives.
- Coordination and management of simultaneous projects with multiple contingent deadlines.
- Committed to accurate and reproducible data, which can only be attained by organized, well-planned, and detail-oriented experimental design. Highly organized.
- Coordination and facilitation of academic retreats, workshops, panels and informational sessions, with experience facilitating student and faculty focus group discussions.
- Ability to independently outline research strategies, and refine research strategies with a team or senior scientist.
- Able to read and keep up to date on scientific literature, novel tools, and technological findings in order to bring new ideas to a project and design new research techniques.
- Excellent oral and written communication skills and presentation expertise.
- Ability to troubleshoot and maintain a level head when delays or errors occur
- Honest, self-responsible, and forthright with potential negative results.
- Reliable, consistent, hardworking, and unafraid of time commitments.
- Physically experienced in farm labor and fruit and vegetable harvest.
- Proficiency in Microsoft Word, Excel and PowerPoint, Adobe Illustrator and Photoshop.
- Experience in accounting and budgeting from starting my own business/farm.
- Strong teaching and mentoring skills.

General Introduction

The rhythms produced by the endogenous circadian clock allow biological systems to anticipate and respond to daily and seasonal environmental cycles. The circadian clock regulates many fundamental processes in plants including metabolism, growth, development, and defense responses (Bendix, Marshall, and Harmon 2015; Greenham and McClung 2015). An appropriately functioning circadian system confers an adaptive advantage to *Arabidopsis thaliana* (Dodd et al. 2005; Rachel M Green et al. 2002). The clock is set, or entrained, by environmental cues that include light and temperature. The circadian clock can be conceptualized as a group of input pathways that signal and integrate environmental cues to the core circadian clock. The core circadian clock is a complex molecular system of genes and proteins that regulate each other to create a robust, self-sustaining, ~24 hour rhythm circadian clock. This circadian clock then regulates a multitude of output processes (Nohales and Kay 2016; Greenham and McClung 2015). Each of these three components (input pathways, the core circadian clock, output processes) can feed back on each other in order to respond to changes in the environment or the organism's needs to ensure maximal fitness.

The Arabidopsis molecular circadian clock comprises interlocked regulatory feedback loops that incorporate transcriptional, post-transcriptional, and post-translational control mechanisms (Nagel and Kay 2012; Nohales and Kay 2016; Fogelmark and Troein 2014; Hsu and Harmer 2014; Greenham and McClung 2015). Beginning at dawn of a single circadian cycle, clock-driven and light-stimulated expression of *CIRCADIAN CLOCK ASSOCIATED1 (CCA1)* and *LATE ELONGATED HYPOCOTYL (LHY)* rises (Schaffer et al. 1998; Z.-Y. Y. Wang and Tobin 1998). These MYB-like transcription factor genes belong to the larger *REVEILLE* (*RVE*) gene family (Chaudhury et al. 1999). CCA1 and LHY repress expression of *TIMING OF CAB EXPRESSION1* (*TOC1*) and the Evening Complex, a complex composed of *LUX ARRHYTHMO* (*LUX*)*, EARLY FLOWERING3* (*ELF3*)*,* and *EARLY FLOWERING4* (*ELF4*) (Nusinow et al. 2011). Dawn also promotes *RVE8* expression (Rawat et al. 2011), and afternoon accumulation of RVE8 activates expression of *PSEUDO-RESPONSE REGULATOR9* (*PRR9)*, *PRR5*, *TOC1*, *GIGANTEA (GI), LUX,* and *ELF4* (Farinas and Mas 2011; Rawat et al. 2011; Hsu, Devisetty, and Harmer 2013). The *NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED (LNK) 1* and *2* genes are expressed during the day and, similar to *RVE8*, activate afternoon and evening genes (Nohales and Kay 2016) by binding the *PRR5* and *TOC1* promoters. Together, the action of the LNK1/2 and RVE8, and sequential expression of *PRR9*, *PRR7*, and *PRR5* throughout the day suppress the morning transcriptional program (Matsushika et al. 2000; Farré et al. 2007; Rawat et al. 2011; Nakamichi et al. 2012). Evening accumulation of TOC1 maintains repression of *CCA1*, *LHY*, and earlier-expressed *PRR*s, and feeds back to repress its own expression (Gendron et al. 2012; W. Huang et al. 2012). The Evening Complex also represses expression of *PRR9, PRR7,* and *GI* at this time (Dixon et al. 2011; Helfer et al. 2011; Chow et al. 2012; Herrero et al. 2012), as well as its own expression through inhibition of *LUX* (Helfer et al. 2011). The LUX homolog, BROTHER OF LUX ARRHYTHMO (BOA, also known as NOX), is not redundant with LUX, but it also forms a complex with ELF3 and ELF4 to recruit the Evening Complex to the *PRR9* promoter (Dai et al. 2011). As dawn approaches,

reduced TOC1 and PRR5 activity, together with repression of *PRR9* and *PRR7* expression, allows *CCA1* and *LHY* expression to rise and begin the next circadian clock regulatory cycle.

Circadian clocks share three fundamental properties: entrainment by environmental cues, self-sustaining rhythms, and temperature compensation (Johnson et al. 2004). Entrainment is the capacity of the circadian clock to adjust the timing, or phase, of rhythms to synchronize daily internal processes with external cues, such as light and temperature (Salomé et al. 2005). The rhythms generated by the circadian clock persist with an ~24-hour period in free running conditions of constant temperature and continuous light or dark. Free running rhythms are temperature compensated to maintain a uniform period over a range of ambient temperatures (Pittendrigh 1954), in contrast to the general rule that the rate of enzymatic processes changes with temperature (Garcia-Viloca et al. 2004). Circadian biologists assess these three properties when determining that a gene is a fundamental part of regulating circadian rhythms.

Regulation within the circadian clock itself occurs at many levels, including chromatin modification, transcriptional, co-/post-transcriptional, translational, and post-translational, as well as at the level of tissue specific regulation (Hsu and Harmer 2014; Greenham and McClung 2015; Joon and Mas 2014; Nohales and Kay 2016). Of these, transcriptional control is best understood, but recent work elucidates other regulatory pathways contributing to plant circadian rhythms (Más and Yanovsky 2009). In fact, it is becoming apparent that each form of regulation on the molecular players that create circadian rhythms is interconnected, and often happens concurrently. This manuscript focuses on regulation at the transcriptional to post-translational level; therefore, the focus of this introduction is on these mechanisms.

Transcriptional regulation

Transcription is the process by which RNA polymerase creates a messenger RNA (mRNA) molecule from a section of a DNA template, otherwise known as a gene (Bentley 2005). This process is initiated by the binding of a transcription factor complex to a specific DNA sequence upstream of the gene coding sequence (CDS), and the subsequent RNA polymerase molecule (Bentley 2005). Transcription factors can also be repressors, and block the RNA polymerase molecule from binding to the promoter (Bentley 2005).

Transcription plays a dominant role in regulation of the circadian clock and downstream processes (Nohales and Kay 2016). Greater than 33% of the Arabidopsis transcriptome is directly regulated by the circadian clock (S L Harmer et al. 2000; Covington et al. 2008; T P Michael and McClung 2002), which includes regulation of many output processes including metabolism, growth, stress response, and hormone signaling (Greenham and McClung 2015). Regulation of circadian clock components at the transcriptional level is also evident in the current working model of circadian clock regulation. Many of the players acting within the core circadian clock are transcription factors that directly regulate other circadian clock genes to create appropriate circadian clock rhythms and responses, in addition to controlling the expression of output genes (Nohales and Kay 2016). For example, the evening element (EE) is a motif found in the promoters of many evening-phased genes including the *PRR*s, *GI*, *LUX*, and *ELF4* (Hsu

and Harmer 2014). RVE8 directly binds the EE to activate expression of these evening phased genes, while other RVE proteins, LHY and CCA1, bind the EE to repress expression of genes such as *TOC1* (Hsu and Harmer 2014).

Co-/post-transcriptional regulation

In order for an mRNA molecule to become mature, it must undergo a series of pre-mRNA processing steps, including 5' end capping, alternative splicing, and 3' end maturation by cleavage/polyadenylation (Bentley 2005). Pre-mRNA processing occurs co-transcriptionally, therefore the mechanisms that carry out theses processes are tightly regulated by a host of protein and RNA molecules, and the efficacy of regulation affects each step in this pathway and subsequent gene expression (Bentley 2005).

The first step in pre-mRNA processing is 5' end capping, which is the addition of a 7-methyl guanosine cap at the 5' end of the pre-mRNA molecule during elongation (Bentley 2005). During capping, splicing factors are recruited to the pre-mRNA molecule to begin splicing of the transcript to remove introns and join exon sequences (Bentley 2005). Transcript splicing is performed by the spliceosome, a large molecular machine composed of proteins and small nuclear ribonuclearproteins (snRNPs) (Matera and Wang 2014). Regulation of the spliceosome is through *cis*-acting splicing regulatory elements, which can either enhance or silence exonic or intronic splicing in the pre-mRNA (Matera and Wang 2014). The *cis*-acting splicing regulatory elements recruit *trans*-acting splicing factors to assist or inhibit the recognition of splice sites in the transcript (Matera and Wang 2014). Once the splice sites have been recognized by the spliceosome complex, it is believed that conformational changes between the complex and the pre-mRNA allow the snRNP to catalyze the removal of introns through the hyrodlysis of ATP (Matera and Wang 2014). Alternative splicing occurs when spliceosomes use different 5'- and 3'-splice site combinations to yield two or more structurally different transcripts, or splice variants, from the same gene (Matera and Wang 2014; Bentley 2005). The fate of these transcripts is dependent on the resulting open reading frame (Staiger and Green 2011).

Following alternative splicing and transcription, the pre-mRNA molecule is cleaved and polyadenylated resulting in mature mRNA with a 5' guanosine cap, a spliced mRNA, and a poly(A) tail (Bentley 2005). If the open reading frame does not contain a premature termination codons (PTC), the mRNA is sent for translation (Staiger and Green 2011). The stability of mRNA is dependent on the transcript, and turnover of RNA is an important step in gene regulation (Staiger and Green 2011), including regulation of circadian clock genes. Rhythmic transcription of circadian clock genes is dependent on the turnover of mRNA molecules throughout the day (Staiger and Green 2011). For example, degradation of *CCA1* mRNA is light-dependent, and the *CCA1* transcript is relatively stable in the dark. Since *CCA1* peak expression occurs at dawn, the subsequent turnover of the mRNA molecules in the light is important for rhythmic circadian rhythms (Staiger and Green 2011).

Transcript splicing is of particular interest to circadian clock biologists, since recent discoveries show that alternative splicing of transcripts from circadian clock genes is an important regulatory mechanism for circadian rhythms in Arabidopsis (Staiger and Green 2011; Cui, Xu, and Wang 2014; S. Filichkin et al. 2015) as well as in clocks in animals and fungi (A. Diernfellner et al. 2007; Chen et al. 2007). Mutations in

putative spliceosome- associated proteins interfere with Arabidopsis circadian clock function and cause accumulation of alternative splice variants of transcripts from circadian clock genes (Schlaen et al. 2015; Jones et al. 2012; X. Wang et al. 2012; Hernando et al. 2015; Sanchez et al. 2010; Marshall et al. 2016). These spliceosome component mutants and their circadian clock phenotypes will be further discussed in Chapter 2.

Translational and post-translational regulation

Following transcription and alternative splicing, the resulting transcript has the potential to affect translation depending on its sequence (G. Wang, Guo, and Floros 2005; Staiger and Green 2011). Splice variants frequently have alternate upstream open reading frames with PTCs that can lead to multiple fates for that transcript (Lareau et al. 2007; S. A. Filichkin and Mockler 2012). PTCs encountered during translation cause ribosome stalling which induces transcript degradation by nonsense-mediated decay (NMD) (Lewis, Green, and Brenner 2003; McGlincy and Smith 2008; Kalyna et al. 2012). NMD is a conserved mechanism throughout eukaryotes, including the proteins involved in this process (Syed et al. 2012). At the core of the NMD machinery are UPF factors, which target the transcript for degradation once a ribosome has stalled on a PTC (He and Jacobson 2015). In Arabidopsis, the NMD machinery recognizes transcripts to be targeted for degradation by the location of the PTC relative to the start codon, the length of the 3' untranslated region (UTR), as well as any overlapping of open reading frames in the 5'UTR (Kalyna et al. 2012). Some transcripts containing PTCs are not targeted for NMD, which suggests that these transcripts may affect translation of the pool of productive, protein encoding mRNA (Syed et al. 2012). Splice variants can also encode for truncated, partially functional, or inactive protein isoforms that interfere with full-length protein activity through the formation of differentially active or nonfunctional protein complexes (Seo et al. 2012; Syed et al. 2012; S. Filichkin et al. 2015).

Examples of interference with protein activity by the production of protein variants from alternatively spliced transcripts have been described for the Arabidopsis circadian clock component CCA1 (Syed et al. 2012; S. Filichkin et al. 2015; Hsu and Harmer 2014). A splice variant of *CCA1,* which includes a PTC due to the retention of intron 4, is not degraded by NMD (A. B. James et al. 2012). Instead the *CCA1 I4R* splice variant is translated to produce a truncated CCA1 protein that lacks the MYB-like DNA binding domain (Seo et al. 2012). The truncated CCA1 protein is capable of forming homo- and hetero-dimers with full-length CCA1 and LHY to interfere with their DNA-binding activity (Seo et al. 2012; S. Filichkin et al. 2015). Alternative splicing of circadian clock transcripts is often temperature dependent (A. B. James et al. 2012; Marshall et al. 2016), which can affect the activity and expression of circadian clock genes in different temperatures. As an example, CCA1 activity in the context of freezing tolerance is modulated through controlled production of *CCA1 I4R* (Seo et al., 2012). Warm temperatures sharply increase, and cold temperatures suppress, *CCA1 I4R* accumulation (James et al. 2012; Seo et al. 2012; Filichkin et al. 2015). Titration of CCA1 protein activity in this way ensures robust CCA1 activity in cold temperatures.

Beyond alternative splicing another crucial form of post-translational regulation within the clock is modulation of protein activity through protein-protein interactions and

protein turnover (Nohales and Kay 2016). An important regulatory step dependent on protein-protein interactions is the assembly of the Evening Complex: ELF3 is necessary to bring ELF4 and LUX together, while ELF4 targets the complex to the nucleus (Nusinow et al. 2011). The Evening Complex activates the transcription factor activity of LUX to control expression of its target circadian clock genes, including the *PRR9, PRR7,* and *GI* (Chow et al. 2012). Degradation of circadian clock proteins also is crucial for appropriate circadian regulation. TOC1 and PRR5 protein accumulation is controlled by the protein complex containing ZTL and GI (Demarsy and Fankhauser 2009; Ito, Song, and Imaizumi 2012). ZTL is an F-box protein that functions as a clock-specific blue light photoreceptor (Samach et al. 2000; Kim et al. 2007), and activation of ZTL by light promotes interaction with GI (Kim et al. 2007). In this complex, GI and ZTL mutually stabilize each other to prevent degradation, and ZTL controls the nucleocytoplasmic partitioning of GI (Kim et al. 2007). Once the ZTL-GI complex associates with TOC1 or PRR5, the PRR protein is degraded by the 26S proteasome (Mas et al. 2003; Kiba et al. 2007; Abe et al. 2008). Circadian clock proteins therefore play a role in regulating not only circadian clock output functions, but also the actual fine-tuning of the molecular circadian clock.

Clearly, the plant circadian clock relies on intertwined regulatory mechanisms ranging from control of transcription to modification of protein activity. Of these forms of regulation, the role alternative splicing plays will be of focus in this thesis, but one cannot consider alternative splicing without considering all the many other forms of circadian clock regulation. Furthermore, the exact mechanism and effect splice variants of circadian clock transcripts play within circadian clock function remains unclear. We know that regulation of alternative splicing is important to circadian clock function, and this is a temperature-dependent process, but little else. In addition, temperature plays an important role in regulating and entraining circadian rhythms, but little is understood of the mechanism that signals changes in temperature to the circadian clock, and how the circadian clock elicits a proper response. This thesis undertakes the study to understand these two phenomena and how they are connected. We describe characterization of a new mutant allele of the *SICKLE* (*SIC*) gene, *sic-3*, and the existing *sic-1* allele with respect to their effects on the Arabidopsis circadian clock temperature response, alternative splicing of circadian clock genes, LHY protein accumulation, and the potential mechanism of SIC in controlling these processes.

Chapter 1: The *warp2/sic* mutant has circadian clock phenotypes in response to cool temperature changes

The following text is modified from an article published in *The Plant Cell*, by Marshall *et al.*, 2016.

Abstract

Here, we describe characterization of a new mutant allele of the *SICKLE (SIC)* gene, *sic-3*, and the existing *sic-1* allele with respect to their effects on the Arabidopsis circadian clock. SIC is needed for the circadian clock to integrate cool temperature cycles to achieve entrainment and to produce temperature-compensated rhythms. *sic* mutant seedlings have impaired temperature compensation. In addition, exposure of *sic* to cool temperature cycles during entrainment produces low amplitude rhythms and causes lasting negative effects on clock period and rhythm amplitude. In contrast, *sic* has high amplitude rhythms with similar periods both during thermocycles of LL|28°-22°C and afterward in free running conditions of LL|22°C. On the other hand, *SIC* is not needed for entrainment in photocycle conditions, even under photocycles at LD|18°C. These observations indicate that circadian clock function is impaired under cool temperature cycles in *sic*. However, *sic* is not a general low temperature-sensitive mutant, since it does not exhibit significantly altered thermoresponsive flowering.

Introduction

Temperature plays an important role in regulating the fundamental properties of the circadian clock. Temperature cues can entrain circadian rhythms, even in the absence of light cues (Stacey L. Harmer 2009; Johnson et al. 2004). Two ways in which temperature interacts with the clock are through temperature compensation and temperature entrainment, but the molecular mechanisms that dictate these processes are not clear due to the inherent complexity of studying the circadian clock.

Most research on entrainment of the Arabidopsis circadian clock is focused on light entrainment. Light plays a very important role on setting and maintaining the pace of the circadian clock, and many processes such as transcription of circadian clock genes, mRNA stability, translation, and protein stability are affected by light (Nohales and Kay 2016) (see general introduction above). Despite all these examples of light's role in regulating the circadian clock, how this information is transmitted and used to entrain the clock is still not fully understood. To add to this complexity, the expression of photoreceptors is circadianly regulated (Fankhauser and Staiger 2002), and the circadian clock gates its response to light signals (Wenden et al. 2011; Li et al. 2011). Therefore light cues from the environment can be sensed by photoreceptors and used to set the circadian clock to the appropriate time of day, but the circadian clock also dictates when those photoreceptors function, and if the clock will respond or reset to light cues. Furthermore, there are a multitude of photoreceptors that function in the sensing of light cues, and light entrainment cannot be ascribed to a single molecule. One of the most obvious photoreceptors that affect the circadian clock is the aforementioned ZTL, a blue light photoreceptor and an important molecular player in the

core circadian clock (Demarsy and Fankhauser 2009; Ito et al. 2012; Kim et al. 2007). Other clock components are light sensitive even if they are not photoreceptors, such as CCA1 and LHY (see general introduction above), and allow another mechanism by which the circadian clock senses light (Staiger and Green 2011). The phytochromes (phy), red light sensing photoreceptors, also play a role in transmitting light signals to the circadian clock and determining period length (Nohales and Kay 2016). A well-defined pathway is phyB's interaction with ELF3 (Liu et al. 2001), which mediates ELF3's interaction with other light-signaling components (H. Huang et al. 2016). PhyB even interacts with core circadian clock components in a red- (R) and far-red- (FR) dependent manner (Yeom et al. 2017). The actual mechanism and downstream regulation of the circadian clock through these interactions is yet to be fully understood, but the interaction is indicative of phyB's role in light signaling to the circadian clock. Other photoreceptors such as the cryptochromes play a role in light signaling to the circadian clock, but their role is less defined (Nohales and Kay 2016).

Entrainment of the circadian clock to ambient temperatures represents a large gap in current knowledge of circadian clock biology. Studies to date indicate that temperature compensation and temperature entrainment are difficult to entirely separate. Originally the *prr7 prr9* double mutant was described as necessary for temperature entrainment to thermocycles, but it was later shown that *prr7 prr9* actually had a defect in temperature compensation (Salome, Weigel, and McClung 2010). To date, the core circadian clock genes *ELF3*, *LUX, PRR7, PRR9, CCA1, LHY,* and *GI* have been shown to interact between the oscillator and ambient temperature signals (K D Edwards et al. 2006; Gould et al. 2006; Salome, Weigel, and McClung 2010; Thines and Harmon 2010; Mizuno et al. 2014). Although our current knowledge indicates that these temperature sensitive genes are simultaneously regulated by light, there is also evidence that these pathways are separate. Michael et al. 2003 suggested the presence of two circadian oscillators that are differentially regulated by light or temperature (Todd P Michael, Salome, and McClung 2003). This indicates that there may be a separate pathway that is specifically sensitive to temperature and able to transmit this information to the clock.

Studies on the molecular mechanism of temperature compensation in Arabidopsis collectively indicate that certain circadian clock genes act to lengthen the circadian period at higher temperatures, and others shorten the circadian period at lower temperatures. Appropriate levels of *FLOWERING LOCUS C (FLC)*, *TOC1*, *PRR7, PRR9, GI, CCA1,* and *LHY* are needed for a plant's response to changes in ambient temperature (Edwards et al. 2006; Gould et al. 2006; Salome, Weigel, and McClung 2010). In Arabidopsis, mutants in the core circadian clock genes *PRR7, PRR9, CCA1*, and *LHY* exhibit altered temperature compensation behavior (Kieron D. Edwards et al. 2005; Gould et al. 2006; Salome, Weigel, and McClung 2010). Repression of *LHY* and *CCA1* by PRR7 and PRR9 is important for maintenance of the circadian clock period under varying temperatures (Gould et al. 2006; Salome, Weigel, and McClung 2010). The circadian clock in a *prr7 prr9* double mutant overcompensates at warm temperatures so that the period is long at 30°C and close to that of WT at 12°C (Salome, Weigel, and McClung 2010). Overcompensation in *prr7 prr9* depends on *LHY* and *CCA1*, since the addition of either a *cca1* or *lhy* mutant to *prr7 prr9* suppresses the long period at 30°C (Salome, Weigel, and McClung 2010). Temperature compensation

also requires *CCA1* at cool temperatures and *LHY* at high temperatures (Edwards et al. 2005; Gould et al. 2006). Despite these findings, an integrated understanding of the molecular mechanisms providing temperature compensation and temperature entrainment to the Arabidopsis circadian clock is still unclear.

Together this data indicates an overall lack of knowledge of the mechanism behind light and temperature signaling to the Arabidopsis circadian clock. In particular, temperature's effect on the circadian clock and how temperature cues are sensed are very poorly understood. This chapter characterizes the novel circadian clock mutant, *sic-3*, and an existing *sic-1* allele, in the context of its effect on circadian clock temperature responses. *sic* mutants exhibit diminished cool temperature responses of the Arabidopsis circadian clock in terms of temperature entrainment, maintenance of rhythms, and temperature compensation.

Results

warp2 is a mutant allele of *SIC* that alters circadian clock temperature responses

To identify genes necessary for correct temperature perception by the Arabidopsis circadian clock, ethyl methanesulfonate (EMS) mutagenized seedlings were screened for alterations in temperature-induced activation of the *PRR7* promoter (Figure 1A; Methods). In wild type (WT) seedlings, a 28°C temperature pulse of 3 hours generates a strong and reproducible induction the endogenous *PRR7* gene (Thines and Harmon, 2010). The temperature responsiveness of the *PRR7* promoter is faithfully reported by the *ProPRR7:LUC* transgenic reporter construct (Figure 1A). The *warm acute response of PRR7 2* (*warp2*) mutant was selected from the screen based on a weak response to the 28°C temperature pulse (Figure 1B). Subsequent experiments revealed that *warp2* lengthens circadian clock period, indicating that this mutant affects circadian clock function (see below).

The location of the mutation in *warp2* was mapped based on its long circadian period phenotype. Next Generation Mapping (Austin et al. 2011) employing whole genome resequencing of pooled genomic DNA from mutant individuals indicated the mutation was within a 400 kilobase region of chromosome 4 (Figure 2; Methods). The *warp2* long period phenotype segregated 1:3 in the F_2 mapping population, indicating *warp2* is a fully recessive allele in a single nuclear gene. Sequencing of candidate genes in the mapping interval identified the *warp2* mutation as a G → A transition at position +552 (relative to the transcriptional start site) in the AT4G24500 gene (Figure 3A). This mutation eliminates the GU pair of the 5'-splice site for intron 2 in AT4G24500 transcripts (Figure 3A), which causes missplicing of exon 2 to produce two different transcripts (Figure 3B,C). One transcript potentially encodes a protein with an internal deletion of 76 amino acids near the N-terminus, and the second potentially encodes a truncated protein made up of the N-terminal 61 amino acids (Figure 3D). In agreement with the mapping results, qPCR showed AT4G24500 transcript levels are substantially reduced in *warp2* (Figure 3E). The AT4G24500.1 transcript is the predominant transcript in WT (Figure 3A,B). AT4G24500 expression increases somewhat in the evening under photocycles, indicating potentially weak diurnal regulation (Figure 3E).

Two types of genetic tests confirmed that the mutation found in AT4G24500 is responsible for the *warp2* phenotype. First, F_1 seedlings from crosses between *warp2* and *sic-1* (*sickle-1*), a previously identified mutation in AT4G24500, maintain the long period phenotype characteristic of *warp2* (Figure 4A, Table 1). Further, *sic-1* has a long circadian period (Figure 4A; Table 1). *sic-1* is an EMS-induced G → A transition at position +1190 of AT4G24500 that changes Trp143 to a premature stop codon (Figure 3A-D), and it behaves as a recessive allele (Zhan et al. 2012). *warp2* has no obvious growth- or development-related phenotypes (Figure 3F,G), while *sic-1* plants have reduced rosette diameter, as well as serrated and downward curled leaves (Figure 3F) (Zhan et al. 2012). Second, transgenic introduction of a 2,296 base pair genomic region of AT4G24500 from WT into the *warp2* mutant background complemented the *warp2* period phenotype (Figure 4B; Table 1). These results conclusively demonstrate that the mutation found in AT4G24500 is the cause of the *warp2* long period phenotype and indicate that *warp2* is a mutant allele of *SICKLE (SIC)* and, therefore, will be referred to as *sic-3*.

sic mutation lengthens the circadian clock period

Compared to WT, *sic-1* and *sic-3* exhibit longer periods of *ProPRR7:LUC* rhythms after release into free running conditions of constant light and constant 22°C (LL|22°C) from either LL|22°-18°C thermocycle (constant light and 12-hour long cycles of 22°C and 18°C) or LD|22°C photocycle (constant 22°C and 12-hour long cycles of light and dark) entrainment (Figure 5; Figure 6A,B). Similarly, according to qPCR, peak expression of eight core circadian clock genes (*PRR7, CCA1, LHY, PRR9, TOC1, GI, ELF3,* and *LUX*) and three output genes (*CAB [CHLOROPHYLL A/B BINDING PROTEIN], CAT3 [CATALASE 3],* and *CCR2 [COLD-CIRCADIAN RHYTHM-RNA BINDING 2]*) is delayed in *sic-3* seedlings during the 24-48 hours (Zeitgeiber Time (ZT) 24-ZT 48) after release into LL|22°C (Figure 7; Figure 8A,B) following either thermocycle or photocycle entrainment conditions. Quantitative estimation of period length for *ProPRR7:LUC* rhythms with FFT-NLLS (Fast Fourier Transform Non-Linear Least Square) curve fitting analysis (Plautz et al. 1997) confirmed that the period for *sic-3* and *sic-1* is an average of 1 hour and 2 hours longer, respectively, than WT (Table 2; Figure 5B,D). Thus, *sic* mutants have altered circadian clock periods.

sic mutation alters circadian clock temperature responses

ProPRR7:LUC rhythms were assessed in *sic-3*, *sic-1*, and WT during LL|22°-18°C thermocycles and LD|22°C photocycles to assess circadian rhythms during entrainment. Individual *sic-3* and *sic-1* seedlings lacked clear rhythms for *ProPRR7:LUC* expression during five days under LL|22°-18°C (Figure 4A,B; Figure 6A,C). Approximately 40%-50% of *sic-1* and *sic-3* individuals had arrhythmic *ProPRR7:LUC* expression (Figure 9B; Figure 6C; Table 2), defined as *ProPRR7:LUC* expression waveforms with a relative amplitude error (RAE) value >0.6 from FFT-NLLS analysis (McWatters et al. 2000). In addition, periods for rhythmic individuals differed by as much as 12 hours between *sic* individuals (Figure 9B), which is reflected by large standard deviations for the mean period of *sic-1* and *sic-3* (Table 2). Moreover, under LL|22°-18°C, the time at which peak expression of *ProPRR7:LUC* occurs, or phase, was also highly variable across the *sic* populations (Figure 6C; Table 2). By comparison,

populations of WT seedlings under LL|22°-18°C exhibited few arrhythmic individuals, and rhythms were in close agreement in terms of both period length and phase (Figure 9A,B; Figure 6C; Table 2). In contrast to thermocycle conditions, *sic-1* and *sic-3* seedlings under LD|22°C had robust *ProPRR7:LUC* rhythms similar to those of WT in terms of both period and phase, with >95% of the individuals exhibiting rhythmic behavior (Figure 9C,D; Figure 6B,E; Table 2).

To determine whether cycling temperature, and not overall cool temperature, was the cause of arrhythmia in *sic*, rhythms for WT, *sic-3*, and *sic-1* seedlings were tested under LD|18°C. All three genotypes showed comparable period length and no arrhythmic individuals (Figure 10A,B; Table 3). Rhythms of *sic-3* and *sic-1* seedlings entrained under warm LL|28°-22°C thermocycles were comparable to those of WT (Figure 10C,D; Table 3). These observations show that *sic* interferes with the establishment of circadian rhythms only when 1) temperatures cycle and 2) thermocycles are cool.

To confirm that *ProPRR7:LUC* rhythmicity in *sic-3* individuals during thermocycles reflected general clock dysfunction, the expression level and waveform for eight core circadian clock genes and three output genes over a 24-hour time course were evaluated by qPCR in *sic-3* and WT exposed to LL|22°-18°C and compared to behavior under LD|22°C (Figure 11; Figure 8C,D). Under LL|22°-18°C, *PRR7* transcript expression in *sic-3* was characterized by a low and broad peak of expression compared to WT, together with high variability at each time point (Figure 11A). The *sic-3* mutation similarly changed the expression profiles of *LHY*, *TOC1*, (Figure 11B,C) *CCA1*, *PRR9*, *GI*, *ELF3*, *LUX, CAB, CAT3* and *CCR2* (Figure 8C). Low-amplitude rhythms and variable transcript expression in pooled *sic-3* seedlings indicated the same variable behavior observed for *ProPRR7:LUC* activity under LL|22°-18°C (Figure 9A; Figure 11A). Under LD|22°C, however, the expression of the same genes was comparable between *sic-3* and WT (Figure 11D-F; Figure 8D).

sic mutation causes sensitivity to cool temperature cycles

To better understand the nature of the temperature signaling defect in *sic*, the next experiments evaluated whether *sic* was impaired in perceiving specific ambient temperatures (*i.e.*, cool vs. warm temperatures). Regardless of prior thermocycle composition, *sic-3* and *sic-1* mutants always had longer and more variable free-running periods at 22°C than WT (Figure 12A-C; Table 3), as indicated by the larger standard deviation for mutant populations. Exposure of *sic* seedlings to cool thermocycles during entrainment (LL|22°-12°C, LL|22°-16°C, LL|22°-18°C) accentuated the variation in period across the population, as indicated by standard deviation (Figure 12A,B; Figure 5B; Table 3). In addition, *sic* individuals were more likely than WT to exhibit arrhythmic *ProPRR7:LUC* expression in free-running LL|22°C conditions after cool temperature cycle entrainment: 13% of *sic-3* and *sic-1* seedlings were arrhythmic after LL|22°-12°C entrainment, while <1% of mutant seedlings were arrhythmic following LL|28°-22°C entrainment (Table 3). Thus, exposure of *sic* to cool thermocycles during entrainment had lasting negative effects on the period and amplitude of free-running circadian rhythms. These results indicate that the Arabidopsis circadian clock requires *SIC* activity to appropriately respond to cool temperature cues.

***sic* interferes with circadian clock temperature compensation**

A potential explanation for the imprecise rhythms in *sic* during cool thermocycles and the long period under free running conditions was the presence of an underlying problem with temperature compensation. To test temperature compensation in WT, *sic-1*, and *sic-3* seedlings, free-running period was determined at constant 28°C (LL|28°C), 22°C (LL|22°C), and 16°C (LL|16°C), after either LL|22°-18°C or LD|22°C entrainment. Following either type of entrainment, the free-running period of WT exhibited temperature compensation, while those of *sic-3* and *sic-1* did not (Figure 12D; Figure 13; Table 2). The period of *sic-3* was directly correlated with ambient temperature, so that the mean period under LL|16°C was 2-3 hours longer than that under LL| 22°C, and the mean period under LL| 22°C was 1 hour longer than that under LL|28°C. In fact, the period of *sic-3* under LL|28°C approached that of WT, but remained significantly different from WT (Figure 12D; Figure 13A,F,G; Table 2). The average period of a *sic-1* population under LL|16°C was 3-5 hours longer than that under LL| 22°C (Figure 12D; Figure 13A-C; Table 2). However, *sic-1* responded differently to LL|28°C: its average period at this condition was similar to that under LL|22°C (Figure 12D; Figure 13A,D-G; Table 2). Together, these findings indicate that both *sic* alleles interfere with the circadian clock temperature compensation mechanism.

In addition, *sic* individuals entrained with cool thermocycles were more likely to be arrhythmic under cool temperature free-running conditions that under warm conditions. When entrained with LL|22°-18°C thermocycles, 33% of *sic-3* and 15% of *sic-1* seedlings were arrhythmic under LL|16°C (Figure 13B; Table 2); on the other hand, fewer than 4% of individuals from either mutant genotype were arrhythmic when released into LL|28°C conditions (Figure 13F; Table 2). Each *sic* allele continued to show broad dispersion in period length across the population regardless of the temperature used for free-running conditions, which is clear from the larger standard deviation values for mutant populations compared to WT (Figure 13; Table 2). Thus, a combination of cool temperature thermocycles and cool free-running conditions promotes arrhythmicity in *sic*, and warm temperature free-running conditions suppress this arrhythmicity.

Discussion

The *warp* screen intended to further understand the Arabidopsis circadian clock's relationship with temperature. The *warp2* (*sic-3*) mutant is a unique temperature entrainment mutant because of its compromised rhythms only in thermocycles and not photocycles (Figure 9; Figure 11). Identifying the mutated gene in *warp2* (*sic-3*) as *SIC* did not provide obvious clues as to the pathways controlling temperature entrainment. In fact, the exact mechanism by which SIC affects the circadian clock in cycling temperature conditions remains incompletely understood, but preliminary work and hypotheses in this area are discussed in Chapter 4.

The *sic-1* allele is a preexisting mutant identified by others (Zhan et al. 2012). Compared to *sic-3*, *sic-1* has a more severe clock phenotype and pleiotropic effects on growth and development not seen in *sic-3* (Figure 3F; Figure 6). Two additional alleles of *sic*, *ron3-1* and *ron3-2* were identified since our discovery of *sic-3* (Karampelias et al. 2016). These two alleles exhibit reduced apical dominance, primary root length, and

lateral root emergence (Karampelias et al. 2016). This study concludes RON3/SIC is involved in intracellular movement of the phytohormone auxin (Karampelias et al. 2016). The *sic-1* allele and two *ron3* alleles show similar growth and development phenotypes (Zhan et al. 2012; Karampelias et al. 2016). Interestingly, the mutations in *sic-1* and the two *ron3* alleles occur in exon 3: *sic-1* and *ron3-1* are PTC mutations while *ron3-2* is a T-DNA insertion (Zhan et al. 2012; Karampelias et al. 2016). In contrast, the *sic-3* mutation lies at the 5' splice site junction of exon 2 and intron 2 (Figure 3C). As a consequence, the *SIC* transcript in *sic-3* is alternatively spliced, which produce two different mutant transcripts: band A that introduces a PTC early in the transcript, and band B that encodes for a protein with an internal deletion of 76 amino acids (Figure 3D). The deleted portion of SIC in band B is highly conserved (see Chapter 4 for more detail) and it is predicted to be a protein-binding domain. We propose that the differences in phenotype between *sic-3* and *sic-1*, and possibly the *ron3* alleles, may be due to production of a partially functional SIC protein derived from the band B transcript.

Collectively, these studies connect SIC to development, abiotic stress, microRNA biogenesis, RNA splicing, auxin signaling, and circadian clock temperature responses (Zhan et al. 2012; Karampelias et al. 2016; Marshall et al. 2016). SIC is localized both in the nucleus and the cytoplasm (Zhan et al. 2012; Karampelias et al. 2016), implying that separate functions could be carried out in different parts of the cell. Different domains of the SIC protein also may be involved in its diverse set of activities. If the *sic-3* allele does accumulate a partially active protein, then the section missing from this protein may be important for the circadian clock functions of SIC but not its other activities. Ultimately, *sic-3* decreases expression of the *sic* transcript (Figure 3E) and causes a circadian clock phenotype consistent with *sic-1* (Figure 6).

The hallmark circadian clock phenotype of *sic* mutant seedlings is low amplitude or arrhythmic expression of core circadian clock genes under cool thermocycle entrainment conditions (Figure 6; Figure 8; Figure 9; Figure 5). Also, circadian rhythm expression across a *sic* mutant population is disorganized and incoherent (Figure 9). Importantly, these clock-associated phenotypes are not observed in warm thermocycles (Figure 10). Furthermore, *sic* mutants only show the poor and incoherent rhythms phenotypes when temperatures cycle. In fact, circadian rhythms in *sic* mutants are organized, albeit long period, in constant temperatures of 16°C, 22°C, and 28°C (Figure 7; Table 2). Thus, SIC is important for control of circadian rhythms when 1) temperature cycles, and 2) temperatures are below 22°C. The mechanisms of processing cycling temperature cues in the 22°-28°C range appears not to require SIC. Potentially, separate pathways are activated to process different temperature ranges.

Comparison of thermocycle- vs photocycle-entrained *sic* seedlings also show differences with WT in rhythm quality (RAE), and phasing during and after entrainment. In the course of thermocycle entrainment, the population of *sic* seedlings is varied compared to WT in rhythm quality and phase (Figure 6C). In contrast photocycle-entrained seedlings match across genotypes in rhythm quality and phasing, and it can be concluded that *sic* plants correctly entrain to photocycles (Figures 6E). Upon release into free run at LL|22°C, the phase of photocycle-entrained seedlings is on average two hours earlier than thermocycle-entrained seedlings for WT and *sic-3,* and *sic-3* seedlings have an approximately one hour phase delay compared to WT (Figure 6D vs. Figure 6F). *sic-1* defaults to a late phase near ZT 14 in both entrainment conditions

(Figure 6D vs. Figure 6F). This indicates several points: 1) thermocycle-entrained WT seedlings have a later phase than photocycle-entrained WT seedlings, perhaps because it takes more time for the temperature cue to be processed in thermocycle-entrained seedlings, 2) regardless of entrainment conditions *sic* mutants display a phase delay in free-running conditions due to their long period, and 3) because the phase of *sic-3* adjusts closer to WT in the different entrainment conditions, and *sic-1* defaults to a late phase, then *sic-3* seedlings may undergo a level of entrainment even during thermocycles.

Exposure to cool thermocycles during entrainment has a lasting negative effect on the quality of circadian rhythms in free-running conditions (Figure 4; Figure 12; Table 3). The colder the temperature range of thermocycles, the worse the quality of *sic* circadian rhythms will be in free run (Figure 12). Therefore, the function SIC carries out to correctly transmit and process temperature cues to the circadian clock determines how well the plant will entrain to its environment. Studying the molecular pathway of SIC in regulating circadian rhythms will elucidate the pathways that underlie temperature entrainment in the circadian clock.

SIC is not needed for photocycle entrainment under any temperature condition. For example, *sic* mutants show coherent and correctly entrained circadian rhythms under LD|18°C conditions (Figure 10). Therefore, the pathways for photocycle entrainment are likely separate and override the temperature entrainment pathways that rely on SIC. Perhaps this evidence supports the findings of Michael et al. 2003, and SIC plays a role in regulating the temperature-regulated circadian oscillator (Todd P Michael, Salome, and McClung 2003). This data also indicates that *sic* is not an overall low temperature-sensitive mutant, since it is possible for circadian rhythms to be normal in cool temperature conditions. If *sic* mutants were compromised in all aspects of processing low temperature, then the circadian rhythms of *sic* under LD|18°C would be expected to have characteristics similar to those in thermocycles. Also, temperature-related processes outside of the clock appear intact in *sic* mutants, including thermoresponsive flowering (Figure 3G).

sic mutants alter circadian clock temperature compensation, particularly in the cold (Figure 13; Table 2). After entrainment, when *sic* is released into LL|16°C, circadian period, rhythm quality, and period reproducibility across a population, is highly compromised compared to *sic* seedlings released into LL|28°C (Figure 13C-H). This behavior reinforces the conclusion that SIC is important for circadian rhythms in cool temperatures. Furthermore, the defect in temperature compensation provides insight into the underlying cause of poor rhythms under thermocycle entrainment. If *sic* plants have difficulty maintaining period length in different temperatures, then the changes in pace caused by temperature transitions under entrainment conditions are likely to produce incoherent and arrhythmic rhythms. Findings by Salome et al. 2010 support this idea, as they too found defects in thermocycle entrainment and temperature compensation in *prr7 prr9* mutants (Salome, Weigel, and McClung 2010).

The *sic* thermocycle-specific phenotype of *sic* mutants is unique among circadian clock mutants tested in this condition. Other circadian clock mutants that impair temperature compensation may also display thermocycle-entrainment phenotypes, but this remains to be investigated for most of the known circadian clock mutants (Schlaen et al. 2015; Hong et al. 2010; Sanchez et al. 2010; X. Wang et al. 2012; Jones et al.

2012; Gould et al. 2006). Study of *sic* mutants demonstrated the importance of SIC for processing temperature cues in the Arabidopsis circadian clock. Future in depth characterization of the molecular function of SIC is certain to provide valuable information on the regulatory pathways that dictate these processes for the circadian clock.

Materials and Methods

Plant Material and Growth Conditions

All experiments used the *Arabidopsis thaliana* Columbia-0 accession, except for the *warp2* mapping population created by an outcross to the Landsberg erecta-0 accession. All lines carried a *ProPRR7:LUC* reporter construct that consisted of the *PRR7* promoter driving expression of firefly *luciferase* (Salome and McClung, 2005b). *sic-1* seeds were obtained from The Arabidopsis Information Resource (TAIR) and this mutant is described in (Zhan et al., 2012). Crossing of *sic-1* to *sic-3 ProPRR7:LUC* generated plants homozygous for *sic-1* and the *ProPRR7:LUC* construct.

For all experiments, seeds were surface sterilized by 10 minute treatment with a solution of 30% bleach and 0.01% SDS, followed by extensive washing with sterile water, and were then sown on MS plates composed of 1X Murashige and Skoog basal salt medium (pH 5.7-5.8) with 0.8% Type I micropropagation agar (Caisson Labs). After stratification in the dark at 4°C for 3 days, the plates were transferred to constant light and 22°C to promote germination. After 3 days, germinating seedlings were transferred to entrainment for 5 days in the indicated light and temperature conditions; all entrainment treatments were divided into alternating 12-hour cycles of either light:dark or different temperatures. Free-running conditions were always constant light and the indicated temperature. Surface sterilization of seeds and all transfers (*i.e.*, germination, entrainment, and free run) occurred at ZT 0. For RNA expression analysis, seedlings were grown in closely spaced groups on MS plates. Conditions for flowering time experiments were 12 hours light and 12 hours dark, along with the indicated constant temperature. Total number of rosette leaves was counted when the inflorescence was 1 cm tall. Cool white fluorescent bulbs supplied light at 50 $\mu mol \cdot m^{-2} \cdot sec^{-1}$ in growth chambers from Percival Scientific Inc.

Bioluminescence Assays

Each MS plate was sprayed with 1 mL of 5 mM firefly luciferin (Biosynth; Gold Biotechnology) prepared in 0.01% (vol/vol) Triton X-100 (Sigma Aldrich) applied 24 hours before imaging with a luciferase imaging system built by BioImaging Solutions, employing an ORCAII camera (Hamamatsu Photonics) housed in a Percival incubator (Percival Scientific Inc.) for temperature control. Each experiment had 3-9 plates together in the camera. Halogen bulbs provided 50 $\mu mol \cdot m^{-2} \cdot sec^{-1}$ of white light. Seedlings were imaged every 2.5 hours, and bioluminescence of individual seedlings was collected using MetaMorph software (Molecular Devices). Bioluminescence data from images were extracted with MetaMorph software and compiled in Microsoft Excel (Microsoft Corporation) with the Biological Rhythms Analysis Software System 3.0 (BRASS), an Excel workbook for the analysis of rhythmic data series (Locke et al.,

2005; Southern and Millar, 2005). Within BRASS, Fast Fourier Transform-Nonlinear Least Squares (FFT-NLLS) was used to estimate circadian period, RAE, and phase values from the rhythmic bioluminescence data (Plautz et al., 1997). Phase values were corrected to a 24-hour time scale by dividing the FFT-NLLS calculated phase value of each seedling by its estimated period and multiplying this value by 24. Corrected phase values were plotted with a modified version of an R script employing the "polar.plot" function written by Dr. Michael Covington (Harmer and Kay, 2005).

warp2/sic-3 Mutant Discovery.

Col-0 *ProPRR7:LUC* seeds were mutagenized with EMS. The M_2 seeds were germinated for 3 days and entrained under LL| 22:18°C for 4 days. The M_2 seedlings were then transferred to DD| 22°C in the camera and administered a 28°C pulse 28 hours later. 5,328 M_2 seedlings were screened for a fold-change (FC) in bioluminescence lesser or greater than twice the standard deviation of WT mean (FC = signal at 31 hours)/signal at 28 hours) (Figure 1A). Putative M_2 seedlings were rescreened twice, and the phenotype was confirmed in the M_3 generation. One mutant with attenuated FC was designated *warp2*/*sic-3* (Figure 1B).

Mapping of _warp2/sic-3_.

warp2l/sic-3 was mapped with Next Generation EMS Mapping (Austin et al., 2011), using the default settings (Figure 2). The mapping population consisted of the F_2 generation from *sic-3 ProPRR7:LUC* in the Col-0 accession backcrossed to WT plants of the Ler-0 accession. A pool of genomic DNA from 80 F_2 individuals with the *warp2*/*sic-3* phenotype was single end sequenced on a single lane of Illumina HiSeq2000 with 50 base pair read length by the QB3 Vincent J. Coates Genomics Sequencing Laboratory at University of California, Berkeley (http://qb3.berkeley.edu/qb3/gsl). Genomic DNA was randomly sheared genomic by Covaris S2 and used to prepare a sequencing library with the Apollo 324 NGS Library Prep System (WaferGen Biosystems). The average library size was 482 base pairs. Reads were mapped to the Arabidopsis TAIR10 genome sequence with bwa (Burrows-Wheeler Alignment Tool) (Li and Durbin, 2009) using the "index" command (-a bwtsw option), followed by the "aln" command (with default options), and alignments generated in the SAM (Sequence Alignment/Map) format with the "samse" command (with default options). Alignments were converted to the BAM format with SAMtools (Li et al., 2009; Li, 2011). This BAM file is available at the NCBI Sequence Read Archive with BioProject accession PRJNA314711. The mpileup command (-E -ugf options) in SAMtools followed by BFCtools was used to create a VCF (Variant Call Format) file from the BAM file for input into the Next Generation EMS Mapping pipeline (Austin et al., 2011). Candidate genes in the *sic-3* mapping interval were Sanger sequenced at the UC Berkeley DNA sequencing facility (mcb.berkeley.edu/barker/dnaseq/home) to confirm predicted EMS-induced mutations.

Construction of Transgenic Plants

Phusion High-Fidelity DNA Polymerase (New England Biolabs) was used to amplify the desired PCR product with the appropriate primers (Supplemental Data Set 2). PCR products were cloned into the pENTR/D-TOPO vector (Thermo Fisher Scientific) and

transformed into One Shot TOP10 Chemically Competent *E. coli* (Thermo Fisher Scientific). Binary vectors for plant transformation were generated by site-specific recombination with Gateway LR Clonase II enzyme mix (Thermo Fisher Scientific). Binary constructs were transformed into *Agrobacterium* strain GV3101 by electroporation and *Arabidopsis* plants were transformed by the floral dip method (Clough and Bent, 1998). T_1 transformants were selected on 1X MS plates (pH 5.7-5.9) supplemented with either 50 µg/mL kanamycin (Sigma Aldrich) or 16 µg/mL phosphinothricin (Gold Biotechnology). Transgenic lines were genotyped by PCR with appropriate primers (Supplemental Data Set 2).

For the complementation tests, a 2,296 base pair amplicon, which included the full *SIC* (AT4G24500) gene and 134 base pairs of sequence upstream encompassing the predicted promoter region, was PCR amplified from WT genomic DNA and cloned into the pMDC100 binary vector (Curtis and Grossniklaus, 2003). T_1 transformants in the *sic-3* background were identified as kanamycin resistant individuals.

Gene Expression Analysis with qPCR

Seedlings were grown as described. For analysis under entrainment conditions, seedling tissue was collected on day 5 of entrainment (8 days-old) at ZT 0, ZT 3, ZT 6, ZT 9, ZT 12, ZT 15, ZT 18, and ZT 21 in the indicated entrainment conditions. For analysis in free running conditions, seedlings were transferred from entrainment to LL|22°C free run on day 4 (7 days-old) and tissue was collected on day 2 of free run (8 days-old) at ZT 24, ZT 27, ZT 30, ZT 33, ZT 36, ZT 39, ZT 42. After collection, the tissue was immediately placed in liquid N_2. Total RNA was extracted with Plant RNA Reagent according to the manufacturer's recommendations (Thermo Fisher Scientific). Genomic DNA was removed from total RNA with the Ambion TURBO DNA-*free* Kit (Thermo Fisher Scientific). First-strand cDNA was synthesized from 10 µg of total RNA with oligo-dT primers and the Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). qPCR was performed for 3 independent experiments with a CFX Real-Time system (Bio-Rad Laboratories) as described previously (Harmon et al., 2008). Transcript levels of target genes were calculated from the average C_q of two technical replicates using the equation $2^{[C_q(\text{normalization control})-C_q(\text{experimental})]}$, where C_q for each amplification curve was calculated by the "Regression" mode in Bio-Rad CFX Manager 3.0 software (Bio-Rad Laboratories). The normalization control was the geometric mean of C_q values for *IPP2* (AT3G02780) and *PP2A* (AT2G42500), calculated as $\sqrt{C_q(\text{AT3G02780})*C_q(\text{AT2G42500})}$. Primers sequences are shown in Data File 2.

Chapter 1 Figures

A

Screen for *warp* (*warm acute response of PRR7*) mutants. Evaluate M2 generation of EMS mutagenized seedlings for changes (± 2 SD) in *ProPRR7:LUC* response to 3-hour pulse of 28°C:

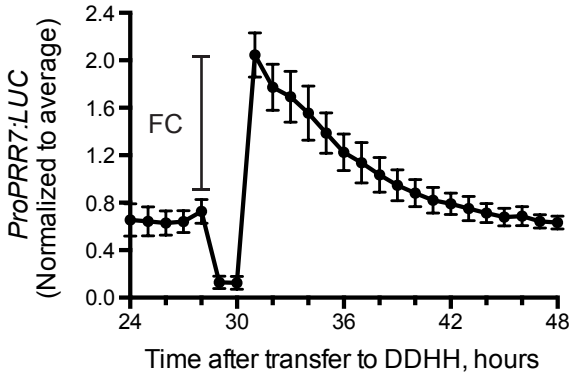
1) Transfer 3 day-old seedlings to LL| 22:18°C for 4 additonal days.

▼

2) Transfer to DD| 22°C in camera; administer pulse 28 hours later.

▼ **5,328 M2 seedlings**

3) Select individuals with fold-change (FC) in bioluminescence lesser or greater than twice the standard deviation of WT mean. (FC = signal at 31 hours)/signal at 28 hours)



▼ **131 primary M2 putative mutants retested**

4) Return to LL| 22:18°C for 2 days and rescreen for FC phenotype.

▼ **74 secondary M2 putants retained**

5) Confirm phenotype in M3 families.

▼ **12 mutants confirmed in M3**

6) *warp* mutants

B *warp2* exhibits severely attenuated FC

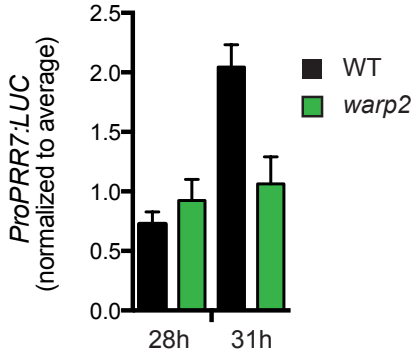


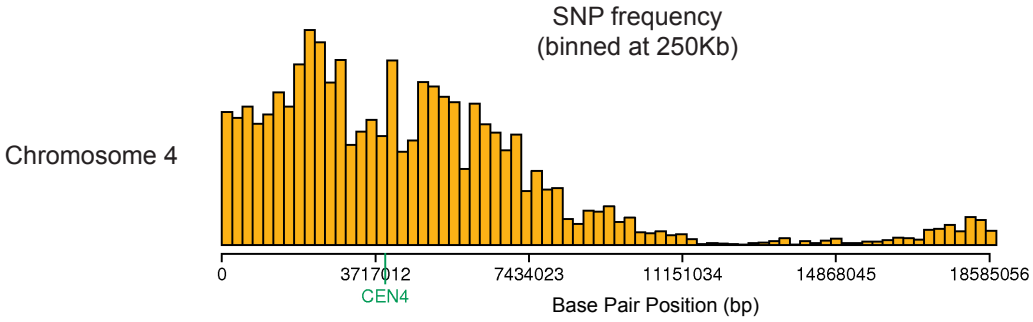
Figure 1. Screen for *warp* (*warm acute response of PRR7)* mutants.

A) Screen design that identified *warp2*.

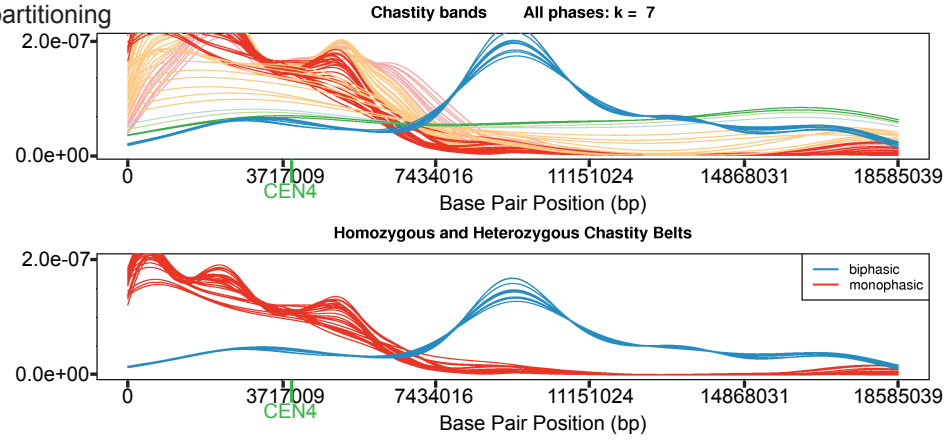
B) *ProPRR7:LUC* activity in WT and *warp2* seedlings before and after a 3 hour treatment of 28°C. Data are from 3 independent experiments, and error bars are standard deviation.

Mapping *warp2/sic* with bulk segregant analysis (BSA) and Next Generation EMS Mapping

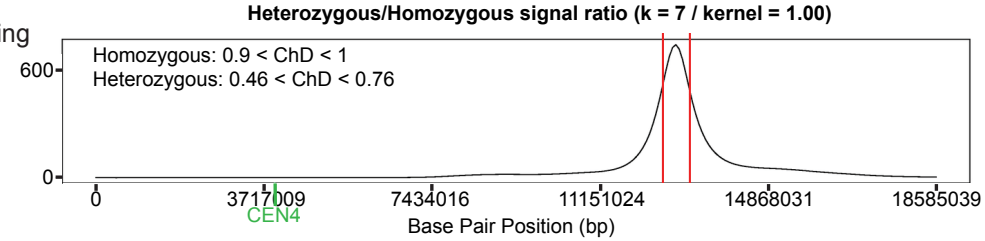
1) Generate *warp2* (Columbia) x Landsberg mapping population
2) BSA pool of genomic DNA from 80 F2 individuals with *warp2* phenotype
3) DNA-seq of BSA pool with illumina HiSeq2000 (single lane SE50 - X total reads)
4) Next Generation EMS Mapping with default settings
5) Map to chromosome



6) Chastity belt partitioning



7) Fine mapping



8) Candidate EMS mutations

SNPs between position 12526042 and 13117663

Chrom.	Position	Ref. Base	SNP base	Depth	Discordant chastity	Accession	Tag	Strand	Ref. codon	SNP codon	AA change	BLOSUM 100
4	12603385	G	A	23	1	AT4G24320	CDS	-	GCT	GTT	A -> V	-2
4	**12659045**	**G**	**A**	**21**	**1**	**AT4G24500**	**cryptic splice site**					
4	12893156	G	A	35	1	AT4G25120	CDS	+	GCA	ACA	A -> T	-1
4	13028319	G	A	20	1	AT4G25515	CDS	+	GGA	GAA	G -> E	-6
4	13035398	G	A	19	1	AT4G25520	CDS	-	CTT	TTT	L -> F	0

Figure 2. Mapping *warp2*/*sic-3* with Bulk Segregant Analysis (BSA) and Next Generation EMS Mapping.

warp2/*sic-3* was mapped with bulk segregant analysis (BSA) and Next Generation EMS Mapping as described in (Austin et al., 2011). The *sic-3* interval was identified with Next Generation EMS Mapping employing the default settings. The mutant region was refined with chastity belt partitioning, which identifies the region within the mapping interval having the greatest divergence between heterozygous and homozygous sites (Austin et al., 2011). Candidate genes in the interval on chromosome 4 were sequenced to identify EMS mutations, which confirmed the presence of a G → A transition at nucleotide position 12,659,045 in AT4G24500.

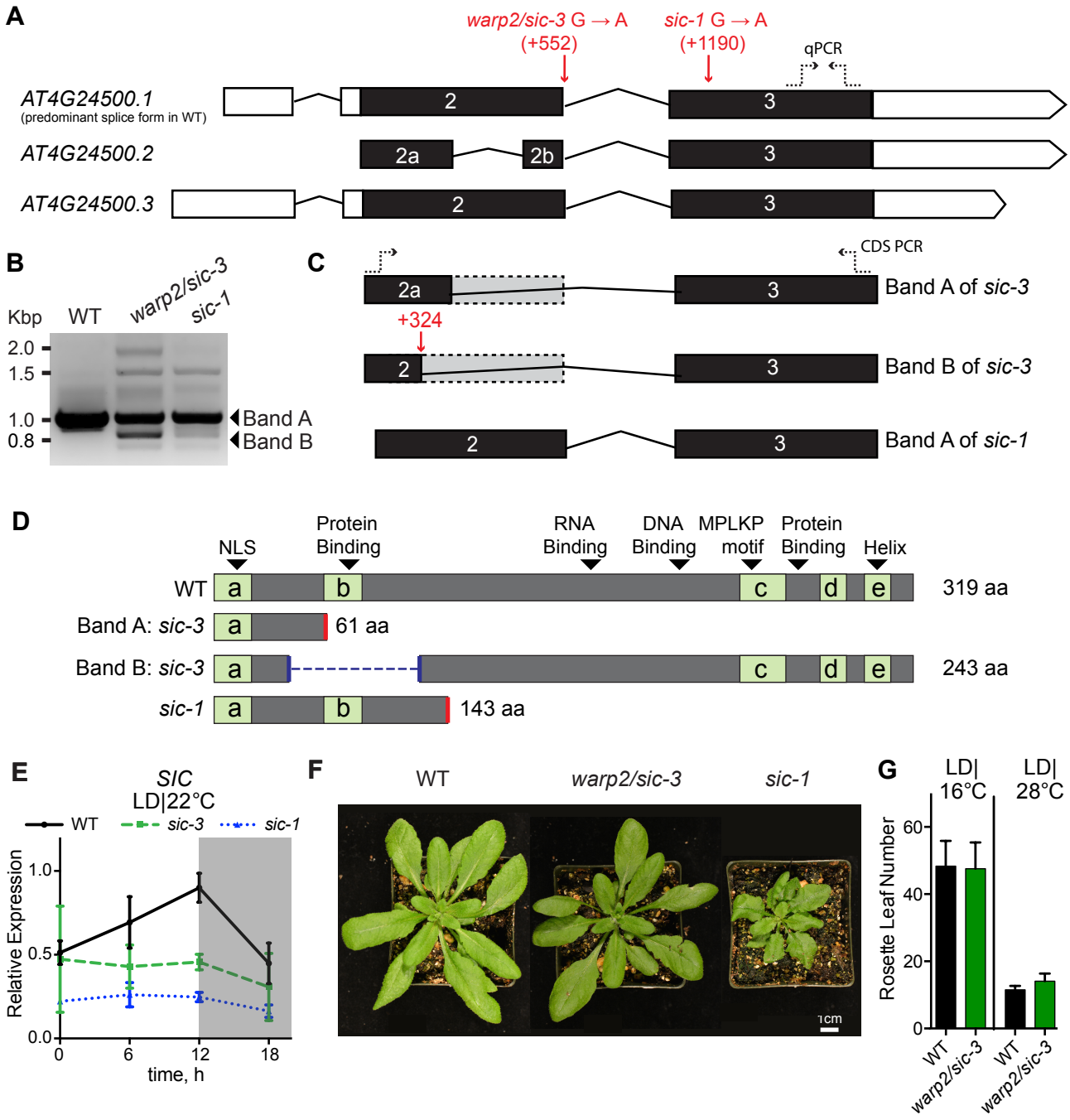


Figure 3. *sic* mutants affect AT4G24500, which encodes a conserved proline/serine rich protein.

A) Structure of transcripts for *SIC* (AT4G24500) from TAIR10 annotation (arabidopsis.org). Red text and arrows indicate the type and location of mutation present in *sic-3* and *sic-1*. Numbers are relative to the transcriptional start site (+1). Dotted arrows at 3'-end indicate location of qPCR primers to quantify *SIC* transcript. Lines indicate introns and boxes indicate exons where white boxes are UTR regions

and black boxes are coding regions. AT4G24500.1 is the predominant *SIC* transcript in WT.
B) PCR products generated by amplification of complete CDS from *SIC* transcript in WT, *sic-3*, and *sic-1*. Arrows labeled “A” and “B” indicate bands from the large and small PCR products, respectively.
C) Structure of *SIC* transcripts in *sic-3* and *sic-1* determined by sequencing of PCR products in **(B)**. Red text and arrow indicate cryptic 5’ss employed in *sic-3* to produce PCR product labeled “band B” in **(B)**. Grey boxes with dotted lines indicate exons removed by missplicing in *sic-3*.
D) Predicted proteins encoded by the transcripts found in *sic-3* and *sic-1*. For the protein encoded by band “B” in *sic-3*, the blue areas and dotted line indicate internal deletion that removes domain b. Red areas are premature stop codons introduced by either *sic-3* encoded by band “A” and *sic-1*. Light green boxes labeled a-e correspond to highly conserved amino acid regions in SIC (see Figure 6).
E) Expression of *SIC* in WT (black bar), *warp2*/*sic-3* (green bar), and *sic-1* (blue bar) in LD|22C determined by qPCR as in Supplemental Figure 5. Each time point is the mean of 3 independent experiments, and error bars are standard deviation. Grey shaded area represents the time of darkness.
F) Rosettes of mature WT, *warp2*/*sic-3*, and *sic-1* grown in LD|22°C. Bar is 1 cm.
G) Rosette leaf number until flowering in WT and *warp2*/*sic-3* for plants grown in either LD|16°C or LD|28°C (corresponding to 12 hours light, 12 hours dark and constant temperature). n=18 for all. Each bar is the mean of 3 independent experiments, and error bars are standard deviation.

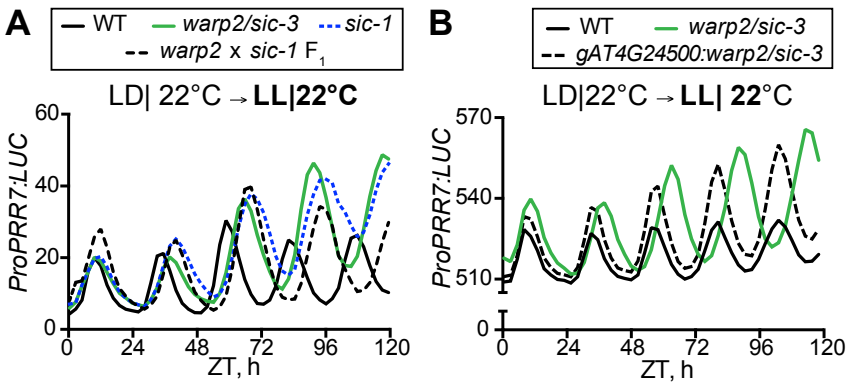


Figure 4. Mutation of AT4G24500/*SIC* is responsible for the *warp2* phenotype.

(A) Non-complementation of long period *ProPRR7:LUC* rhythms in an F_1 cross between *warp2 sic-3* and *sic-1* [*warp2*/*sic-3* x *sic-1* F_1 (dashed black line)] compared to WT (solid black line), *warp2*/*sic-3* (solid green line), and *sic-1* (dotted blue line). Seedlings entrained under LD|22°C and released into LL|22°C. Data are representative of 3 independent experiments.
(B) Complementation of long period rhythms in *warp2*/*sic-3* with a stable transgene carrying the AT4G24500/*SIC* genomic region from WT (*gATG24500:warp2*/*sic-3* (dashed black line), compared to WT (solid black line) and *warp2*/*sic-3* (solid green line). Seedlings were first entrained in photocycles (LD|22°C) and then released into free run at LL|22°C. *ProPRR7:LUC* activity was monitored for 5 days.

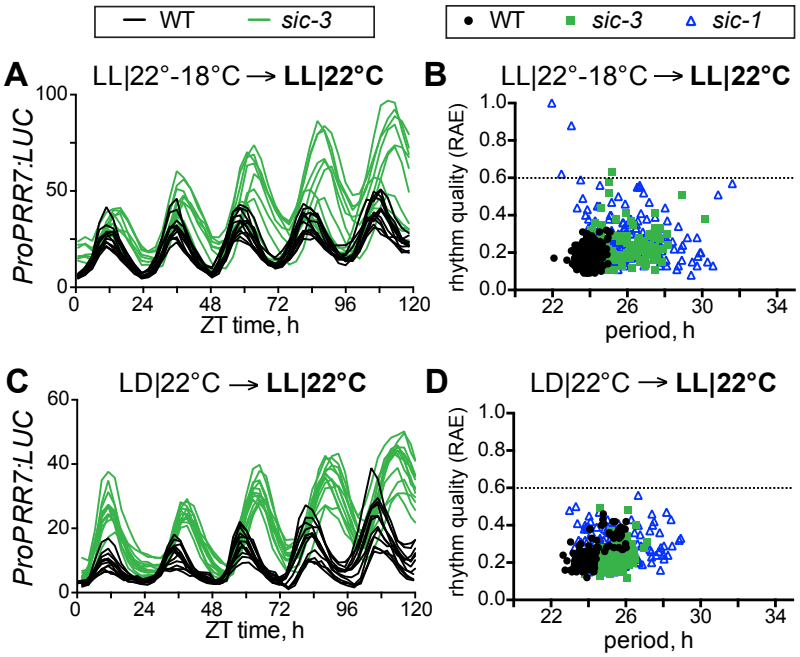


Figure 5. *sic* lengthens the circadian clock period in free-running conditions.

A) and **(C)** Individual traces of *ProPRR7:LUC* activity in WT (black lines) and *sic-3* (green lines) seedlings under LL|22°C beginning at ZT0 after entrainment with **(A)** LL|22°-18°C or **(C)** LD|22°C (n=8 for all). Data are representative of 3 independent experiments.
B) and **(D)** RAE as a function of period (in hours) of *ProPRR7:LUC* activity for WT (black circles), *sic-3* (green squares), and *sic-1* (blue triangles) seedlings under LL|22°C after entrainment under **(B)** LL|22°-18°C (n total: WT=277, *sic-3*=220, *sic-1*=141) and **(D)** LD|22°C (n total: WT=314, *sic-3*=305, *sic-1*=257). Dotted horizontal line indicates RAE (relative amplitude error) = 0.6 threshold, above which traces of *ProPRR7:LUC* activity are considered arrhythmic. Data are from 3 independent experiments.

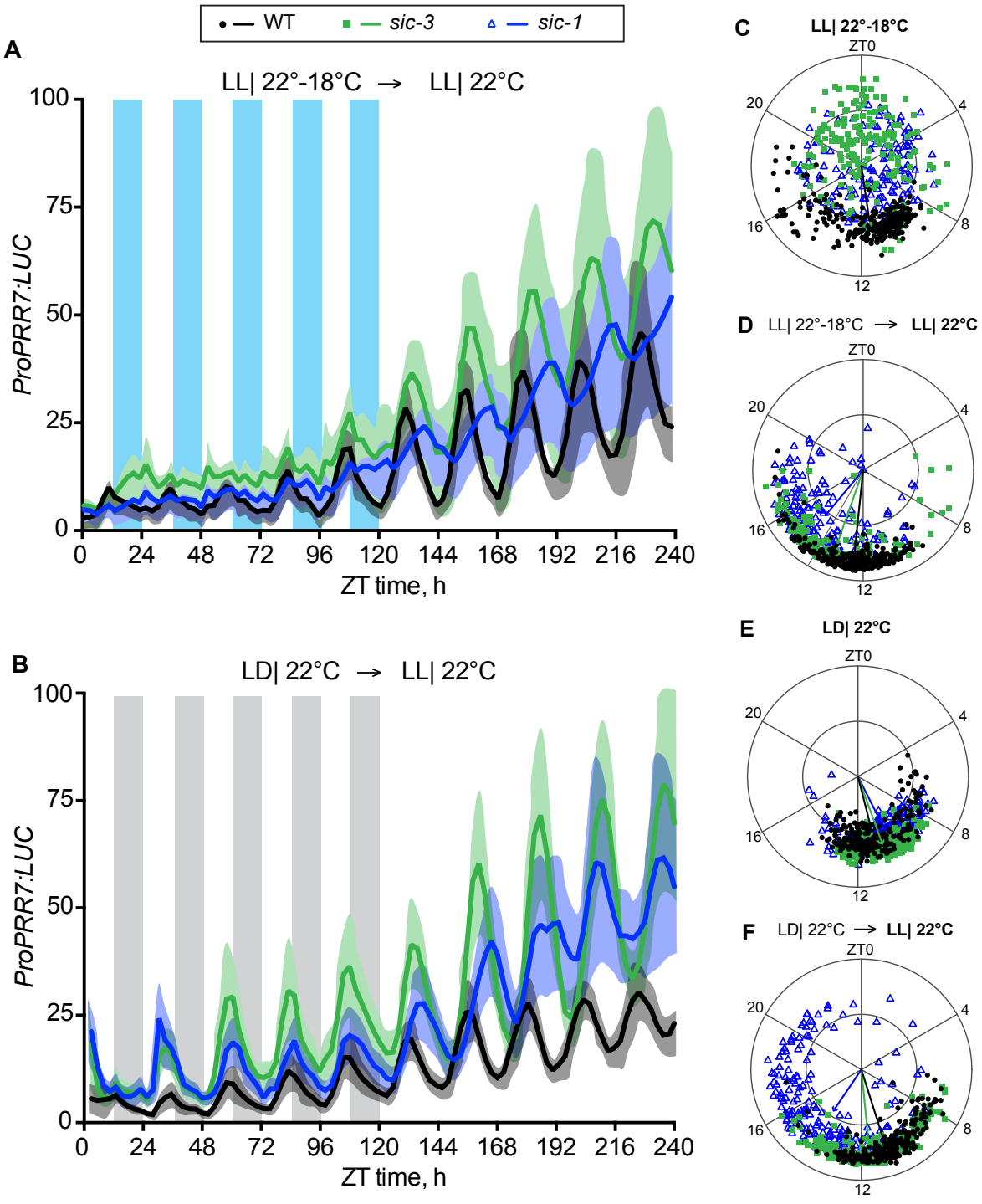


Figure 6. *sic* mutants perturb the accuracy and precision of circadian clock period and phase.

A) and **B)** Mean *ProPRR7:LUC* activity in WT, *sic-3*, and *sic-1* in **(A)** LL|22°-18°C or **(B)** LD|22°C (time 0 to 120 hours) and then release into free run LL|22°C (time 120 to 240 hours). Shadows indicate standard deviation range for *ProPRR7:LUC* activity for WT

(grey), *sic-3* (green), and *sic-3* (blue). Blue bars in **(A)** indicate periods of 18°C and grey bars in **(B)** indicate periods of darkness. Data are from 3 independent experiments. **C-F)** Polar plot of corrected phase for *ProPRR7:LUC* activity in WT, *sic-3*, and *sic-1* seedlings during **(C)** LL|22°-18°C (n: WT=277, *sic-3*=220, *sic-1*=141), **(E)** LD|22°C (n: WT=314, *sic-3*=305, *sic-1*=257), **D)** free run in LL|22°C after LL|22°-18°C entrainment or **(F)** free run in LL|22°C after LD|22°C entrainment (n total in free run identical to entrainment). Phase in ZT (as hours) is plotted around the circumference of the circle and radial position indicates RAE value for each individual where the outermost point is RAE = 0 and the innermost point is RAE = 1. For each genotype arrow position around the circumference of the circle indicate average phase, and length of arrow indicates average RAE.

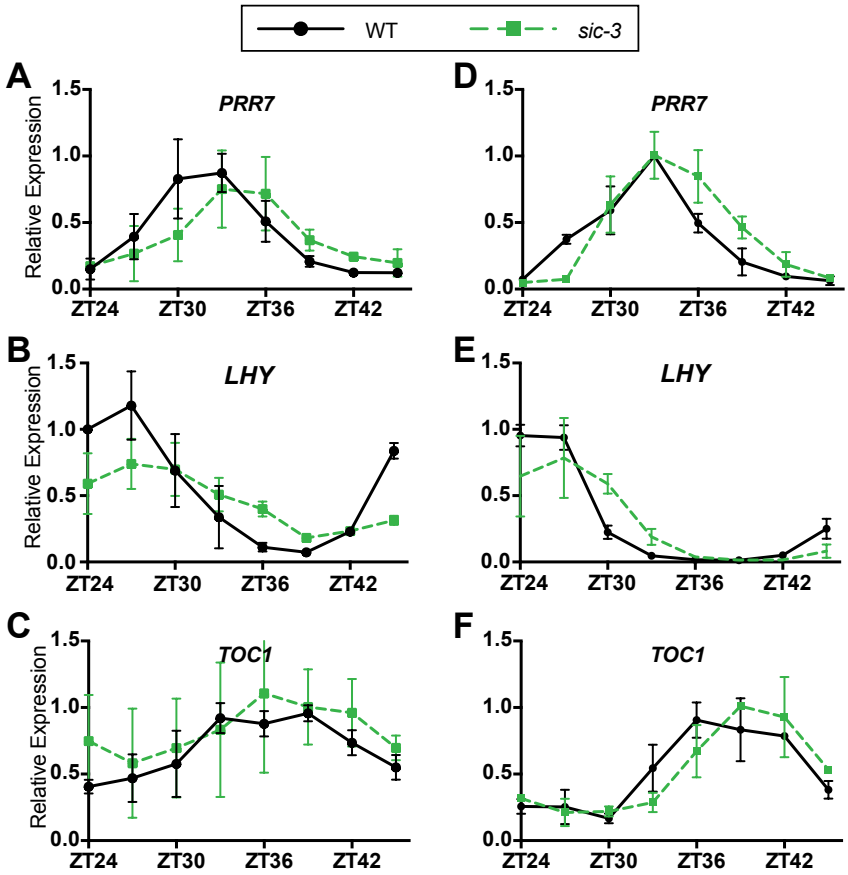


Figure 7. *sic* lengthens expression waveforms of core circadian clock genes in free running conditions.

A) to **(F)** Expression levels of the indicated transcripts in WT (black circles, solid line) and *sic-3* (green squares, dotted green line) grown under LL|22°C after entrainment under **(A-C)** LL|22°-18°C or **(D-F)** LD|22°C. Transcript levels were determined with qPCR. In each biological replicate, Relative Expression for each transcript was calculated by normalization to the highest expression value for WT. Each point is mean of 3 independent experiments, and error bars are standard deviation.

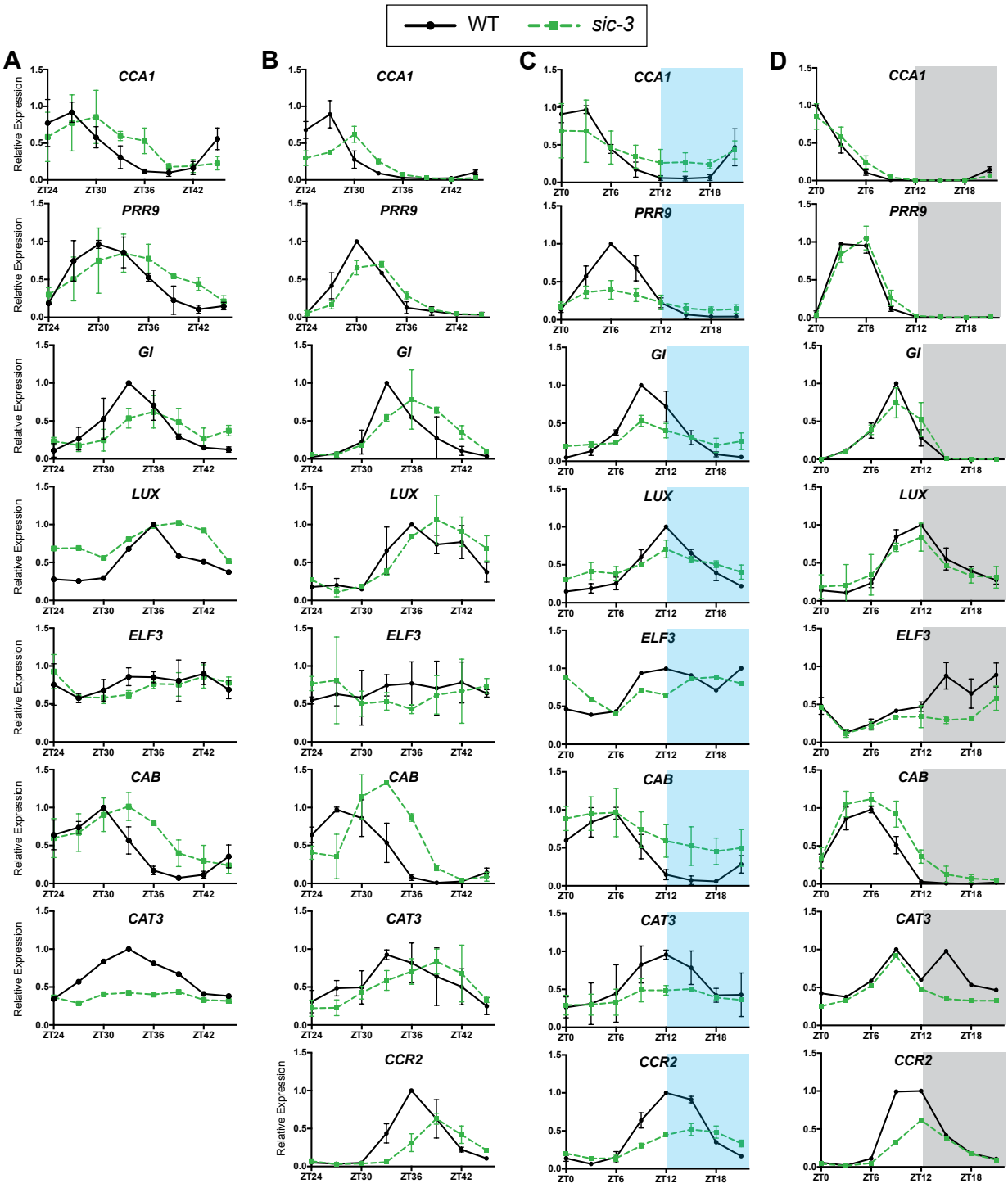


Figure 8. *sic* disrupts expression waveforms of core circadian clock genes and clock-regulated output genes.

A) to **(D)** Relative expression levels of the indicated transcripts in WT (black circles, solid line) and *sic-3* (green squares, dotted green line) grown in **(C)** LL|22°-18°C or **(D)** LD|22°C, or 24 hours after release into LL|22°C free run from **(A)** LL|22°-18°C or **(B)**

LD|22°C. Blue regions in **(C)** indicate periods of 18°C and grey regions in **(D)** indicate periods of darkness. Transcript levels were determined with qPCR. In each biological replicate, relative expression for each transcript was calculated by normalization to the highest expression value for WT. Each point is mean of 3 independent experiments, and error bars are standard deviation.

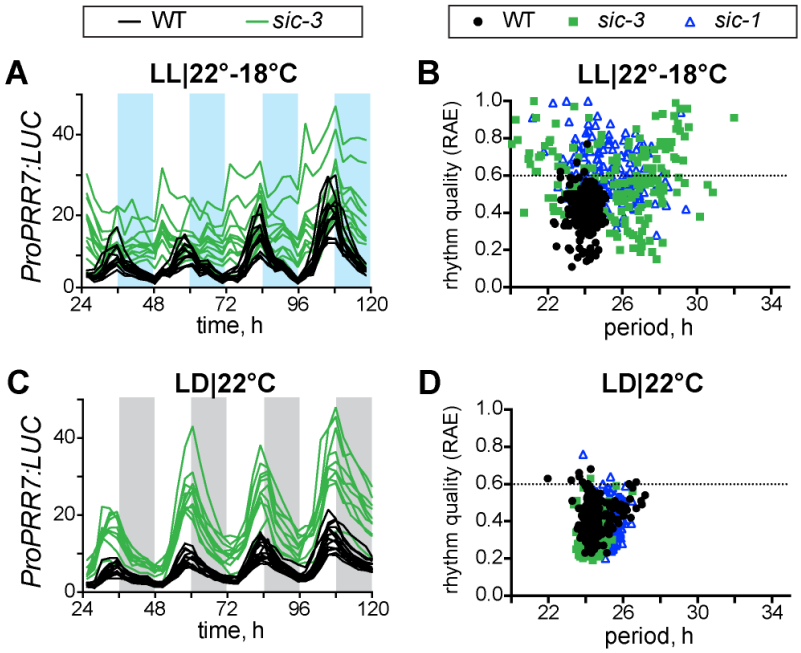


Figure 9. *sic* weakens clock function in cool temperature thermocycles.

A) and **(C)** Individual traces of *ProPRR7:LUC* activity in WT (black lines) and *sic-3* (green lines) seedlings during **(A)** LL|22°-18°C or **(C)** LD|22°C over 4 days (n=8 for all)**.** Blue shading represents periods of 18°C, and grey shading represents periods of dark.
B) and **(D)** RAE as a function of period (in hours) of *ProPRR7:LUC* activity for WT (black circles), *sic-3* (green squares), and *sic-1* (blue triangles) seedlings during **(B)** LL|22°-18°C (n total: WT=277, *sic-3*=220, *sic-1*=141) and **(D)** LD|22°C (n total: WT=314, *sic-3*=305, *sic-1*=257). Dotted horizontal line indicates RAE = 0.6 threshold, above which traces of *ProPRR7:LUC* activity are considered arrhythmic. Data are from 3 independent experiments.

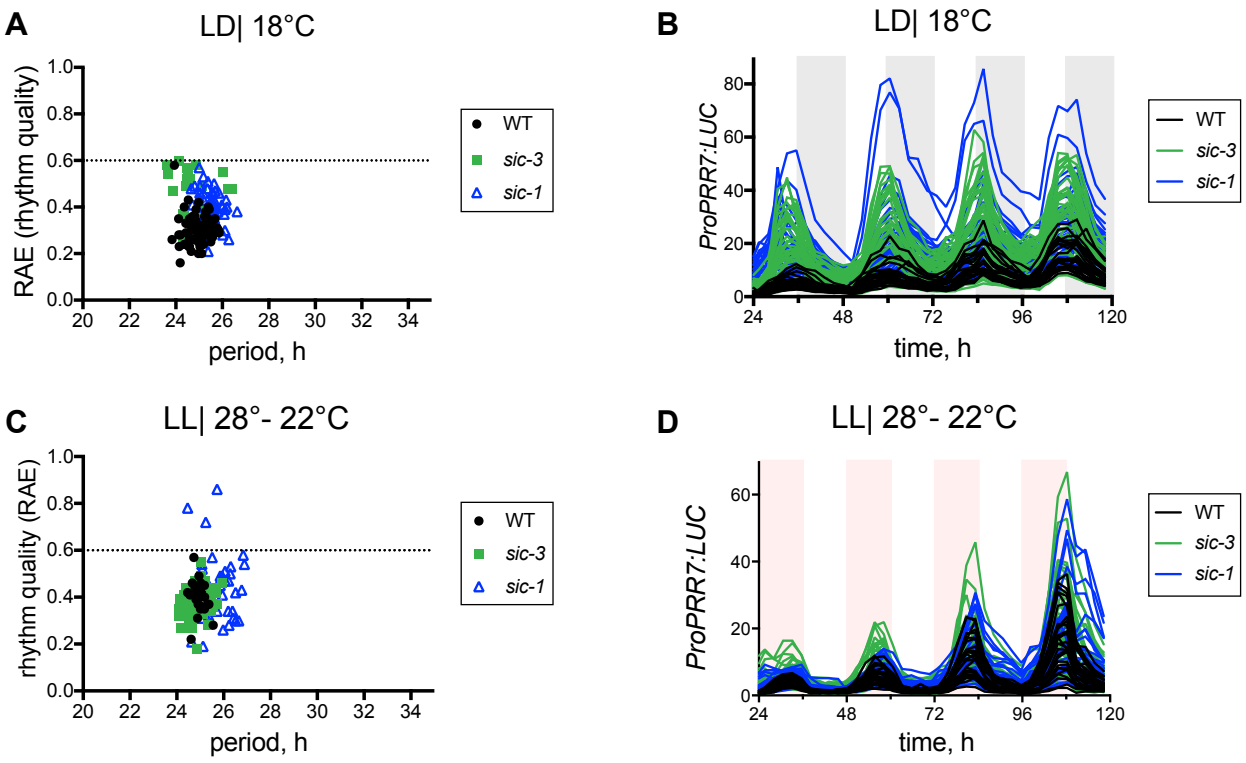


Figure 10. *sic* does not affect rhythms in warm thermocycles and in cool photocycles.

A) RAE as a function of period (in hours) of *ProPRR7:LUC* activity for WT (black circles), *sic-3* (green squares), and *sic-1* (blue triangles) under LD|18°C entrainment (n for WT=55, *sic-3*=41, *sic-1*=44). Dotted horizontal line indicates RAE = 0.6 threshold, above which traces of *ProPRR7:LUC* activity are considered arrhythmic. *sic-3* and *sic-1* are not significantly different from WT based on p-value <0.01 from ANOVA followed by Tukey's multiple comparison test.
B) Individual traces of *ProPRR7:LUC* activity in WT (black lines), *sic-3* (green lines), and *sic-1* (blue lines) seedlings during LD|18°C (12 hour dark intervals) over 4 days (n for WT=55, *sic-3*=41, *sic-1*=44)**.** Grey shading represents periods of darkness.
C) RAE as a function of period (in hours) of *ProPRR7:LUC* activity for WT (black circles), *sic-3* (green squares), and *sic-1* (blue triangles) under LL|28°-22°C entrainment (n for WT=26, *sic-3*=54, *sic-1*=47). Dotted horizontal line indicates RAE = 0.6 threshold, above which traces of *ProPRR7:LUC* activity are considered arrhythmic. No genotypes are significantly different from WT based on p-value <0.01 from ANOVA followed by Tukey's multiple comparison test.
D) Individual traces of *ProPRR7:LUC* activity in WT (black lines), *sic-3* (green lines), and *sic-1* (blue lines) seedlings during LL|28°-22°C (12 hour temperature intervals) over 4 days (n=16 for all)**.** Red shading represents periods of 28°C**.**

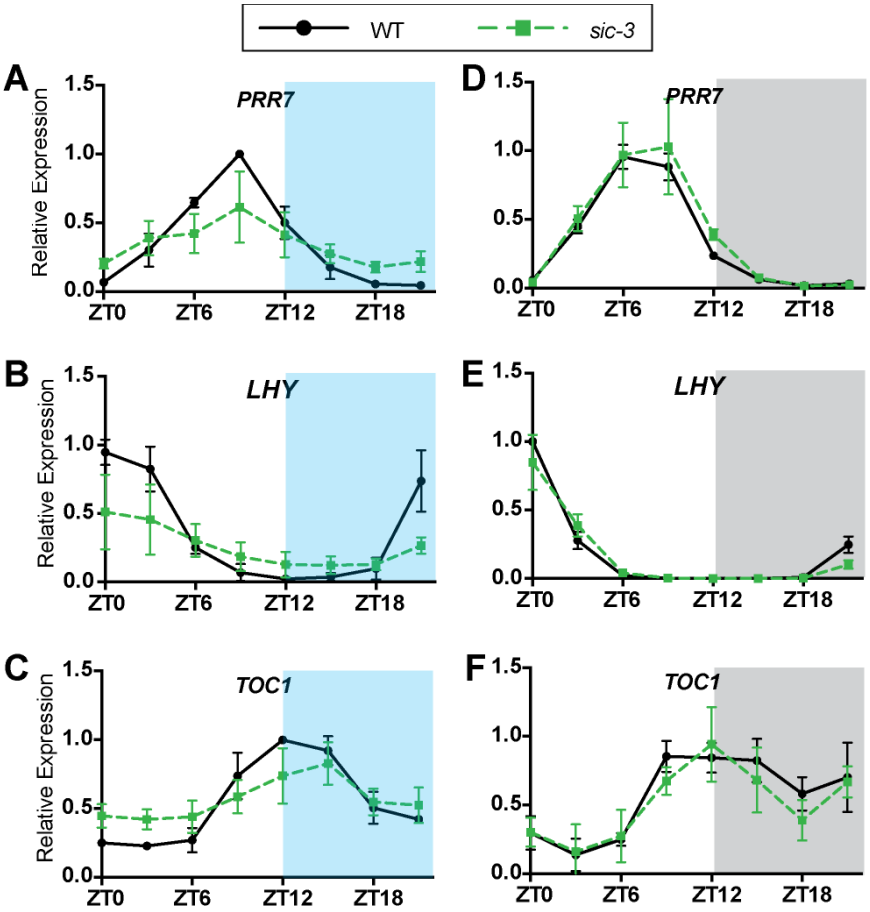


Figure 11. *sic* disrupts the expression waveforms of core circadian clock genes in cool temperature thermocycles.

A) to **(F)** Expression levels of the indicated transcripts in WT (black circles, solid line) and *sic-3* (green squares, dotted green line) grown under **(A-C)** LL|22°-18°C or **(D-F)** LD|22°C. Blue regions in **(A-C)** indicate periods of 18°C, and grey regions in **(D-F)** indicate periods of darkness. Transcript levels were determined with qPCR. In each biological replicate, Relative Expression for each transcript was calculated by normalization to the highest expression value for WT. Each point is mean of 3 independent experiments, and error bars are standard deviation.

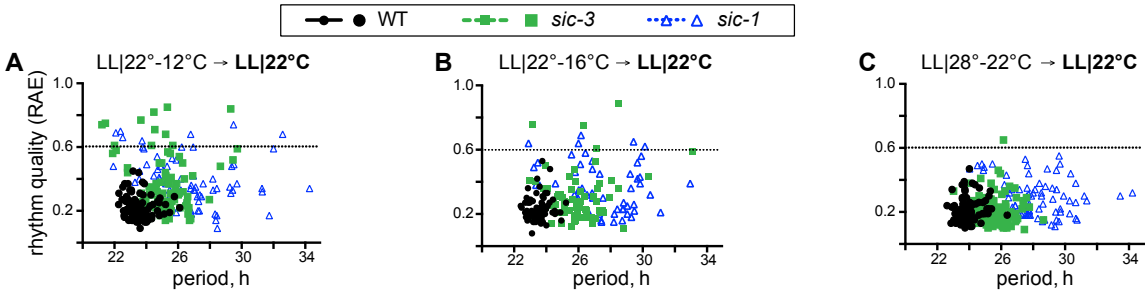


Figure 12. *sic* reduces circadian rhythm quality, particularly in cool thermocycles.

A) to **(C)** RAE as a function of period at 22°C for individual WT, *sic-3*, and *sic-1* seedlings after entrainment with **(A)** LL|22°-12°C (n for WT=90, *sic-3*=90, *sic-1*=53), **(B)** LL|22°-16°C (n for WT=67, *sic-3*=63, *sic-1*=47), or **(C)** LL|28°-22°C (n for WT=102, *sic-3*=98, *sic-1*=80). Points above dotted horizontal line at RAE=0.6 are considered arrhythmic. Data are from 3 independent experiments.

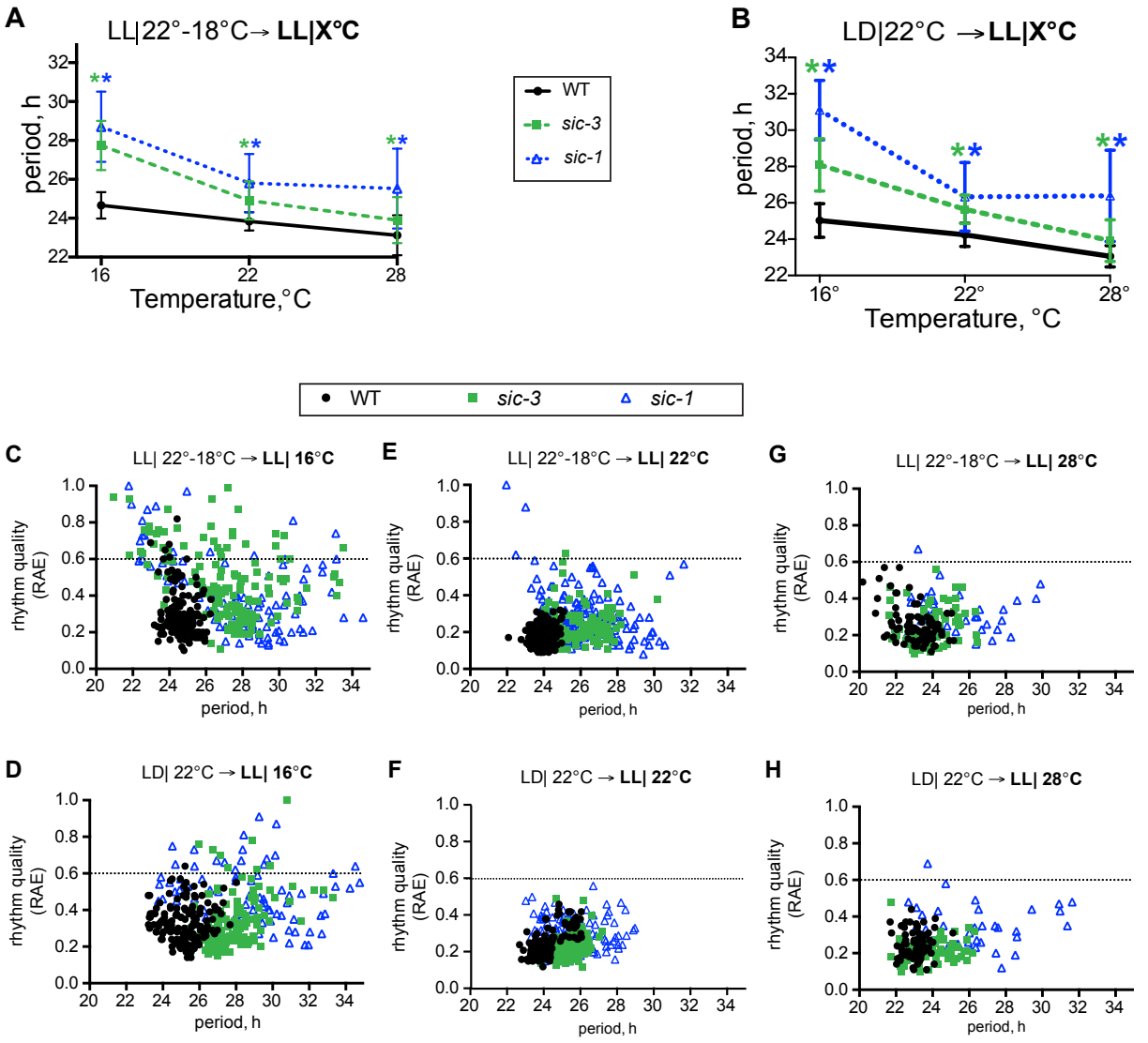


Figure 13. Exposure of *sic* to cold temperature cycles impairs temperature compensation and rhythm quality.

A) Mean period length for *ProPRR7:LUC* activity in WT, *sic-3*, and *sic-1* seedlings under free running conditions of LL|X°C, where X stands for either 16°C, 22°C, or 28°C as indicated. Entrainment was either **(A)** LL|22°-18°C or **(B)** LD|22°C for 5 days prior to release into free running conditions. Each point is the mean of all rhythmic individuals (RAE<0.6) from three independent biological replicates (**(A)** n total at 16°C: WT=106, *sic-3* =97, *sic-1* =73; at 22°C: WT=277, *sic-3*=220, *sic-1*=141; at 28°C: WT=83, *sic-3*=77, *sic-1*=34, and **(B)** n at 16°C: WT=108, *sic-3*=94, *sic-1*=28; at 22°C: WT=314, *sic-3*=305, *sic-1*=257; at 28°C: WT=83, *sic-3*=80, *sic-1*=35). Error bars are standard deviation and asterisk (*) indicates significance p-value <0.0001 from ANOVA with Tukey's multiple comparison test between all genotypes at the same temperature.

C) to **(H)** RAE as a function of period for *ProPRR7:LUC* activity in WT, *sic-3*, and *sic-1* seedlings first entrained under LL|22°-18°C **(C,E,G)** or under LD|22°C **(D,F,H)**, then released into LL free run at 16°C **(C,D)**, 22°C **(E,F)**, or 28°C **(G,H)**. **E)** and **F)** are the same data as in Figure 2A and 2D, respectively. Dotted horizontal line indicates RAE = 0.6 threshold above which traces of *ProPRR7:LUC* activity are considered arrhythmic. Data are from 3 independent experiments.

Table 1. Circadian clock parameters determined from *ProPRR7:LUC* rhythms in the indicated conditions.

Entrainment condition	Free run condition	Genotype	Mean Period ± SD	Mean RAE ± SD	Mean Phase ± SD	Percent Rhythmic	Total
LD| 22°C	LL|22°C	WT	23.3 ± 0.6	0.23 ± 0.05	10.7 ± 0.9	N/D	16
		sic-3	25.7 ± 0.8	0.21 ± 0.05	10.5 ± 0.8	N/D	16
		gAT4G24500:sic-3	23.5 ± 0.4	0.21 ± 0.04	10.8 ± 0.6	N/D	16
LD|22°C	LL|22°C	WT	24.2 ± 0.6	0.20 ± 0.05	11.0 ± 0.9	N/D	64
		sic-3	26.7 ± 1.0	0.24 ± 0.10	10.4 ± 2.1	N/D	42
		sic-1	27.7 ± 1.4	0.24 ± 0.06	10.8 ± 2.1	N/D	52
		sic-3 x *sic-1* F_1	27.3 ± 1.0	0.21 ± 0.06	10.9 ± 1.4	N/D	36

Rhythmic *ProRR7:LUC* activity assayed for 4-5 days in the indicated condition and the indicated genotypes. Mean period, mean RAE (relative amplitude error), and mean phase is for rhythmic individuals and SD is standard deviation. Rhythmic seedlings are those with RAE<0.6. N/D, not done.

Table 2. Circadian clock parameters determined from *ProPRR7:LUC* rhythms in the indicated conditions.

Entrainment condition	Free run condition	Genotype	Mean Period ± SD	Mean RAE ± SD	Mean Phase ± SD	Percent Rhythmic	Total
LL|22°-18°C	-	WT	24.1 ± 0.5	0.40 ± 0.10	11.6 ± 2.4	98%	277
		sic-3	25.7 ± 1.9	0.56 ± 0.20	11.2 ± 7.7	60%	220
		sic-1	25.5 ± 1.5	0.60 ± 0.15	9.6 ± 6.0	52%	141
	LL|16°C	WT	24.7 ± 0.7	0.3 ± 0.13	16.7 ± 2.7	96%	106
		sic-3	27.7 ± 1.3	0.46 ± 0.19	17.2 ± 5.3	77%	97
		sic-1	28.7 ± 1.8	0.4 ± 0.2	6.3 ± 7.3	85%	73
	LL|22°C	WT	23.8 ± 0.5	0.17 ± 0.06	12.2 ± 1.2	100%	277
		sic-3	25.2 ± 1.0	0.20 ± 0.08	13.4 ± 1.7	99%	220
		sic-1	25.8 ± 1.4	0.34 ± 0.16	14.6 ± 2.8	97%	141
	LL|28°C	WT	23.1 ± 1	0.3 ± 0.2	14.0 ± 2.1	100%	83
		sic-3	23.9 ± 1.2	0.3 ± 0.2	13.3 ± 2.2	100%	77
		sic-1	25.5 ± 2.0	0.36 ± 0.2	12.9 ± 3.8	97%	34
LD|22°C	-	WT	24.5 ± 0.7	0.40 ± 0.09	10.9 ± 1.8	95%	314
		sic-3	24.4 ± 0.4	0.31 ± 0.07	10.7 ± 1.1	99%	305
		sic-1	25.2 ± 0.5	0.39 ± 0.09	10.3 ± 1.6	98%	257
	LL|16°C	WT	25 ± 0.9	0.33 ± 0.1	13.9 ± 2.8	99%	108
		sic-3	28.1 ± 1.4	0.33 ± 0.14	13.0 ± 2.7	94%	94
		sic-1	31.1 ± 1.6	0.49 ± 0.15	13.1 ± 2.8	77%	28
	LL|22°C	WT	24.2 ± 0.6	0.24 ± 0.08	10.8 ± 1.6	100%	314
		sic-3	25.6 ± 0.8	0.23 ± 0.07	11.6 ± 1.8	100%	305
		sic-1	26.3 ± 1.6	0.30 ± 0.10	14.3 ± 3.8	100%	257
	LL|28°C	WT	23 ± 0.6	0.24 ± 0.07	13.3 ± 1.7	100%	83
		sic-3	23.9 ± 1.1	0.2 ± 0.06	13.6 ± 2.0	100%	80
		sic-1	26.3 ± 2.5	0.35 ± 0.11	11.8 ± 4.6	97%	35

Rhythmic *ProRR7:LUC* activity assayed for 4-5 days in the indicated condition and the indicated genotypes. Mean period, mean RAE (relative amplitude error), and mean phase is for rhythmic individuals and SD is standard deviation. Rhythmic seedlings are those with RAE<0.6. N/D, not done.

Table 3. Circadian clock parameters determined from *ProPRR7:LUC* rhythms in the indicated conditions.

Entrainment condition	Free run condition	Genotype	Mean Period ± SD	Mean RAE ± SD	Mean Phase ± SD	Percent Rhythmic	Total
LL|22°-12°C	LL|22°C	WT	25.6 ± 0.8	0.22 ± 0.07	15.9 ± 1.8	100%	90
		sic-3	25.5 ± 1.5	0.36 ± 0.17	15.9 ± 3.0	87%	90
		sic-1	27 ± 2.7	0.4 ± 0.15	15.2 ± 3.9	87%	53
LL|22°-16°C	LL|22°C	WT	23.5 ± 0.6	0.24 ± 0.08	17.4 ± 1.7	100%	67
		sic-3	26 ± 1.7	0.31 ± 0.15	16.4 ± 2.7	94%	63
		sic-1	27.6 ± 2.1	0.37 ± 0.15	15.1 ± 3.3	92%	47
LL|28°-22°C	LL|22°C	WT	23.9 ± 0.6	0.2 ± 0.07	13.2 ± 1.9	100%	102
		sic-3	25.6 ± 1.1	0.21 ± 0.09	15.9 ± 2.6	99%	98
		sic-1	26.1 ± 1.9	0.29 ± 0.11	15.5 ± 3.0	100%	80
	-	WT	25 ± 0.25	0.4 ± 0.07	N/D	100%	26
		sic-3	24.8 ± 0.5	0.37 ± 0.07	N/D	100%	54
		sic-1	25.7 ± 0.7	0.42 ± 0.14	N/D	94%	47
LD|18°C	-	WT	24.8 ± 0.5	0.31 ± 0.06	N/D	100%	99
		sic-3	24.3 ± 0.28	0.28 ± 0.06	N/D	100%	87
		sic-1	25.2 ± 0.6	0.42 ± 0.08	N/D	100%	137

Rhythmic *ProRR7:LUC* activity assayed for 4-5 days in the indicated condition and the indicated genotypes. Mean period, mean RAE (relative amplitude error), and mean phase is for rhythmic individuals and SD is standard deviation. Rhythmic seedlings are those with RAE<0.6. N/D, not done.

Chapter 2: SICKLE mediates temperature-dependent expression and alternative splicing of circadian clock transcripts

The following text is modified from an article published in *The Plant Cell*, by Marshall *et al.*, 2016.

Abstract

Alternative splicing of circadian clock transcripts and the whole transcriptome was evaluated in WT and *sic* mutants grown in contrasting temperature conditions to better understand the contribution of alternative splicing to the circadian clock temperature responses. The *sic* mutant was previously implicated in transcript splicing, microRNA (miRNA) biogenesis, and stress responses (Zhan et al. 2012). According to quantitative reverse transcription-PCR (RT-PCR), alternative splicing of *LHY*, *CCA1*, *PRR7*, and *ELF3* transcripts is elevated at all temperatures in *sic-3* and *sic-1*, but cool temperatures markedly stimulate splice variant production. RNA-seq analysis and qPCR confirm the elevated accumulation of *LHY* and *CCA1* splice variants in *sic-3* at cool temperatures, and revealed novel splice variant accumulation not found via RT-PCR. Large changes in gene expression were apparent for several circadian clock genes between WT and *sic-3*, as well as between cool and warm temperatures in WT. These results show SIC is crucial for appropriate transcript splicing and expression of several circadian clock genes in cool temperatures.

Introduction

Alternative splicing is an important form of gene regulation, and its role in regulating the circadian clock is becoming apparent in many eukaryotes. This is best exemplified in the central temperature compensation mechanism for the filamentous fungi *Neurospora crassa* (A. Diernfellner et al. 2007). In *N. crassa*, core clock protein activity of FREQUENCY (FRQ) is modulated through temperature-induced alternative splicing (A. C. R. Diernfellner et al. 2005). This generates unproductive splice variants with alternate upstream open reading frames and PTCs that are inefficiently translated or encode inactive protein isoforms (A. C. R. Diernfellner et al. 2005). In this mechanism, the ratio of unproductive to productive splice variants of *FRQ* adjusts the length of the circadian period in response to ambient temperature and allows for temperature compensation by slowing down the clock in high temperatures and speeding up the clock in low temperatures (A. C. R. Diernfellner et al. 2005; A. Diernfellner et al. 2007). Temperature also affects alternative splicing of the *period* transcript in *Drosophila melanogaster* (Lim and Allada 2013). These examples of alternative splicing adjusting the pace of the circadian clock in different temperature conditions demonstrate a mechanism that could be found in Arabidopsis.

Temperature influences alternative splicing of circadian clock transcripts in Arabidopsis (Filichkin et al. 2010). The abundance of splice variants from many circadian clock genes changes in response to external temperature cues (A. B. James et al. 2012; Marshall et al. 2016; Seo et al. 2012; S. Filichkin et al. 2015). In warm

temperatures intron retention events in *CCA1* and *PRR7* (see general introduction) create splice variants that lead to truncated proteins, as well as intron retention events in *ELF3* and *TOC1* transcripts that are subject to NMD (Kwon et al. 2014). Cool temperatures promote intron retention in important circadian clock genes such as *LHY* and *RVE8*, which lead to transcripts targeted for NMD (A. B. James et al. 2012; Allan B. James et al. 2012). For example, cool temperatures cause intron retention in the 5'UTR of *LHY* in very significant amounts, but this splice variant does not introduce a PTC (A. B. James et al. 2012), although it can be targeted for NMD. Instead, the added 5'UTR intron increases the number of upstream open reading frames, which may affect the transcripts translatablities (A. B. James et al. 2012). This example adds another layer of complexity to alternative splicing; splice variants may not only affect protein translation by changing a gene's CDS, but it may also affect the efficiency of translation (G. Wang, Guo, and Floros 2005). Despite these observations, it remains unclear how the accumulation of circadian clock splice variants mechanistically affects circadian clock function and its responses to temperature cues.

Further evidence that alternative splicing plays a regulatory role in the Arabidopsis circadian clock can be seen in four spliceosome-associated gene mutants, which each exhibit altered circadian clock period and alternative splicing of transcripts from core circadian clock genes. The *SNW/Ski interacting* (*skip*), *spliceosomal timekeeper locus1* (*stipl1*), and *protein arginine methyl transferase5* (*prmt5*) mutants have a lengthened clock period (Hong et al. 2010; Sanchez et al. 2010; Jones et al. 2012; X. Wang et al. 2012), while the *gemin2* mutant has a shortened clock period (Schlaen et al. 2015). The *skip* and *gemin2* mutants also have impaired circadian clock temperature compensation (Wang et al. 2012; Schlaen et al. 2015). Furthermore, all four mutants exhibit altered alternative splicing of core circadian clock transcripts, which affects the abundance of splice variants from those transcripts (Hong et al. 2010; Sanchez et al. 2010; Jones et al. 2012; Wang et al. 2012; Schlaen et al. 2015). Thus, the Arabidopsis circadian clock requires full spliceosome activity and alternative splicing of circadian clock transcripts to establish period and temperature compensation.

Recent findings indicate that functional splice variants of spliceosome components, specifically the U1 snRNP, are present in limited levels in cool temperatures (Schlaen et al. 2015). To compensate for this loss Arabidopsis produces higher levels of a functional splice variant of the U1 snRNA in cool temperatures, which results in higher levels of functional snRNP, which ensures spliceosome efficiency (Schlaen et al. 2015). The GEMIN2 protein is necessary for this transition to occur, as *gemin2* mutants show changes in alternative splicing in the cold and are impaired in U1 snRNP assembly (Schlaen et al. 2015). Furthermore, cool temperatures trigger high expression of RNA processing genes, including genes involved in pre-mRNA processing (Schlaen et al. 2015). Together, this indicates that Arabidopsis activates several pathways to enhance spliceosomal activity in cool temperatures, perhaps because spliceosome activity is compromised in cool temperatures.

Alternative splicing plays an undoubtedly important role in regulating Arabidopsis circadian rhythms, and this regulation is subject to temperature influence. While recent data observes changes in the pattern of splice variants of circadian clock genes in changing temperatures, we do not understand the mechanistic significance these splice variants play in regulating circadian rhythms. Putative spliceosome mutants such as

gemin2 and *skip* indicate that changes in circadian clock splice variant accumulation affect temperature compensation, but they do not explain how the splice variants alter circadian clock temperature response (Schlaen et al. 2015; Wang et al. 2012). It is still undetermined if these splice variants are either a signal of changes in external temperature, or a negative outcome due to decreased spliceosome efficiency which must be overcome by the plant.

In this chapter we tested temperature-dependent alternative splicing of circadian clock genes, and the whole transcriptome in WT and *sic* mutants grown in different temperature conditions. Our results reveal SIC's fundamental role in regulating alternative splicing and gene expression in cool temperatures.

Results

sic mutation changes alternative splicing of circadian clock transcripts, particularly under cool temperatures

The effect of temperature on the alternative splicing of several circadian clock transcripts was evaluated in WT, *sic-3,* and *sic-1* seedlings. RT-PCR was used to detect specific splice variants from *CCA1*, *LHY*, *PRR9*, *PRR7*, *GI*, *ELF3*, *TOC1*, *PRR5,* and *PRR3* transcripts under conditions of LD|28°C, LD|22°C, and LD|16°C (Figure 14; Figure 15). Splice variants in cDNA generated from a pool of time points taken over a 24-hour period were detected with splice variant-specific primers. Labeling of PCR products with the FAM fluorophore allowed relative quantification of splice variant accumulation (Figure 14F-J). Bulk levels of each transcript were also assessed to confirm that changes in splice variant levels were explained by alternative splicing instead of changes in overall transcript expression (Figure 16).

A splice variant of *LHY* previously shown to arise in cool (12°C) to cold (4°C) temperatures is derived from retention of intron 1 (*LHY I1R*) (James et al. 2012). In WT seedlings, *LHY I1R* accumulated to higher levels as ambient temperature was reduced from LD|28°C to LD|16°C (Figure 14A,F). Increased amounts of *LHY I1R* were visually apparent in *sic-3* and *sic-1* at all temperatures. *LHY I1R* reaches its highest levels in mutants grown under LD|22°C and LD|16°C (Figure 14A,F). Bulk *LHY* transcript levels in the *sic* mutants were somewhat lower than in WT, indicating the observed changes represent authentic elevation of *LHY I1R* accumulation (Figure 15).

Retention of *CCA1* intron 4 (*CCA1 I4R*) was proportional to ambient temperature in WT, so that the most *CCA1 I4R* transcript appeared at LD|28°C and the least occurred at LD|16°C (Figure 14B,G), in agreement with previous observations (James et al. 2012; Seo et al. 2012 ref). *sic-3* and *sic-1* exhibited clear accumulation of *CCA1 I4R* at all temperatures, and levels were significantly higher in *sic-1* (Figure 14B,G). It is notable that *CCA1 I4R* levels in *sic-1* were significantly higher than in WT at LD|16°C, although this was not the condition where *CCA1 I4R* reaches peak levels for either genotype. The increase in *CCA1 I4R* levels was not due to higher bulk *CCA1* transcript levels (Figure 16D).

Both *sic* alleles permitted the accumulation of a *PRR7* splice variant from retention of intron 4 (*PRR7 I4R*) and an *ELF3* splice variant from retention of intron 2 (*ELF3 I2R). PRR7 I4R* levels in WT did not vary in a temperature-dependent manner

(Figure 14C,H), but the level of this splice variant was elevated in *sic-3* at LD|22°C and LD|16°C (Figure 14C,H), and it appeared at all temperatures in *sic-1*. *ELF3 I2R* in WT had a cool temperature-dependent accumulation pattern in which levels were highest under LD|16°C (Figure 14D,I). *sic-3* and *sic-1* accumulated *ELF3 I2R* under LD|22°C, and *sic-1* generated significantly higher levels of this splice variant at LD|16°C (Figure 14 D,I). Like the bulk transcripts from the other genes tested, increased expression of *PRR7* and *ELF3* does not explain the splice variant increase in the *sic* mutants (Figure 16B,C). Strikingly, accumulation of other described splice variants for *CCA1*, *LHY*, *PRR7*, and *ELF3*, as well as known splice variants for *PRR9*, *GI*, *TOC1*, *PRR5* and *PRR3*, appeared unchanged in either *sic* allele (Figure 15). It is possible, however, that the pooling strategy used here masked the presence of low abundance splice variants. Nevertheless, these results indicate that *sic* mutants have increased abundance of splice variants for *LHY, ELF3, CCA1,* and *PRR7*. In addition, *sic* alleles have a broader range of temperature conditions under which these splice variants occur, particularly at cool temperatures.

sic-3 substantially changes the normal transcriptional response to cool temperatures

Whole transcriptome sequencing by RNA-seq was performed on WT and *sic-3* plants grown under LD|22°C then transferred to either LL|28°C or LL|16°C; after 8 hour acclimatization to the new temperature, tissue was collected and pooled over 24 hours (Figure 17A). These conditions are different from the RT-PCR experiment above (Figure 14) because the RNA-seq tissue was collected from plants in constant light (vs. LD cycles for RT-PCR) and after transfer from 22°C to either warm 28°C or 16°C (Figure 17A). The goal of this experiment was to assess the contribution of *SIC* to transcriptional responses occurring upon a change in temperature conditions. We expected the immediate transcriptome reaction to contrasting temperature changes would reveal novel aspects of the circadian clock response to temperature changes, as well as more global aspects of how Arabidopsis reconfigures transcription to adapt to a shift in ambient temperature.

To assess gene expression, comparisons were made between genotypes (WT vs. *sic-3)* and temperatures (LL|16°C vs. LL|28°C) to identify significant changes in gene expression (false discovery rate (FDR) of ≤0.05) caused by temperature and/or *sic-3* (Figure 17B). Between WT seedlings transferred to LL|16°C and to LL|28°C, 13% of expressed genes exhibit increased or decreased expression, while 16.1% of the transcriptome is differentially expressed in *sic-3* (Table 5). When both *sic-3* and WT were at 16°C, a total of 13.5% of genes were differentially expressed between the genotypes (Table 5). The effect of *sic-3* is somewhat weaker at 28°C, since only 9.5% of genes are different between the mutant and WT under this condition (Table 5).

Several circadian clock genes were among those differentially expressed between 16°C and 28°C in WT. Both *LHY* and *CCA1* were substantially up-regulated at LL|16°C compared to LL|28°C (Figure 18A). Other circadian clock genes were down-regulated at LL|16°C compared to LL|28°C (Figure 18A), including other members of the *RVE* family *RVE4* and *RVE7*, three members of the *PRR* family *TOC1*, *PRR3*, and *PRR7,* and Evening Complex associated genes *ELF3*, *LUX*, and *BOA* (Figure 18A).

Similar to WT, *LHY* and *CCA1* are also up-regulated at LL|16°C in *sic-3* plants compared to LL|28°C, and *RVE7*, *TOC1*, *PRR7*, *LUX*, and *BOA* were down-regulated (Figure 18B). In contrast to WT, many more circadian clock genes were differentially expressed in *sic-3* at LL|16°C compared to LL|28°C (Figure 18A vs B), indicating that SIC is important for temperature-dependent regulation circadian clock genes.

Comparing *sic-3* to WT within each temperature condition revealed a more profound effect of the mutant at cool temperatures (Figure 18C vs. D). At LL|16°C, most differentially expressed circadian clock genes were up-regulated in *sic-3*. Notably, the circadian clock genes that were differentially expressed under LL|16°C *sic-3* plants compared LL|28°C were also differentially expressed in *sic-3* at LL|16°C compared to WT at LL|16°C; these were *CCA1, PRR5, RVE8, RVE2, RVE1,* and *LNK1* (Figure 18B vs. C). Therefore, these genes are not only up-regulated in the cold in *sic-3* mutants, but they are also up-regulated to an even greater extent than WT. Missing from this list was *LHY*, which can be interpreted as *LHY* having a similar induction at LL|16°C in both WT and *sic-3* plants. At LL|28°C, comparison of *sic-3* to WT showed less genes are differentially expressed between the genotypes than at LL|16°C, namely *RVE7*, *PRR7, LNK1, CCA1,* and *CHE* (Figure 18D). All of these genes, aside from *CHE,* were also up-regulated in *sic-3* at LL|28°C, indicating that the *sic-3* mutation causes an overall up-regulation of *CCA1, PRR7, RVE7,* and *LNK1* across both warm and cool temperature transitions which is not present in WT (Figure 18D). *sic-3* therefore shows larger changes in differential expression of circadian clock genes than WT, which is more apparent in cool temperatures.

qPCR was used to confirm several of the gene expression changes observed in the RNA-seq experiment (Figure 19; Figure 20; Figure 21; Figure 21; Figure 22; Figure 23). Indeed the induction of *LHY* (Figure 19) and *CCA1* (Figure 20), and the down-regulation of *TOC1* (Figure 21) and *LUX* (Figure 22) at LL|16°C was confirmed in both WT and *sic-3*. Furthermore, the increased expression of *CCA1* (Figure 20) and *PRR7* (Figure 23) at LL|16°C in *sic-3* compared to WT was also observed by qPCR, but the same was not true for the *CCA1* induction in LL|28°C *sic-3* plants (Figure 20).

sic-3 substantially changes the normal alternative splicing response to cool temperatures

The same RNA-seq experiment was employed to analyze the differential accumulation of splice variants in circadian clock genes between WT and *sic-3* plants transferred to either LL|16°C or LL|28°C. Comparisons were made between genotypes (WT vs. *sic-3)* and temperatures (LL|16°C vs. LL|28°C) to identify significant changes (FDR ≤0.1) in alternative splicing caused by temperature and/or *sic-3* (Figure 17B). In WT, transcriptome-wide only 3.9% of expressed transcripts experienced changes in splicing between LL|16°C and LL|28°C (Table 5). Notably, a higher percentage of transcripts were differentially spliced in *sic-3*: 16.4% of the transcriptome demonstrated differential alternative splicing between these two conditions (Table 5). A similar effect of *sic-3* was apparent when comparing levels of splice variants between *sic-3* and WT at LL|16°C (Table 5); under this condition 19.3% of the splice variant pool was different between the two genotypes. The same comparison between *sic-3* and WT under LL|28°C revealed a weaker effect of the mutant since only 6.5% transcripts were

differentially spliced (Table 5). Clearly, *sic-3* causes elevated levels of alternative splicing and its effect is more substantial at LL|16°C than at LL|28°C.

WT plants transferred to LL|16°C versus transferred to LL|28°C showed differential accumulation of splice variants for *RVE4, RVE8, LHY, CCA1,* and *LNK1* (Figure 24A). Strikingly, the *LHY I1R, LHY L-intR (I5),* and *LHY E5R* splice variants, which have been described as induced by cold in previous studies (A. B. James et al. 2012; S. A. Filichkin et al. 2015), were less abundant at LL|16°C than at LL|28°C (Figure 24A). This observation is also at odds with our previous RT-PCR and qPCR findings that showed an increase in the accumulation of these splice variants in the cold (Figure 14; Figure 26). *LHY* showed a decrease in accumulation of three other novel splice variants at LL|16°C, *LHY I2R, LHY I3R,* and *LHY I4R*, which makes *LHY* the circadian clock gene with the greatest number of differentially accumulated splice variants in WT (Figure 24A). Two other previously reported splice variants that were suggested to have a temperature-dependent accumulation (Allan B. James et al. 2012), *RVE8 I3R* and *RVE8 I7 Alt 3'ss*, had an increased accumulation in WT at LL|16°C (Figure 24A). Of the genes that have differential accumulation at LL|16°C versus LL|28°C in WT, only *RVE4, LHY,* and *CCA1* also show differential expression in the same conditions (Figure 18A).

Comparison of splice variant accumulation within *sic-3* plants transferred to LL|16°C versus LL|28°C revealed an increase in splice variant accumulation at LL|16°C for *BOA, LKP2, PRR3, RVE4, and RVE1,* and a decrease in splice variant accumulation for *LHY, CCA1, LNK1,* and *LNK2* (Figure 24B). Similar to WT, *sic-3* had decreased 16°C-accumulation of *LHY I2R, LHY I3R, and LHY I4R*, together with the novel splice variant *LHY I8R*. Contrary to our expectations based on Figure 14A, levels of *LHY I1R* splice were not significantly different in *sic-3* between temperatures (Figure 24C), and neither were levels of the previously reported splice variants *LHY L-intR (I5), LHY E5R,* and *CCA I4R* (Figure 24C). In both WT and *sic-3* there was decreased accumulation of *CCA1 I4R Alt3'ss,* a *CCA1* splice variant that uses an alternative 3'ss within intron 4 causing a partial intron retention. The TAIR and Ensembl databases predict the *CCA1 I4R Alt3'ss* splice variant shifts the start codon upstream in the gene, which would ultimately eliminate the MYB-like domain of CCA1 if translated. Two new splice variants arise as differentially accumulated in *sic-3* at LL|16°C, *CCA1 I1R* and *CCA1 I2R* (Figure 24B). Furthermore the two *LNK1* splice variants with decreased accumulation at LL|16°C in WT were also present in the *sic-3* comparison (Figure 24B). Overall, *sic-3* had a larger number of individual genes and circadian clock gene splice variants that were differentially accumulated at LL|16°C compared to LL|28°C than WT (Figure 24A vs. B). In addition, these were not always the expected splice variant one sees in previously reported studies. Finally, of the genes that had differential splice variant accumulation at LL|16°C versus LL|28°C in *sic-3,* overall differential gene expression was significantly different in all genes but *RVE4* and the *LNKs* (Figure 18B vs. Figure 24B). Therefore differential gene expression may affect the accumulation of splice variants, or vice versa.

When comparing the differentially accumulated splice variants of circadian clock genes between WT and *sic-3* within temperatures, multiple circadian clock genes were differentially accumulated at 16°C and none were differentially accumulated at 28°C (Figure 24C). Under LL|16°C, *sic-3* had an increase in *LHY I1R* in *sic-3* compared to

WT, which is consistent with our findings in Figure 14A . Furthermore, *LHY L-intR* and *CCA1 I4R Alt 3'ss* were also more abundant in *sic-3* than WT (Figure 24C). A novel splice variant of *LHY, LHY I5R Alt 3'ss*, was more abundant in *sic-3* (Figure 24C); this splice variant results from a partial retention of intron 5 (I5 or the large intron(L)) and introduces a PTC early in the CDS. The *RVE4* and *RVE1* splice variants found in the *sic-3* 16°C/28°C comparison (Figure 24B) were also more abundant in *sic-3* than WT at LL|16°C. Novel splice variants in *RVE2, RVE6, TOC1,* and *TIC1* showed differential accumulation in *sic-3* versus WT at 16°C (Figure 24C). Notably, *RVE4, CCA1, RVE1,* and *RVE2* expression was also altered in *sic-3* at 16°C.

To validate the RNA-seq findings, qPCR was used to assess the accumulation pattern of *LHY* splice variants in identical samples (Figure 25; Figure 26). The accumulation pattern for *LHY I1R* was compared to the *LHY* splice variant that removes intron 1 (*LHY I1S)*. Sets of qPCR primers were designed to specifically quantify removal or retention of intron 1 in *LHY*. This approach is in contrast to the *LHY* primers qPCR used to determine *LHY* expression levels in previous experiments (Figure 19). The prior set of primers reports expression of all forms of the *LHY* transcript by virtue of targeting a shared sequence within the 3' end of the CDS. The qPCRs measuring spliced intron 1 retention, *LHYI1S,* confirmed an induction of *LHY* under LL|16°C in both genotypes, and showed an insignificant difference in the accumulation of the *LHY I1S* splice variant between WT and *sic-3* at all temperatures (Figure 25B). Interestingly *LHY I1S* was decreased in *sic-1* compared to WT (Figure 25B). The qPCRs measuring retained intron 1 did not confirm the RNA-seq results of a decrease in the accumulation of the *LHY I1R* splice variant under LL|16°C compared to LL|28°C for both WT and *sic-3* (Figure 26B). The higher accumulation of *LHY I1R* in *sic-3* under LL|16°C was confirmed, while under LL|28°C there was no difference in accumulation between genotypes (Figure 26B). *sic-1* was not measured in the RNA-seq experiment, but it showed a more drastic accumulation of *LHY I1R* in the cool compared to WT and *sic-3* (Figure 26B).

sic mutants delay the phase of circadian clock rhythms

In addition to confirming the RNA-seq data, qPCR tests were performed to determine the temporal expression behavior of *LHY, CCA1, TOC1, LUX,* and *PRR7* expression in WT, *sic*-3, and *sic-1* seedlings transferred from LD|22°C to either LL|16°C, LL|22°C, or LL|28°C (Figure 19; Figure 20; Figure 21; Figure 22; Figure 23). These qPCRs were performed with the same time points that were combined for the RNA-seq experiment, but they were not pooled. Each transcript maintained rhythmic expression but peak expression for all was substantially delayed in *sic-3* and *sic-1* compared to WT (Figure 19; Figure 20; Figure 21; Figure 22; Figure 23). For example, *LHY* peaked between ZT 24-28 in WT in all three temperatures, but in *sic-3* and *sic-1* *LHY* peaked at ZT 28 and ZT 32, respectively, under LL|16°C, and both peaked at ZT 28 under LL|28°C (Figure 19). Furthermore, the phase delay in *sic-3* and *sic-1* was accentuated in the cool temperature range for *LHY, TOC1,* and *PRR7* (Figure 19; Figure 21; Figure 23). This phase delay is consistent with the long period phenotype of each *sic* mutant. Also, the effect of temperature on rhythmic amplitude and expression level was apparent in WT. As predicted from the RNA-seq experiment, *CCA1* and *LHY* transcript levels and amplitude increased as temperature decreased (Figure 19; Figure 20), while the opposite was true for *TOC1, LUX,* and *PRR7* (Figure 21; Figure 22;

Figure 23). Finally, transfer of WT to either LL|16°C or LL|28°C did not cause any detectable resetting of the clock, which would have appeared as obvious shifts in peak gene expression earlier or later than observed under LL|22°C (Figure 19; Figure 20; Figure 21; Figure 22; Figure 23).

Discussion

In Arabidopsis, alternative splicing is influenced by ambient temperature conditions to modify the composition of the splice variant pool for most circadian clock genes, often making nonproductive transcripts with PTCs (James et al. 2012). These changes in the splicing of clock transcripts were observed in response to external temperatures, but the ultimate cause behind their formation, and the effect they play on the circadian clock is not known (James et al. 2012). While a mechanism explaining the link between alternative splicing of clock genes and temperature compensation is well defined in *N. crassa*, no such explanation exists for plants. One mechanism plants could employ to perceive temperature and regulate circadian rhythms may involve alternative splicing of clock transcripts. Because *sic-1* was previously shown to alter transcript splicing (Zhan et al. 2012), SIC is potentially important for this aspect of temperature perception and response. To test this idea, the two *sic* mutant were evaluated for changes in temperature-dependent alternative splicing.

The combination of circadian clock and alternative splicing phenotypes present in *sic* demonstrate that the Arabidopsis circadian clock requires SIC to maintain appropriate levels of alternative splicing for circadian clock transcripts. The temperature-dependent alternative splicing phenotype of *sic* was first identified using RT-PCR and revealed changes in splice variant accumulation for *LHY, CCA1*, *ELF3,* and *PRR7* transcripts brought on by cool temperature conditions (Figure 14). PTCs can be predicted for *ELF3 I2R* and *PRR7 I4R*, but there is no evidence that these splice variants are subject to NMD or result in the production of truncated proteins. The *CCA1 I4R* splice variant is not degraded by NMD (James et al. 2012); instead, it is translated to produce a truncated CCA1 protein that interferes with the activity of full-length protein (Seo et al. 2012; Filichkin et al. 2015). *LHY I1R* is degraded by NMD under some conditions (James et al. 2012). Unlike the other splice variants prevalent in *sic*, *LHY I1R* does not introduce a PTC or change the coding potential of the transcript because the retained intron interrupts the normal 5'UTR. It is possible that this alternate 5'UTR interferes with ribosome loading or has reduced translation initiation that ultimately triggers NMD degradation (G. Wang, Guo, and Floros 2005).

Previously, the *sic-1* allele was shown to have reduced tolerance to cold temperature chilling (*i.e.*, 4°C) (Zhan et al. 2012), which may be related to, but distinct from, the cool ambient temperature phenotypes observed here. In addition, *sic-1* exhibits reduced accumulation of certain miRNAs and, therefore, was implicated in miRNA biogenesis (Zhan et al. 2012). Whole genome-tiling array analysis of *sic-1* in the same study found intron retention splice variants for many transcripts, but not those observed here (Zhan et al. 2012). Together, these observations indicate that the miRNA accumulation defect reported for *sic-1* is more likely a consequence of altered transcript splicing than a direct effect on miRNA biogenesis. Furthermore, it is doubtful that the

circadian clock phenotypes in *sic* arise from changes in miRNA biogenesis, since the circadian clock transcripts with disrupted expression in *sic* are not targeted by miRNAs.

RNA-seq analysis was conducted to capture the whole of transcriptomic changes in gene expression and alternative splicing, including circadian clock transcripts (Figure (Figure 17; Figure 18; Figure 24; Table 5). This experiment was designed to reveal the rapid transcriptomic response of WT and *sic-3* seedlings that were transferred from LL|22°C to either LL|16°C or L|28°C. Different results in alternative splicing were observed between the RNA-seq experiment and the RT-PCR experiment when testing temperature-transferred seedlings compared to seedlings that were grown in constant cool or constant warm temperatures (as in the first RT-PCR experiment) (Figure 14 vs. Figure 24). As expected, the RNA-seq experiment revealed a broader spectrum of genes modified by the *sic* mutation both in a temperature dependent manner and in comparison to WT. Loss of SIC affected 10-20% of the transcriptome, not just circadian clock genes (Table 5), depending on the temperature conditions. Under LL|16°C, 13.5% of genes showed differential expression between *sic-3* and WT, while only 9.5% were differentially expressed under LL|28°C (Table 5). The number of transcripts experiencing alternative splicing in *sic-3* was greater than in WT. Under LL|16°C, 19.5% of transcripts in *sic-3* exhibited alternative splicing compared to WT, while fewer than 7% of transcripts were alternatively spliced under LL|28°C (Table 5). Therefore, *SIC* activity is necessary to keep alternative splicing in check when Arabidopsis experiences a decrease in ambient temperature. The absence of *sic* has a negative impact on the transcriptional response of plants to temperature changes, particularly in the cold.

An interesting response of WT seedlings was strong temperature dependence for *LHY* and *CCA1* expression (Figure 18). A 8-fold difference in expression was apparent for these genes between LL|28°C and LL|16°C. By comparison to expression levels at 22°C in the qPCR experiment, *LHY* and *CCA1* appear to be repressed at 28°C and induced at 16°C (Figure 19). Similar regulation of *LHY* and *CCA1* expression has been reported previously (Gould et al. 2006; A. B. James et al. 2012). Thus, the circadian system appears to respond to temperature conditions by modulating expression of these two core circadian clock genes.

Known regulators of *LHY* and *CCA1* expression are obvious candidates to enforce the temperature-dependent expression on these genes. TOC1 is a repressor of *LHY* and *CCA1* (Nohales and Kay 2016). While *TOC1* expression does decrease under LL|16°C relative to LL|28°C, the magnitude of this change is quite small (~2-fold), which is not consistent with the strong repression of *LHY* and *CCA1* (Figure 18; Figure 19; Figure 20). Furthermore, *PRR7* is another target of TOC1-mediated repression, but *PRR7* expression actually decreases under LL|16°C, which is not in line with a large reduction in TOC1 activity at 16°C (Figure 18; Figure 23). Therefore, *TOC1* may contribute to the induction and repression of *LHY* and *CCA1* in a temperature dependent manner, but other factors most likely contribute to this phenomenon. In the future this could be tested by measuring *LHY* and *CCA1* expression under LL|16°C and LL|28°C in a *toc1* mutant.

The most notable difference between the RNA-seq and RT-PCR experiments was the decrease in *LHY I1R* splice variant accumulation in WT under LL|16°C relative to LL|28°C shown by the RNA-seq experiment (Figure 24A). A potential explanation for

this different result is in the way in which the two techniques determine transcript abundance. RT-PCR and qPCR depend on normalization to an external standard gene within the same sample, which is often a constitutively expressed gene. The approach used for RNA-seq was to normalize abundance only to the average expression of the same gene across all samples. With these different normalization approaches, RNA-seq may be more effective at revealing the actual contribution of *LHY I1R* to the overall transcript pool arising from each gene in a specific condition or genotype. Therefore *LHY I1R* may actually increase under LL|16°C, as reported by RT-PCR and previous experiments (A. B. James et al. 2012), but it represents a smaller proportion of the overall *LHY* transcript pool compared to LL|28°C due to the large induction of *LHY* in cool temperatures (Figure 18A). Discrepancies between the experimental techniques used to quantify splice variant accumulation must be considered in the future, and we must determine which technique most accurately represents the true changes in splice variant accumulation.

The enhancement of *LHY* splice variants in *sic*, indicates *SIC* is crucial for control over *LHY* alternative splicing. At least three *LHY* splice variants show increased accumulation in *sic-3* under LL|16°C compared to WT (Figure 24C): *LHY I1R,* which modifies the 5'UTR, *LHY L-intR,* and *LHY I5 Alt 3'ss* which both introduce a PTC. The temperature-dependent *LHY* splice variants that modify the *LHY* CDS introduce either a PTC or relocate the start codon and eliminate crucial domains of the protein, such as the MYB-like DNA binding domain. Clearly, loss of SIC function affects the splice variant accumulation of *LHY* in the cold, which is predicted to change LHY protein expression and, consequently, appropriate circadian clock responses to temperature.

CCA1, like its *LHY* homolog, shows strong temperature-dependent induction in cool temperatures and differential alternative splicing. In WT, alternative splicing of Intron 4 leads to three differentially accumulated splice variants of *CCA1* between LL|16°C and LL|28°C (Figure 24A). The *CCA1 I4R* splice variant accumulates in the expected manner in RNA-seq experiment: *CCA1 I4R* increases under LL|28°C compared to LL|16°C (Figure 24A). This result matches previously reported experiments (Seo et al. 2012; S. A. Filichkin et al. 2010) and our RT-PCR experiment (Figure 24A). Loss of SIC function changes the temperature-dependent *CCA1* splice variant accumulation, of which only *CCA1 I4R Alt3'ss* splice variant is different in *sic-3* compared to WT under LL|16°C (Figure 24C). The *CCA1 I4R Alt3'ss* splice variant is different than the *CCA1 I4R* splice variant, which Seo et al. 2012 reports is translated and interferes with LHY and CCA1 homo- and hetero-dimerization in a temperature-dependent manner (Seo et al. 2012). It is unknown what is the role of the *CCA1 I4R Alt3'ss* splice variant, but if translated it would lead to a truncated CCA1 protein with no MYB-like domain that could also interfere with LHY and CCA1 protein function.

The mechanistic consequences of changed splice variant accumulation in *sic* remain unclear. Splice variants frequently have alternate upstream open reading frames with PTCs (Lareau et al. 2007; Filichkin and Mockler, 2012). PTCs have the capacity to stall ribosomes, which often leads to transcript degradation by NMD (Lewis, Green, and Brenner 2003; McGlincy and Smith 2008; S. Filichkin et al. 2015). On the other hand, splice variants may encode truncated, partially inactive protein isoforms that interfere with full-length protein activity (Seo et al. 2012; Syed et al. 2012; Filichkin et al. 2015). For the splice variants that do not introduce a PTC in the CDS and instead modify the

UTR, such as *LHY I1R,* it is possible that these splice variants have different translation efficiencies. It was shown in humans that splice variants causing changes in the 5'UTR displayed higher or lower translation efficiencies depending on the sequence (G. Wang, Guo, and Floros 2005). Therefore temperature-induced changes to the available splice variant pool may affect the timing and rate of translation of the corresponding proteins. In the case of *LHY*, overall gene expression is strongly induced by the cold, of which a large proportion of the splice variant pool is full length *LHY* transcript (Figure 18), and a small proportion is the *LHY I1R* splice variant. Although RNA-seq analysis showed the *LHY I1R* splice variant decreases in accumulation under LL|16°C compared to the overall *LHY* transcript pool, this does not negate the fact that *LHY I1R* is prevalent in the cold and *sic* has more of the splice variant than WT under LL|16°C (Figure 14C). We predict the difference in splice variant accumulation of *LHY I1R,* other *LHY* splice variants, and the *CCA1 I4R Alt3'ss* splice variant interferes with efficient translation of the full-length *LHY* and *CCA1* transcript which may interfere with circadian clock function in cool temperatures.

Gene expression also is affected in *sic-3* mutants. A total of 15 circadian clock genes were differentially expressed between 16°C and 28°C in *sic-3*, while only 10 had temperature-dependent expression; seven of the genes overlapped between the two genotypes (Figure 18). Furthermore, most genes that increase in expression in *sic-3* compared to WT also show increased splice variant accumulation (Figure 18 vs. Figure 24). The increase in splice variant accumulation is not due to an increase in overall gene expression, because the RNA-seq analysis method normalizes for this variable. These effects of *sic* may represent changes in transcription, splicing, or both these activities.

Our work contributes to the emerging knowledge that upon a shift down in temperature, the Arabidopsis splice variant pool undergoes significant modifications, including the structure of several transcripts from core circadian clock genes (A. B. James et al. 2012; S. A. Filichkin et al. 2015). Since mutants like *sic* amplify alternative splicing of circadian clock genes and exhibit strong circadian clock defects, the expectation is that modifying of the splice variant pool impacts circadian clock activity. Schlaen et al. 2015 propose that spliceosomal activity is compromised in cool temperatures and specific proteins are needed to buffer spliceosomal function against the effects of temperature (Schlaen et al. 2015). It is possible that SIC plays such a role. However, the consequences of changes in splice variant accumulation for the circadian clock, as well as other plant systems, remains unclear. A possibility is that formation of splice variants has negative effects on the clock mechanism, but WT plants are capable of overcoming these problems. Alternatively, splice variants created upon a change in temperature are useful for circadian clock temperature entrainment and temperature compensation, thus serving as a "signal" of the temperature changes. These two ideas are not mutually exclusive: the efficiency of the spliceosome is compromised in cool temperatures, but it is possible that the plant also uses the pattern in splice variant accumulation as a signal and modifier of circadian rhythm. While the molecular function of SIC remains to be uncovered, it is clear that SIC is important for Arabidopsis to produce an appropriate transcriptional response following a change in temperature environment.

Materials and Methods

Identification and Quantification of Transcript Splice Variants

WT, *sic-3*, and *sic-1* seedlings were grown as described in Chapter 1. Entrainment conditions were either LD|16°C, LD|22°C, or LD|28°C and experiments were always done concurrently. Tissue from 8-day-old seedlings was collected at ZT 0, ZT 6, ZT 12, and ZT 18, followed by RNA extraction and cDNA synthesis as described. cDNA pools for each entrainment condition and genotype were made by combining equal amounts of cDNA from each time point. Splice variants were detected by RT-PCR with primers designed with Primer3 software (Rozen and Skaletsky, 1998) or previously published (James et al., 2012). The desired RT-PCR products were amplified from cDNA pools with *Taq* polymerase and 25 to 35 cycles of amplification depending on the primer set. PCR products were separated in 1.2% agarose gels (Sigma Aldrich) in 1X TAE buffer and bands visualized with a Typhoon FLA 7000 biomolecular imager (GE Life Sciences). Qualitative detection was performed by staining with 0.5 µg/mL of ethidium bromide. Quantitative detection employed FAM fluorophore labeling of PCR products as previously described (Schuelke, 2000; Lu et al., 2011), with the indicated modifications. Forward primers had an additional 5' leader corresponding to the M13(-21) universal sequence (5'-TGTAAAACGACGGCCAGT-3'). Immediately after completion of a standard RT-PCR, a mixture of universal FAM-labeled forward primer (5'-[FAM]-TGTAAAACGACGGCCAGT-3') and reverse primer was spiked into the reaction and amplification allowed to proceed for an additional eight cycles. FAM-labeled bands were quantified with ImageQuant TL 8.1 software (GE Life Sciences). Splice variant abundance for each sample was calculated by subtracting background from the band intensity for the FAM-labeled PCR product. Background was the intensity of an empty section of the same lane identical in area to the PCR band.

Growth conditions for RNA collection designated for qPCR and RNA-seq

WT, *sic-3*, and *sic-1* seedlings were grown as described in Chapter 1. Seedlings were stratified for 3 days, then germinated in LL|22°C for 3 days. Entrainment conditions were LD|22°C for 5 days then transferred to LL|22°C free run. At ZT 12 on the 5th day of entrainment, plants were transferred to either LD|16°C, LD|28°C, or kept at LD|22°C. Tissue was collected after 8 hours acclimatization to the new temperature, or at ZT 20, followed by ZT 24, ZT 28, ZT 32, ZT 36, and ZT 40. Experiments were performed concurrently and three replicate experiments were collected. RNA was extracted as described in Chapter 1 and split into two aliquots. One aliquot was used to evaluate alternative splicing and gene expression changes via qPCR; cDNA synthesis and qPCR was conducted as described in Chapter 1. The second aliquot was allocated to library preparation for RNA-seq analysis.

See Ch.1 for qPCR and plant growth conditions

qPCR experiments from Figures 19-23 and Figures 25 and 26 were only normalized to *IPP2*. Equal amounts of total RNA were pooled per timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C).

RNA-seq library preparation, sequencing, and analysis

RNA from WT and *sic-3* plants transferred to either LD|22°C or LD|28°C was evaluated by RNA-seq. Total RNA was extracted as described above (Chapter 1). RNA concentration was measured using the Qubit RNA HS assay and equal amounts of RNA were pooled from each timepoint for each genotype and growth condition. Poly(A) RNA was purified from 3 µg of total RNA using a Dynabeads Oligo(dT) kit (Invitrogen), and this purification was repeated again to exclude rRNA. After elution of poly(A) RNA with 9uL of 10mM Tris-HCl, RNA-seq libraries were prepared using the ScriptSeq v2 RNA-Seq library preparation kit (Epicentre) based on the manufacturer's protocol. Twelve libraries were indexed using ScriptSeq Index PCR primers (Epicenter) according to the manufacturer's recommendations. Paired-end sequencing at 150bp was performed by Illumina HiSeq4000. Untrimmed reads were aligned to the *Arabidopsis thaliana* TAIR10 genome sequence, downloaded from EnsemblPlants (ftp://ftp.ensemblgenomes.org/pub/plants/release-36/fasta/arabidopsis_thaliana/dna/), with HISAT2 utilizing default settings with *--rna-strandness* set to *FR*. Gene expression and transcript splice variants were quantified with the R package ASpli pipeline utilizing default settings from Mancini E, Yanovsky M and Chernomoretz A (2017). *ASpli: Analysis of alternative splicing using RNA-Seq*. R package version 1.2.2. Differential expression and differential exon/intron/junction usage was considered significant at a false discovery rate of ≤0.05 and ≤0.1, respectively.

Chapter 2 Figures

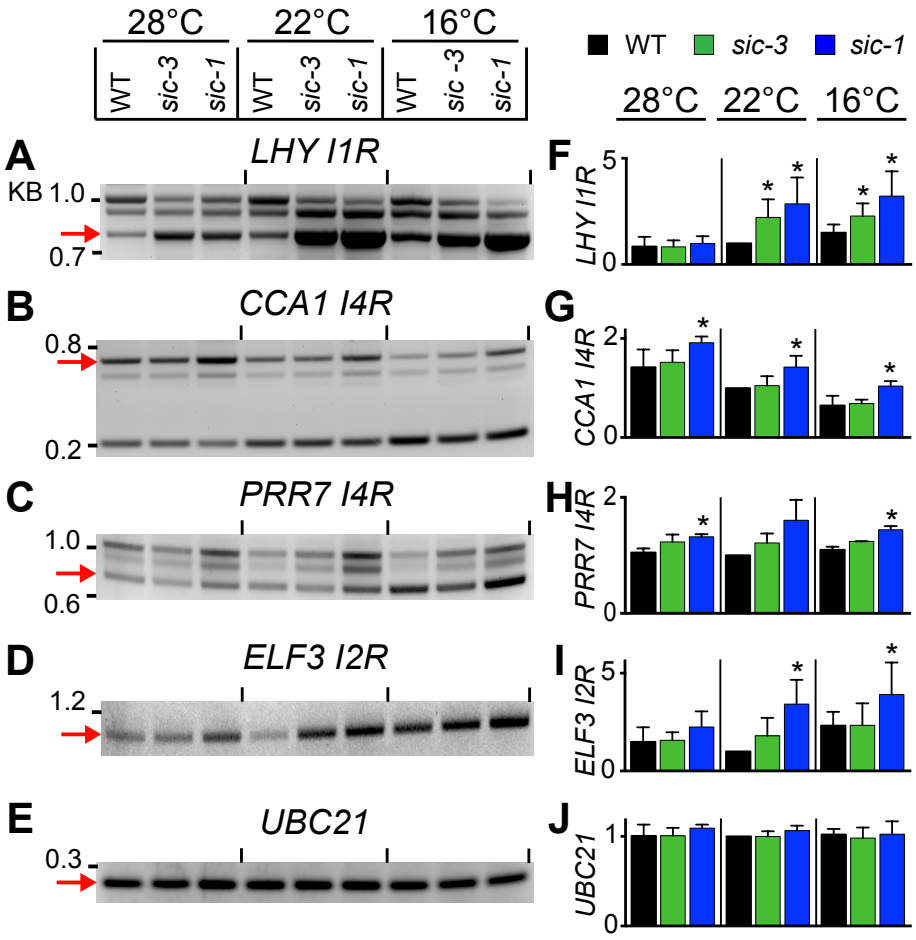


Figure 14. *sic* enhances intron retention in transcripts of *LHY*, *CCA1*, *ELF3*, and *PRR7*.

A) to **D)** Discovery of the splice variants **(A)** *LHY I1R*, **(B)** *CCA1 I4R*, **(C)** *PRR7 I4R,* and **(D)** *ELF3 I2R* in WT, *sic-3*, and *sic-1* seedlings grown under LD|28°C, LD|22°C, or LD|16°C by RT-PCR. Red arrows indicate the specific RT-PCR product.
(E) *UBC21* served as the transcript level reference gene. Bands detected by ethidium bromide staining and images inverted for clarity. Images are representative of 3 independent experiments.
(F) to **(J)** Quantification of FAM-labeled RT-PCR products from WT, *sic-3*, and *sic-1* grown under LD|28°C, LD|22°C, or LD|16°C. Data are the mean of 3 independent experiments. Error bars are standard deviation and asterisk (*) indicates significance p-value <0.05 from two-tailed t-test between WT and the indicated genotype at the same temperature.

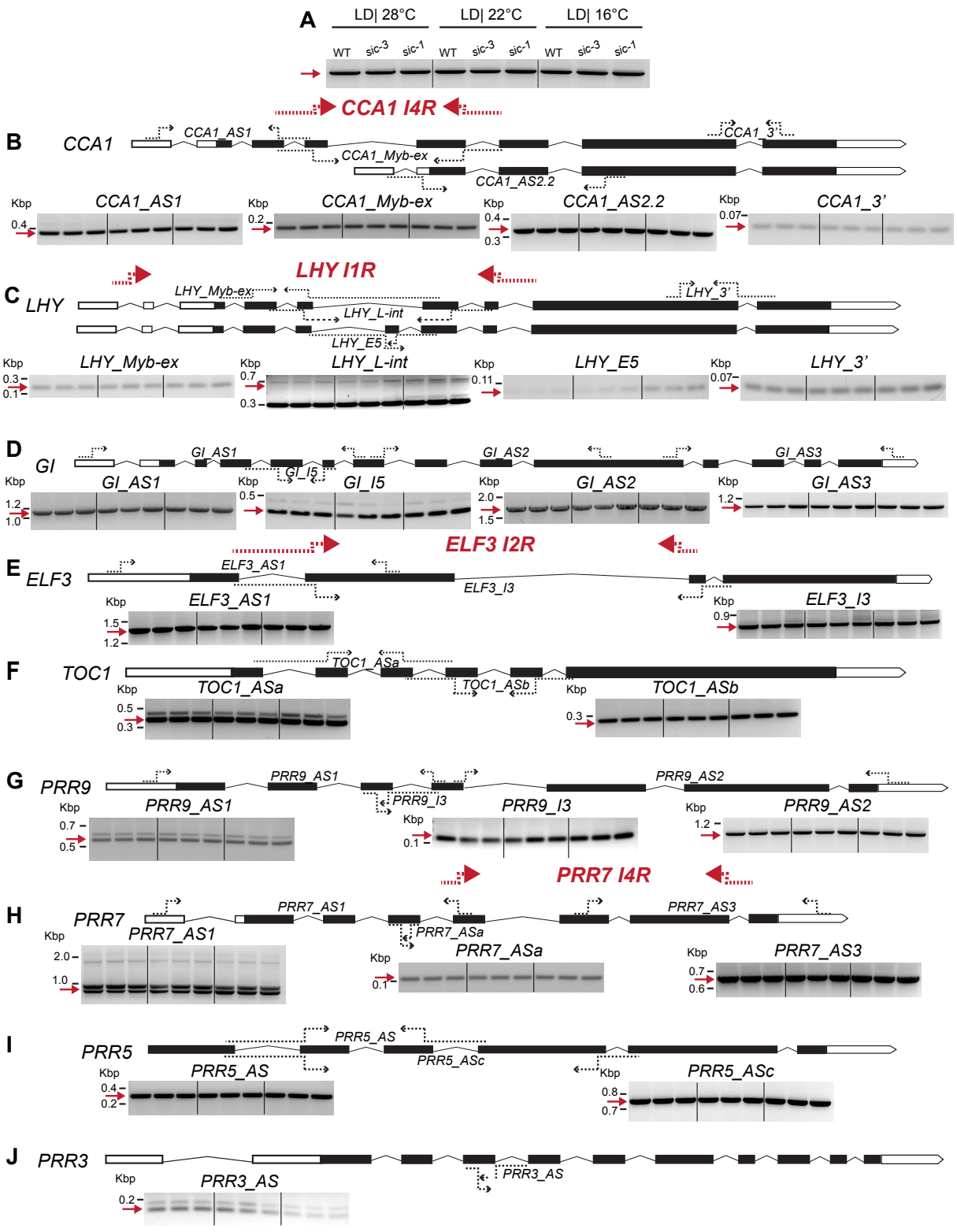


Figure 15. *sic* does not change all splice variants possible for transcripts from *CCA1*, *LHY*, *GI*, *ELF3*, *TOC1*, *PRR9*, *PRR7*, *PRR5* and *PRR3*.

A) Amplification from *UBC21* served as the transcript level reference gene. All other gels loaded identically to this control gel.
B) to **J)** Tests for potential splice variants for **(B)** *CCA1*, **(C)** *LHY,* **(D)** *GI*, **(E)** *ELF3*, **(F)** *TOC1*, **(G)** *PRR9*, **(H)** *PRR7*, **(I)** *PRR5*, and **(J)** *PRR3* in WT, *sic-3*, and *sic-1* seedlings grown in LD|28°C, LD|22°C, or LD|16°C using RT-PCR. Red arrows indicate the specific RT-PCR product. Model of transcript at top of each panel illustrates exon-intron structure and the location of primers (dotted arrows) employed to detect the indicated splice variant. Bold red text and heavy red dotted arrows indicate primer pairs for splice variants shown in Figure 8. White boxes indicate 5'UTR exons, black boxes show coding exons, and thin lines are introns. Bands detected by ethidium bromide staining and images inverted for clarity. Images are representative of 3 independent experiments.

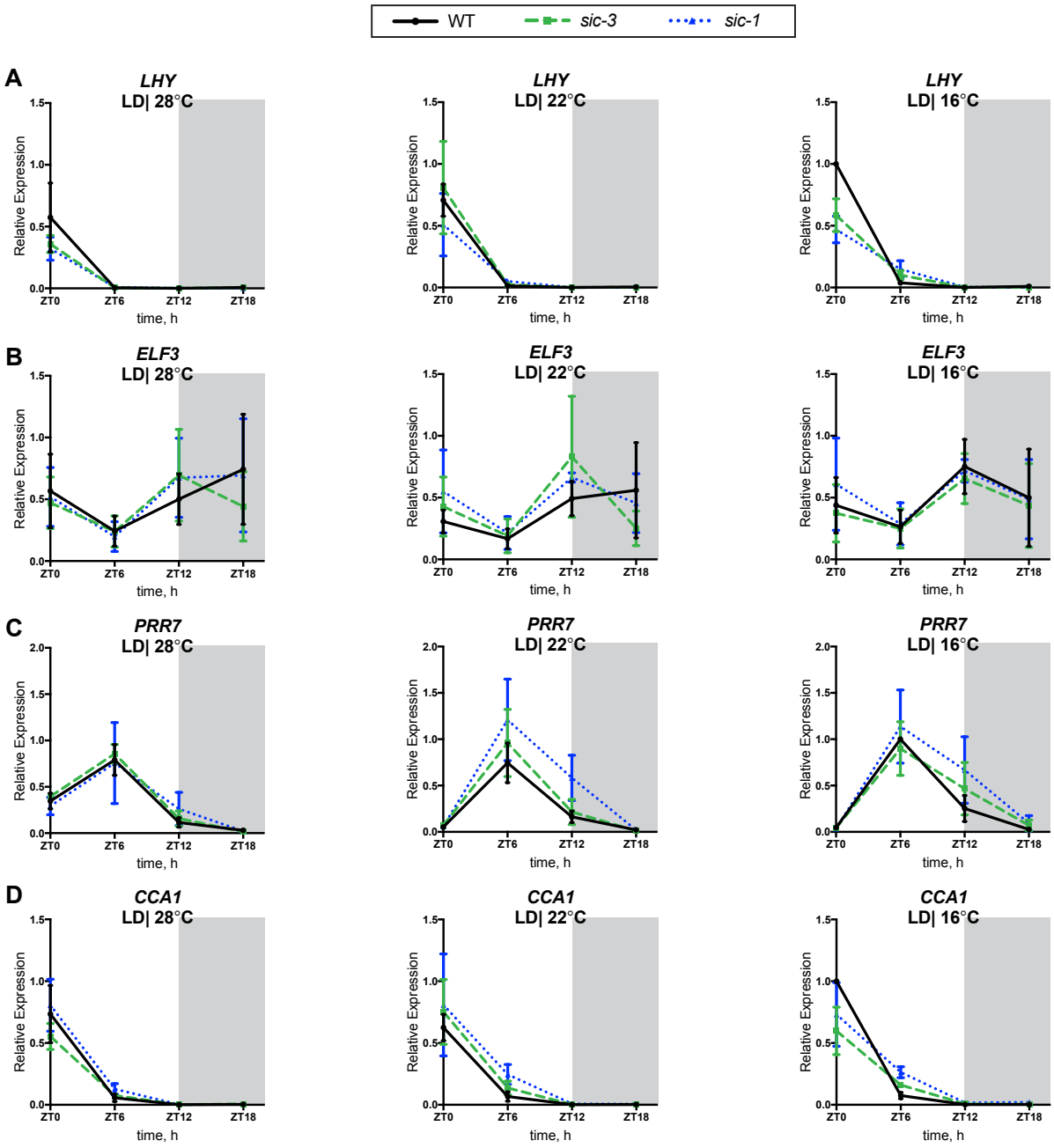
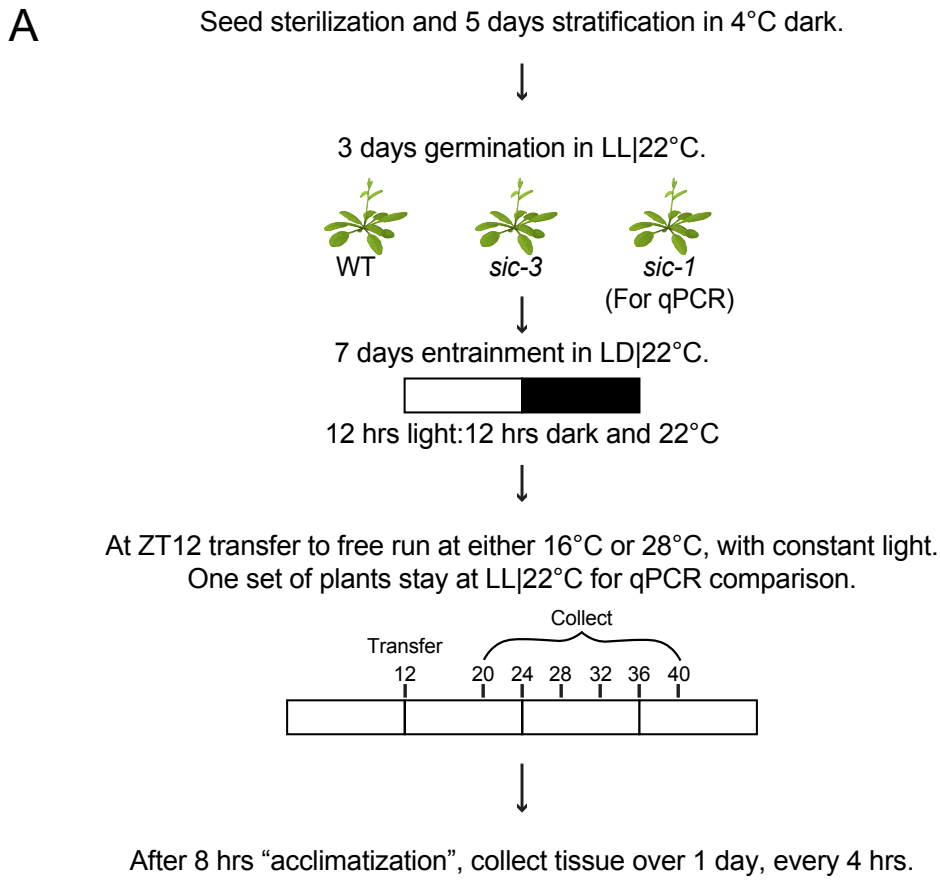


Figure 16. Expression of *LHY*, *ELF3*, *PRR7*, and *CCA1* transcripts in *sic* and WT under conditions for detection of splice variants.

A) to **(D)** Expression levels of **(A)** *LHY*, **(B)** *ELF3*, **(C)** *PRR7*, and **(D)** *CCA1* transcripts in WT, *sic-3* and *sic-1*) grown under LD|28°C, LD|22°C, or LD|16°C as indicated. Grey boxes indicate dark periods. Samples are those pooled for the splice variant analysis in Figure 14 and Figure 15. Transcript levels were determined with qPCR. Each point is mean from 3 independent experiments, and error bars are standard deviation.



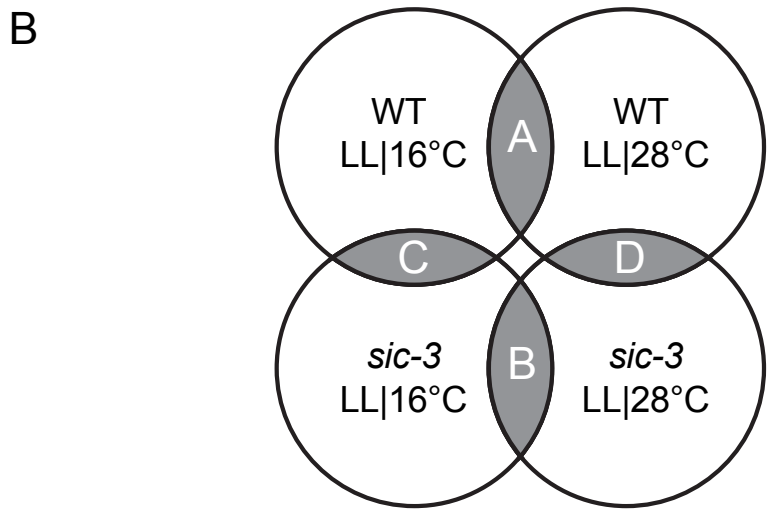


Figure 17. Parameters of RNA-seq experimental growth conditions and RNA-seq analysis comparison.

A) Description of growth conditions for WT, *sic-3*, and *sic-1* seedlings and the times at which seedlings were transferred to entrainment and free-running conditions. Tissue collection occurred during ZT 20-ZT 40 on 11-day-old seedlings. Although tissue from *sic-1* and the seedlings under LL|22°C were not used in the RNA-seq experiment, this tissue was collected concurrently for qPCR analysis.
B) Representation of RNA-seq analysis comparison. Differences in differential alternative splicing and differential expression were measured by comparing WT LL|16°C seedlings to WT LL|28°C seedlings (yielding results A in Figure 18A and Figure 24A), *sic-3* LL|16°C seedlings to *sic-3* LL|28°C seedlings (yielding results B in Figure 18B and Figure 24B), *sic-3* LL|16°C seedlings to WT LL|16°C seedlings (yielding results C in Figure 18C and Figure 24C), and *sic-3* LL|28°C seedlings to WT LL|28°C seedlings (yielding results D in Figure 18D).

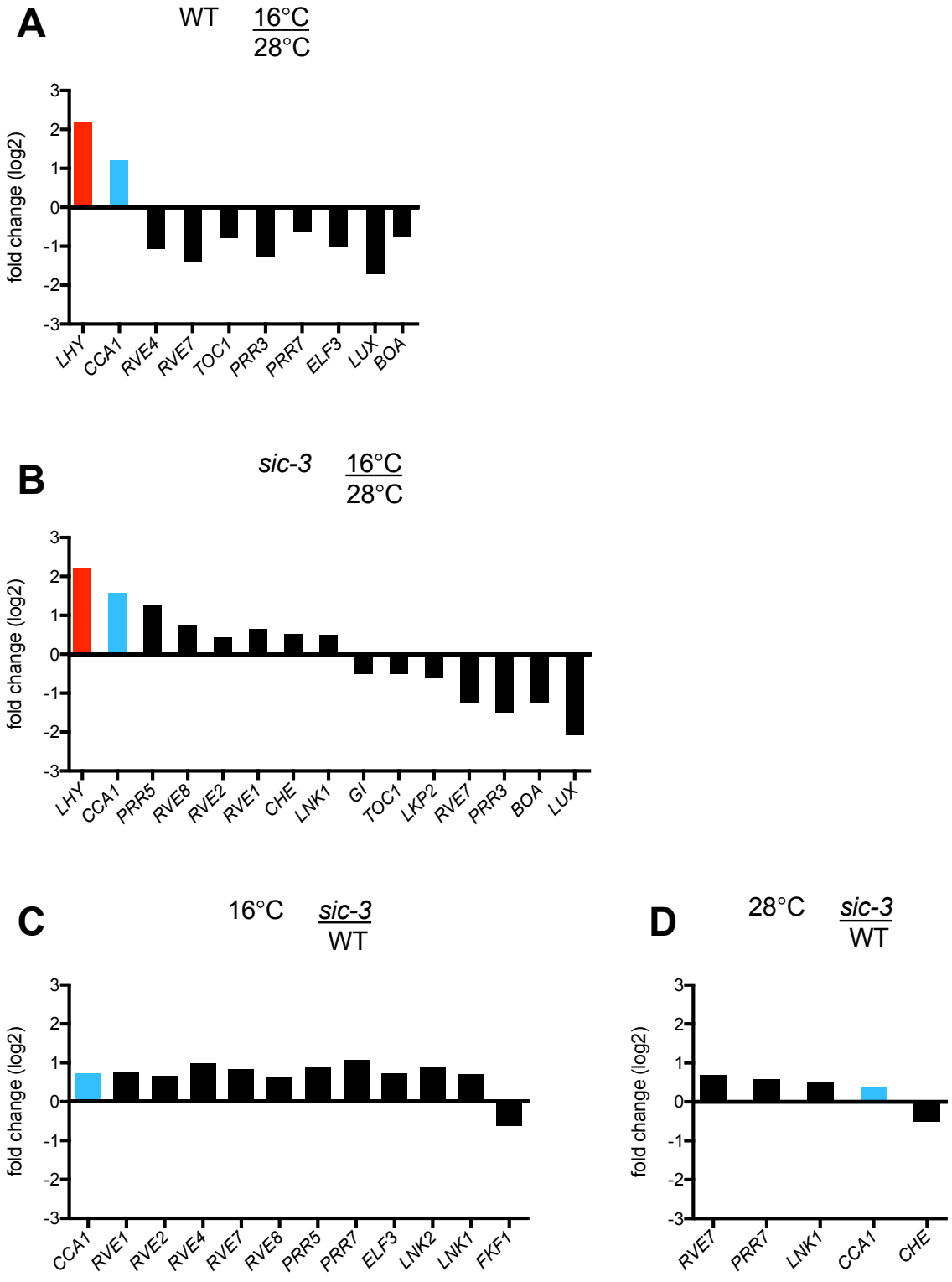


Figure 18. WT and *sic-3* demonstrate differences in temperature-dependent differential expression of circadian clock genes according to RNA-seq.

A) to **(D)** Differential expression of circadian clock genes was identified in whole transcriptome sequencing (false discovery rate (FDR) of ≤0.05) caused by temperature and/or *sic-3* in the following binary comparisons (see Chapter 2 Methods). Y-axis represents the log2 fold change between conditions/genotypes, and X-axis lists the circadian clock genes with significant differential expression. *LHY* is highlighted in red and *CCA1* is highlighted in light blue.
A) Circadian clock genes with differential gene expression between WT seedlings transferred to LL|16°C versus to LL|28°C.
B) Circadian clock genes with differential gene expression between *sic-3* seedlings transferred to LL|16°C versus to LL|28°C.
C) Circadian clock genes with differential gene expression between *sic-3* seedlings transferred to LL|16°C versus WT seedlings transferred to LL|16°C.
D) Circadian clock genes with differential gene expression between *sic-3* seedlings transferred to LL|28°C versus WT seedlings transferred to LL|28°C.

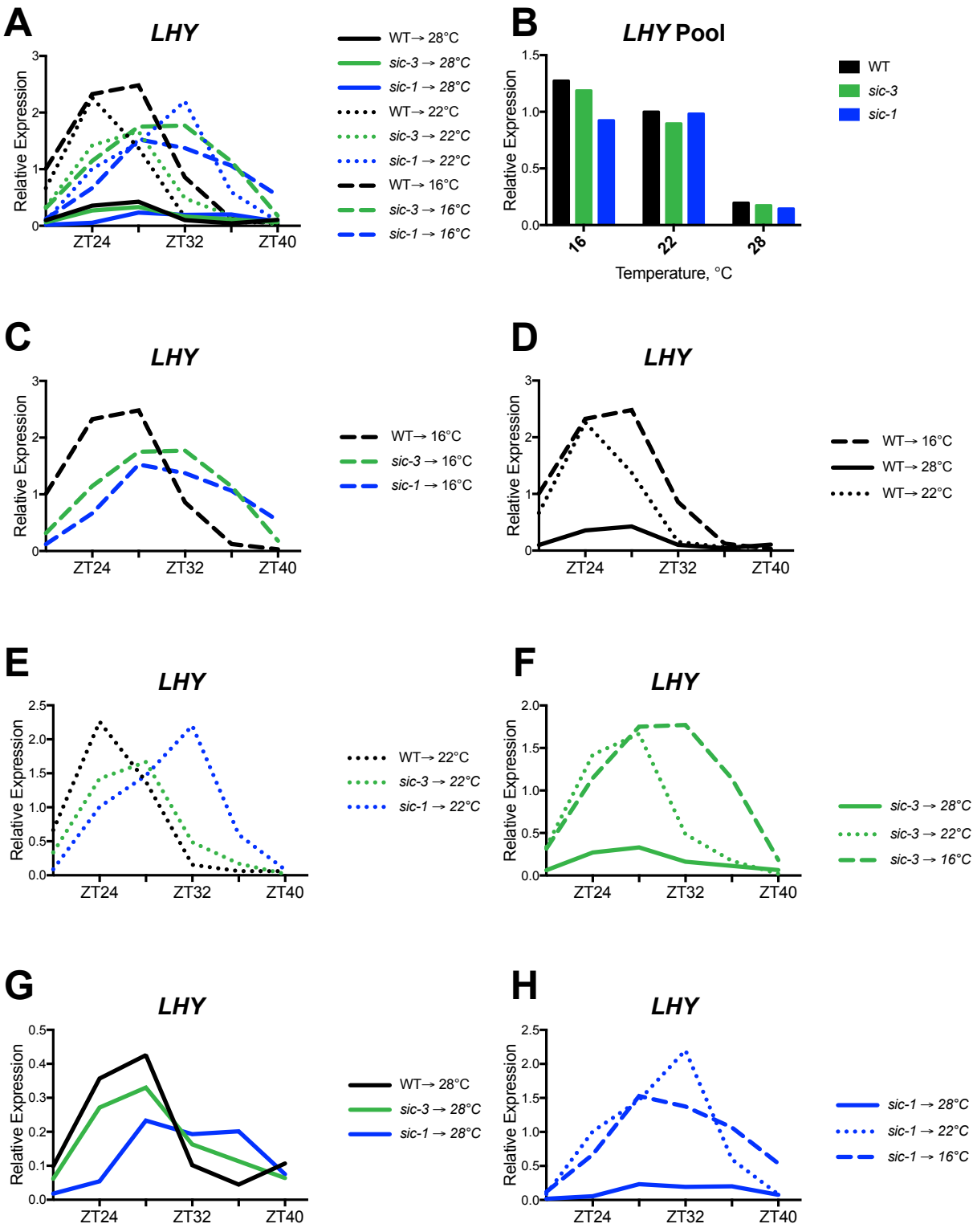


Figure 19. Temperature and the *sic* mutation alter *LHY* expression.

A), C) to **(H)** Expression levels of the *LHY* over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3*

transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

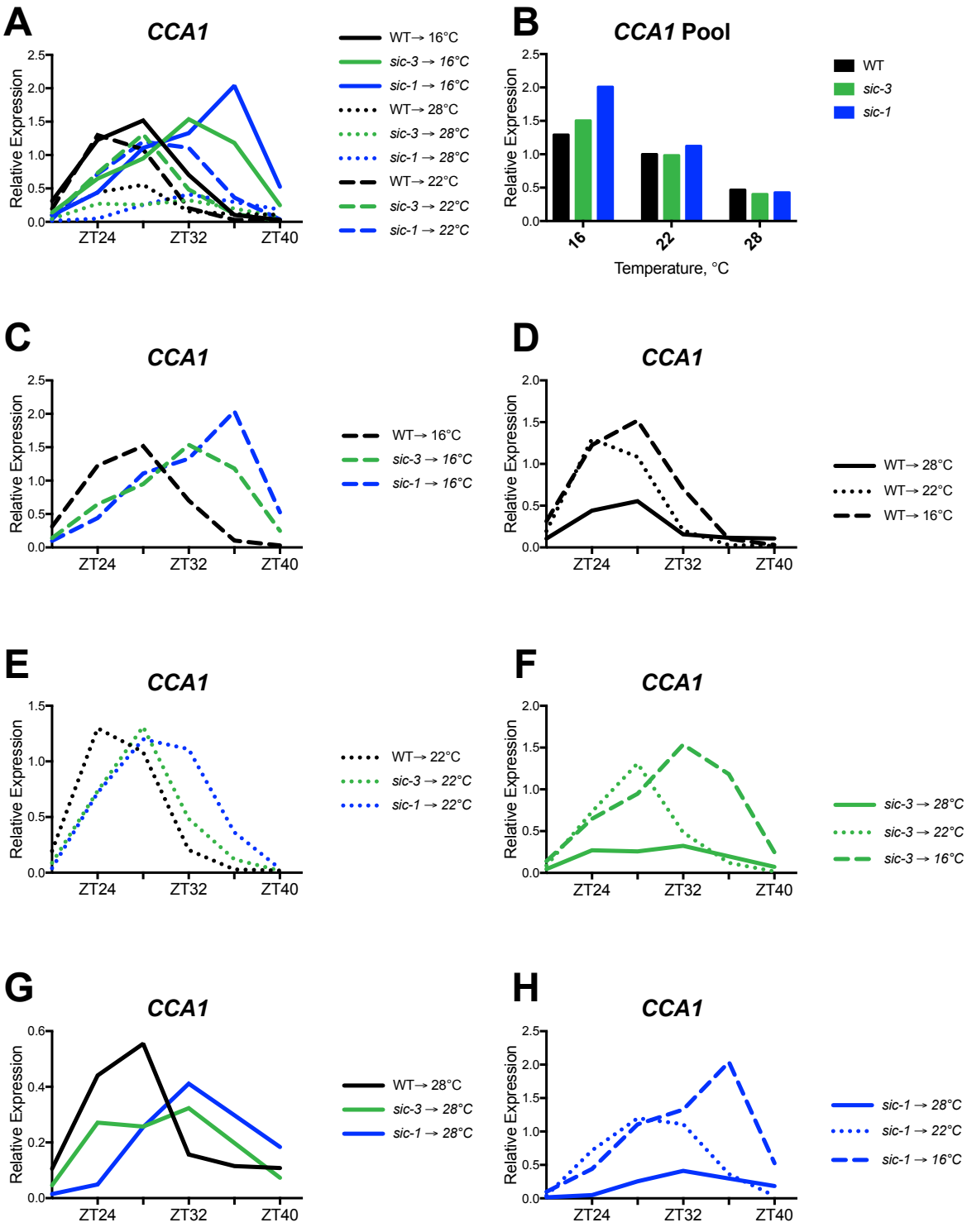


Figure 20. Temperature and the *sic* mutation alter *CCA1* expression.

A), C) to **(H)** Expression levels of the *LHY* over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3*

transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

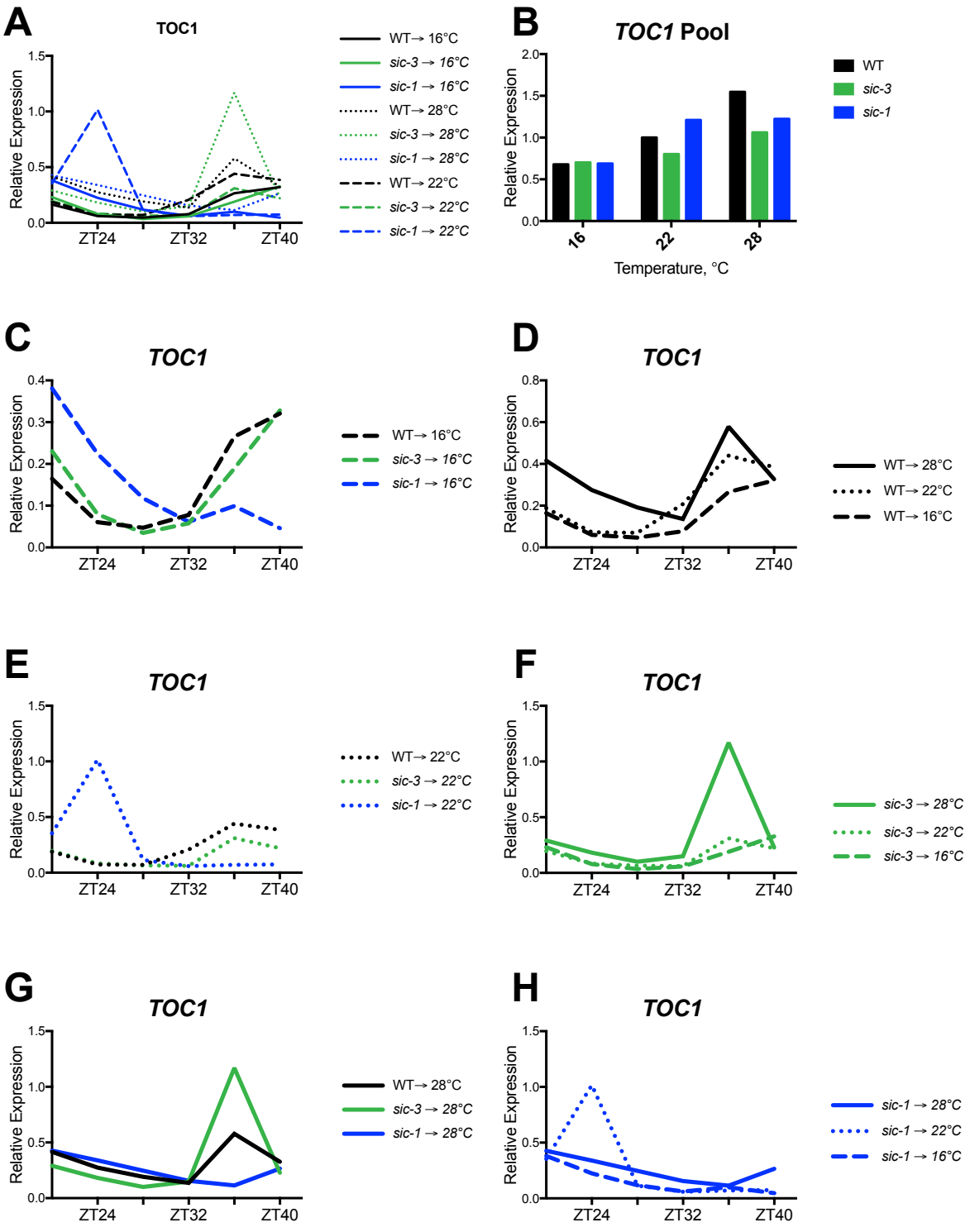


Figure 21. Temperature and the *sic* mutation alter *TOC1* expression.

A), C) to **(H)** Expression levels of the *LHY* over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3*

transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

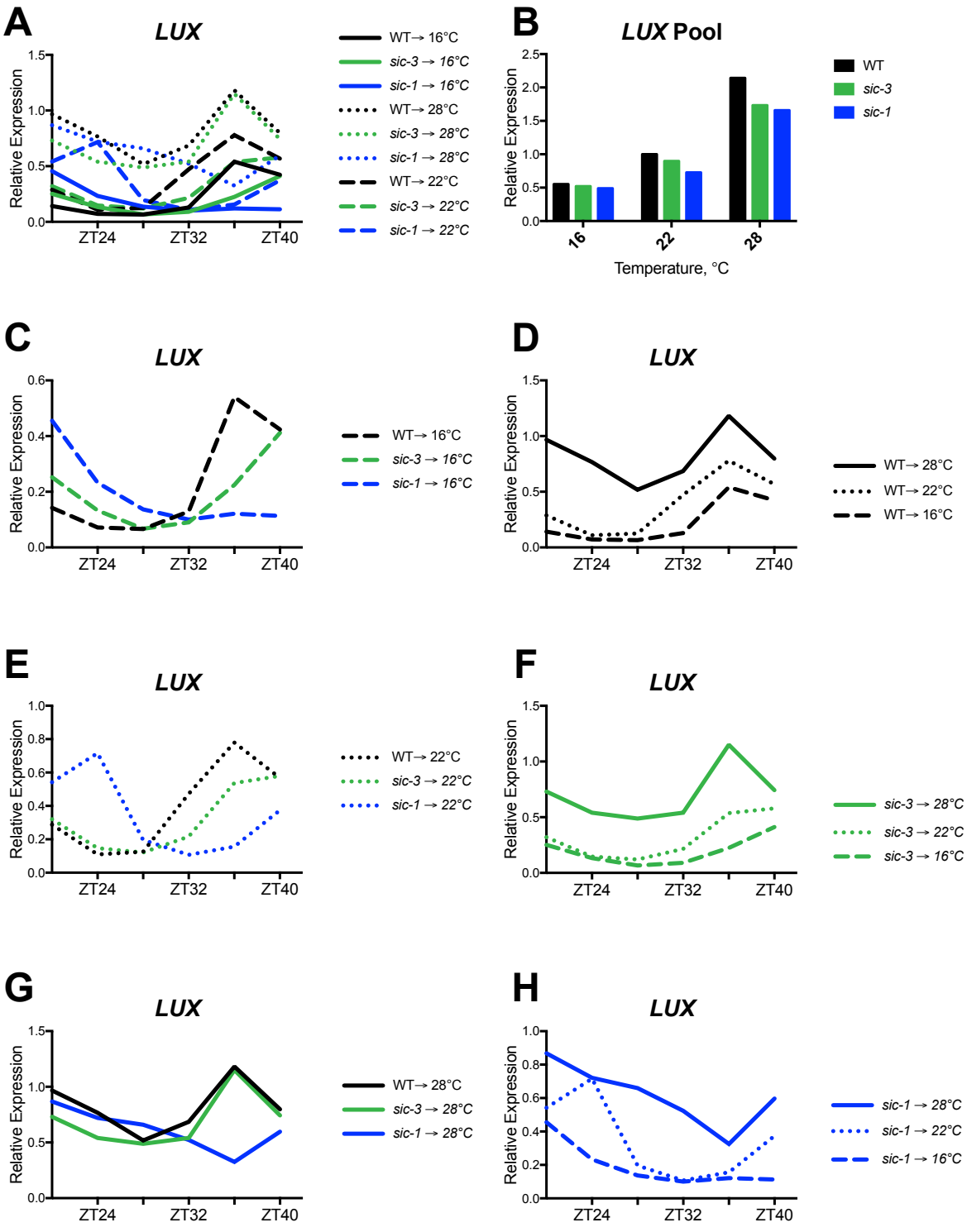


Figure 22. Temperature and the *sic* mutation alter *LUX* expression.

A), C) to **(H)** Expression levels of the *LHY* over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3*

transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

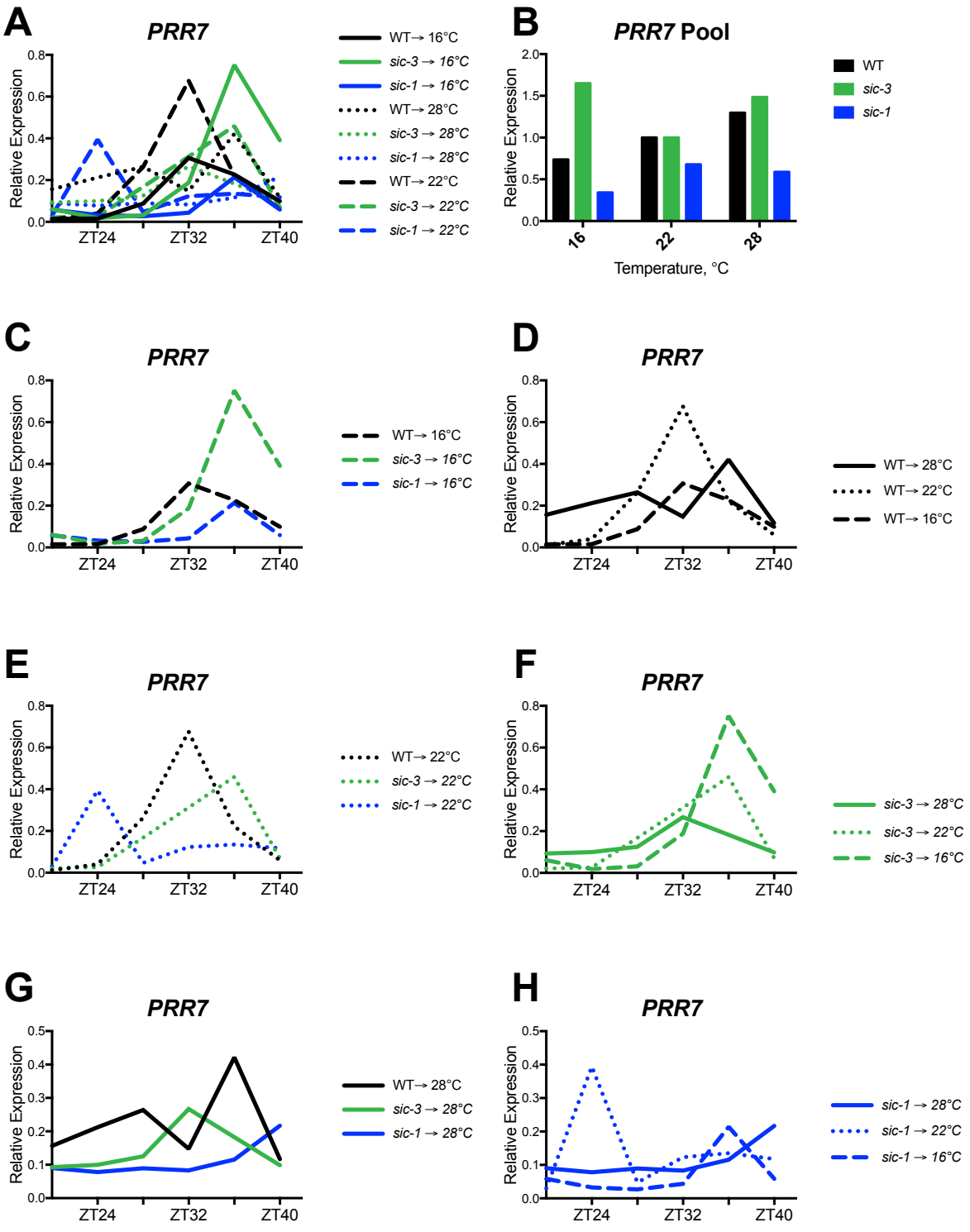


Figure 23. Temperature and the *sic* mutation alter *PRR7* expression.

A), C) to **(H)** Expression levels of the *LHY* over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3*

transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

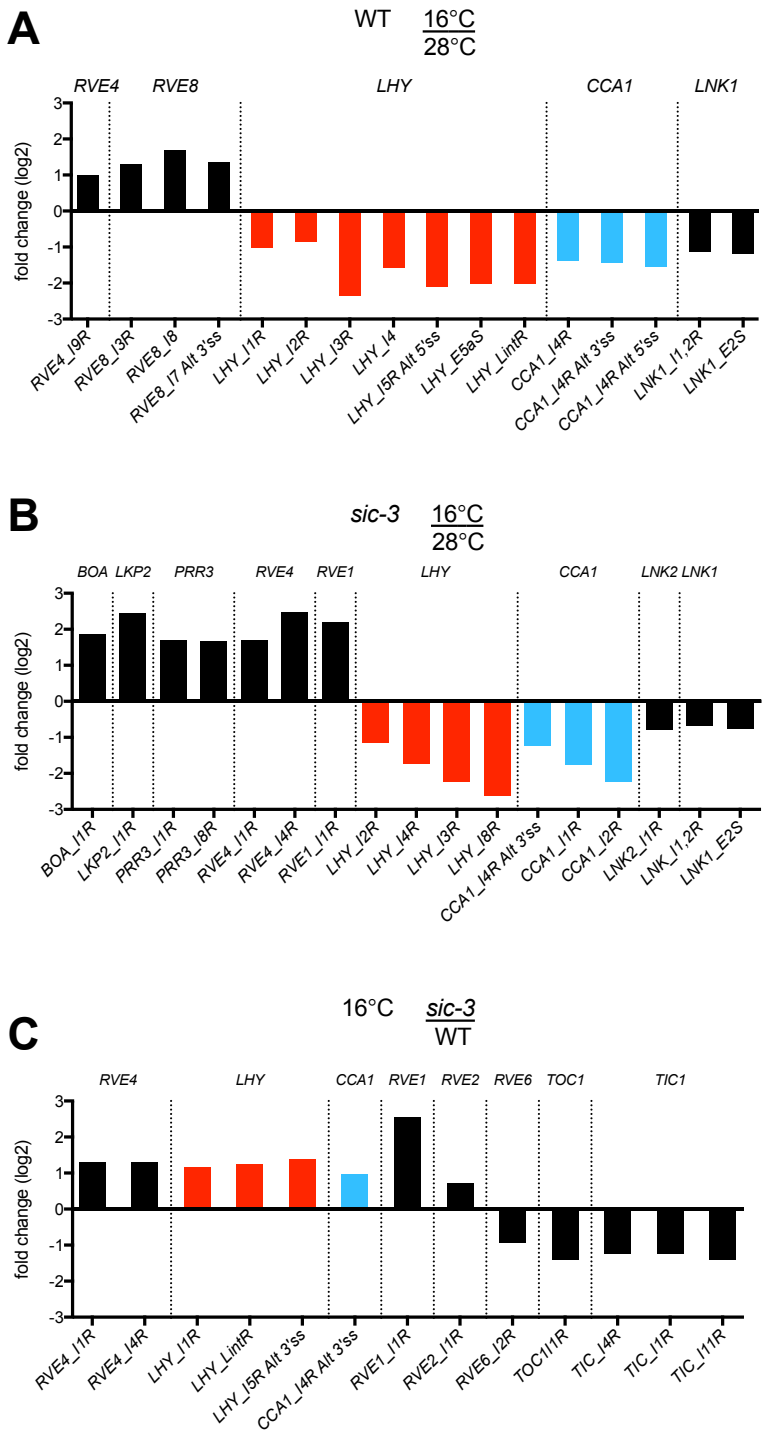


Figure 24. WT and *sic-3* demonstrate differences in temperature-dependent alternative splicing of circadian clock genes according to RNA-seq.

A) to **(C)** Differential alternative splicing of circadian clock genes was identified (false discovery rate (FDR) of ≤0.1) caused by temperature and/or *sic-3* in the following binary comparisons (see Chapter 2 Methods). Y-axis represents the log2 fold change between

conditions/genotypes, and X-axis lists the circadian clock gene splice variants with significant differential alternative splicing. *LHY* is highlighted in red and *CCA1* is highlighted in light blue.
A) Circadian clock genes with differential alternative splicing between WT seedlings transferred to LL|16°C versus to LL|28°C.
B) Circadian clock genes with differential alternative splicing between *sic-3* seedlings transferred to LL|16°C versus to LL|28°C.
C) Circadian clock genes with differential alternative splicing between *sic-3* seedlings transferred to LL|16°C versus WT seedlings transferred to LL|16°C.

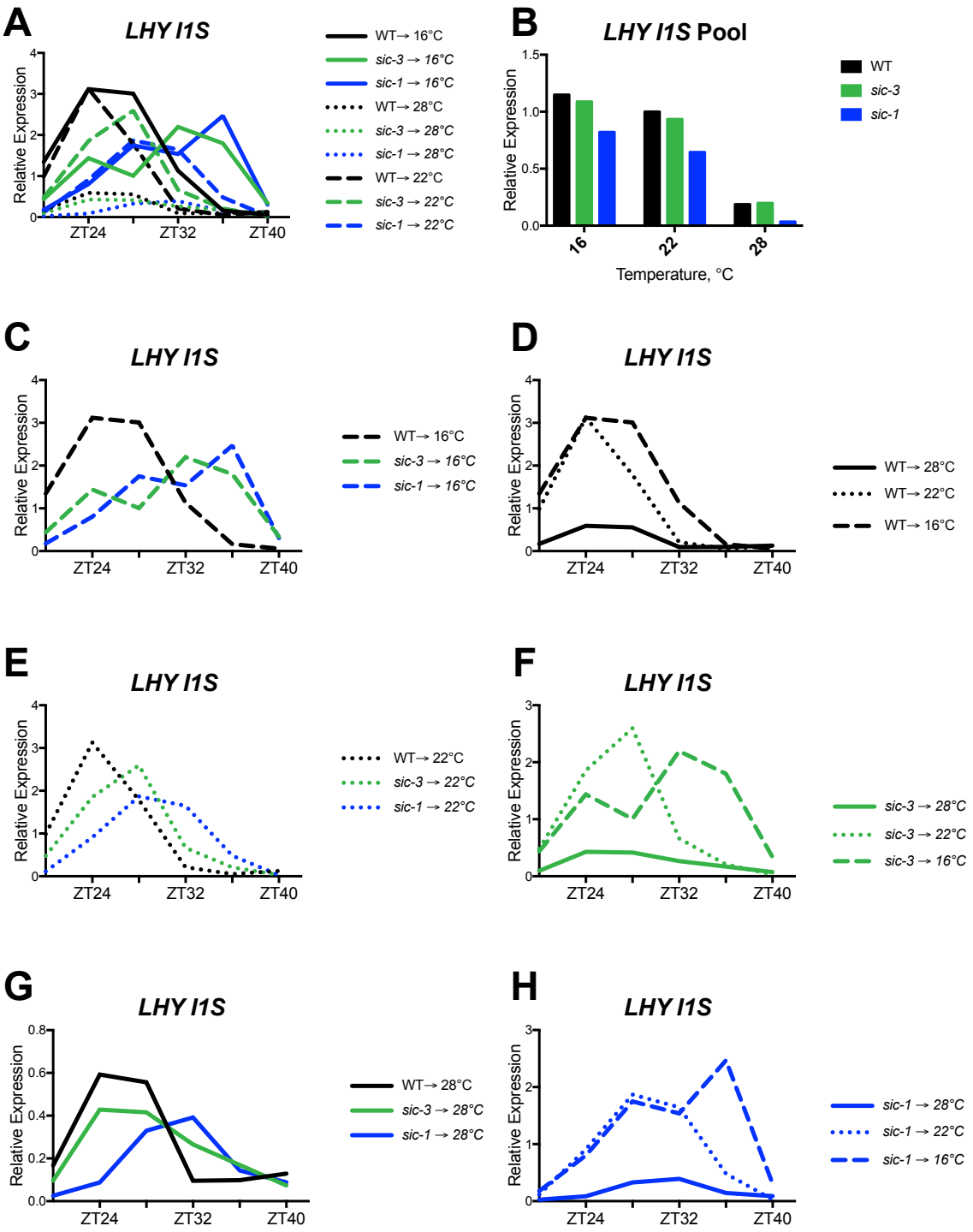


Figure 25. Temperature and the *sic* mutation alter the expression of the fully spliced intron 1 *LHY* splice variant.

A), C) to **(H)** Expression levels of the fully spliced intron 1 *LHY* splice variant over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C (dashed black line) and *sic-3* transferred to LL|28°C (solid green line), LL|22°C

(dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

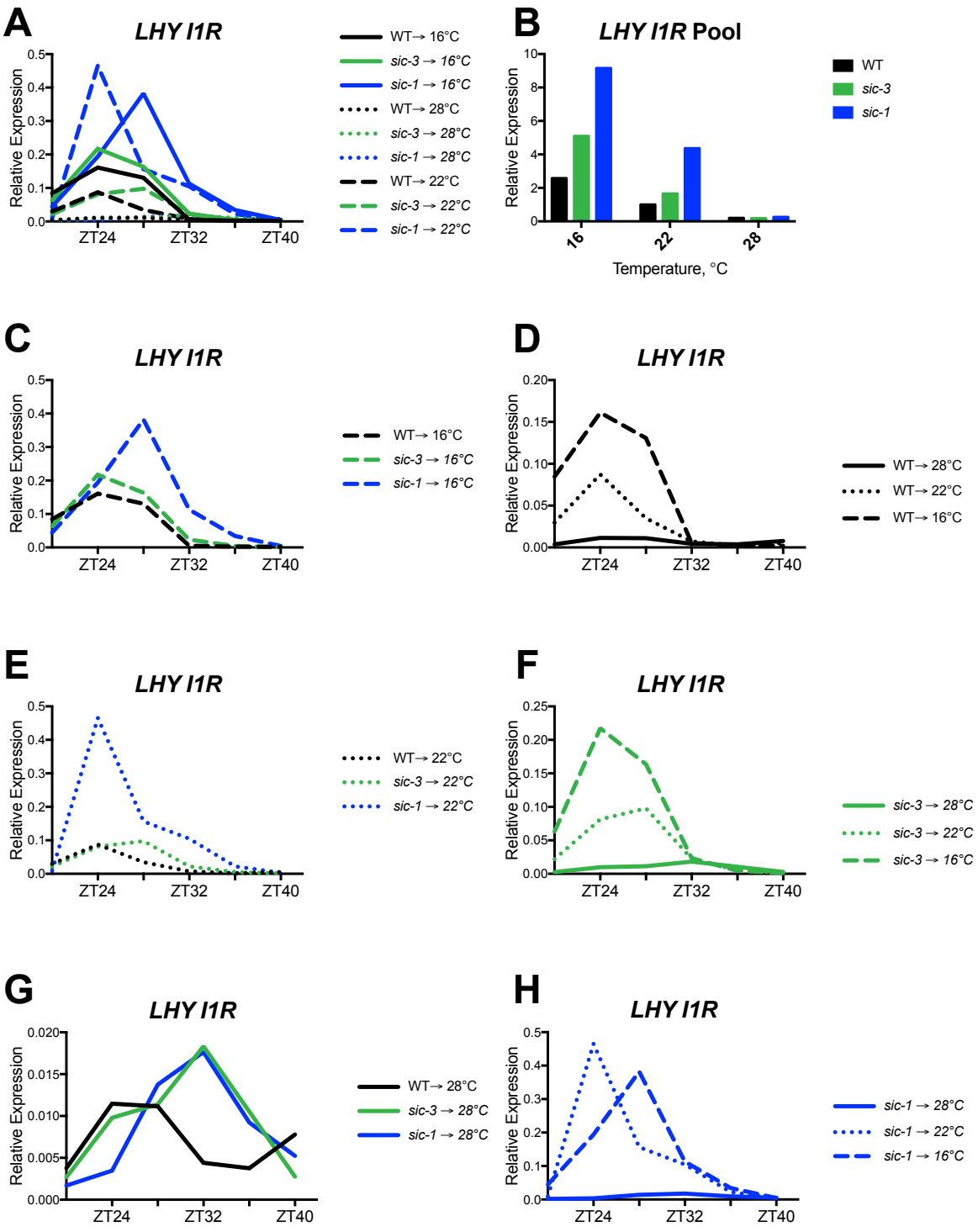


Figure 26. Temperature and the *sic* mutation alter the expression of the intron 1 retention *LHY* splice variant.

A), C) to **(H)** Expression levels of the intron 1 retention *LHY* splice variant over 24 hours in WT transferred to LL|28°C (solid black line), LL|22°C (dotted black line), and LL|16°C

(dashed black line) and *sic-3* transferred to LL|28°C (solid green line), LL|22°C (dotted green line), and LL|16°C (dashed green line), and *sic-1* transferred to LL|28°C (solid blue line), LL|22°C (dotted blue line), and LL|16°C (dashed blue line). Graphs show comparisons between:
A) All genotypes in all temperature conditions.
C) WT, *sic-3,* and *sic-1* seedlings transferred to LL|16°C.
D) WT seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
E) WT, *sic-3,* and *sic-1* seedlings transferred to LL|22°C.
F) *sic-3* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
G) WT, *sic-3,* and *sic-1* seedlings transferred to LL|28°C.
H) *sic-1* seedlings transferred to LL|28°C, LL|22°C, and LL|16°C.
B) A pool of equal amounts of total RNA from each timepoint within a genotype (WT, *sic-3,* and *sic-1*) and within a temperature condition (LL|28°C, LL|22°C, and LL|16°C). Transcript levels were determined with qPCR.

Table 4. Transcriptome analysis of genes showing differential expression and genes showing differential alternative splicing between WT and *sic-3* plants transferred to 28°C or 16°C

Comparison by genotype or by condition	Differential expression		False Discovery Rate	Differential alternative splicing *		False Discovery Rate
	# of genes	% of transcriptome		# of genes	% of transcriptome	
WT 16°C/28°C	3,536	13	≤ 0.05	748	3.9	≤ 0.1
sic-3 16°C/28°C	4,229	16.1	≤ 0.05	3,059	16.4	≤ 0.1
sic-3 16°C/ WT 16°C	3,620	13.5	≤ 0.05	3,614	19.3	≤ 0.1
sic-3 28°C/ WT 28°C	2,557	9.5	≤ 0.05	1,233	6.5	≤ 0.1

* Counted as number of genes with alternative splicing; no gene is counted more than once

Chapter 3: Improper regulation of *LHY* and *CCA1* in *sic* leads to circadian clock dysfunction

The following text is modified from an article published in *The Plant Cell*, by Marshall *et al.*, 2016.

Abstract

The functional dynamic between *LHY, CCA1,* and *SIC* was further explored based on the published role of *LHY* and *CCA1* in regulating the circadian clock temperature response in Arabidopsis, and the alternative splicing phenotypes observed in the *sic* background (see Chapter 2). Combining the *sic-3* mutation with *lhy* and *cca1* mutants revealed that the *cca1-1 lhy-20* double mutant is epistatic to *sic-3*, indicating *LHY* and *CCA1* splice variants are needed for *sic-3* circadian clock phenotypes. Furthermore, the double and triple mutant combinations of *sic-3* with *cca1-1* and *lhy-20* indicate that CCA1 and LHY function redundantly to modify temperature compensation, and that this regulation is dependent on SIC in cool temperatures. In addition, LHY protein accumulation is temperature-dependent, and followed a rapid reduction in expression in warm temperatures that is consistent with the temperature-dependent *LHY* transcript accumulation. LHY protein is broadly expressed in *sic* and the timing of its peak level is significantly delayed relative to WT. Cool temperatures aggravate *sic*'s changes in LHY accumulation. These observations indicate that the *sic* circadian clock phenotype depends on SIC regulation of *LHY* and *CCA1* transcript, because SIC affects alternative splicing of these transcripts, and SIC is needed for correct LHY protein accumulation.

Introduction

LHY and *CCA1* are key core circadian clock genes (Nohales and Kay 2016; Hsu and Harmer 2014; Greenham and McClung 2015). Both *LHY* and *CCA1* are MYB-like transcription factors in the *RVE* gene family, but whereas *RVE4, RVE6,* and *RVE8* are transcription activators in the core circadian clock, *LHY* and *CCA1* are transcription repressors (Hsu, Devisetty, and Harmer 2013). *LHY* and *CCA1* peak gene expression is in the dawn, followed closely by peak protein expression (Nohales and Kay 2016). A key part of LHY and CCA1 protein function is to form hetero- and homo-dimers to bind the *cis*-regulatory evening element (EE) and repress the evening expressed genes *TOC1*, *PRR9/7/5, GI, LUX*, *BOA*, *ELF3,* and *ELF4* (Hsu and Harmer 2014), as well as each other (Z.-Y. Y. Wang and Tobin 1998; Schaffer et al. 1998). The role of *TOC1* as a repressor or activator of *CCA1* and *LHY* has long been contested, but the current standing is that *TOC1* represses *LHY* and *CCA1* (W. Huang et al. 2012; Gendron et al. 2012; Pokhilko et al. 2012). Aside from their regulatory role within the core circadian clock, *LHY* and *CCA1* also play an important role in output processes controlled by the circadian clock, including iron uptake (Hong et al. 2013; Salomé et al. 2013), nitrogen metabolism (Gutierrez et al. 2008), reactive oxygen species (ROS) homeostasis (Lai et al. 2012), and plant defense (W. Wang et al. 2011; Bhardwaj et al. 2011).

Evidence that *LHY* and *CCA1* are crucial for circadian clock function is the short period displayed by single *lhy* and *cca1* mutants under free running conditions (Gould et al. 2006) and the even shorter free running period of the double *cca1 lhy* mutant (Lu et al. 2009). Furthermore, overexpression of *LHY* or *CCA1* drives the clock to arrhythmicity (Schaffer et al. 1998; Z.-Y. Y. Wang and Tobin 1998). *LHY* and *CCA1* have partially redundant functions in the Arabidopsis circadian clock, but the fact that the single mutants do not always display identical phenotypes points to the two proteins having distinct roles (Mizoguchi et al. 2002; Gould et al. 2006).

The regulatory mechanisms for temperature compensation are not well understood in Arabidopsis, but the prevailing evidence points to an important contribution by *LHY* and *CCA1* (Gould et al. 2006). Gould et al. 2006 investigated the role of *CCA1* and *LHY* in temperature compensation by measuring the period of *cca1* and *lhy* mutants transferred to LL|12°C, LL|17°C, and LL|27°C (Gould et al. 2006). They found that the *cca1* and *lhy* mutants had the same short-period phenotype at 17°C, but *lhy* was nearly 2 hours shorter than WT at 27°C and *cca1* was less than one hour shorter than WT (Gould et al. 2006). At 12°C the *cca1* mutant had a period even shorter than *lhy* (Gould et al. 2006). Gould et al. 2006 concluded *CCA1* plays a greater role in temperature compensation at lower temperatures than *LHY,* and *LHY* plays a more important role in warmer temperatures (Gould et al. 2006). In another study, *PRR7* and *PRR9* were proposed to regulate *CCA1* and *LHY* activity in temperature compensation, since the *prr7 prr9* double mutant overcompensates for temperature and hyperactivates *CCA1* and *LHY* (Salome, Weigel, and McClung 2010). Finally, the binding activity of CCA1 to the promoter of circadian clock genes is under control of the protein kinase CK2; CK2 activity counterbalances the increased CCA1 binding activity of CCA1 to promoters by phosphorylating and antagonizing CCA1 binding which in turn regulates temperature compensation (Portolés and Más 2010; Daniel, Sugano, and Tobin 2004). Collectively these findings do not point to a specific pathway in which *LHY* and *CCA1* work to regulate circadian rhythms in different temperature conditions, but they do highlight the importance of these genes in maintaining temperature compensation.

The relationship between *LHY, CCA1*, and temperature goes beyond temperature compensation. As discussed in depth in Chapter 2, *LHY* and *CCA1* are susceptible to temperature-dependent alternative splicing (A. B. James et al. 2012; Seo et al. 2012; Filichkin et al. 2015); and regulation of *CCA1* alternative splicing is important for freezing tolerance (Seo et al. 2012). *CCA1* and *LHY* also directly regulate transcription of the C-Repeat Binding Factor (CBF) cold-response pathway (Dong, Farré, and Thomashow 2011). The CBF pathway is characterized by induction of the master cold acclimation transcription factors *CBF1, CBF2,* and *CBF3* (Dong, Farré, and Thomashow 2011). *CCA1* and *LHY* are therefore involved in the circadian regulation of cold tolerance in Arabidopsis through this circadian output pathway.

In this chapter the effect of *SIC* on *LHY* and *CCA1* function was evaluated by testing the circadian clock activity in double and triple mutant combinations and through Western blot analysis of LHY protein expression. SIC function is necessary in modulating the activity of LHY and CCA1 in cool temperatures, as well as timing of LHY's protein accumulation throughout the day. It is unclear how SIC modifies the function of CCA1 and LHY, but it may be linked to *sic*'s alternative splicing defects observed in Chapter 2.

Results

The period and thermocycle entrainment phenotypes caused by *sic-3* require *LHY* and *CCA1*

To determine whether *LHY* and *CCA1* transcripts are required for clock impairment in the *sic* mutant background, we generated *lhy-20*, *cca1-1*, *lhy-20 sic-3*, *cca1-1 sic-3*, and *cca1-1 lhy-20 sic-3* mutants carrying the *ProPRR7:LUC* reporter. As expected, the *lhy-20*, *cca1-1*, and *cca1-1 lhy-20* mutants had significantly shortened free running period compared to WT under LL|22°C after either LL|22°-18°C thermocycle or LD|22°C photocycle entrainment (Figure 27A,B; Table 5). Addition of *sic-3* to either *lhy-20* or *cca1-1* lengthened period to a degree consistent with an additive phenotype (Figure 27A,B; Table 5). In contrast, the period of the *cca1-1 lhy-20 sic-3* was significantly shorter than the period of either *sic-3, lhy-20 sic-3*, or *cca1-1 sic-3* (Figure 27A,B; Table 5). Thus, the period lengthening effects of *sic-3* were dependent on the presence of *CCA1* and *LHY* transcript. Other hallmark *sic*-associated clock phenotypes also required *CCA1* and *LHY*. For example, *cca1-1 lhy-20* and *cca1-1 lhy-20 sic-3* populations under LL|22°-18°C thermocycles exhibited comparable mean period lengths, limited period variability, and few arrhythmic individuals (Figure 27D,F; Table 5), unlike the broad distribution of periods observed for a *sic-3* population (Figure 27D,F; Table 5). Therefore, the absence *CCA1* and *LHY* suppressed the thermocycle entrainment phenotype of *sic-3*.

LHY and CCA1 function with SIC in cool temperatures to regulate temperature compensation

To further explore the relationship between *SIC*, *LHY, CCA1,* and temperature, temperature compensation was tested in WT, *sic-3, lhy-20*, *cca1-1*, *lhy-20 sic-3*, *cca1-1 sic-3*, *cca1-1 lhy-20,* and *cca1-1 lhy-20 sic-3* (Figure 28). The free-running period of each mutant line was determined at constant 28°C (LL|28°C), 22°C (LL|22°C), and 16°C (LL|16°C), after 5 days of LD|22°C entrainment (Figure 28). As expected, WT maintained an approximately 24-hour period across all three temperatures (Figure 28; Table 5). As well, *sic-3* did not exhibit temperature compensation, instead the clock ran markedly slower under LL|16°C (27.8 ± 1.7 hours) than under LL|28°C (24.0 ± 0.8 hours) (Figure 28; Table 5). The *lhy-20, cca1-1,* and *cca1-1 lhy-20* mutants had short periods across the full temperature range, but all three mutants displayed temperature compensation defects between LL|22°C and LL|28°C, because the pace of their rhythms increased with temperature (Figure 28A; Table 5). Under LL|16°C, *lhy-20* had a period of 22.4 hours and *cca1-1 lhy-20* of 22.8 hours (Figure 28A; Table 5). The period of *cca1-1* under LL|16°C could not be determined due to technical issues, but published data indicate that *cca1-1* has a period of 22.8 hours under LL|17°C (Gould et al. 2006). In contrast, the period of all three mutants shortened significantly under LL|28°C: *cca1-1* to 21 hours, *cca1-1 lhy-20* to 19.7 hours, and *lhy-20* to 20 hours (Figure 28A; Table 5). This indicates that the *cca1-1* mutant is able to temperature compensate more than *lhy-20* and *cca1-1 lhy-20* under LL|28°C.

The behavior of the *lhy-20 sic-3* and *cca1-1 sic-3* double mutants was comparable to the corresponding single *lhy-20* and *cca1-1* mutants at high but not low temperature (Figure 28B,C; Table 5). The *lhy-20 sic-3* mutant had a short period under LL|28°C matching *lhy-20* (20.9 hours) and it lengthened to a pace comparable to the *lhy-20* mutant under LL|22°C (23 hours) (Figure 28B; Table 5). On the other hand, the period of *lhy-20 sic-3* under LL|16°C was 26 hours, which was nearly 4 hours longer than *lhy-20* (Figure 28B; Table 5). The *cca1-1 sic-3* double mutant also had a short period under LL|28°C (22.2 hours) that increased under LL|22°C (23.5 hours); the pace of *cca1-1 sic-3* therefore increases with temperature faster than the pace of the single *cca1-1* mutant (Figure 28C; Table 5). Similar to the *lhy-20 sic-3* mutant, *cca1-1 sic-3* period was two hours longer under LL|16°C (Figure 28C; Table 5). Collectively these data indicate that both the *cca1-1 sic-3* and *lhy-20 sic-3* double mutants are temperature compensation mutants that are particularly sensitive to cool temperature conditions. In comparison to the double mutants, the single mutant *cca1-1* and *lhy-20* counterparts exhibit temperature compensation for their short periods.

CCA1 and *LHY* together regulate SIC temperature compensation. The *cca1-1 lhy-20* double mutant displayed a pattern of temperature compensation comparable to the *lhy-20* single mutant: under LL|28°C the period was short (19.7 hours) and lengthened under LL|22°C (21.3 hours) and under LL|16°C (22.8 hours) (Figure 28A,D; Table 5). The triple *cca1-1 lhy-20 sic-3* mutant period was short at all temperatures, just like its *cca1-1 lhy-20* counterpart: period under LL|28°C was 20 hours, under LL|22°C was 20.8 hours, and under LL|16°C was 22.5 hours (Figure 28D; Table 5). The triple *cca1-1 lhy-20 sic-3* mutant did not show the dramatic increase in period length under LL|16°C seen in the *lhy-20 sic-3* and *cca1-1 sic-3* double mutants, implying that the period lengthening effects that are caused by *sic-3* are dependent on both *CCA1* and *LHY* transcript in the cold.

LHY accumulation is temperature dependent

We sought to determine the effect of temperature on LHY protein accumulation, because *LHY* expression was high when WT seedlings were under LL|22°C, was induced when transferred to LL|16°C, and was suppressed when transferred to LL|28°C (see Chapter 2; Figure 18; Figure 19). LHY protein followed a temperature-dependent pattern of accumulation that was comparable to *LHY* transcript expression (Figure 18; Figure 19; Figure 29D). Seedlings of WT, *lhy-20*, and *LHY-OX* (overexpression) seedlings were first entrained under LD|22°C and then transferred to LL|22°C, LL|16°C, or LL|28°C (Figure 29). These conditions were identical to the RNA-seq experiment in Chapter 2 (Figure 17). Tissue was collected in the subjective morning (ZT 25), day (ZT 30), and night (ZT 35, ZT 40), which correspond in WT to the time of peak LHY accumulation followed by times during which it declines to its trough level (Figure 29). LHY protein was detected in whole cell extracts by western blot with an anti-LHY antibody for each individual time point in each of the three temperature conditions (Figure 29A-C). An extract from *LHY-OX* (Schaffer et al 1998) was used as a positive control for LHY protein and the null *lhy-20* mutant as the negative control. Only one replicate of this experiment was fully completed, but a comparable experiment with fewer timepoints showed identical results. Future work will be completed to validate this experiment. An equal amount of total protein was pooled from all four time points in

each temperature condition to determine total LHY protein accumulation for the entire time course (Figure 29D). LHY in pooled WT samples was much higher under LL|22°C and LL|16°C than under LL|28°C (Figure 29D). LHY protein accumulation is therefore rapidly suppressed upon a shift raised from 22°C to 28°C; this regulation occurs within the first 13 hours after transfer (Figure 29). The individual time points under LL|28°C showed that despite the overall decrease in LHY protein, the pattern of LHY protein accumulation remained rhythmic (Figure 29A). In all temperature conditions LHY protein had peak accumulation in WT at ZT 25 and remained detectable at ZT 30 (Figure 29A-D). At ZT 35 a small amount of LHY protein was observable under LL|28°C and none was seen under LL|22°C and LL|16°C, but this could be due to differences in exposure time between the films (Figure A-D; see Methods). No LHY protein was observable at ZT 40 in all temperature conditions (Figure 29A-D). In WT seedlings, LHY protein accumulation was temperature dependent, and was rhythmic despite overall changes in LHY protein accumulation.

The rhythmic pattern of LHY accumulation is disrupted in *sic* mutants

LHY levels were evaluated in *sic* mutants to determine whether the LHY waveform is influenced by *SIC* activity given the elevated alternative splicing of *LHY* transcript and changed *LHY* expression in *sic* mutants. *sic-3* and *sic-1* seedlings were collected over the same time course and tested for LHY protein as described above (Figure 29). In the samples that were pooled by temperature, both *sic* mutant backgrounds displayed total LHY protein accumulation matching WT: LHY protein was much higher under LL|22°C and LL|16°C than under LL|28°C (Figure 29D). LHY protein accumulation was perhaps higher in the *sic* mutants than WT under LL|28°C, but further experiments are required to corroborate this possibility (Figure 29D). In contrast, the individual timepoints of LHY protein accumulation followed a very different pattern than WT in all temperature conditions (Figure 29A-C).

In general, LHY had a delayed peak and a broad accumulation pattern in *sic* mutants (Figure 29A-C). Under LL|22°C LHY protein accumulation in *sic-3* peaked at ZT 25 and ZT 30, and LHY protein disappeared by ZT 40 (Figure 29B); this pattern of LHY protein accumulation was also similar to WT, but showed a broader time of peak LHY protein accumulation. Under LL|16°C *sic-3* had peak accumulation of LHY at ZT 25 and ZT 30, and a small amount of LHY was still present at ZT 35 and ZT 40 (Figure 29C); under LL|16°C LHY protein had the broadest expression pattern in *sic-3* compared to other temperatures. Under LL|28°C *sic-3* showed peak LHY protein at ZT 25 and LHY protein disappeared by ZT 40; this pattern was similar to WT under LL|28°C (Figure 29A). *sic-1* had very delayed expression of LHY protein across all three temperature conditions (Figure 29A-C). Under LL|28°C there was no LHY protein expression present at ZT 25 in *sic-1*, instead LHY accumulation peaked at ZT 35, a full 10 hours after WT (Figure 29A). Furthermore a small amount of LHY was present at ZT 40 in *sic-1*, indicating that LHY accumulation is very delayed and broad, even in warm conditions (Figure 29A). Under LL|22°C *sic-1* seedlings had a small amount of LHY protein at ZT 25, and peak LHY accumulation at ZT 30 and ZT 35, which disappeared by ZT 40 (Figure 29B); this also indicates that LHY protein accumulation is very broad and delayed in *sic-1* seedlings under these conditions. Under LL|16°C, *sic-1* seedlings had no LHY protein present at ZT 25, a high amount at ZT 30, peak accumulation at ZT 35,

and no LHY at ZT 40 (Figure 29C); this indicates a very delayed accumulation of LHY in *sic-1* under LL|16°C. The *sic* mutation therefore affects the regulation of LHY protein accumulation throughout the day, but does not appear to affect overall LHY protein accumulation in a temperature-dependent manner. The delay in LHY protein accumulation in *sic* mutants is likely due to the transcriptional defects observed in Chapter 2 and not to an additional post-translational effect.

Discussion

While the precise role LHY and CCA1 play in temperature compensation remains unclear, these transcription factors are undoubtedly crucial for circadian clock function. Based on the short period of *cca1, lhy,* and *cca1 lhy* mutants (Gould et al. 2006) (Figure 27; Figure 28; Table 5), the logical assumption is that CCA1 and LHY are necessary to slow down the pace of the clock (i.e. lengthen the period) at all temperatures. This is a counterintuitive point of view when considering the rapid induction of LHY and CCA1 plants in cool temperatures; in cool temperatures the circadian clock does not need to be slowed down, therefore why are high levels of LHY and CCA1 needed in the cool temperature range? Furthermore, based on the results in Chapter 2 (Figure 18; Figure 19), *LHY* and *CCA1* expression is strongly induced in cool temperatures and repressed in warm temperatures, reiterating the important role of LHY and CCA1 in cool temperatures. We propose another perspective: LHY and CCA1 play an important role in moderating the circadian clock's response to all temperatures, and especially in the cold. LHY and CCA1 do not function to specifically lengthen the clock, but instead they are part of a circadian clock regulatory network that orchestrates appropriate circadian rhythms in an environment where temperatures change. The rapid induction of *LHY* and *CCA1* in cool temperatures is one of the many modifications that occur to orchestrate a balance of circadian clock protein activities. When function of either LHY or CCA1 is lost, the entire circadian clock network is disrupted and the clock resorts to a short period. Loss of function *lhy*, *cca1*, and *cca1 lhy* mutants resort to a short period phenotype and temperature compensation is slightly disrupted, but one can consider these mutants to display a degree of temperature compensation to maintain a short period across a range of temperatures. The temperature compensation machinery is therefore not entirely abolished in *lhy*, *cca1*, and *cca1 lhy* mutants.

Based on our findings (Figure 28) and previously published work (Gould et al. 2006), we know that *LHY* and *CCA1* play a role in maintaining the circadian period in different temperatures. Under LL|16°C and LL|22°C, the *cca1-1, lhy-20,* and *cca1-1 lhy-20* mutants were very similar in their short period phenotypes, implying functional redundancy of *CCA1* and *LHY* in this low temperature range (Figure 28). Under LL|28°C, the *lhy-20* mutant was more influenced by the warm temperature than *cca1-1,* which resulted in a shorter period for *lhy-20*; the *cca1-1 lhy-20* mutant was like the *lhy-20* mutant under LL|28°C (Figure 28). This indicates that *LHY* plays a slightly different role than *CCA1* in warm temperatures to maintain circadian rhythms, and that *LHY* is epistatic to *CCA1* in warm temperatures.

The *sic* phenotype is dependent on the presence of both *LHY* and *CCA1* transcripts, since the *cca1-1 lhy-20* double mutant was epistatic to *sic-3* (Figure 27; Figure 28; Table 5). When temperature compensation was tested in *cca1-1 sic-3* and

lhy-20 sic-3 double mutants the period noticeably changed from short period to long period under LL|16°C (Figure 28B,C; Table 5). In the triple mutant *cca1-1 lhy-20 sic-3,* the short period persisted under LL|16°C (Figure 28D; Table 5). These results indicate two points: 1) SIC affects both LHY and CCA1 activity at 16°C, and 2) as expected from the single *cca1-1* and *lhy-20* results, *LHY* and *CCA1* are redundant at 16°C and the presence of their transcripts is necessary for the *sic-3* phenotype. For example, at 16°C in the *cca1-1 sic-3* mutant, *LHY* is still being expressed and the *sic* mutation causes a defect in a pathway that modifies *LHY* function, leading to a lack of temperature compensation at 16°C.

Under LL|22°C and LL|28°C, the *cca1-1 sic-3* mutant was longer than the single *cca1-1* mutant, indicating that SIC affects CCA1 activity in these temperatures but not to as great an extent as under LL16°C (Figure 28). Under LL|22°C and LL|28°C, the *lhy-20 sic-3* mutant period was slightly longer than the *lhy-20* mutant, which suggests that LHY activity at these temperatures is less affected by SIC (Figure 28B). Since the *lhy-20* period under LL|28°C indicates that LHY plays a more important role than CCA1 in warm temperatures, we can conclude that this function does not require SIC and LHY-dependent control of temperature compensation is independent of SIC under warm temperatures. When both *LHY* and *CCA1* were eliminated in the triple *cca1-1 lhy-20 sic-3* background, the *sic* phenotype could not manifest itself and the circadian period resorted to that of a *cca1-1 lhy-20* mutant at all temperatures (Figure 28D; Table 5). Collectively, our data suggests that SIC acts at 16°C to modify the activity of LHY and CCA1, and as temperatures increase the activity of SIC is less important.

Based on the results in Chapter 2, we hypothesized that SIC regulated alternative splicing of *LHY* and *CCA1* in a temperature-dependent manner. We propose that when SIC is eliminated in the *sic* background, the pathway that regulates alternative splicing of *LHY* and *CCA1* is disrupted and this leads to a long period phenotype. In particular, SIC functions to modify the splice variant pool of *LHY* and *CCA1* in cool temperatures, and this regulation is crucial for maintaining temperature compensation in cool temperatures and correct thermocycle entrainment. In an *lhy-20 sic-3* or *cca1-1 sic-3* double mutant, lack of SIC function leads to incorrect accumulation of either *LHY* or *CCA1* splice variants. In a *cca1-1 lhy-20 sic-3* mutant, no *LHY* or *CCA1* transcript is present, therefore the splice variant pool of these genes is non-existent and the triple mutant resorts to a short period like that of *cca1-1 lhy-20*. Furthermore, our data indicates SIC affects the *LHY* and *CCA1* splice variant pool mostly in cool temperatures, and a SIC-independent pathway modifies temperature compensation under LL|28°C.

Changing the expression and behavior of the *LHY* or *CCA1* transcript pool feedbacks and affects the timing of LHY protein, as is evident in the delayed and dispersed expression of LHY in a *sic* mutant (Figure 29). This may also be true for CCA1 protein, as well as other circadian clock gene proteins. The accumulation of *CCA1 I4R Alt 3'ss* and *LHY I1R* splice variants in *sic,* and the requirement for *CCA1* and *LHY* transcripts for the period lengthening effects and cool thermocycle arrhythmia of *sic* (Figure 27), indicate these splice variants likely contribute to clock impairment in *sic*. It is possible that accumulation of *CCA1 I4R Alt 3'ss* and *LHY I1R* splice variants in *sic* slows the pace of the clock, even in the presence of the full-length transcripts (Figure 16A,D; Figure 18; Figure 19). Our temperature compensation data with *sic-3*

and the *cca1-1* and *lhy-20* mutants indicate that the timing of LHY, and perhaps CCA1, are particularly important in moderating temperature compensation in cool temperatures. Of course, many other circadian clock genes demonstrate differential alternative splicing and gene expression patterns in the *sic* mutant (Figure 18; Figure 24), and these undoubtedly contribute to the *sic* phenotype. Based on the historical importance of *LHY* and *CCA1* in the circadian clock's temperature response, the temperature-dependent alternative splicing phenotypes of *LHY* and *CCA1* in *sic*, and the epistatic genetic tests between *LHY, CCA1,* and *SIC,* we propose that SIC specifically acts in cool temperatures to modify splicing of *LHY* and *CCA1,* which is crucial for the circadian clock's temperature response.

Materials and Methods

See Ch.1 for plant growth conditions and bioluminescence assays

Generation of *lhy, cca1,* and *sic* mutant combinations

All experiments used the *Arabidopsis thaliana* Columbia-0 accession. *sic-3* plants carrying a *ProPRR7:LUC* reporter construct that consisted of the *PRR7* promoter driving expression of firefly *luciferase* (Salomé et al. 2005) were crossed to a *cca1-1 lhy-20* double mutant. Subsequent genotyping generated plants homozygous for the *ProPRR7:LUC* construct and *cca1-1*, or *lhy-20*, or *cca1-1 lhy-20,* or *cca1-1 sic-3*, or *lhy-20 sic-3*, or *cca1-1 lhy-20 sic-3.*

Protein detection by Western blot analysis

Tissue was grown as described and collected at the indicated times. The *LHY-OX* mutant is the *lhy* mutant from the Landsberg accession (Schaffer et al. 1998). After collection, the tissue was immediately placed in liquid N_2. Whole-cell extracts were prepared by vortexing ground tissue with an equal volume of extraction buffer (50 mm HEPES/NaOH, pH 7.5, 100 mm NaCl, 10% glycerol, 5 mm EDTA, 5 mm DTT, 1% NP-40, 0.5% deoxycholate, 0.1% SDS, 5 mm benzamidine, EDTA-free protease inhibitor (Roche, http://www.roche.com, 1 mm PMSF (phenylmethanesulfonyl fluoride), 50 μm MG132 (Peptides International, http://www.pepnet.com) and 1% polyvinylpolypyrrolidone). The lysate was cleared by centrifugation at 17 000 ***g*** for 30 min at 4°C. Total protein concentration was determined using the reducing agent-compatible BCA protein assay (Pierce, http://www.piercenet.com). For pooled samples, 40 μg of total crude protein was pooled from each timepoint for each growth condition and each genotype. 120 μg of total protein was resolved in 10% SDS-PAGE gel as described previously (Sambrook and Russell, 2001). Proteins were transferred to a 0.2-nm nitrocellulose membrane. LHY was detected using anti-LHY antibody (1:4000) (R M Green and Tobin 1999), followed by HRP-conjugated goat anti-rabbit IgG-HRP secondary antibody (1:10,000) (Santa Cruz Biotechnology; sc-2030). Loading control was detected using an anti-Tubulin (B-5-1-2) antibody (1:200) (Santa Cruz Biotechnology; sc-23948). HRP activity was determined using SuperSignal West Pico chemiluminescent substrate (Pierce) and BioMax film (Kodak, http://www.kodak.com). Immunoreactive protein was visualized on X-ray film (Kodak, Rochester, NY) using

SuperSignal West Pico Substrate (Thermo Scientific, Waltham, MA). Exposure time of the film to the Western blot varied depending on strength of LHY expression.

Chapter 3 Figures

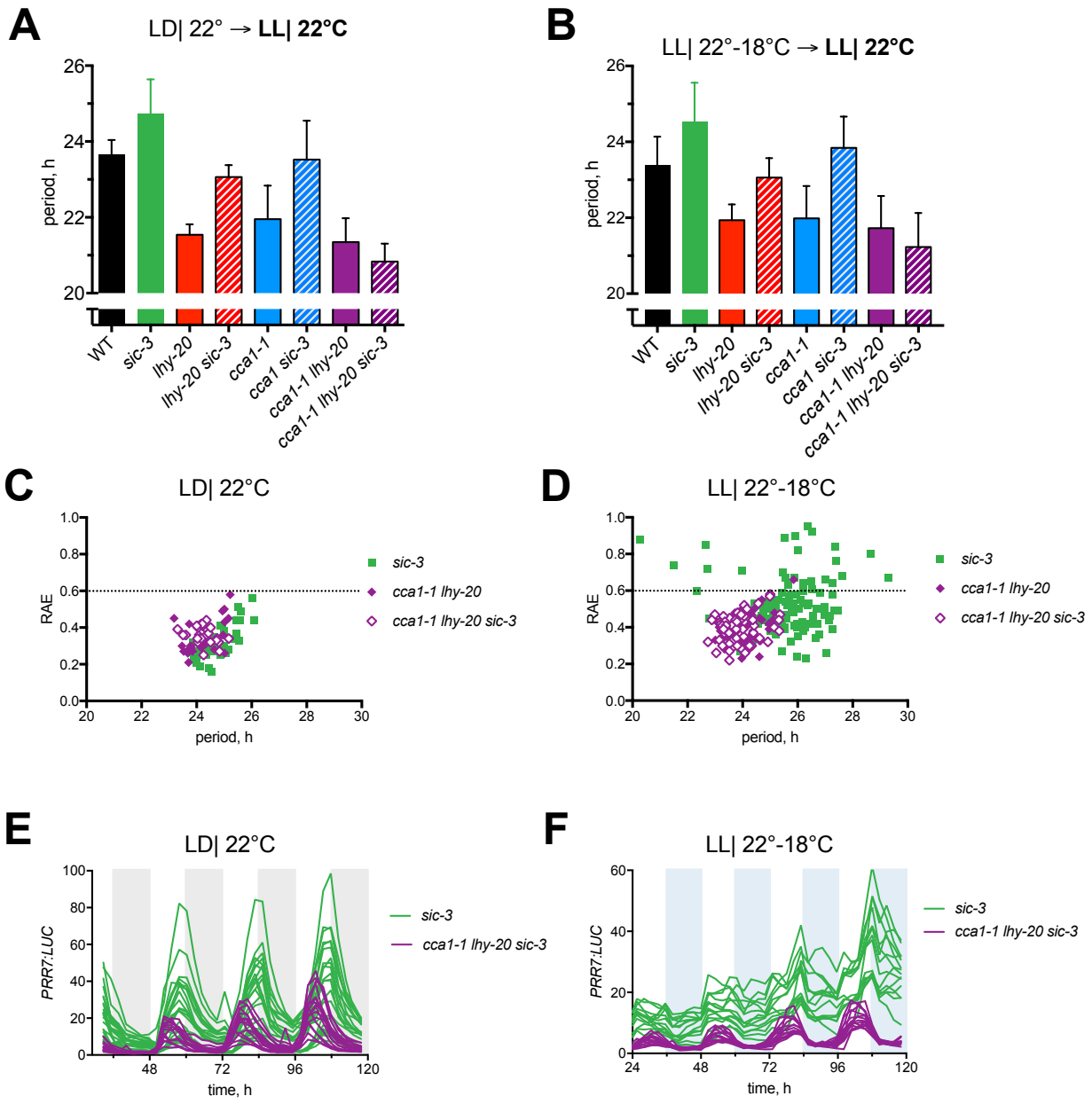


Figure 27. Long period and weakened rhythms under thermocycles in *sic* depend on *CCA1* and *LHY*.

A) and **(B)** Mean free running period of *ProPRR7:LUC* activity for WT, *sic-3*, *lhy-20, lhy-20 sic-3, cca1-1, cca1-1 sic-3, cca1-1 lhy-20,* and *cca1-1 lhy-20 sic-3* under LL|22°C following entrainment under **(A)** LD|22°C and **(B)** LL|22°-18°C. Error bars are standard deviation. All genotypes are significantly different from *sic-3* based on p-value <0.01 from ANOVA followed by Tukey's multiple comparison test.
C) and **(D)** RAE as a function of period (in hours) of *ProPRR7:LUC* activity for *sic-3* (green squares), *cca1-1 lhy-20* (filled purple diamonds), and *cca1-1 lhy-20 sic-3* (open purple diamonds) under **(C)** LD|22°C (n for *sic-3*=48, *cca1-1 lhy-20*=23, *cca1-1 lhy-20*

sic-3=30) and **(D)** LL|22°-18°C entrainment (n for *sic-3*=126, *cca1-1 lhy-20*=57, *cca1-1 lhy-20 sic-3*=64). Dotted horizontal line indicates RAE = 0.6 threshold, above which traces of *ProPRR7:LUC* activity are considered arrhythmic. Data are from 3 independent experiments.
E) and **(F)** Individual traces of *ProPRR7:LUC* activity in *sic-3* (green lines) and *cca1-1 lhy-20 sic-3* (purple lines) seedlings during **(E)** LD|22°C and **(F)** LL|22°-18°C over 4 days (n=16 for all)**.** Blue shading represents 12 hour periods of 18°C and grey shading represents 12 hour periods of dark.

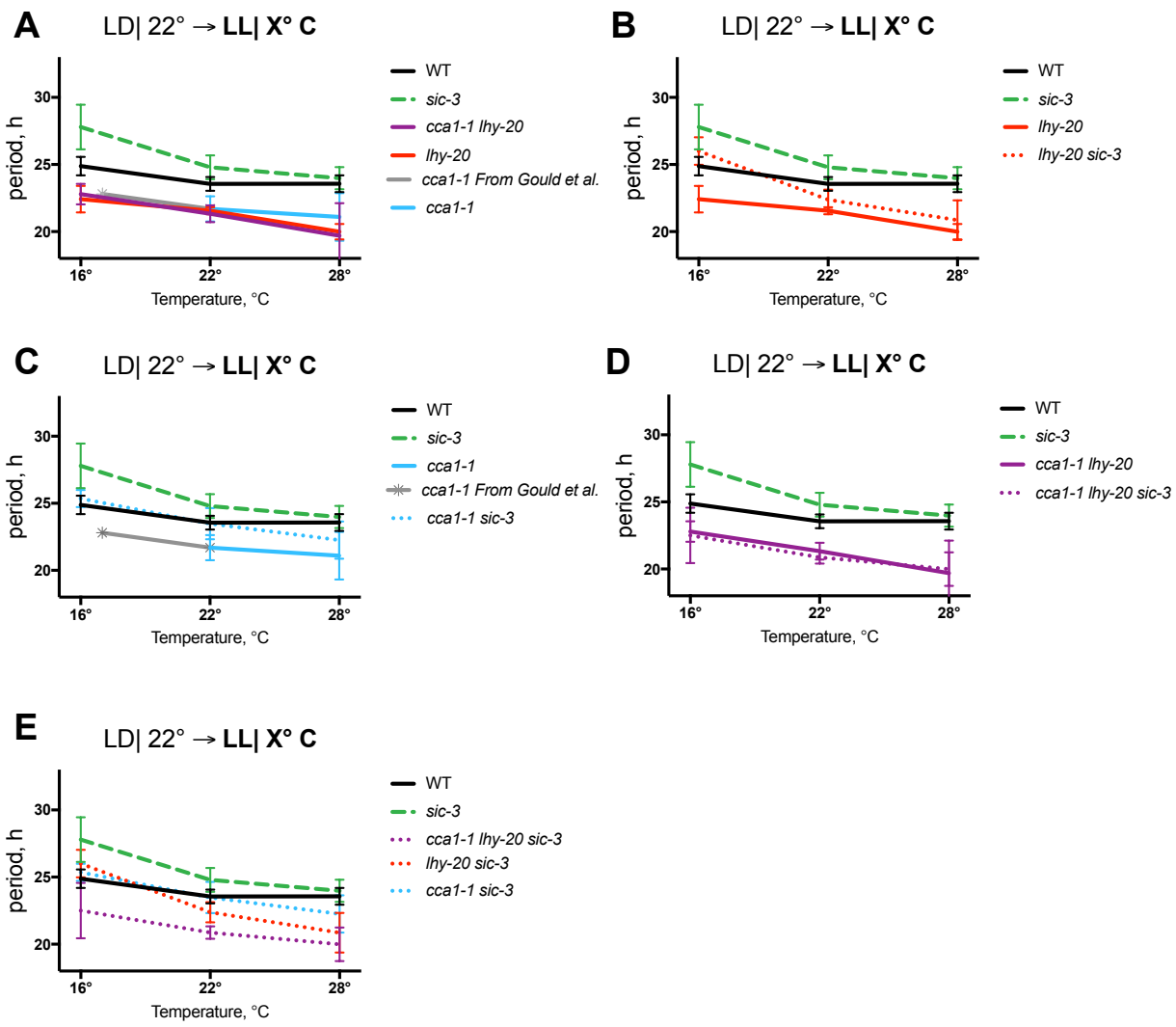


Figure 28. LHY and CCA1 function with SIC in cool temperatures to regulate temperature compensation.

A) to **(E)** Mean period length for *ProPRR7:LUC* activity in WT (solid black lines), *sic-3* (dashed green line), *lhy-20* (solid red line), *lhy-20 sic-3* (dotted red line), *cca1-1* (solid light blue line), *cca1-1 sic-3* (dotted light blue line), *cca1-1 lhy-20* (solid purple line), and *cca1-1 lhy-20 sic-3* (dotted purple line) seedlings under free running conditions of LL|X°C, where X stands for either 16°C, 22°C, or 28°C as indicated. Entrainment was LD|22°C for 5 days prior to release into free running conditions. Each point is the mean of all rhythmic individuals (RAE<0.6) from three independent biological replicates **(A)** n total at 16°C: WT=52, *sic-3*=48, *lhy-20*=21, *lhy-20 sic-3*=27, *cca1-1*=Not Determined, *cca1-1 sic-3*=60, *cca1-1 lhy-20*=40, and *cca1-1 lhy-20 sic-3*=25; at 22°C: WT=48, *sic-3*=48, *lhy-20*=38, *lhy-20 sic-3*=8, *cca1-1*=13, *cca1-1 sic-3*=30, *cca1-1 lhy-20*=28, and *cca1-1 lhy-20 sic-3*=28; at 28°C: WT=43, *sic-3*=47, *lhy-20*=20, *lhy-20 sic-3*=29, *cca1-1*=24, *cca1-1 sic-3*=27, *cca1-1 lhy-20*=34, and *cca1-1 lhy-20 sic-3*=19. Error bars are standard deviation.

A) Comparison between WT, *sic-3, cca1-1 lhy-20, lhy-20, cca1-1* (for LL22|°C and LL|28°C), and *cca1-1* (from Gould et al. 2006 under LL|17°C).
B) Comparison between WT, *sic-3, lhy-20,* and *lhy-20 sic-3*.
C) Comparison between WT, *sic-3, cca1-1* (for LL22|°C and LL|28°C), *cca1-1* (from Gould et al. 2006 under LL|17°C)*,* and *cca1-1 sic-3*.
D) Comparison between WT, *sic-3, cca1-1 lhy-20,* and *cca1-1 lhy-20 sic-3.*
E) Comparison between WT, *sic-3, cca1-1 lhy-20 sic-3, lhy-20 sic-3*, and *cca1-1 sic-3.*

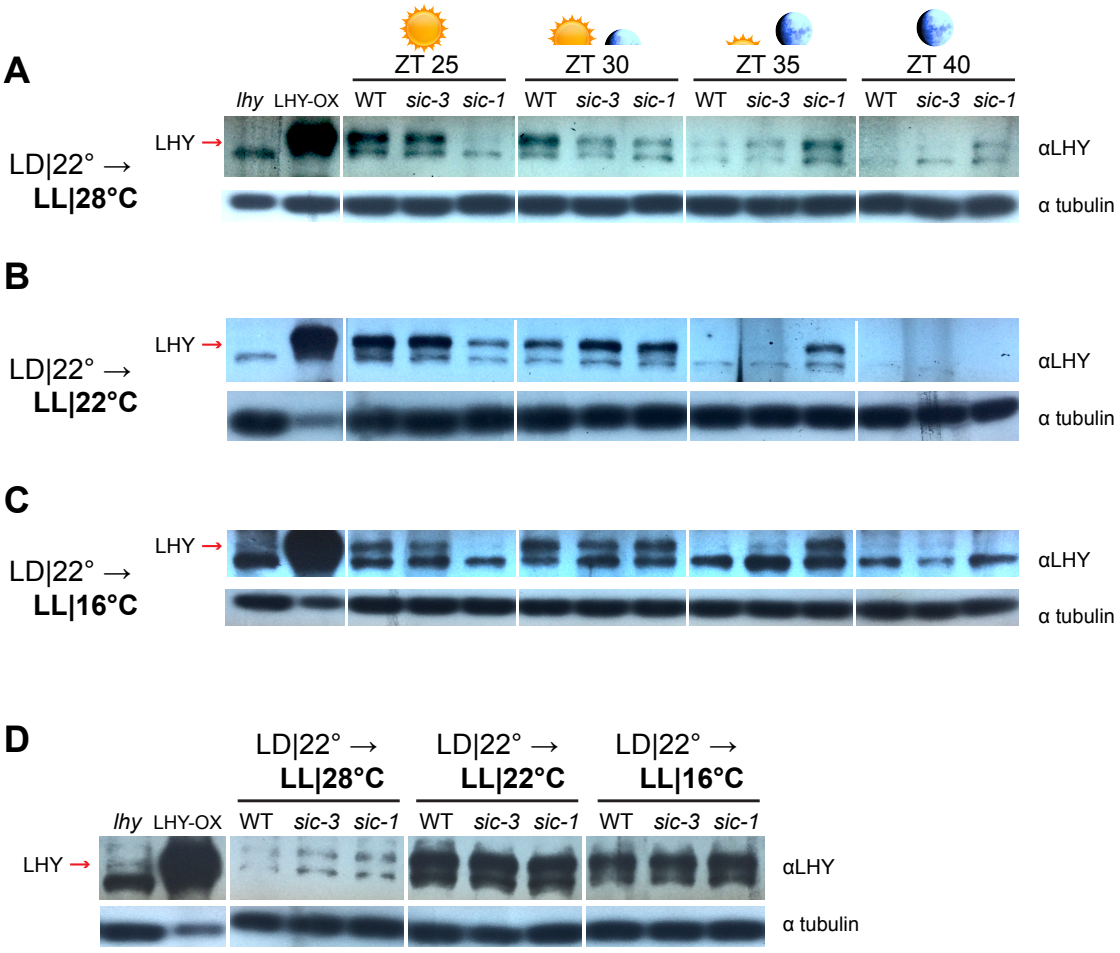


Figure 29. LHY accumulation is temperature-dependent and the rhythmic pattern of LHY accumulation is disrupted in *sic* mutants.

A) to **(D)** Western blots probed with anti-LHY (upper panel), and the anti-tubulin loading control (lower panel) in WT, *sic-3, sic-1, lhy-20*, and *LHY-OX* seedlings entrained in LD|22°C and transferred to **(A)** LL|28°C, **(B)** LL|22°C, and **(C)** LL|16°C. These conditions were identical to the RNA-seq experiment in Chapter 2 (Figure 17). Tissue was collected at ZT 25, ZT30, ZT 35, and ZT 40 for WT, *sic-3*, and *sic-1*, and at ZT 25 for *lhy-20* and *LHY-OX.* **(D)** Represents an equal amount of total protein pooled from all four timepoints in each temperature condition. Red arrows indicate the LHY protein band. Only one replicate of this experiment was fully completed, but a comparable experiment with fewer timepoints showed identical results.

Table 5. Circadian clock parameters determined from *ProPRR7:LUC* rhythms in the indicated conditions.

Entrainment condition	Free run condition	Genotype	Mean Period ± SD	Mean RAE ± SD	Mean Phase ± SD	Percent Rhythmic	Total
LD|22°C	-	WT	25 ± 0.6	0.3 ± 0.1	N/D	100%	47
		sic-3	24.8 ± 0.5	0.3 ± 0.09	N/D	100%	48
		sic-1	25 ± 0.8	0.45 ± 0.2	N/D	81%	47
		lhy-20	24.4 ± 0.5	0.35 ± 0.07	N/D	100%	38
		lhy-20 sic-3	24.5 ± 0.5	0.3 ± 0.1	N/D	100%	16
		cca1-1	25 ± 0.4	0.4 ± 0.13	N/D	100%	13
		cca1-1 sic-3	24.5 ± 0.5	0.3 ± 0.07	N/D	100%	31
		cca1-1 lhy-20	24.2 ± 0.6	0.3 ± 0.1	N/D	100%	23
		cca1-1 lhy-20 sic-3	24.3 ± 0.4	0.4 ± 0.05	N/D	100%	30
	LL|16°C	WT	24.9 ± 0.7	0.3 ± 0.1	N/D	88%	52
		sic-3	27.8 ± 1.7	0.3 ± 0.1	N/D	100%	48
		lhy-20	22.4 ± 1	0.4 ± 0.1	N/D	100%	21
		lhy-20 sic-3	26 ± 1	0.3 ± 0.09	N/D	84%	27
		cca1-1	22.8*	N/D	N/D	N/D	N/D
		cca1-1 sic-3	25.3 ± 0.6	0.2 ± 0.07	N/D	94%	60
		cca1-1 lhy-20	22.8 ± 0.7	0.3 ± 0.09	N/D	97%	40
		cca1-1 lhy-20 sic-3	22.5 ± 2	0.35 ± 0.13	N/D	83%	25
	LL|22°C	WT	23.7 ± 0.4	0.2 ± 0.03	N/D	100%	48
		sic-3	24.7 ± 0.9	0.2 ± 0.08	N/D	100%	48
		sic-1	26.3 ± 1.1	0.2 ± 0.1	N/D	100%	47
		lhy-20	21.5 ± 0.3	0.3 ± 0.08	N/D	100%	38
		lhy-20 sic-3	23 ± 0.3	0.2 ± 0.04	N/D	100%	8
		cca1-1	22 ± 0.9	0.3 ± 0.07	N/D	100%	13
		cca1-1 sic-3	23.5 ± 1	0.2 ± 0.04	N/D	100%	30
		cca1-1 lhy-20	21.3 ± 0.6	0.2 ± 0.1	N/D	100%	28
		cca1-1 lhy-20 sic-3	20.8 ± 0.5	0.2 ± 0.07	N/D	100%	28
	LL|28°C	WT	23.6 ± 0.6	0.2 ± 0.09	N/D	100%	43
		sic-3	24 ± 0.8	0.2 ± 0.07	N/D	100%	47
		lhy-20	20 ± 0.6	0.3 ± 0.1	N/D	95%	20
		lhy-20 sic-3	20.9 ± 1.5	0.3 ± 0.1	N/D	82%	29
		cca1-1	21 ± 1.8	0.3 ± 0.08	N/D	100%	24
		cca1-1 sic-3	22.2 ± 1.4	0.3 ± 0.14	N/D	87%	27
		cca1-1 lhy-20	19.7 ± 2.4	0.3 ± 0.1	N/D	94%	34
		cca1-1 lhy-20 sic-3	20 ± 1.2	0.3 ± 0.13	N/D	73%	19

Table 5. (continued) Circadian clock parameters determined from *ProPRR7:LUC* rhythms in the indicated conditions.

LL|22°-18°C	-	WT	25 ± 0.5	0.4 ± 0.1	N/D	96%	113
		sic-3	25.6 ± 0.8	0.51 ± 0.1	N/D	81%	125
		sic-1	26.8 ± 1.1	0.56 ± 0.2	N/D	61%	119
		lhy-20	23.9 ± 0.3	0.38 ± 0.07	N/D	100%	42
		lhy-20 sic-3	24.7 ± 0.8	0.5 ± 0.1	N/D	95%	94
		cca1-1	24.7 ± 0.5	0.3 ± 0.1	N/D	94%	18
		cca1-1 sic-3	24.6 ± 1.3	0.5 ± 0.1	N/D	63%	57
		cca1-1 lhy-20	23.9 ± 0.57	0.4 ± 0.07	N/D	98%	57
		cca1-1 lhy-20 sic-3	23.9 ± 0.61	0.4 ± 0.1	N/D	98%	64
	LL|22°C	WT	23.3 ± 0.7	0.2 ± 0.07	N/D	100%	111
		sic-3	24.5 ± 1	0.2 ± 0.07	N/D	100%	129
		sic-1	25.9 ± 1.3	0.27 ± 0.1	N/D	100%	126
		lhy-20	21.9 ± 0.4	0.27 ± 0.1	N/D	100%	43
		lhy-20 sic-3	23 ± 0.5	0.2 ± 0.05	N/D	100%	94
		cca1-1	22 ± 0.8	0.3 ± 0.1	N/D	100%	22
		cca1-1 sic-3	23.8 ± 0.8	0.2 ± 0.08	N/D	100%	59
		cca1-1 lhy-20	21.7 ± 0.9	0.26 ± 0.09	N/D	100%	57
		cca1-1 lhy-20 sic-3	21.2 ± 0.9	0.26 ± 0.1	N/D	100%	64

Rhythmic *ProRR7:LUC* activity assayed for 4-5 days in the indicated condition and the indicated genotypes. Mean period, mean RAE (relative amplitude error), and mean phase is for rhythmic individuals and SD is standard deviation. Rhythmic seedlings are those with RAE<0.6. N/D, not done.

Chapter 4: Identifying interacting partners of SICKLE to understand its molecular function

The following text is in part modified from an article published in *The Plant Cell*, by Marshall *et al.*, 2016.

Abstract

SIC is a protein of unknown function that localizes to both the cytoplasm and to the nucleus. In the nucleus, the protein accumulates in punctate foci. SIC has five highly conserved amino acid regions that are shared with homologous proline/serine rich proteins in angiosperms, but these domains do not point to a specific function for SIC. Co-immunoprecipitation-mass spectrometry (IP-MS) and bimolecular fluorescence complementation (BiFC) experiments identified multiple interacting partners of SIC, including PRMT4a, PRMT4b, and DBR1. PRMT4a and PRMT4b are paralogous protein arginine methyltransferases that methylate splicing factors. DBR1, lariat debranching enzyme protein 1, is a member of the NineTeen Complex (NTC). The NTC is a multiprotein complex involved in transcription, transcription coupled DNA repair, and alternative splicing. Thus, SIC may participate in a protein complex(s) that modifies spliceosome function, possibly to maintain robust transcript splicing in all ambient temperature conditions.

Introduction

While SIC contains several conserved domains, none of these have previously characterized functions. As a result the molecular function of the SIC protein is difficult to predict a priori. A recently published IP-MS experiment pulled down SIC together with several proteins predicted to have roles in pre-mRNA metabolism (Karampelias et al. 2016). The proteins include a putative RNA binding protein involved in transcription-coupled DNA repair protein (AT5G28740), a protein predicted to have intron binding activity (AT2G38770), the lariat debranching enzyme protein (DBR) 1, and the two protein arginine methyltransferase (PRMT) proteins PRMT4a and PRMT4b (Karampelias et al. 2016). The protein encoded by gene AT5G28740 is homologous to the human XPA-binding protein 2 (to be referred to as AtXAB2 for *Arabidopsis thaliana* XAB2) (Kuraoka et al. 2008), and the AT2G38770 protein is homologous to the human Aquarius protein (to be referred to as AtAQR for *Arabidopsis thaliana* AQR) (Kuraoka et al. 2008). AtDBR1, AtXAB2, and AtAQR are all a part of the Prp19 complex (PRP19C) or NineTeen Complex (NTC) (Kuraoka et al. 2008; Chanarat and Sträßer 2013).

The NTC or Prp19C is a multiprotein complex whose assembly in the nucleus is crucial for progression of transcript splicing (Koncz et al. 2012; Chanarat and Sträßer 2013). Studies in yeast, *Drosophila*, mice, and humans show the NTC regulates co-transcriptional spliceosome assembly; because splicing occurs concurrently with transcription, the sequential formation of spliceosome conformations is necessary for completion of transcript splicing, and splicing efficiency (Chanarat and Sträßer 2013; Hogg, Mcgrail, and Keefe 2010; Guilgur, Prudêncio, and Sobral 2014; Yonemasu et al. 2005). Three of the nineteen proteins in the NTC were identified as SIC protein

interactors via IP-MS: AtDBR1, AtXAB2, and AtAQR (Karampelias et al. 2016). Only AtDBR1 has been characterized in Arabidopsis (H. Wang, Hill, and Perry 2004).

Human XPA-binding protein 2 (hXAB2), like its yeast homolog Synthetic Lethal with Cdc4 (Syf1), is a fundamental NTC component that is the core of a conserved multiprotein assembly composed of at least five splicing-associated proteins (Kuraoka et al. 2008). The function of the Syf1/hXAB2 complex in NTC remains unclear; however, Syf1 function is essential in yeast (Ben-Yehuda et al. 2000) and mice (Yonemasu et al. 2005). In humans hXAB2 was originally identified as an interactor with the hXPA protein, which is involved in transcription coupled DNA repair (Kataoka 2015). hXAB2 stabilizes the interaction between RNA polymerase and hXPA, a key protein that verifies lesions in DNA during transcription (Hanawalt and Spivak 2008). The hXAB2, hXPA, and RNA polymerase interaction is transient and triggered by DNA damage: *in vivo* studies showed the interaction is enhanced when DNA damage agents are added to HeLA cells (Kuraoka et al. 2008). Transcription coupled DNA repair targets DNA damage by removing DNA lesions in actively transcribed genes that block RNA polymerase during transcription (Hanawalt and Spivak 2008). Mutations in genes regulating transcription coupled DNA repair cause defects in DNA repair, but also transcription efficiency and alternative splicing (Hanawalt and Spivak 2008; Yonemasu et al. 2005). In addition to its role in transcription coupled DNA repair, hXAB2 also binds directly to RNA and interacts with several proteins known to be involved in mRNA splicing, including hAquarius (hAQR) (Kuraoka et al. 2008). Knockdown of hXAB2 results in a reduced rate of RNA splicing, indicating it is a pre-mRNA splicing factor (Kuraoka et al. 2008).

hAQR is a component of the spliceosome that binds to a region between small nucleolar RNA and a branch point in the splicing reaction complex (Stein et al. 2003); hAQR therefore plays a key role in assembly of the snRNPs necessary for pre-mRNA splicing (Stein et al. 2003). DBR are crucial for completion of transcript splicing to allow degradation of free spliced introns (Moore 2002). DBR remove the RNA lariat structure made during the first catalytic step of intron removal, formed as a unique 2'–5'-phosphodiester bond between the 5' end of the intron and an internal intron site known as the branch point (Keller 1984; Ooi et al. 2001). Correct removal of lariat intermediates in the nucleus is required for high fidelity splicing (Burgess and Guthrie 1993). hDBR1 associates with the NTC through interaction with hXAB2 (Masaki et al. 2015). hDBR1 shuttles between the nucleus and cytoplasm (Masaki et al. 2015). AtDBR1 is the only lariat debranching enzyme in Arabidopsis and its activity is essential (H. Wang, Hill, and Perry 2004), like hDBR1 (Findlay et al. 2014). Alternative splicing phenotypes in Arabidopsis *dbr1* mutants have never been tested (H. Wang, Hill, and Perry 2004). Previous work shows that hDBR1 accumulates in both the nucleus and cytoplasm (Masaki et al. 2015).

Arabidopsis PRMT4a and PRMT4b are two type I Protein Arginine Methyltransferases (PRMTs) redundantly required for control of flowering time, photomorphogenic responses, and salt stress tolerance (Hernando et al. 2015; Niu et al. 2008). PRMT4a and PRMT4b are thought to be important for the regulation of RNA splicing, similar to PRMT5 (Bedford and Richard 2005), possibly through direct methylation of spliceosomal proteins or by influencing the coupling between transcription and RNA processing(Cheng et al. 2007; Kuhn et al. 2011). As type I

enzymes, PRMT4a and PRMT4b catalyze the addition of asymmetric dimethyl-arginine moieties onto target proteins (Niu et al. 2008). The *in vivo* targets of the Arabidopsis PRMT4s are not well characterized, but *prmt4a;b* double mutants cause a global reduction in asymmetric dimethylation of histone H3 *in vivo* (Niu et al. 2008). In agreement, PRMT4a and PRMT4b each asymmetrically dimethylate histone H3 and a nonspecific protein substrate *in vitro* (Niu et al. 2008). Homodimers and heterodimers of PRMT4a/4b accumulate in the cytoplasm and nucleus (Niu et al. 2008). *prmt4a;b* double mutants exhibit major changes in transcription and alternative splicing for regulatory genes associated with abiotic stress responses (Hernando et al. 2015). Interestingly, circadian clock activity is unchanged in *prmt4a;b* (Hernando et al. 2015); however this has only been tested under optimal environmental conditions.

In an effort to understand the activity of SIC, the protein was characterized based on its amino acid sequence. We also sought to confirm interacting partners of SIC previously discovered in the Karampelias et al. 2016 IP-MS experiment (Karampelias et al. 2016). As predicted by this experiment, SIC was found to interact with DBR1 and both PRMT4a and PRMT4b in BiFC tests. These results indicate SIC has a role in a multiprotein complex, possibly the NTC, responsible for modifying spliceosome activity.

Results

SIC encodes a conserved proline/serine-rich protein found in nuclear foci

SIC is a 319 amino acid protein of unknown molecular and biochemical function (Zhan et al. 2012). The protein is enriched in proline and serine residues, which constitute 23% of all amino acids. Stable transgenic expression of a translational fusion of SIC and YELLOW FLUORESCENT PROTEIN (YFP) by the 35S promoter (*35S:SIC-YFP*) in WT Arabidopsis plants localized exclusively to punctate foci within the nucleus (Figure 30A). These foci are distributed throughout the nucleoplasm and are distinct from the nucleolus. Transient expression of *35S:SIC-YFP* from the same *35S:SIC-YFP* vector in *Nicotiana benthamiana* leaves results in the expected nuclear protein accumulation, but also protein in the cytoplasm (Figure 32H).

Sequence and phylogenetic analysis of SIC identified it as an angiosperm-specific protein with no homologs in gymnosperms, ferns, bryophytes, or algae (Figure 31). Alignment of 70 SIC protein homologs based on amino acid identity (Data File 1) identified five highly conserved amino acid regions (regions a-e) located at the N- and C-terminal portions of the protein (Figure 30B), while the central part of the protein is largely variable. The 70 known SIC homologs share 30% overall amino acid similarity, while the conserved amino acid regions together share 61.5% similarity. Individually, region a has 60% similarity, region b has 72% similarity, region c has 61% similarity, region d has 49% similarity, and region e has 72% similarity (Data File 1). A phylogenetic tree of SIC protein homologs shows it is a singleton in most species except for in maize (*Zea mays*), California poplar (*Populus trichocarpa*), desert poplar (*Populus euphratica*), banana (*Musa acuminate*), eucalyptus (*Eucalyptus grandis*), flax (*Linum usitatissimum*), Jatropha (*Jatropha curcas*), cocoa (*Theobroma cacao),* soybean *(Glycine max*), and *Brassica rapa* (Figure 31).

SIC has several predicted amino acid motifs (Figure 30B). At the N-terminus is a likely nuclear localization signal (NLS) sequence, consistent with the observed nuclear accumulation of SIC (Figure 30B). At the C-terminus is a short amino acid motif, SMxxDPWxxLxP (Figure 30B), which is present in a family of proline-rich proteins that are widely distributed across plant and animal proteomes. The Pfam database predicts this motif as a conserved MPLKIP (M-phase-specific PLK1-interacting protein) protein motif (Finn et al. 2014) (SerMetXGluAspXXLeuXPro) near the C-terminus (Zhang et al. 2007) (Figure 30B). The remaining conserved domains of SIC harbor additional short amino acid motifs predicted to mediate interactions with DNA, RNA, or other proteins (Figure 30B). Proteins in the MPLKIP family are not well studied, except for human TTDN1 (trichothiodystrophy nonphotosensitive 1) (Zhang et al. 2007). A mutation in TTDN1 is the cause of the human disorder nonphotosensitive trichothiodystrophy (TTD), which causes intellectual impairment and developmental defects (Zhang et al. 2007). The invariant M in the SMxxDPWxxLxP motif of TTDN1 is changed to a V in nonphotosensitive TTD patients (Zhang et al. 2007), indicating the importance of this motif in protein function. The function of TTDN1 is undefined, but it is known to physically interact with and is a substrate of the protein kinase PLK1 (Polo Kinase 1) (Zhang et al. 2007), a key regulator of the mitotic cell cycle.

SIC interacts with DBR1, a component of the NineTeen Complex

XAB2, DBR1, and AQR are three presumed members of the Arabidopsis NTC that potentially interact with SIC (Karampelias et al. 2016). BiFC was used to gain *in vivo* evidence that SIC directly interacts with these proteins. BiFC is based on reconstitution of YFP activity through the interaction of two test proteins, each of which carries either the N- or C-terminal half of YFP; interaction between the test proteins brings the two halves of YFP in close enough proximity to restore YFP fluorescence (Walter et al. 2004). The fusion proteins are transiently expressed by Agrobacterium infiltration of *N. benthamiana* leaves. *XAB2* and *AQR* could not be tested because the CDS of each could not be cloned, possibly due to very low transcript expression. The CDSs of *SIC* and *DBR1* were each cloned into four possible BiFC vectors: *35S:proteinCDS-YFPn* (N-terminal half of YFP), *35S:proteinCDS-YFPc* (C-terminal half of YFP), *35S:YFPn-proteinCDS*, and *35S:YFPc-proteinCDS* (Figure 32D; Table 6).

When the *35S:YFPn-DBR1* and *35S:YFPc-SIC* fusions were combined, obvious fluorescence appeared in the cytoplasm and in the nucleus, indicating a SIC-DBR1 interaction (Figure 32D), consistent with published results that hDBR1 localizes to the nucleus and cytoplasm (Masaki et al. 2015). If SIC and DBR1 interact, these two proteins were expected to colocalize within cells. To test this possibility, a *35S:SIC-YFP* fusion was coexpressed with a *35S:DBR1-CYAN FLUORESCENT PROTEIN (CFP)* fusion in *N. benthamiana* leaves. As expected, *35S:SIC-YFP* appeared in the cytoplasm and within the nucleus as punctate foci (Figure 32A). *35S:DBR1-CFP* exhibited a subcellular distribution similar to *35S:SIC-YFP*, including the appearance of nuclear foci (Figure 32B). Overlay of the YFP and CFP channels demonstrated that the signal from SIC and DBR1 overlapped, consistent with colocalization (Figure 32C). Therefore, SIC directly interacts with DBR1 in a complex that occurs in the cytoplasm and within discrete nuclear foci. Cloning of the *DBR1* CDS, colocalization, and BiFC tests were primarily conducted by Emma Kovak, a graduate student from the Harmon Lab.

SIC interacts with the methyltransferases PRMT4a and PRMT4b

BiFC was employed to test for a direct interaction between SIC and either PRMT4a or 4b (Figure 32E-G; Table 6). The CDSs of *SIC*, *PRMT4a*, and *PRMT4b* were each cloned into four possible BiFC vectors: *35S:proteinCDS-YFPn, 35S:proteinCDS-YFPc, 35S:YFPn-proteinCDS,* and *35S:YFPc-proteinCDS* (Table 6). All possible combinations of SIC, PRMT4a, and PRMT4b were evaluated for YFP signal. PRMT4a together with PRMT4b served as a positive control (Figure 32G), since these proteins are known to interact with one another in the cytoplasm and nucleus (Niu et al. 2008). SIC together with either PRMT4a or PRMT4b generated YFP signal (Figure 32E,F), indicating direct SIC interaction with both PRMT proteins (Figure 32E,F) in the cytoplasm and nucleus (Figure 32E,F). Thus, SIC interacts with both PRMT4a and PRMT4b *in vivo*.

Discussion

SIC has five highly conserved amino acid regions (regions a-e) that are shared with homologous proline/serine rich proteins in angiosperms (Figure 30B; Figure 31; Data File 1). The function of the SIC protein is unknown, but these five highly conserved domains may indicate the locations of the protein that interact with proteins or possibly with RNA. For example, the MPLKIP motif, which is not present in *sic-1,* may play a role in binding proteins involved in the auxin signaling pathways, while the conserved domain B of SIC may be involved in circadian clock pathways (Figure 3D; Figure 30B). Due to the conserved nature of SIC in angiosperms (Figure 31), it is possible that this protein plays an important role in the ability of flowering plants to adapt and process different temperature conditions.

SIC accumulates in punctate nuclear foci distributed throughout the nucleus (Figure 30A; Figure 32A,H). Previously, SIC foci were shown to co-localize with foci formed by HYL1 (Zhan et al., 2012), an RNA-binding protein important for the generation of miRNAs (Song et al. 2007). Colocalization with HYL1 foci, the DBR1 foci (Figure 32C), together with the shape and location of the SIC-containing foci (Figure 32A,H), indicates these are RNA processing-associated nuclear bodies, possibly “dicing bodies” involved in microRNA (miRNA) biogenesis or sites of spliceosome assembly/activity (Shaw and Brown 2004). Our data determined that SIC interacts with PRMT4a, PRMT4b, and DBR1 in the nucleus and the cytoplasm (Figure 32; Table 6). Stable transgenic lines expressing *35S:SIC-YFP* in Arabidopsis do not show obvious accumulation outside of the nucleus (Figure 30A), but transient expression in *N. benthamiana* shows SIC accumulation in the nucleus and cytoplasm (Figure 32A,H). Further microscopy work must be done to determine the cause of this discrepancy.

It is possible that SIC plays multiple roles that are determined by its protein interacting partners and by its location in the cell. For example, Karampelias et al. 2016 determined that SIC interacts and regulates the activity of Protein Phosphatase 2A (PP2A) in the cytoplasm, which phosphorylates PIN proteins for PIN recycling and polarity establishment in auxin-mediated development (Karampelias et al. 2016). Furthermore, while PRMT4a and PRMT4b are the obvious candidates for explaining a SIC pathway involved in alternative splicing, the *prmt4a;4b* mutant displays no circadian

clock phenotypes (Hernando et al. 2015), in contrast to the long period phenotype of *prmt5* mutants (Hong et al. 2010; Sanchez et al. 2010). It is possible that *prmt4a;4b* mutants would show circadian clock phenotypes in non-optimal conditions, such as in cool temperature cycles, but it is also possible that SIC's role with PRMT4a and PRMT4b is outside of the circadian clock and alternative splicing pathway. In particular, SIC and either PRMT4a or PRMT4b interaction appears most strongly in the cytoplasm (Figure 32E,F). Another possibility is that SIC acts to suppress PRMT4a and 4b activity, therefore future experiments to test PRMT4a and 4b overexpression may reveal circadian clock phenotypes. Nonetheless, it is likely that the circadian clock phenotypes caused by *sic* are not simply due to a loss of PRMT4 activity. Clearly, a more complete understanding of the biochemical and molecular activities of SIC will be critical to rationalize these paradoxical observations.

Another model for SIC function is that it participates in modifying the activity of the NTC. SIC is not homologous to a protein in humans, mice, *Drosophila*, or yeast that is known to modify the NTC, and of the NTC proteins identified as SIC interactors (XAB2, AQR, DBR1) only DBR1 has been briefly studied in Arabidopsis (H. Wang, Hill, and Perry 2004). Therefore, no obvious pathway presents itself to define SIC's function in modifying the NTC. Nevertheless, SIC's interaction with DBR1 represents the intriguing possibility that SIC contributes to transcription, transcription coupled DNA repair, and alternative splicing. It is well understood in humans that disruption of hXAB2, hDBR1, or hAQR affects the splicing efficiency of the spliceosome, DNA repair that occurs during transcription, and transcription of genes (Yonemasu et al. 2005; Kuraoka et al. 2008; Masaki et al. 2015). When any of the NTC proteins are mutated in humans, the pathway is disrupted and all subsequent processes are affected, including alternative splicing (Yonemasu et al. 2005; Kuraoka et al. 2008; Masaki et al. 2015; Hanawalt and Spivak 2008). We propose that in Arabidopsis, a function of SIC is to modify the NTC through interaction with DBR1, XAB2, and/or AQR, and this regulation is crucial to proper splicing of transcripts, including circadian clock transcripts, as well as gene expression. Without SIC, transcription of circadian clock genes and the splice variant pool of circadian clock transcripts may be modified due to a change in NTC activity, ultimately leading to improper timing of proteins such as LHY and CCA1. Furthermore, transcription is a temperature-sensitive process (Roe, Burgess, and Record 1985), and we hypothesize that SIC is important for maintaining NTC activity in cool temperatures.

SIC is unlike predictable regulators of the spliceosome or spliceosomal components that affect the circadian clock when absent, such as PRMT5, SKIP, STIPL1, and GEMIN2 (Hong et al. 2010; Sanchez et al. 2010; X. Wang et al. 2012; Jones et al. 2012; Schlaen et al. 2015) because it is not a putative spliceosomal component. Instead of acting specifically within the core clock mechanism, SIC appears to be necessary at cool temperatures to maintain control of splice variant production for specific clock-associated transcripts (Figure 33). In this model (Figure 33), appropriate levels of certain splice variants, such as *LHY* and *CCA1*, are critical for temperature entrainment and temperature compensation. Ambient temperature influences spliceosomal activity in such a way as to change levels of these splice variants, as previously proposed (Schlaen et al. 2015). Another explanation for the splice variant phenotype in *sic* is that SIC is necessary for NMD-promoted removal of clock-

associated splice variants that are an undesirable consequence of cool temperatures. Due to the strong localization of SIC in the nucleus and its interaction with splicing-related proteins, it is more likely to contribute to processes within the nucleus, as NMD is a translation-dependent process that occurs outside the nucleus (He and Jacobson 2015).

In summary, our model can be described as the following: alternative splicing converts circadian clock pre-mRNA transcripts to mature mRNA in a temperature-dependent manner, and the resulting splice variant pool determines the proper accumulation and timing of circadian clock proteins (Figure 33). This form of post-transcriptional circadian clock regulation is crucial for temperature compensation and thermocycle entrainment in Arabidopsis (Figure 33). Where SIC fits into this model is unknown, but we propose two possibilities (Figure 33): 1) SIC modifies PRMT4a and PRMT4b activity, which then affects spliceosome activity, or 2) SIC modifies the NTC to regulate transcription, transcription coupled DNA repair, and alternative splicing. All of these processes are temperature-dependent, and it is clear that the activity of SIC is most important in cool temperatures. Discovery of the biochemical and molecular activities of SIC is certain to reveal important details about the functional link between alternative splicing and temperature responses within the Arabidopsis circadian clock, and holds the potential to provide new insights into the molecular processes involved in pre-mRNA processing.

Materials and Methods

Construction of Transgenic Plants

To generate *35S:SIC-YFP* lines, a 957 base pair amplicon of the SIC CDS (lacking the endogenous stop codon) was PCR amplified from WT cDNA and cloned into the pEarleyGate 101 binary vector (Earley et al. 2006). T_1 transformants in the WT background were identified as phosphinothricin-resistant individuals. Expression of the SIC-YFP fusion protein in the T_3 generation was confirmed by protein gel blot with anti-GFP Living Colors A.v. Monoclonal Antibody (JL-8) (Clontech) at a 1:20,000 dilution, and goat anti-mouse IgG HRP antibody (Santa Cruz Biotechnology) at a 1:10,000 dilution, as described in Chapter 3. Crossing of the *35S:SIC-YFP* construct into the *sic-3* background rescued the long period phenotype. To determine SIC-YFP subcellular localization, *35S:SIC-YFP* lines were grown under LD|22°C, and YFP signal was visualized in 13 day-old seedlings. Fluorescence and bright field microscopy were performed on roots with a Leica DM4000B microscope (Leica Microsystems). Cloning and sequencing of the *SIC* transcript in WT and *sic* alleles employed *Taq* DNA Polymerase to amplify the desired PCR product from cDNA (see Data File 2 for primers). PCR products were cloned into the pCR2.1-TOPO vector (Thermo Fisher Scientific) and sequenced at the UC Berkeley DNA sequencing facility (mcb.berkeley.edu/barker/dnaseq/home).

Phylogenetic Analysis

Arabidopsis thaliana SIC homologs were identified using BLASTP to identify homologous amino acid sequences (Altschul et al. 1990) in the UniProt Database (Consortium 2015), Phytozome 11 (Goodstein et al. 2012), and EnsemblPlants (Kersey et al. 2016). Homologs were compiled using Geneious V. 9.0.5 software with the default settings (Kearse et al. 2012). This analysis generated a multiple sequence alignment with the MUSCLE tool based on a BLOSUM 62 matrix (Edgar 2004), which was manually edited for obvious mismatched amino acid alignment errors in Mesquite V.3.04. Mesquite was also used to trim the alignment to each separate domain, and re-uploaded to Geneious to determine percent similarity. A maximum likelihood tree was built using PhyML 3.0 (Guindon et al. 2010) with the aBayes Fast likelihood-based method with 1,000 bootstrap replicates. Data were visualized in FigTree v1.4.2. Predicted protein domains were identified using the PredictProtein tool (Yachdav et al. 2014). Alignment may be viewed in Data File 1.

Plant material and growth conditions

Nicotiana benthamiana plants were grown in a growth chamber at 25°C with 9 h light (~130 µEinsteins m-2 s-1) and 15 h dark cycles. Seeds were sowed on Sunshine Mix soil in individual pots, and used 3 to 5 weeks after transplanting. 20-20-20-fertilizer mix was used every 2 weeks to supplement the soil.

Cloning of transient expression vectors

A 1587 base pair amplicon of the PRMT4a CDS (with the endogenous stop codon) was PCR amplified from WT cDNA and cloned into the pB7GWYn and pB7GWYc binary vectors with Q5 DNA polymerase. A 1605 base pair amplicon of the PRMT4b CDS (with no endogenous stop codon) was PCR amplified from WT cDNA and cloned into the pB7GWYn, pB7GWYc, pK2GWYn, and pK2GWYc binary vectors with Q5 DNA polymerase.

Amplicons of 1257 and 1254 base pairs of the DBR1 CDS (with the endogenous stop codon and without, respectively) were PCR amplified from WT cDNA using Q5 DNA polymerase, and cloned into either pB7GWYn and pB7GWYc (for CDS with stop codon), or pK2GWYn and pK2GWYc (for CDS without stop).

A 2112 base pairs amplicon (with no endogenous stop codon) of the SIC genomic DNA sequence and a 960 base pair amplicon (with endogenous stop) of the SIC CDS were PCR amplified from WT genomic DNA and WT cDNA using Q5 DNA polymerase, respectively. The SIC genomic sequence was cloned into pB7GWYn and pB7GWYc, and the SIC CDS was cloned into pK2GWYn and pK2GWYc.

The pB7GWYn, pB7GWYc, pK2GWYn, and pK2GWYc binary vectors were obtained from Jennifer Lewis at the Albany USDA Plant Gene Expression Center.

Colocalization experiments used the *35S:SIC-YFP* construct used to make the stable Arabidopsis transgenic lines (see above). The *35S:DBR1-CFP* construct was created by cloning the DBR1 CDS (with no endogenous stop), that was PCR amplified above, into the pEarleyGate 102 binary vector (Earley et al. 2006).

Agrobacterium-mediated transient expression

Agrobacterium tumefaciens GV3101 (pEG101 and BiFC vectors) or GV2260 (P19 silencing suppressor) cultures were grown overnight at 28°C in LB broth with antibiotics. The next day, the cultures were resuspended in 10 mL of induction medium (50 mM MES pH 5.6, 0.5% (wt/vol) glucose, 1.7 mM NaH2PO4, 20mM NH4Cl, 1.2 mM MgSO4, 2mM KCl, 17 µM FeSO4, 70 µM CaCl2, and 200 µM acetosyringone) and incubated at 28°C for ~4 hours. Cultures were spun down and resuspended in 10 mM MES pH 5.6 with 200 µM acetosyringone to an optical density of 600 nm (OD600) of 1.0. The cultures containing each plasmid were mixed in equal volumes to a final OD600 of 0.25 per construct. P19 silencing suppressor culture was used in all infiltrations. The underside of the leaves of 5- to 7-week-old *Nicotiana benthamiana* plants were infiltrated by hand with a needleless syringe. Fluorescence microscopy was performed on leaves with a Leica DM4000B microscope (Leica Microsystems). We thank Mael Baudin from Jennifer Lewis' lab for the P19 silencing suppressor strain and for the transient expression assay.

Accession Numbers

Gene models from this thesis can be found in the Arabidopsis Genome Initiative or GenBank/EMBL data libraries under the following accession numbers: *SIC*, AT4G24500; *LHY*, AT1G01060; *CCA1*, AT2G46830; *CHE,* AT5G08330; *RVE1,* AT5G17300*; RVE4,* AT5G02840; *RVE7,* AT1G18330; *RVE8*, AT3G09600; *TOC1*, AT5G61380; *PRR9*, AT2G46790; *PRR7*, AT5G02810; *PRR5*, AT5G24470; *PRR3*, AT5G60100; *GI*, AT1G22770; *LUX*, AT3G46640; *BOA,* AT5G59570; *ELF3*, AT2G25930; *ELF4*, AT2G40080; *TIC1,* AT3G22380*; LNK1,* AT5G64170; *LNK2,* AT3G54500*; LKP2,* AT5G62430; *FKF1,* AT1G68050; *CAB2*, AT1G29920; *CAT3*, AT1G20620; *CCR2*, AT1G06820; *SKIP*, AT1G77180; *STIPL1*, AT1G17070; *PRMT5*, AT4G31120; *DBR1,* AT4G31770; *XAB2,* AT5G28740; *AQR,* AT2G38770; *GEMIN2,* AT1G54380; *PRMT4A,* AT5G49020; *PRMT4B,* AT3G06930; *PP2A*, AT5G60100; *IPP2*, AT3G02780; *UBC21*, AT5G25760; AT4G31770; AT2G38770; AT3G18790; AT1G13690; and AT5G28740. Sequence data from this article can be found in the NCBI Sequence Read Archive under BioProject accession number PRJNA314711.

Primers

Primers for this thesis can be found in Data File 2.

Chapter 4 Figures

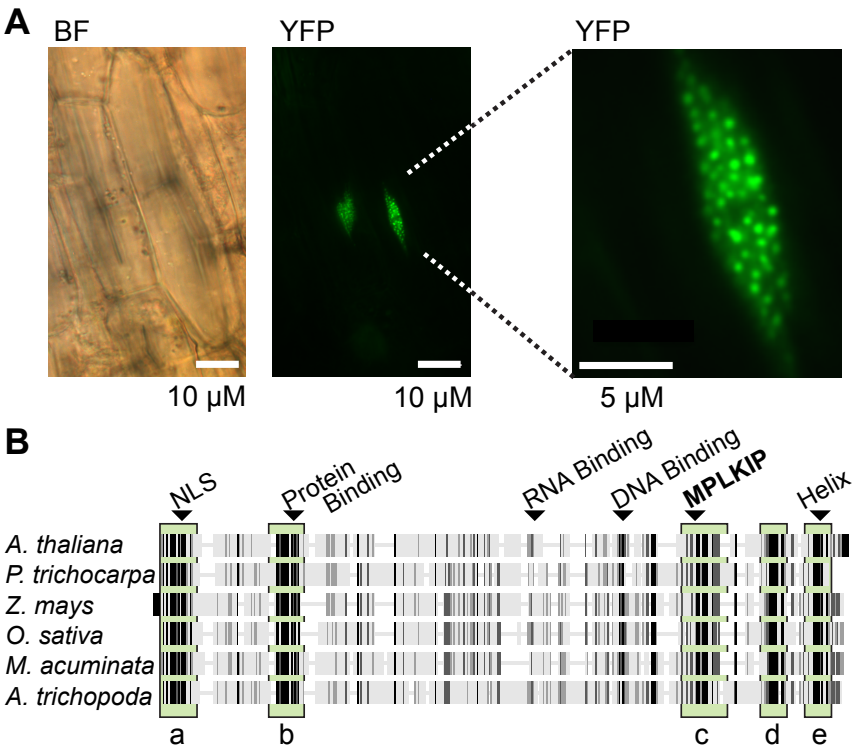


Figure 30. SIC is a conserved proline rich protein that accumulates in nuclear foci.

(A) Subcellular localization of a SIC-YFP fusion in nuclei of whole root cells detected by fluorescence microscopy. Images are bright field (BF) and false colored YFP fluorescence (SIC-YFP). Bars indicate scale as indicated.
(B) MUSCLE alignment of *Arabidopsis thaliana* SIC to SIC-like proteins from *Populus trichocarpa*, *Oryza sativa japonica*, *Zea mays*, *Musa acuminata*, and *Amborella trichopoda*. Conserved amino acid regions are highlighted in green, corresponding to Arabidopsis SIC amino acids 1-20 (region a), 58-70 (region b), 239-260 (region c), 273-284 (region d), and 300-319 (region e). Short vertical bars indicate the level of percent amino acid similarity corresponding to 100% (black), 80-100% (dark grey), and 60-80% (light grey).

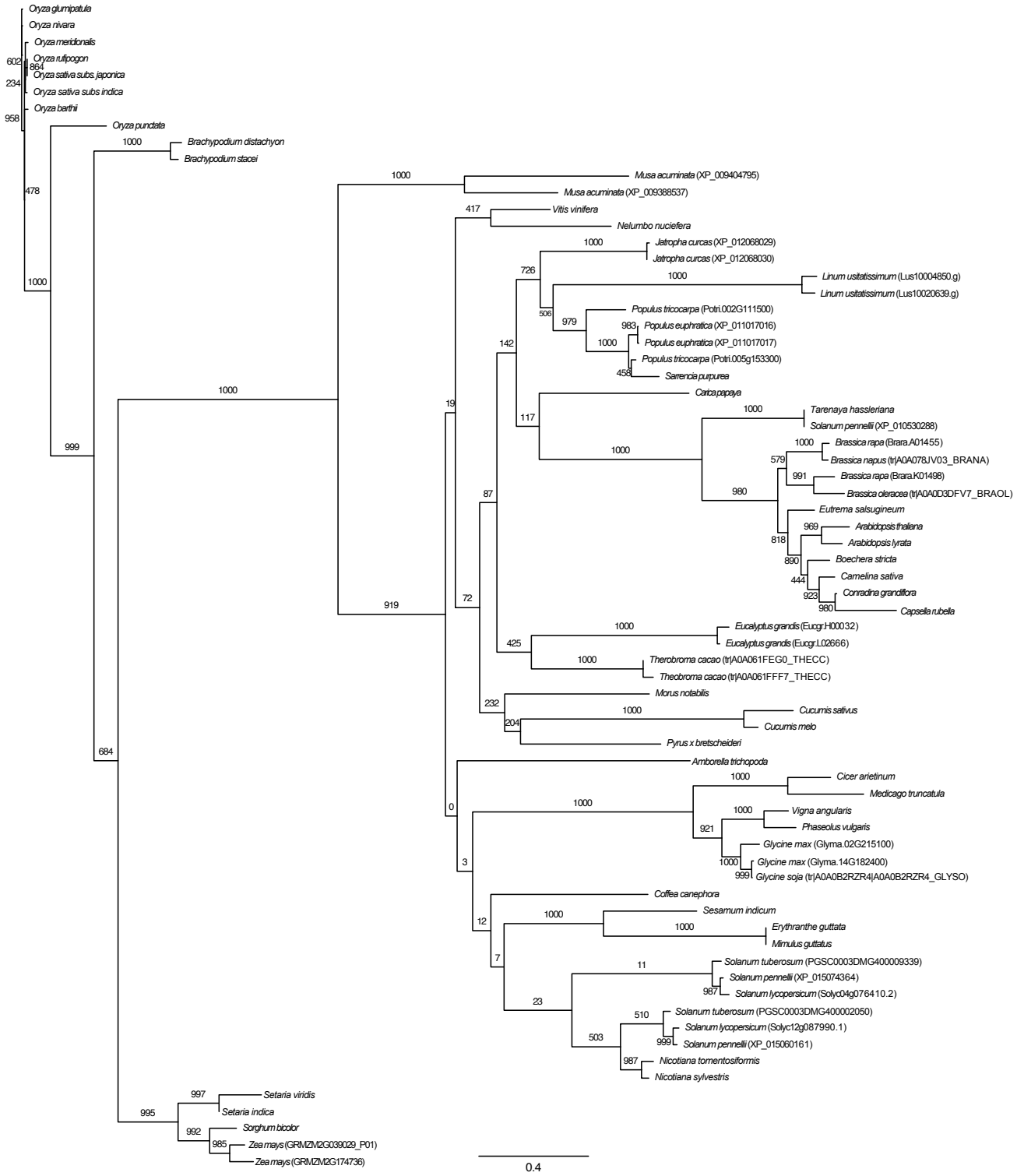


Figure 31. Phylogenetic tree of SIC homologs in angiosperms.

Phylogenetic tree based on maximum likelihood for the amino acid sequence of SIC homologs in angiosperms. Tree is unrooted as it is an angiosperm-specific protein, and belongs to no putative protein family. Tree is drawn to scale with branch lengths

measured in number of substitutions per site. Bootstrap values at each node are for 1,000 trials.

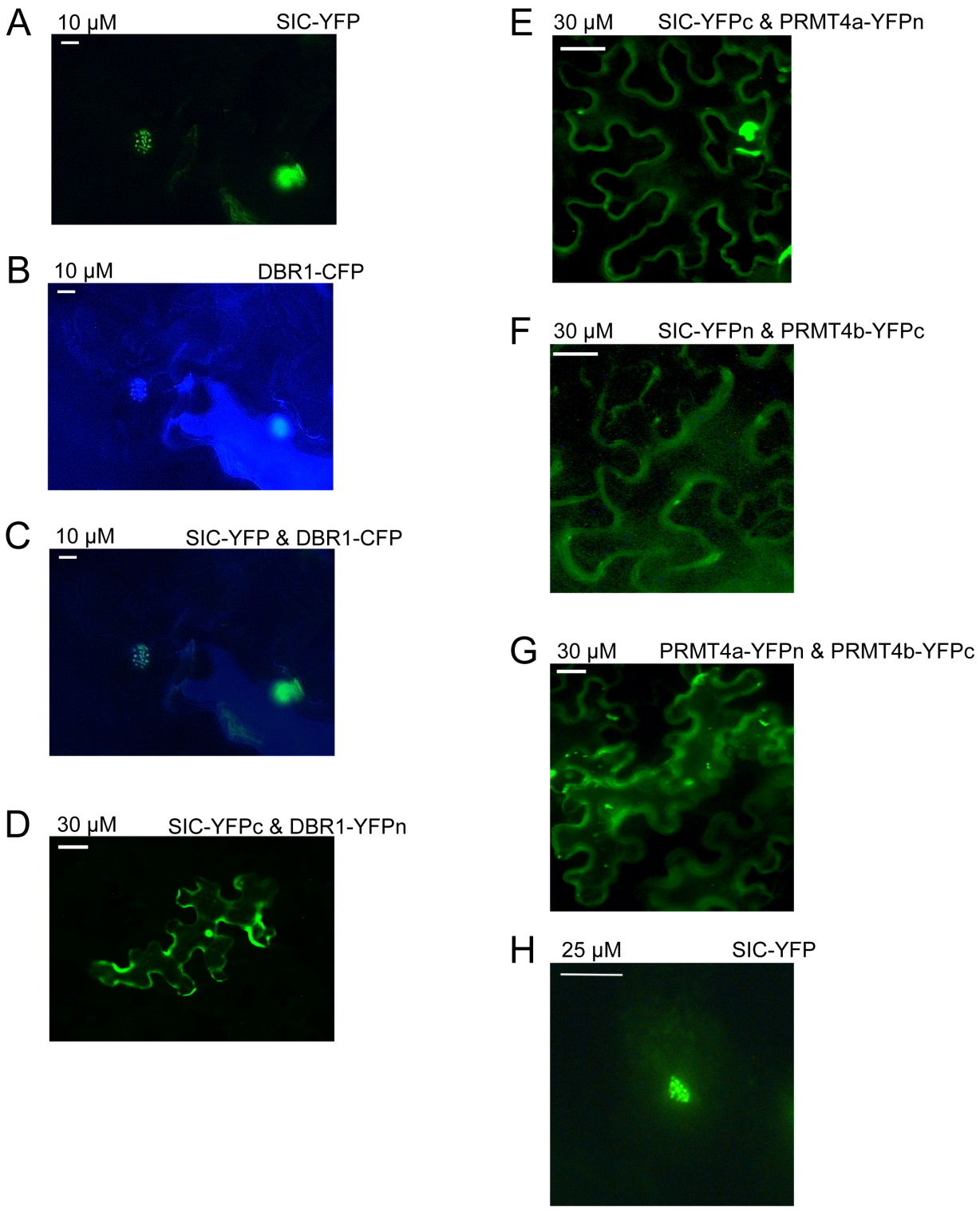


Figure 32. SIC interacts and co-localizes with DBR1, and SIC interacts with PRMT4a and PRMT4b.

A) to **(H)** Transient expression of fluorescent tagged Arabidopsis proteins in *N. benthamiana*.

A) to **(D)** represents work completed by Emma Kovak.
A) to **(C)** Subcellular co-localization of **(A)** *35S:SIC-YFP* and **(B)** *35S:DBR1-CFP*. **(C)** is an image overlay of **(A)** and **(B).**
(D) BiFC test showing interaction of SIC-YFPc (carrying the C-terminus half of YFP) and DBR1-YFPn (carrying the N-terminus half of YFP).
E) to **(G)** BiFC tests showing interaction of **(E)** SIC-YFPc and PRMT4a-YFPn, **(F)** SIC-YFPn and PRMT4b-YFPc, and **(G)** PRMT4a-YFPn and PRMT4b-YFPc.
H) Experimental positive control for BiFC tests of *35S:SIC-YFP*.
Bars indicate scale as indicated.

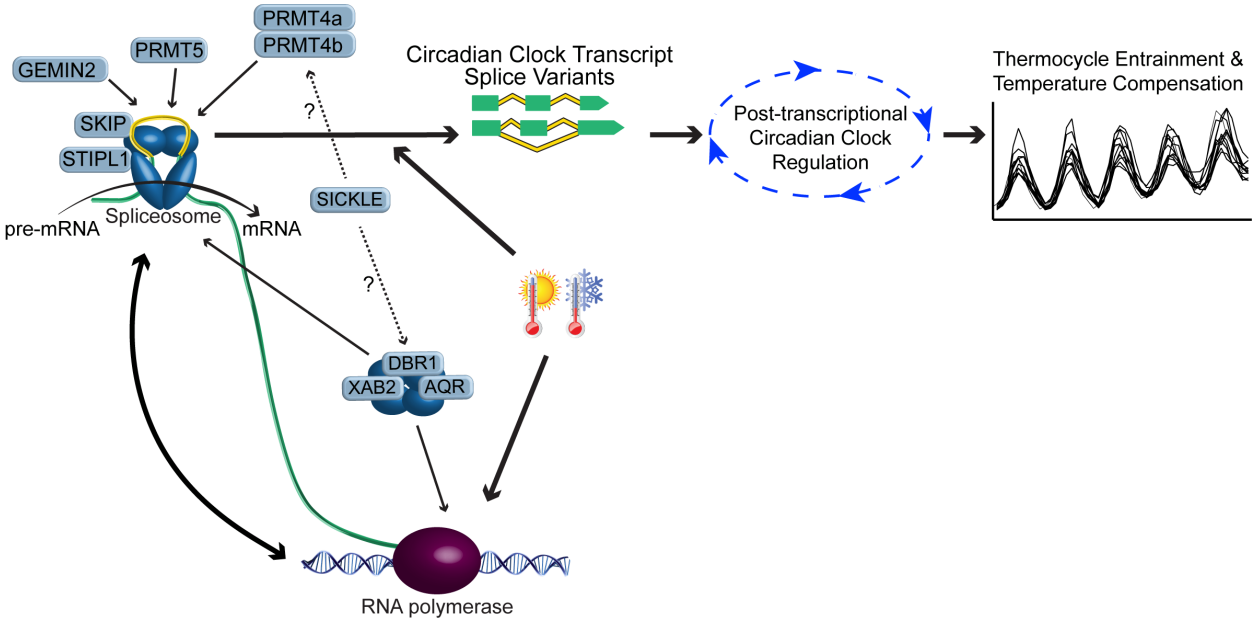


Figure 33. Model for how SIC affects circadian clock temperature responses.

See text for details. Blue boxes represent proteins. Blue spheres indicate spliceosome component proteins. Thin arrows represent indirect protein interaction with the spliceosome. Green bars represents exons, and yellow bars represent introns for mRNA. Thick arrows represent sequential order of events. Dotted lines represent potential functions of SIC as a regulator of splice variant accumulation or a promoter of nonsense-mediated decay.

Table 6. Bimolecular fluorescence complementation of SIC with PRMT4a, PRMT4b, and DBR1

	35S:SIC -YFPn	35S:SIC -YFPc	35S:YFPn -SIC	35S:YFPc -SIC	35S:PRMT4b -YFPn	35s:PRMT4b -YFPc	35S:YFPn -PRMT4b	35S:YFPc -PRMT4b
35S:YFPn -PRMT4a	N/A	none	N/A	Strong interaction Membrane and cytosol	N/A	none	N/A	Strong interaction Membrane and cytosol
35S:YFPc -PRMT4a	Strong interaction Membrane and cytosol	N/A	Strong interaction Membrane and cytosol	N/A	none	N/A	Strong interaction Membrane and cytosol	N/A
35S:PRMT4b -YFPn	N/A	none	N/A	none	N/A	N/A	N/A	N/A
35s:PRMT4b -YFPc	none	N/A	none	N/A	N/A	N/A	N/A	N/A
35S:YFPn -PRMT4b	N/A	none	N/A	none	N/A	N/A	N/A	N/A
35S:YFPc -PRMT4b	none	N/A	Interaction Membrane and cytosol	N/A	N/A	N/A	N/A	N/A
35S:DBR1 -YFPn	N/A	none	N/A	none	N/A	N/A	N/A	N/A
35s:DBR1 -YFPc	none	N/A	none	N/A	N/A	N/A	N/A	N/A
35S:YFPn -DBR1	N/A	none	N/A	Interaction Nucleus and cytoplasm	N/A	N/A	N/A	N/A
35S:YFPc -DBR1	none	N/A	none	N/A	N/A	N/A	N/A	N/A

Data File 1. SIC alignment within Angiosperms

>Sevir.7G033000 (Setaria viridis)
----------MRQPPSPGAP------LPD----EVAATPTRRQAARLPSR--AD---EPRALKLCGNYGDNN----------
----------PPHQHHLAPSPIHSPPLI-PRGVPGSCP-WRSPMQFQDPMSGY--QGAP----------
------PGAPPPW-GPH----------SGPP--------ARGSY----PNSPR?G-FRHP--------NPGRGGSPM----NYG----------
PRGSLNSSYGRGRG----------PNNYGSSG----------SRGR----------GGRCG-FGWQ----------D----------
RTYFIKSMVDDPWLGLQPIV----------GNILIPKGD--S-ESWLPKSLREKKETPAQGQ--IKSTS-GLSLAEYLDLSFNEVSNKET-------
>tr|K3Y8V1|K3Y8V1_SETIT (Setaria italic)
MDGGEHEPEESSSQRRERLLALRSA---ASASPAGAPPPAPAGSL-----LPDPDL---AGDEATCQRPRPPQR--FDYYTNPAA-
AFSSSYSGGA----------TNPTWSHKRKSPPACYY-PRPAPPPP-AYGNY----------GDNN----------PPHQHHLAPSPIHSPPLI-
PRGVPGSCP-WRSPMQFQDPMSGY--QGAP----------PGAPPPW-GPH----------SGPP--------ARGSY----
PNSPRFG-FRHP--------NPGRGGSPM----NYG----------PRGSLNSSYGRGRG----------PNNYGSSG----------
SRGR----------GGRCG-FGWQ----------D----------RSYFIKSMVDDPWLGLQPIV----------GNILIPKGD--S-
ESWLPTSLREKKETPAQGQ--IKSTS-GLSLAEYLDLSFNEVSNKET-------
>tr|C5YB92|C5YB92_SORBI (Sorghum bicolor)
MDGGEHEA-ESSSQRRERLLALRSA---ANASPAGDPPQAPAGSL-----LPDP---ELPGDQAASVCPPPPQR--FDYYTNPAA-
AFTSSYSGGA----------TNPTWSHKRKSPPACYA-----PPPP-AYGNY----------GSNYP----------PRQQHIPSQVHSPSLM---
PQDAPGNSP-WRSPMQFQDPMSGY--QGAP----------PRAPPSW-GSH----------CGP--------RGRGPY----
SNSPNFG-FRHP--------NPGRGGSPM----NYG----------PRGGPYSSYGRGRE----------PNYFGNPG-------
SRGRGGRGG-----V--------GFQNHPGWQGRS----------YFNKSMLDDPWLDLQPAV----------
GNILIPRAEYDSNKSWLPESLRKRKET-PAQGQ-IKTTS-GLSLAEYLDLSFNEVSDKEK-------
>GRMZM2G039029_P01 (Zea mays)
MDGGEHEAEESSYKRRERLLALRSA---ANASPAGAPPPTAAGSI-----LPDPDL---PGDEAASVCPTPPQR--FDYYTNPAA-
AFTSSSGGAT----------NPTWSHKRKSPPACYA---PPPPPP-AYGNY----------GSDY----------HRPHQQPTSWQVHSTSPM-
PQDAARSSP-WQSPMQFQDPMSGY--QGGP----------PSAPRPW-GSH----------SGP--------GGRGSY----
SNPPNFG-FRHP--------NPSRGGSPM----NYG----------PRGGPYSSYGRGRG----------PNYFGSPG----------
SRGR----------GGR---VGCH----------G----------RSYFNKSMLDDPWLDLQPGV----------
GNILIPRGEYDSNKSWLPESLRKKKETPPAQGQ-IKPTP-GLSLAEYLDMSFNEVSDKEK-------
>GRMZM2G174736 (Zea mays)
MDGGEHEAEESSSKRRERLLALRSA--------ANVSPAGAPPPTAAGSLLPDPDL---PDDEAASVCPPPPQR--FDYYTNPAA-
AFTSSSGGAT----------NPTWSHKRKSPPACYA-----PPPP-AYGNY----------CSNYH----------PPHQQPTSSWQVHSPSPM-
PQDAPRSSP-WQRPMQFQDPMSGY--QGGP----------PSAPPPW-VSH----------SGP--------RGRGSY----
SNPPNFG-FRHP--------NPSRGGSTM----NYG----------PRGGPYSSYGRGRG----------PNYFGSPG-------
SRGRGGRGG--WQGQ----------SYMNKSMLDDPWVDLQPGV----------
GNILIPRGEYDSNKSWLPESLRKKKETPPAHGQ-IKPAP-GLSLSEYLDLSFNEVSDKEK-------
>Bradi2g21130 (Brachypodium distachyon)
MDTGEQ-AEESSASRRERLLALRSAAASASSSASSPSAAPPPPPGTAAWDLPEPDL-------MPSSAPRPPPR--FDFYTNPGA-AFSSA--
----------AAPHKRKSAHSPPRSPA---PAPPQA-GSGNY----------GNNYP----------PPHQQHMPRSPIHSPFPM-
APGTPGNSQ-WQSHMQFQTPMSGY--RGTH----------PGPPPQW-NPH----------SASP--------AQVYH----
PHPPSYG-FRGP--------NVGRGGTPM----NFV----------PRGSPYPSPGQGRG----------GNYYSNPG-------
SRGKGGRGF----------QNHSGWQDRR----------NYYSKSMVDDPWQDLQPIV----------GNIMIPVGA--
ASQSWLPKSLRAKKDT-SDQGP-QGSGS-RLSLAEYLDLSFNETSNET--------
>Brast08G076100 (Brachypodium stacei)
MDTGEQ-AEESAASRRERLLALRSA-A-AAAS-SSPSAAPPPPPGTAAWDLPEPDL-------MPSSAPRPPPR--FDFYTNPGA-AFSSA-----
----------AAPQKRKSPYSPSHSP---APAPPQA-GSGNY----------GINYP----------PPHQQHMPRSPIHSPSPM-APGTPGNSQ-
WQSHMQFQTPMSGY--RGTP----------PGPPPQW-NPH----------SASP--------AQVYH----PHSPSYG-FRGP------
--NVGRGGSPM----NFV----------PRGSPYSSPGRGRG----------GNYYSNPG-------PRGRGGRGF----------
QNHSGWQDRR----------NYYSKSMVDDPWRDLQPIV----------GNIMIPVGA--TSQSWLPKSLRAKKDT-SDQGP-
QVSGS-RLSLAEYLDLSFNETSNET--------
>tr|A0A0E0K081|A0A0E0K081_ORYPU (Oryza punctate)
MEGGEH--EESSAQRRERLLALRSSSS-AAAASSPSAAPHPPPPAADVWDLPVRDLMDASA---SAAAPRPPPR--FDYYTNPTA-
AFASSSAASH----------KRKVAESPPPGNY----------GSGY----------PPHQHQQFQAPMSGY----
RGPPPGTPPPWSPHSSAPPPWSPH--SGAR----------PPW-SPH----------SAP--------SQGPY----PHPPSYG-
PRNY--------NPGQRGGRM----NYG----------PRGSPYSSYGRGRG----------QNNYNNPG-------SRGRGGRGG-----
S--------GTQNYSGRQDGI----------VCYNKSMTDDPWRDLQPIV----------GNIMIPRDG--S-KSWLPESLRAKKDT-
SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|A0A0E0CKZ6|A0A0E0CKZ6_9ORYZ
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHLPPPAADAWDLPVRDL-MDASTTTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPAPIHTPSPL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP--------PPWNPH---
SAPPSQGPYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG--------
-SR-GRGG--RDGS--------GTQNYSGWQDGR----------VR------YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------

>tr|A0A0D3F5J3|A0A0D3F5J3_9ORYZ (Oryza barthii)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHPPPPAADAWDLPVRDL-MDASATTSAAARRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GLGNY----------GSGYP----------PPHQHHMAPPPIHTPSPL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPWNPH---
SAPPSQGPYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG-------
SRGRGGRDG-GRDGS--------GTQNYSGWQDGR----------VR------YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|A0A0E0NFD1|A0A0E0NFD1_ORYRU (Oryza rufipogon)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHPPPPAADAWDLPVRDL-MDASATTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPPPIHTPSRL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPWNPH---
SAPPSQGLYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG--------
-SR-GRGG--RDGS--------GTQNYSGWQDGR----------VR----YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|Q6K2E7|Q6K2E7_ORYSJ (Oryza sativa subsp. japonica)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHPPPPAADAWDLPVRDL-MDASATTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPPPIHTPSRL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPWNPH---
SAPPSQGLYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG--------
-SR-GRGG--RDGS--------GTQNYSGWQDGR----------VR----YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|A2X594|A2X594_ORYSI (Oryza sativa subsp. indica)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSVAPHPPPPAADAWDLPVRDL-MDASATTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPPPIHTPSPL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPRNPH---
SAPPSQGPYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG--------
-SR-GRGG--RDGS--------GTQNYSGWQDGR----------VR----YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|A0A0D9YSQ7|A0A0D9YSQ7_9ORYZ (Oryza glumipatula)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHPPPPAADAWDLPVRDL-MDASATTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPPPIHTPSPL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPWNPH---
SAPPSQGPYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG-------
SRGR-GGRDGGRDGS--------GTQNYSGWQDGR----------VR--YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
RSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>tr|A0A0E0IRA0|A0A0E0IRA0_ORYNI (Oryza nivara)
MEGGEQ--EESSAQRRERLLALRSS---AAAAPSPSAAPHPPPPAADAWDLPVRDL-MDASATTTAAAPRPPPR--FDYYTNPAA-
AFASSSAASH----------KRKVAEPPPP----------GSGNY----------GSGYP----------PPHQHHMAPPPIHIPSPL-
SHDSPGGSP-WRSPMQFQAPMSGY--RGPP----------PGAPPPW-SPH----------SGVP---------PPWNPH---
SAPPSQGPYPHPPSYGPRNYNPGQGGGRM----NYG----------PRGRPDSPYGRGRG----------QNNYNNPG-------
SRGR-GGRDGGRDGS--------GTQNYSGWQDGR----------VR--YHKSMTDDPWRDLQPIV----------GNIMIPRDG--S-
KSWLPESLRAKKDT-SDRGQ-VKPPS-GLSLAEYLDLSFNEASNDT--------
>XP_009404795 (Musa acuminate)
--------MEESTKRRERLRAMSLQ-A-SRAMVTTLEPSYSLLTVPPPL-LLLPQISNSPIQPPPALECQPSISNSDFFITDPLS-
AFTSSKRMRGVGAIGFNSSSHGVGTIGFTPSTPYHPAP----------SS-FPTVI----------AGR----------GNNQFHMSQPPDCISYKM-
PSCGLHSSP-WRSPVRLLTPFSGH--QGTP----------ASGRGSW-NR----------SGGI----------
RGFPTNSPISVL--------S-----GPH-----FG----------PWGSP----------NSIAQSSGTPQ------SISVR-
GSSLQISSGS--------SLKPATC---GR-G-QSSN----------D---KASVQRKPKRYYCKSMLEDPWRNLEPIV----------GNILEPMAG--P-
GYWLPESISGTKRKVSETQSINKLNS-KWSLAEVLADSFEEAI----------
>XP_009388537 (Musa acuminate)
--------MEDSAKRRERLQAMRTEAS-RAAASPYSSPASAPLP------LPLPLLSEPPVPPPGGLESPPPPAPRFEFYSDPMS-
TFSAAKRMRG----------VGTGAAGFSPSSPYLPHPPT----PPPPPPQ-STFPS----------GIRNTYMDSSYA--------
SVQQFHMSQPPDSSSYGT-PSHVPYNSP-WRSPVRSPTPFSGY--RGSP----------ARGPGSW-NRS----------
SG----------SGFPTYSSNSVP--------R-----SPH-----FG----------PVESP----------RSIAFSNSPQ------
SITGR-GRGR----QF--------CGGPSPRPSGGR-G-RSFS----------G---KASAREDIERYYLKSMIENPWHDLEPVV----------
GNILEPMAG--Q-GYWLPESISGKKSEVSETERRNQFSS-KSSLAEFLAASLEEAVNDD--------
>Glyma.02G215100 (Glycine max)
--------MEDSEQRKKRLKQMRVQ-A-DQAEVSGGREGSVVP----GF-LSNPLI-EAPSTMPSRDTSYAAPR--FDYYTDPMS-
AFSSKRNNAS----------TQAAPDNFPPSKFGG---PPMAQY-SSPHP----------ESKNPQMTP----------
HPIQASPAAYRNPVWSGP-GGPAHYNFP-LHPSSGGTYPSPRFEPSGGP----------LY-NTA----------QGIA--
----HQPSYSPN---PPYPGYVNSPRP--------SYSPNPSPG-----YS----------NCPMPSYSP----------NPSPGYRNSPS-
-----PGQGR-GRGF--WRN---------TGSPVSGWGSGQ-G-PNFH----------GHRSNENTVHGPDRFYKRSMVEDPWEHLEPIIW-------
KANDGY-LNTSRVPLN--S-QPWI-SKASSTKGEGSSAAS-VKSSS-EPSLAEYLASAFNEAANDAENV-----
>Glyma.14G182400 (Glycine max)

--------MEDSEQRKKRLKEMRVQ-A-DQAEVSGGVEGSVVP----GF-LSNPLM-EAPSTMPSQDKSYAAPR--FDYYTDPMS-
AFSSNRRNIA----------SIQAALDNFPPSKFGG---SPMAQY-SSPHPGT----------ESTNPQRTP----------HPIQASPAAYRKPIWSGP-
GGPAHYNFP-IHPSSGGTYPSPRFEPSGGP----------LY-NSGQGI----------AHP--------PSYSPN---
PPYPGYGDSPRP--------SYNPNPSPG-----YG----------NSPRPSYNPNPSPG----------YGNSPRPSYRPNPSPGYRNSPS------
PGQGR-GRGF--WR-N--------TGSPVSGRGSGQ-G-PNFH----------GHRSNENTARGPDRFYNRSMVEDPWEHLEPIIW-------
KANDGY-LNTSRIPLN--S-QPWI-SKATSTKGEGSSAAS-VKSSS-EPSLAEYLASAFNEAANDAENV-----
>tr|A0A0B2RZR4|A0A0B2RZR4_GLYSO (Glycine soja)
--------MEDSEQRKKRLKEMRVQ-A-DQAEVSGGVEGSVVP----GF-LSNPLM-EAPSTMPSQDKSYAAPR--FDYYTDPMS-
AFSSNRRNIA----------SIQAAPDNFPPSKFGG---SPMAQY-SSPHP----------ESTNPQRTP----------HPIQASPAAYRKPVWSGP-
GGPAHYNFP-IHPSSGGTYPSPRFEPSGGP----------LY-NSGQG----------IAHP--------PSYSPN---
PPYPGYGDSPRP--------SYNPNPSPG-----YG----------NSPRPSYRP----------NPSPGYRNSPS------PGQGR-
GRGF--WRN---------TGSPVSGRGSGQ-G-PNFH----------GHRSNENTARGPDRFYNRSMVEDPWEHLEPIIW-------KANDGY-
LNTSRIPLN--S-QPWI-SKATSTKGEGSSAAS-VKSSS-EPSLAEYLASAFNEAANDAENV-----
>tr|A0A0S3RMQ5|A0A0S3RMQ5_PHAAN (Vigna angularis var. angularis)
--------MEDSEQRKKRLREMRMQ-A-DHADVSGAVEGSAMPGS-----LSNPLI-EAPLIMPSWDKSNAAPR--FDFYTDPMS-
AFSSNKRNNT----------SIQAAPDSFPPNVG----PSPMAQY-SSPHP----------ESTNPRMTL----------PPFQVSAPAYGNLDFSGP-
RGPAHDNFL-FHPSSVGRHPYPRFEPSGGP----------LYNSTQGITHRASYSPNPSPGYRNSP----------
GPSYSPNPSPGYSNSHRPSYSP----NPSPGYSNSPRP--------SYSSNPSPG-----YS----------NSHRPSYSPNPSPG----------
YSNSPRPSYSPNPSPGYRSNPR------PGQGR-GRGV--WHN---------PGSSVSGRGSGR-G-PNFH----------
GHLSNENPGPGPNRFYKSSMVEDPWNHLEPKIW-------EAIDGS-LHSSRIPEK--V-KPWI-SKSKSTIGEGSSAAS-VKFRS-
EPSLAEYLATAFNEAANDAENV-----
>tr|V7B5T5|V7B5T5_PHAVU (Phaseolus vulgaris)
--------MEDSEQRKKRLREMRTQ-A-DHAEVSGGVEGFAMPGS-----LSNPLI-ETPLSMPSGDKSYAAPR--FDFYTDPMS-
AFSSNKRNNT----------SIQAAPDSFPPNVG----GSLMAQY-SSPHP----------ESTNPQMTL----------PPFQVSAAAYRNLDLSGP-
RGPAHDNFL-FHPPSGGRYPNPRFEPSGGPLYNSALQGIAHQPSYSINPSPGYRNSPRPSYSPNPSPGDW-NS----------PRP--
------SYSPN----PSAGDWD-SPRP--------SYSPNPSPG-----YR----------NSPRPSYGPNPSPG----------
YRNSPRPSYSSNPSPGYRSSPRPCYSPSPGQAR-GRGI--WHN---------PRSSVSEQGSGR-G-PNFH----------
GRWSNENAGPGPNRFYKSSMVEDPWNRLEPKIW-------EAITGS-LHTSHIPEK--V-KPWI-SKSKSTIGEGSSAAS-VKSSS-
EPSLAEYLAAAFNEAANDAENV-----
>gi|502095430|ref|XP_004490453.1| (Cicer arietinum)
--------MEDSEQRKKRLKEMRLE-A-DLAEDSGGGGGAAEGSGIPSL-LSNPLAEEAPSNMLSQNKSYAAPR--FNYYTDPMN-
AFSSDKRSSV----------NVHPRPDYLPPMPYF----------SSPHL----------ESTNRQMPP----------SPTQALPTSYRNPVWNGP-
RGPPQYNFP-FRPSGGGIYPSPRF--ESPS----------CPSF-NSA----------PVMNHW----------
----------PNQNP----------NPAPAYMNNPH------FSQGR-GRGF--WNNN--------TRSPDSGRGRGR-
G-SSSH----------GRWSNEDRGHVPNRFYKMSMVEDPWKFLKPIIW-------NSTYSSYSNTAYTPEN--S-
KPWIPSKSTSTKQEESPAVF-SKVSS-GPSLAEYLASALNEASSTEENV-----
>Medtr5g066450 (Medicago truncatula)
--------MEDSEQRKKRLKEMRLQ-A-DRVQDSGAAAGDGSGT---PSVLANPLA-EAPSAVLSTD---SAPR--FNYYTDPMN-
AFSDKRGAVD----------VRPASEYLPPPPPLP--LNFVPEF-SSQRS----------ATTNPQMSP----------SPTQALPTPYRNPVWNGP-
RGPPQYNFP-FRPSGGGTYPSPRF--ESPG----------SPY-NSA----------PGMNQWP-NHNP-
-------NISSEYIPI----------PTSEYSP----------NQSPAFRNNPN------SSRGR-GRGF--WH-N--------
NRSHVSGRGSGQ-G-SSSH----------GRWSNEDRGCGPERFYKRSMVEDPWISLKPIIW-------YSTISS-SNTSYAREN---------
SKSTGTKREGPSAVF-SKSNS-GPSLAEYLASAFNEASNTE--------
>Lus10004850.g (Linum usitatissimum)
-------MEENSESRSDRLKAMRAV-A-ADAA----SSGLPAIP---GV-LANPML-DGPSEAETRRGG-GVPR--FDFYTDPMA-AFSSTRNRNP--
----------PLHNTPGNFIPPAGRL---PPPPAW-IPPSPSPGYCLKIPDLVTGPWNPTPRS----------TPPHMQGDY-PVN------
QGMYQGNIP-YGSQGSIRSPFPFN--QRAP----------HPQPW-SEF----------ADP--------AANY-----
GSHPRGGMPGVS--------PRYHQGNPN-----YA----------PVRSPS----------SGYGRHQF------PSSGR-G--------
---------NRSSVFGRGNGR-G-RGFH----------NHGSTSSEKQDPWKFYHDSMVKDPWEGLEPIEW-------RGITTG----------
SSDSWMPSSVTTKRPRGGPVSSFNQSNS-GQSLADYLASSFNEASNAEQNT-----
>Lus10020639.g (Linum usitatissimum)
-------MEENSESRSDRLKAMRAV-A-ADAA----SSGLPAIP---GV-LANPML-DGPSETETRGGG-GDPR--FDFYTDPMA-AFSSTRNRNP-
----------PLHNTPSNFIPPAGRL---PPPPAW-NPPSPSP----------GPWNPAPRS----------TPPHMQGHYHPVN------QGMYQGNRP-
YGSQGGIRSPFPFN--QRAP----------HPQRW-SEF----------ANP--------ASNY-----GSHPRGGMPGVG--------
PRYHQGNPN-----YA----------PARSPSVNH----------HPNSGYGRHQF------PSSGR-G----------
NRSSVCGRGSGR-G-RSFH----------NHGSTSSEKQDPWKFYHESMVEEPWEGLEPIEW-------RGITAG-SP----------
DSWMPSSISMKRPRGGPMSSFNRSSNSGQSLADYLASSFNEASNAEQNT-----
>gi|729292597|ref|XP_010530288.1| (Tarenaya hassleriana)
--------MEDPAKRKERLRAMQTE-A-AQINASAETSMNPSQ-------LSNPFA-ETSGPLPSQEESYQTQR--FDYYTDPLS-AFSSNKRNKT--
----------HQRIEQHYSSSP---------VSQF-PPSIP----------GPVG----------NTYQPHNSHGGFQA-----EGQYHSGSP-
HIDPRGMTNPSFVY--RGTP----------QESW-NNF----------RHP-------APNFRP----FHSPHP--VPGT--------
FLNHQAIPD-PANRFG----------GQGNYNRPPIPGYA----------RPFSNYG-RQNSNWAGRMG------SGSGR-GRGD---GGR---
-----NMNSSFGRNGGR-G-QKYD----------D---GGYRAQGPERFYDKSMSEDPWKNLEPVLW-------KSCP---FSCTSSSTG----
QAWLPMSIVRKKPTVAETSS-NKSSSAHQSLAEYLAASLDEATRDDASQ-----
>XP_010530288 (Solanum pennellii)

--------MEDPAKRKERLRAMQTE-A-AQINASAETSMNPSQ-------LSNPFA-ETSGPLPSQEESYQTQR--FDYYTDPLS-AFSSNKRNKT--
-------------HQRIEQHYSSSP---------VSQF-PPSIP------------GPVG---------------NTYQPHNSHGGFQA-----EGQYHSGSP-
HIDPRGMTNPSFVY--RGTP-----------------------------QESW-NNF------------------RHP-------APNFRP----FHSPHP--VPGT--------
FLNHQAIPD-PANRFG-------------GQGNYNRPPIPGYA-----------------------RPFSNYG-RQNSNWAGRMG------SGSGR-GRGD---GGR---
-----NMNSSFGRNGGR-G-QKYD-----------D---GGYRAQGPERFYDKSMSEDPWKNLEPVLW-------KSCP---FSCTSSSTG----
QAWLPMSIVRKKPTVAETSS-NKSSSAHQSLAEYLAASLDEATRDDAS------
>Brara.K01498 (Brassica rapa)
--------MENSEKRKQMLKAMRMEAA-AQNDGSTEPETSVNM----GH-LSNPLA-ETST---QQQESCEAPR--FDYYTDPMS-
AYSSFKRNKT-------------------SKQQDISSPGYQM--SPPVPHF-PQSVQ-------------GSLG----------------SDYQAHANHGGFQ------
EAHCGGDNL-HAEPRGVA---PSH--RAPP-----------------------------------VPWINNY-----------------RPPP--------PPFNH----LGPPQW--
VPRP--------FPFFQGNPDMGNNRFG-------------GRGGSYNNTAPQFP---------------------------QYGPRQTSNWAGNTY------PSSAR-
GRSR---------------GRNTSFGRGGGRRH-----------------------MEQGAERFYSNSMAEDPWKYLKPVLW-------KSCSVT-SSSNSTG------
QAWRPNSIAPKKPVIIEA-S-HKPSNNQQSLAEYLAASLDEATSDDPSN-----
>tr|A0A0D3DFV7|A0A0D3DFV7_BRAOL (Brassica oleracea var. oleracea)
--------MEDSEKRKQMLKAMRME-A-AAAAAQNDGSTEPETSMNRGH-LSNPLA-ETST---HQQESCEAPR--FDYYTDPMS-
AYSSFKRNKT-------------------SKQEHISCPPSYQ--MMPLAPEFPPSVP-------------GSLG----------------SDYQAHANHGGFQ------
EAHCGGDDL-HTEPRGVALSH-----RGSP-----------------------------------VPW-NDN-----------------FRPP-------PPRFNH----LGPPQW--
VPRS--------FPFSQGNHDMGDNRFG-------------GRGGSYNNTAPQFP---------------------------QYGPRQASNWAGNTY------PSSAR-
GRSR---------------GRNTSFGRGGGR----------------------RHMEQGADRFYSNSMAEDPWKYLKPVLW-------KSCSVT--SSSNS-TG----
QAWRPNSIAPKKPMIIE-AS-HKPSNNQQSLAEYLAASLDEATSDDPSN-----
>Brara.A01455 (Brassica rapa)
--------MEDSEKRKEMLKAMRMEAA-AAAAAENDGSTELGTSMNTSH-LSNPLA-EASI---HQQESYDKPR--FDYYTDPMA-
AYSSFKRNKS-------------------PKQQYISSPSHQM--SPPVPQF-PPSAP-------------GLVG----------------NDYQAHPNHGGFQ------
EAHYGGDNL-HTEPRGMAPSY-----RGPP-----------------------------------APWNNNN-----------------FRPP---------PPVNH----LGPPQW--
VPRP--------YPFPQGNHDMGDNRYG-------------GRGPRGGNYNNNPP-----------------------QFPHYG-RQNSNWAGNTY------PNSGR-
GRGR---GGR--------GMNTSYGRGGGRRP-----------------------MEQGAERFYSNSMAEDPWKYLQPVLW-------KSCSDASSSNSTG------
-QAWRPNSIAPKKPMISEASH-NPSSN-QQSLAEYLAASLDEATCDESSN-----
>tr|A0A078JV03|A0A078JV03_BRANA (Brassica napus)
--------MEDSEKRKEMLKAMRME-A-AAAAAQNEPETSMNASH-----LSNPLA-ETST---HQQESYDKPR--FDYYTDPMA-AYSSFKRNKS-
-----------------PKQQYISSPSHQM--SPPAPQF-PPSAP-------------GLVG----------------NDYQAHPNHGGFQ------EAHYGGDNL-
HTEPRGMAPSY-----RGPP-----------------------------------APW-NNN-----------------FRPP--------PPVNH----LGPPQW--VPRP--------
YPFPQGNHDMGDNRYG-------------GRGPRGGNYNNNPP-----------------------QFPHYG-RQNSNWAGNTY------PNSGR-GRGR---
GGR--------GMNTSYGRGGGRRP-----------------------MEHGAERFYSNSMAEDPWKYLQPVLW-------KSCSDASSSNSTG-------
QAWRPNSIAPKKPMISEASH-KPSTN-QQSLAEYLAASLDEATCDESSN-----
>tr|V4MEG9|V4MEG9_EUTSA (Eutrema salsugineum)
--------MEDSEKRKQMLKAMRMEAA-AQDDGSTGPETSMNTSH-----LSNPLA-ETST---HQQESYETPR--FDYYTDPMS-
AYSSFKRNKT-------------------PKQQHISSPSCQI--SPPVPPF-PPSVP-------------GSLG----------------CDYQAHANHAGFQ------
GPQYEGDNL-HTVPRGMA---PSH--RGSP-----------------------------------VAW-NNN-----------------FRPP--------PVNH-----LGPPQW--
VPRP--------FPFSQESPDMGNNRFN-------------GRGSYNYTA-----------------------------PQYPNYG-RQNSNWVGNTY------PSSGR-
GRGR---GGR--------GINTSFGRDGGR-----------------------RPMELGAERFYSNSMAEDPWKYLKPVLW-------KSCSD--GASSNSTGQ----
----ANSVAPKKPMISEASH-TPSKN-QQSLAEYLAASLDEATCDEPSN-----
>gi|727552180|ref|XP_010448556.1| (Camelina sativa)
--------MEDSEKRKQMLKAMRMEAA-AQNDDDSIGLETSVNT---GH-LSNPLA-ETST---HQQESFEKPR--FDYYTDPMS-
AYSSFNRNKT-------------------PKQQYISSPHHHQTSSPVPPQF-PPSVP-------------GSLG----------------SEYQAHTNHGGFQ------
AAHYDGHNL-HTGPRGMAHLSPSH--RGSP-----------------------------------AAW-NNN-----------------YRPP--------PVNH-----LGPPQW--
VPRP--------LPFSPEYPDMGNNRFG-------------GRGNYNNTA-----------------------------PQYSHYG-RQNSNWAGNTY------PNSGR-
GRSR----GR--------GMNTSFGRDGGR-----------------------RPMELGAERFYSNSMTEDPWKYLKPVLW-------KSCSDSSSSNSTTS------
QAWLPNSIAPKKSVTSEASH-KASNN-QQSLAEYLAASLDEATCDESSN-----
>Bostr.7867s0257 (Boechera stricta)
--------MEDSEKRKQMLKAIRMEAA-AQNDDSTGLETSLNT----GH-LSNPLA-ETST---HQQESFETPR--FDYYTDPMS-AYSSFKRNKT--
-----------------PKQQYISSPNHQT--SSPVPTQ-FPPSVP------------GSLG----------------SEYRAHTNPGGFQ------AAHYDGNNL-
HAEPRGMAHLSPSN--RGSP-----------------------------------AAW-NNN-----------------YKPP--------PVNH-----LGPPQW--VPRP--------
SPFSPEIPDMGNNRFS-------------GRGSYNNTA-----------------------------PQFSNYG-RQNSNWVGNTY------PNSGR-GRSR----GR--------
GMNTSFGRDGGR-----------------------RPMELGAERFYSNSMAEDPWKYLKPVLW-------KSYSDASSSNSTS-------
QAWLPNSIAPKKSATSEASH-KASNN-QQSLAEYLAASLDEATCDESSH-----
>Cagra.2171s0011 (Conradina grandiflora)
--------MEDSEKRKQMLKAMRMEAA-AQNDDGSTGLETSGDTAH----LSNPLA-ETST---HQQDSFETPR--FDYYTDPMS-
AYSSFKRNKT-------------------PKQQYVSSPHHQTNSPVPPPQF-PPSVP-------------GSLG----------------SEYQVHTNHGGFQ------
ASHYDGNNL-HTGPRGMGHLSPSH--RGSP-----------------------------------AAW-NYN-----------------YRPP--------PVNH-----LGPPQW--
VPRP--------FPFSQEIPDMGNDRFG-------------GRGSYNNTA-----------------------------PQFSHYG-RQNSNWVGNTY------PNSGR-
GRSR----GR--------GMNTSFGRDGGR-----------------------RPMELGAERFYSNSMAEDPWKYLKPVLW-------KGCSDA-CSSNSTS------
QSWLPNSIAPKKSVNSEASH-KDSNN-QQSLAEYLAASLDEATGDESSN-----
>Carubv10006986m.g (Capsella rubella)
--------MEDSEKRKQMLKAMRMEAA-AQNDDGSTGLETSGDTAH----LSNPLA-ETST---HQQDSCETPR--FDYYTDPMS-
AYSSFKRNKT-------------------PKQQYVSSP---HHQTSSPVP------PPQFPPSVP--

EIPDMGNNRFGG---RGSY----------NNT----------A----------PQFSHYG----------
----------RQNSNWVGNTY------PNSGR-GRSR----GR--------GMNTSFGRDGGR----------
RPMELGAERFYSNSMAEDPWKYLKPVLW-------KGCSDA-CSSNSTS------QSWLPNSIAPKKSVNSEA-S-
HKASNNQQSLAEYLAASLDEATGDESSN-----
>AT4G24500 (Arabidopsis thaliana)
--------MEDSEKRKQMLKAMRME---AAAQNDDDATTGTETSMSTGH-LSNPLA-ETSN---HQQDSFETQR--FDYYTDPMA-
AYSSFKKNKT----------PKQQYISSPSHQGSSPVPPQFP-PSVPP----------GSLC----------SEYQAQTNHGGFHAA----------
--HYEPRGMAHLSPSH--RGPP----------AGW-NNN----------FRPP--------PVNH-----SGPPQW--VPRP--------
FPFSQEMPNMGNNRFG----------GRGSYNNTP----------PQFSNYG-RQNANWGGNTY------PNSGR-GRSR----GR------
--GMNTSFGRDGGR----------RPMEPGAERFYSNSMAEDPWKHLKPVLW-------KNCSDA-SSSSS-TG-----
QAWLPKSIAPKKSVTSE-AT-HKTSSNQQSLAEYLAASLDGATCDESSN-----
>AL329212 (Arabidopsis lyrata)
--------MEDSEKRKQMLKAMRME-A-AAQNDNDDSTTDPETSMNTGH-LSNPLA-ETST---QHQDSFETSR--FDYYTDPMS-
AYSSFKKIKT----------PKQQYISSPSHQA-SSPVPPQFPPSVPP----------GSLG----------SEYQAHTNHGGFQAA----------
HYEPRGMSHLSPPY--RGSP----------ASW-NNN----------FRPP--------PVNH-----PGPPQW--VPRP--------
FPFSQEIPNMGNNRFG----------DRGSYNNTA----------PHFSNYG-RQNANWVGNTY------PNSGR-GGGR----GR--------
GMNTSFGRDGGR----------RPTELGAERYYSNSMADDPWKYLKPVIW-------KSCSDA-SSSNS-TG-----
QAWLPNSTAPKKSVTSE-AT-HKPSNNQQSLAEYLAASLDEATCDESSS-----
>gi|747073497|ref|XP_011083710.1| (Sesamum indicum)
--------MEESEKRRERLKAMRIE-A-ARAGVNTDSESSGLPS---RG-LSNPLI-ESETTSASPLQ-YTSPR--FDYYTDPLA-AFSGSRRNNT--
----------SHQVSQGHYNTPS----------RPVYHGMTPT----------PVYQNAPPNPPGQIMFHP-PGAHLSHGP-
FGSPMEANSPLSRPHFTHGPL----------RSQMEANSSLSRPQGNP-PSAW-GGS----------GAA-------LSYSSP----
PNLYRGGNFTRP--------DF----------GGDSPYVKYGRGRG----------QRYSSSSH------YDSGR-GRGQ--
STAN--------CMSPGSGHSSSR-G-RGSR----------D---SVSAELRPDLYYSKGMMEDPWKNMTPVVW-------KGVPAS-DSD----------
KSWLPKSVQMKKAKVSSEAS-QPSIS-QQSLAEYLAASFNDSNDEAVDEERGT-
>gi|848854091|ref|XP_012846656.1| (Erythranthe guttata)
--------MEESEKRRERLNAMRTE-A-AQTRETTHPHNSLGG-------LSNPLI-ESESASTAPLH-HTSQR--FDFYTDPMA-AFSGSKRRNN---
----------ISPQFPQTRYNTPPQ----------RPLYEGMLPPP---------PVYQNTHVNAPGQRILPV-PEPYFNQGS---
PPHFNQGPPPHFN-QGPP----------PHFNQGP----------HFNQSP-------RPNFYQ----SPQPHSNQSPRP--------
HFNQSPRPNFNQNHVN----------YATPPNFPRGGNNFT----------SPSFRQGDSSH------TNYGR-GRGE--
YNSNSHNN----SGRGGRGRSQGNRG-AGSH----------N---SVSADQRPDLYYRKEMVEDPWKMLTPVIW-------KGVDAL-GSD----------
QSWLPKSIAVKRAKVSPEAA-KTSIS-QQSLAEYLAASFNDSTDEAVDEQPDL-
>Migut.C00706 (Mimulus guttatus)
--------MEESEKRRERLNAMRTE-A-AQTRETTHPHNSLGG-------LSNPLI-ESESASTAPLH-HTSQR--FDFYTDPMA-AFSGSKRRNN---
----------ISPQFPQTRYNTPPQ----------RPLYEGMLPPP---------PVYQNTHVNAPGQRILPV-PEPYFNQGS---
PPHFNQGPPPHFN-QGPP----------PHFNQGP----------HFNQSP-------RPNFYQ----SPQPHSNQSPRP--------
HFNQSPRPNFNQNHVN----------YATPPNFPRGGNNFT----------SPSFRQGDSSH------TNYGR-GRGE--
YNSNSHNN----SGRGGRGRSQGNRG-AGSH----------N---SVSADQRPDLYYRKEMVEDPWKMLTPVIW-------KGVDAL-GSD----------
QSWLPKSIAVKRAKVSPEAA-KTSIS-QQSLAEYLAASFNDSTDEAVDEQPDL-
>evm_27.TU.AmTr_v1.0_scaffold00033.225 (Amborella trichopoda)
-------MDESSEKRRERLKAMRLE-A-EQTPNLPPQTFSSSSS---SSSLSNPFN-DLSATLCASTQ---ASR--FDFYTNPVS-
AFSSMNRKGE----------RYNDSNQHSLSAQSP----------MAPTP----------GQSNAAMTHA----------TAYPMQISYN-DQRQYET---
PPHRSGS-WGGPMMTGSPFS----HGSPI----------PGSW-TNNNARPV----------HTSP--------
PNFHGGIRTPNFGRGS-SPNP--------NFGRGSSPS---SNYG----------SRSSPHPNYGRGSS----------PSPSYGRGSSP-
SPNYGRGSSPN------FSSGM-GRGR--HFSW--------SPGHSSGRGGGR-G-RGYH----------L---
DVSAKEHPERFYNKSMVEDPWSSLIAVTKRLGIAGDRSVVSD-SNKSGTPDS--L-RSWLPKSIS-KKARISDI-K-DEFNS-
ETSLADSLVLAFEDAANDRTDV-----
>tr|A0A0A0LQW8|A0A0A0LQW8_CUCSA (Cucumis sativus)
--------MEESEKRRERLRAMRME-A-AQADVANYVETSLPNH------LSNPLV-ESSATMVGQLAPCTAPR--FDYYTNPMA-
AFSTSKKKGK----------IENQPVSDNFV----------PYHHNTSSTTYFPPTFPGDSEAGGHGRP------
GMPRPYAVN--QGDL----------HMW-RG----------PRGPFVNQFPTQ--------PPREMNSPS---
HVSG----------PRGNPYTNPTQNRA----------NYRSSSPNPGFRGSFS------PGRGSYGHHG----------
NMTPSPRFGYGR-A-TGSH----------G-RHSSSDKSGPEQFYNISMLEDPWKVLQPCIW-------TTIAPL-SNSAKPS------EYWI-
SKFGTKKARVSDSSS-SRSSSQQPSLAEYLAASFKEAIEEAPNA-----
>XP_008442005
--------MEESEKRRERLRAMRME-A-AQADVANYVETSLPNH------LSNPLV-ESSATMMGQLAPCTTPR--FDYYTNPMA-
AFSTSKKKGK----------IENQLVSDNFVPYHHNTS---------SPTFP----------GLRNPEMSS----------ASTHQFHQCSPDRRMFYA-
RGDSEAGG--HGSP-GMPRPYAVD--QGDP----------HMW-RG----------SKRPFVNQYPTH--
------PPREMNSPS----HVSR----------PRGNSYTNPTQDRA----------NYRSSSPNPGFLGSFS------PGRGSHGHHG----------
----NMTPSPRFGYGR-G-TGSH----------GRHSSLDKSPGPEQFYNVSMLEDPWKVLQPCIW-------TTIAPS-SNSTEPS------
ESWISTKFGTKKARVSDSSS-GRSNS-QPSLAEYLAASFKEAIED---------
>Eucgr.H00032 (Eucalyptus grandis)
--------MEESEKRRERLRAMRAE-A-ARAGLSDSPEASAEAS---C--LANPLI-ETRTPESLHEMPRASPR--FDYYTDPMA-
AFSGNRARGN----------AGFQPPQANFTPP----------PPHFP----------GPRNLVTSPS----------PAYQWRPRNPPNSGIYQA-
EGPRPRFPP-HHSPRGNVHHFPMQ--DGAS----------PQAW-NR----------SSSTTS-------

-YFSPSSTPGSGTHNF----------NQMQDRSHWRGNSP----------GPNLGQGGNYA------TSPGR-GRGRGQWF--------
--GGSPRSGRSGGG-G-PGHR----------GHGSTLARPLGPERFFHESMLEDPWKDLKPVLW-------RGVDSP-FVGSSSRGS--S-
SPWMPESIRPKKPRVAES-P-NKFGS-GLNLAEFLAASFNETVKETSSP-----
>Eucgr.L02666 (Eucalyptus grandis)
--------MEESEKRRERLRAMRAE-A-ARAGLSDSPEASAEAS---C--LANPLI-ETRTPESQHEMSRASPR--FDYYTDPMA-
AFSGNRARGN----------AGFQPPQANFTPP----------PPHFP----------GPRNLVTSPS----------PAYQWRPRNPPNSGIYQA-
EGPRPRFPP-HHSPRGNVHHFPMQ--DGAS----------PQAW-NR----------SSSTTS--------
-YFSPSSTPG----------SGTHNFNQMQDRS----------HWRGNSP-GPNLGQGGNYA------TSPGR-GRGRGQWF--------
--GGSPRSGRSGGG-G-PGHR----------GHGSTLARPLGPERFFHESMLEDPWKDLKPVLW-------RGVDSP-FVGSSSRGS--S-
SPWMPESIRPKKPRVAES-P-NKFGS-GLNLAEFLAASFNETVKETSSP-----
>evm.TU.contig_29936.2 (Carica papaya)
--------MEESEKRKERLKELRVD-A-AQAEVSK-NGQISAIP---GL-LPNPLI-DTSTT--MQAQSSPAPR--FDFYTDPMS-AFSANRKKSK-----
----------VENQFQQNCFSSPGYNS--GSPSAQI-LSSIS----------GPRSDELTPS----------RPQQMQREEGMYQG-----QGIY-GGTP-
LMGPRGVSSSSPLPH-QRTS----------PVPW-NGS----------GSS---------TGYYTHVQQNYPRG--TSSP--------
FPLHQETPESWKRAVGRGNRNFPSNLVHVPPCSPVFHRGQGRA----------HWFGRSQ-SPGSGRGASPG------GGSGR-
GRSS--WYGG--------SASHGSVRSGRR-L-QSPH----------SRNLSEGDMLGPECFYDKSMEEDPWLNLKPVIW-------RGLDIP-
ESCSNTLGS--S-NSWLPKSISTNKPRVAES-S-KKFNS-QQSLAEYLAASFNEAVKDLPNE-----
>PGSC0003DMG400009339 (Solanum tuberosum)
--------MEESEKRKERLKAIRKE-A-AEAGDNNEEQNSIGGPLDHG--LTNPLI-ETPSAASGKDEPR--PR--FDYYTDPMA-
AFSANNKMNN----------LSPQVSQPCNTSP----------RPMNDGS----------PVYHAQCNYNSAQRTYRP-RGV--
NAIP--LGIRRKTNPFCTP--PGNS----------TL-DSS----------LGTP--------NNYSL----PNSPQIG----------
GVSSHGSP----------QVSGAGSQYGQ----------GSPY----------Q-GSGF----------RSKAYQGSRGGK-
G----------RFKFYYHKSMVEDPWKALMPVIW-------KPRGDT-QDCL---------KSCLPNSISAKRAKLGET-P-TKSTP-
QKSLAEYLAAAFNEAAGEDTVNNQSNT
>Solyc04g076410.2 (Solanum lycopersicum)
--------MEESEKRKERLKAIREQ-A-AEAGDNNEEQNSIGGPLDHG--LTNPLI-ETPSASSRKDEPR--PR--FDYYTDPMA-
AFSANNKMNN----------LSPQVSQPCNTPP----------RPMNAGS----------FAYHAQGNYNSAQRTYWP-RGV--
NAIP--LGIRRNTNPFCMP--QGDS----------TL-GSS----------LGTP--------NNYSL----PNSPQIG----------
GISGHGSP----------QVSGAGSQYGQ----------GSPY----------Q-GSGF----------RSKAYQGSRGGK-
G----------RFKFYYNESMMEDPWKALKPVIW-------KPRGDT-QDCL---------KSRLPNSISAKRAKLGET-P-TKSTP-
QKCLAEYLAAAFNEAAGEDCE------
>XP_015074364 ((Solanum pennellii)
--------MEESEKRKERLKAIREQ-A-AEAGDNNEEQNSIGGPLDHG--LTNPLI-ETPSASSRKDEPR--PR--FDYYTDPMA-
AFSANNKMNN----------LSPQVSQPCNTPP----------RPMNAGS----------LAYHAQGNYNSAQRTYRP-RGV--
NAIP--LGIRRNTNPFCTP--QGNS----------TL-GSS----------LGTP--------NNYSL----PNSPQIG----------
GISSHGSP----------QVSGGCSQYGQ----------GSPY----------Q-GSGF----------RSKAYQGSRGGK-
G----------RFKFYYNESMVEDPWKALKPVIW-------KPRGDT-QDCL---------KSRLPNSISAKRAKLGET-P-TKSTP-
QKSLAEYLASAFNEAAGEDTVN-----
>PGSC0003DMG400002050 (Solanum tuberosum)
--------MDESEKRKERLKAMRME-A-SQSGDYNEAVGHGG--------LTNPLT-DVPS---GNVESYAMPRPRFDYYTDPMA-
AFSANKRSNN----------QPHVSPQISQQCYTPP----------RATNPQS----------PICTPRGNYSVDQRS----QGVH--
YNP-LGNP-GQNSPFGTPQ-RGSP----------SAW-NNS----------FGTP--------NNYLP----PNSSMGGNFASP----
----GIHQGGRPG---FHYGQ----------GSGQPGSGYG-GSPYQGSGYRGNPYQDSGHRGSPYQHSGNRGSP-YQRPGNRGIPY------
QGSGQ-GRSQ--WRGN--------SSSPISFRGGRRGG-RGSH----------G---GTSGESRPDLYYSKSMVEDPWKELKPVIW-------KAFSD----
----------KSWLPHSISAKKAKFPDA-P-VKSIP-QQSLAECLAASFNEAASSEAATDGSGT
>Solyc12g087990.1 (Solanum lycopersicum)
--------MDESEKRKERLNAMRME-A-SQSGDYNEAVGYGG--------LTNPLT-DVPS---GNVESYAMPRPRFDYYTDPMA-
AFSANKRSNN----------QPHVSPQVSQQCYT----------RATNPQS----------PICTPRGNYSVDQRS----
QGVHHTFNP-LGNP-GQNSPFGIPQ-RGSP----------SAW-NNS----------FDTP--------KNYLP----
PNSSMGGNFASP--------GIQRGGRPG---FHYGQ----------GSGQPGSGYGGSPY-QGSGYRGNPYQDSGHRGSPSQGSGHRGSP-
YQHSGNRGSPY------QGSGQ-GRSQ--WRGN--------SSSPFSFRGGRRGG-RGSH----------G---GTSGESRPDLCSSK----------
----------I--H---FS-AKP----------C----------
>XP_015060161 (Solanum pennellii)
--------MDESEKRKERLKAMRME-A-SQSGDYNEAVGYGV--------LTNPLT-DVPS---GNVESYAMPRPRFDYYTDPMA-AFSANRRSNN-
----------QPHVSPQVSQQCYT----------RATNHQS----------PICTPRGNYSVDQRS----QGVHHTFNP-LGNP-
GQNSPFGIPQ-RGSP----------SAW-NNS----------FDTP--------KNYLP----PNSSMGGNFASP--------
GIQQGGRPG---FHYGQ----------GSGQPGSGYGGSPY-QGSGYRGNPYQDSGHRGSPSQGSGHRGSP-YQHSGNRGSPY------
QGSGQ-GRSQ--WRGN--------TSSPFSFRGGRRGG-RGSH----------G---GTSGESRPDLYYSKSMVEDPWKELKPVIW-------KAFPE---
----------KPWLPHSISAKKAKFPDA-P-VKSIP-QQSLAECLAASFNEAASSEAA------
>gi|697107224|ref|XP_009607446.1| (Nicotiana tomentosiformis)
-------MEEESEKRKERLKAMRME-A-SECGNYNETENQLQG-------LSNPLV-ESPS---GQAEFCAAPR--FDYYTDPMA-AFSANKKRNN-
----------VSPQVSQACYTPP----------RPRNPQS----------PIYTAQDNYSLDQRSQS--QGVHHTFNP-LGNP-
GQNSPFGTPQ-RSSP----------NAW-GSS----------FGTP--------NNYFP----PNSSIGGNFASP--------
GIHRGGRPG---FHYGQ----------GRGNPGSGYRGSPS----------QGSGYRGSPNQGSGYRGSP-YQGPGHRGSPY------QGSAQ-

GRSQ--WMGN--------SSSPVSVQRGRR-G-LGSH----------G---CTSAESRPDLYYNKSMVEDPWKEMKPVIW-------KPLNAP-
SNNLDTPES--EKSSWLPKSISAKKAKIPDA-P-LKSTP-QQSLAEYLSASFNEAAGNESGKDGSGT
>gi|698479741|ref|XP_009786944.1| (Nicotiana sylvestris)
-------MEEESAKRKERLKAMRME-A-SQSGDYNETQNHLQG-------LSNPLI-ESPS---GQAEFHAAPR--FDYYTDPMA-AFSANKKWND--
----------VSPQVSQPCYTPP---------------------------------RPRNPQS----------PIYTARGNYSVDQRSQS--QGVHHTFNP-LGNP-
GQNSPFGTPQ-RSSP-----------------------------------NAW-DSS----------FGTS--------NNYFP----PNSSIGGNFASP--------
GIHQGGRPG---FHYGQ----------GRGNPGSGYRGSPS---------------------QGSGYRGSP-YQGPGHRGSPY------QGSAQ-GRSQ--
WMGN--------SSSPVSVQRGRR-G-LGSH----------G---CTSAESRPNLYYNKSMVEDPWKELKPVIW-------KPLNAP-SNKLDTPES--
EKSSWLPKSISAKKAKIPDAPP-LKSTP-QQSLAEYLSASFNEAASNEAGKDGSGI
>XP_012068029 (Jatropha curcas)
----------LKAMRTV-A-AQAEASSHVQTSSGYI---GF-LANPLL-ESPE--LTQEPSHATPR--FDFYTDPMA-AFYSNKKRSG----------
--VGNQAPQGYLTPPSDSA---SSMSQF-SSPHPA----------GPRNPDMTPF----------PSNQMQHNYSPYQIMDQT-QVAYNSIPP-
CTSPR--AGPFPMH--QGMP---------------------------------YAQGGSS----------GVA----------YYH----NNAPHRG-MTSQ--------
YHVRSRNPN-----FQ----------PEGNHSFNYGQGRP---------------------LSPRIGNNP-YFGSGRGGSST-------HSGR-GQGQ--WHGS----
----SRSQVSGRNGGR-G-RGFY----------SHGIGSDAALRAESFYDKSMVEDPWQRLEPVLW-------KGLD----------GS--S-
DSWLPKSASMKKPRVSES-S-NKSSS--QNLAEYLAAAFNESVNDAPS------
>XP_012068030 (Jatropha curcas)
----------LKAMRTV-A-AQAEASSHVQTSSGYI---GF-LANPLL-ESPE--LTQEPSHATPR--FDFYTDPMA-AFYSNKKRSG----------
--VGNQAPQGYLTPPSDSA---SSMSQF-SSPHP----------GPRNPDMTPF----------PSNQMQHNYSPYQIMDQT-QVAYNSIPP-
CTSPR--AGPFPMH--QGMP---------------------------------YAQG---------GSS-------GVAYYH----NNAPHRG-MTSQ--------
YHVRSRNPN-----FQ----------PEGNHSFNYGQGRP---------------------LSPRIGNNP-YFGSGRGGSST-------HSGR-GQGQ--WHGS----
----SRSQVSGRNGGR-G-RGFY----------SHGIGSDAALRAESFYDKSMVEDPWQRLEPVLW-------KGLD----------GS--S-
DSWLPKSASMKKPRVSES-S-NKSSS--QNLAEYLAAAFNESVNDAPS------
>Potri.005G153300 (Populus trichocarpa)
--------MEDSEKRRERLKAMRSIAA-AQAETSN-NVETSAPP---GL-LAYPLL-GTPATLLAQGESSAIPR--FDFYTDPSA-AFSANRKGAA--
----------GNQAARGYFTSPSN----NSSVPQL-SSPHP----------GQRNLEVTPPHAYQMQNSYPHANQMQSNHLPNQRMYRG-
QGPYHNAAS-YRSLRGFLCPFPMN--QGAP---------------------------PEMW-SGP----------GFP--------ASYF-----SSTVHGG-
LSSP--------YPIRQGNPG-----FG----------PVGSSP---------------------------------SPVSGYGGSPA--------ISQTGQGH--W----------
HSSSGFGQSGGR-G-RGFH----------SRGFAPNEAQGPECFYDNSMVEDPWQHLEPVLW-------SGLDDW-GNNLNGPGS--S-
NSLLPKSISMKKSSVAES-S-NKSTS-GLSLAEYLAA----------
>SapurV1A.0004s0580 (Sarracenia purpurea)
--------MEDAEKRSGRLKAMRAIAS-AQAETCNNNVETSVVP---GL-LANPLL-ENPATQPAQEELSATPR--FDFYTDPSA-
AFSANRKRTA----------TANQVARGLTPPNN----ISSTPQFSSSPRP----------GQRIPEVTPS----------
SAYQMQSNYSPNQRMYPG-QRPDHNAAF-YGIPRGFAGPFTMN--QGTP---------------------------------SEMW-SGS----------
GGP--------GSYH-----SSTPYRG-ISRP--------YPIHQGNPG-----FG----------PVGSRP---------------------------------SPVSGYGGSPA--
-----SSGR-GQGY--W----------NNSSGRGQSGGR-G-RGFR----------SRGFGLNETQEPECFYDNSMVEDPWQHLKPVLW-------
RGLGDL-GNNLNGPVS--S-NSWLPKSISVKKPKMSES-S-TKSTS-GQTLAEYLSAAFTEAANDEPNV-----
>Potri.002G111500 (Populus trichocarpa)
--------MEDAEKRSERLKAMRAVAS-AQAETCNDNVETSAVP---GL-LANPLL-ENPATQPALEELSATPR--FDFYTDPSA-
AFSSDRKRTA----------TANQVARGFRPPNN----ISSMPQF-SSPHP----------
GQRNPEVTPSSAYQMQNNYSPANQMQSNYSPNQRMYPG-QGPYHNAAF-YRTPSNFARPFTMN--QGTP---------------------------------
---EMW-NG----------PGGP--------ASNH-----SSTPYRG-ISRP--------YPIHQGNPG-----FG----------PVGSSP---------------------------------
----------SPVSGYGGSPA-------SSGR-GQGY--W----------DSSSGLGQSGGR-G-RGFR----------
SRGFALNETQEPECFHDNSMVEDPWQHLKPVLW-------RGLDDP-GNNLNGPVS--S-NSWLPKSISVKKPRISES-S-NKSTS-
GQTLAEYLSAAFTEATNDAPNV-----
>XP_011017016 (Populus euphratica)
--------MEDAEKRSERLKAMRAVAS-AQAETCNDNVETSAVP---GL-LANPLL-ENAATRPALEESRATPR--FDFYTDPSA-
AFSANRKRTA----------TANLVARSFTPPNN----ISSMPQF-SSPRP----------
GQRNPEVTPSSAYQLQNNYSHANQMQSNYSPNQRMYPG-QGPYHNAAF-YRTPSNFARPFTMN--QGTP---------------------------------
--EMW-NGP----------GGP--------ASYQ-----SYTPYRG-ISRP--------YPIHQGNPG-----FG----------PVGSSP---------------------------------
----------SPVSGYGGSPA-------SSGR-GRGY--W----------DSSSGLGQSGGR-G-RGFR----------
SRGFAPNETQEPECFYDNSMVEDPWQHLTPVLW-------RGLDDP-GNNLNGPVS--S-NSWLPKSISVKKTRISES-S-NKSTS-
GQTLAEYLSAAFTEATND---------
>XP_011017017 (Populus euphratica)
--------MEDAEKRSERLKAMRAVAS-AQAETCNDNVETSAVP---GL-LANPLL-ENAATRPALEESRATPR--FDFYTDPSA-
AFSANRKRTA----------TANLVARSFTPPNN----ISSMPQF-SSPRP----------GQRNPEVTPS----------
SAYQMQSNYSPNQRMYPG-QGPYHNAAF-YRTPSNFARPFTMN--QGTP---------------------------------EMW-NGP----------
GGP--------ASYQ-----SYTPYRG-ISRP--------YPIHQGNPG-----FG----------PVGSSP---------------------------------SPVSGYGGSPA--
-----SSGR-GRGY--W----------DSSSGLGQSGGR-G-RGFR----------SRGFAPNETQEPECFYDNSMVEDPWQHLTPVLW-------
RGLDDP-GNNLNGPVS--S-NSWLPKSISVKKTRISES-S-NKSTS-GQTLAEYLSAAFTEATND---------
>tr|A0A068VH56|A0A068VH56_COFCA (Coffea canephora)
--------MEESEKRKERLNALRRE-A-DQAGVHGDSENFMGA----SYTLANPLI-EASVPPSVHAELQAAPR--FDYYTDPMS-
AFSANKRGKV----------SHQISPDYFTPP---------------------------------RPVRPDMTNPWGH-------PAYQAQATYCADQSMFQS-
PRPYHRPGP-FRSPRGTPSPFGTP--EG---------------------------------------GGS----------SGTP--------PYVS-----
SNFSRGGSVGSP--------GFAPGGSPSFNDGHIR----------GYGCINSP---------------------------------ESGFGNFGSPY------PNSGR-

GQSR--WLGN--------NSPQASGRGRGR-G-QGYNAPVSANSGRRVGSHDFLSAEHNPGRFYNKSMVEDPWATLKPVIW-------
RRKDAP---TLGTHHS--K-KSWLPESIRGKKAQVSEV-S-HKYSS-GRSFAEDLAASFHAATNYESADDNDDG
>XP_010106355 (Morus notabilis)
--------MEESEKRRERLRAMRHEAA-AQSVNSD-NNEAPAMP---CY-LSNPLV-ETSAAAPPPEQSHGTSR--FDFYTDPMA-
AFSANKRRNN----------TSDPISSHHVTPPANSGSPMLRSPSPFS----------GPRYAGMS----------
PAHQFQSNYSPNPRMYQP-QGFGHDPIS-QSGELGMSRPFNMH--QGNM----------DPSI-GP----------GSA-
--------AGYYNFP--SNQPRGSRFPSP--------RIGPTGSFF----------NAGQGRAHWHNHSP----------
NPGLGRGGSPS------PSLGR-GGGR--WHGG--------STSPGSGRRGGR-G-PGSA----------
GRHFTMDRQLGPERFYDESMIEDAWKFLEPVVW-------REVDAS-LSSLSTPDS--S-KSWITRSLGAKKAKVSDS-T-SKSGS-
QPSLAEYLAASFDEANKDESS------
>tr|A0A061FEH0|A0A061FEH0_THECC (Theobroma cacao)
--------MDESEKRKERLKAMRLE-A-AQSEVPN-NVATPSVP---GH-LSNPLS-ETSSTAAVQEDFCSTPR--FDYYTDPMA-
AFSANKKRGK----------ADNQSTQNYFTPPTTSGWPVARVSPS-HP----------GPRNYDMNP----------
PVRHMQSQYSLDQRMYHQ-QGPHSNFAA-HRSPI-TRSPSHMH--HGNS----------DAW-
NGSQAFGNYYSSASDGSPGGMFGTP--------LMHPG----TTPRFWNPSNAS--------RYSNSPTPG-----FS----------PADIP---YGRGRP-
----------QQFGNYP-LPSPGHGGSLG------LSSGR-GRGR--GYGG--------SITHGIGRSGGR-G-LGFH----------
GHSSASNRMMGPESFYDESMLEDPWQHLKPVLW-------RRREAG-MDSLSNPDS--S-NSWFPKSISAKKVKVSEA-S-NKFNS-
QLSLAEYLAASFNKAVEDTKNE-----
>tr|A0A061FFF7|A0A061FFF7_THECC (Theobroma cacao)
--------MDESEKRKERLKAMRLE-A-AQSEVPN-NVATPSVP---GH-LSNPLS-ETSSTAAVQEDFCSTPR--FDYYTDPMA-ATSG----------
----------WPVARVSPSHP----------GPRNYDMNP----------PVRHMQSQYSLDQRMYHQ-QGPHSNFAA-HRSPI-
TRSPSHMH--HGNS----------DAW-NGSQAFGNYYSSASDGSPGGMFGTP--------LMHPG----TTPRFWNPSNAS--
------RYSNSPTPG-----FS----------PADIP---YGRGRP----------QQFGNYP-LPSPGHGGSLG------LSSGR-GRGR--GYGG-----
---SITHGIGRSGGR-G-LGFH----------GHSSASNRMMGPESFYDESMLEDPWQHLKPVLW-------RRREAG-MDSLSNPDS--S-
NSWFPKSISAKKVKVSEA-S-NKFNS-QLSLAEYLAASFNKAVEDTKNE-----
>XP_009376205 (Pyrus x bretschneideri)
--------MDESEKRKERLRAMRIE-A-EETEASL-KAATSAVP---VY-LSNPLA-EDTTAIPVPEEPCAPFR--FDFYSDPMA-AFSSDNKRIK----
----------VGDQIAQENFRHSNTGGFPGARLPSP-LSG----------GPRNPQMTAS----------PAHQFQRSYSPDQRMYQA-
QGSYQNFSP-QRSPVGMERPFPMH--HGNR----------PEVW-NG----------AEFRP----
PASPGYGPQGSP--------RFRPQGSPG-----FR----------PPGSPRFQPPGSPG----------FRPPTSPGFRPPGSPG------
SNIGQ-GRGH--WFSHTPRPQSVHGGSPSPGSSSGRGGWRGSH----------G--RAMDRQLGPERFYNATMVEDPWKFLEPVIW-------
KGVDTP-LRCLNTHGS--S-KLSIGRSSSTNNASISEA-L-NKSMP-QPSLAEFLAASLNEAVDDAPS------
>gi|720010143|ref|XP_010259151.1| (Nelumbo nucifera)
--------MEEAEKRRERLKAMRLE-A-AQSEDSCKVDSSTAQ----GY-LSNPLI-EPTATPLAKENSHVVSR--FDYYTDPMS-
AFSGNKRRSN----------SNQMSQSYTSPSISSG--SPMTRPS-LSPSA----------GPRSPFMSVS----------
PAHQFQANNSPDHRTNET-HMAFHGPDS-WRSPMGMTSPFPGY--RETP----------GPR-NRS----------
PGMP--------GYGFP----FNSPRGGGFPSP--------GLGRGGSPS----------PGSGIGGSY----------RFGNSP-
RNRAGWGGNPN------VSSGR-GCNY---HFN--------SPRNGSGRGRGR-G-RGSH----------A---
YVSARDRPELFYNKSMMDDPWKFLKPVI---------RRPIL-SNCQDTPDS--L-RSWLPKSINMKKARISET-S-NEFSS-
QSSLAECLSASFEEAANDATGI-----
>GSVIVG01009188001 (Vitus vinifera)
--------MEESEKRRERLKAMRME-A-AQTKVSD-TVDTSAMP---GY-LSNPLV-EGSATLPVQEDSCVTPR--FDFYTDPMS-
AFSSNKRRSK----------VGNQIQQDYLTPSSNSGYTATMARMS-SSLSA----------GPRNCEMTPS----------
PNPPFQPNFSPGQGINQA-QGLYHSSGP-YRSPIEMASPFPAH--QGTP----------GVW-NGS----------NGMP-
-------RYGVP----SNSPRGGNFP----------SPG-----FR----------PVGSPSFRSGRGRG----------HWFNNSP-
SPVSGRGGSSS------PNSGR-GRSG--WFGN--------SMSPGSGRGRGR-G-LGFH----------A---
HVSAQDRPELFYNKSMVEDPWKFLKPVIW-------SREKAL-GKMGNASDS--P-KSWLPKSINMKKTRVSEA-T-NESSS-
QQSLAEYLAASFNEAVNDASGT-----

Data File 2. Primers used in this study.

Target	Primer Name	Sequence, 5'-3'	Reference
CCA1	CCA1-QF	CCGCAACTTTCGCCTCAT	(Pruneda-Paz et al., 2009)
	CCA1-QR	GCCAGATTCGGAGGTGAGTTC	(Pruneda-Paz et al., 2009)
LHY	LHY-QF	GACTCAAACACTGCCCAGAAGA	(Hazen et al., 2005)
	LHY-QR	CGTCACTCCCTGAAGGTGTATTT	(Hazen et al., 2005)
LHY I1R	LHY_I1R-QF	GGCTACTCTCAAGGGTATAACAGTT	This study
	LHY_I1R-QR	CAGCAGCCAAACAGAGATCTTAGTAT	This study
LHY I1S	LHY_I1S-QF	TCAAGGGTTTTGGCTGCG	This study
	LHY_I1S-QR	ATGGCTTTCTTGCCTTAGCTAATAATTCT	This study
PRR7	PRR7-QF	CCACGAGCGGTATCTCTATGG	(Harmon et al., 2008)
	PRR7-QR	ACTTGGAAACTCAGGGTTAGAA	(Harmon et al., 2008)
PRR9	PRR9-QF	GCGTCGGCCTTCTCAAGATTT	(Harmon et al., 2008)
	PRR9-QR	TCCTCTAAAGCAACGACGGC	(Harmon et al., 2008)
TOC1	TOC1-QF	TCTTCGCAGAATCCCTGTGAT	(Hazen et al., 2005)
	TOC1-QR	GCTGCACCTAGCTTCAAGCA	(Hazen et al., 2005)
GI	GI-QF	GGGTAAATATGCTGCTGGAGA	(Harmon et al., 2008)
	GI-QR	CAGTATGACACCAGCTCCATT	(Harmon et al., 2008)
CAB2	CAB2-QF	CCGGAAAGGCTGTGAACCT	This study
	CAB2-QR	CACACGGCCGCTTCCA	This study
ELF3	ELF3-QF	AGGCAAGGGAGCACAGGAA	(Thines et al., 2014)
	ELF3-QR	GCCGAAAGGACTTGCTACCA	(Thines et al., 2014)
LUX	LUX-QF	TAACGTGGAGGAGGAAGATCGA	(Thines et al., 2014)
	LUX-QR	TCCATCACCGTTTGATGTCTTT	(Thines et al., 2014)
CAT3	CAT3-QF	GGATCCTTACAAGTATCGTCCTTCA	This study
	CAT3-QR	AGGGAAGAGATGTTGTTGGAGACT	This study
CCR2	CCR2-QF	TAGGGCGACGTTATTGATTCC	This study
	CCR2-QR	CCCTCAATCGCATCCTTCA	This study
PP2A	PP2A-QF	ACGTGGCCAAAATGATGCAA	(Goh et al., 2012)
	PP2A-QR	TCATGTTCTCCACAACCGCT	This study
IPP2	IPP2-QF	GTATGAGTTGCTTCTCCAGCAAAG	(Hazen et al., 2005)
	IPP2-QR	GAGGATGGCTGCAACAAGTGT	(Hazen et al., 2005)
SIC	SIC-QF	GGTGGAAATGAACCGCATCC	This study
	SIC-QR	AAACCAAGTTCAGTTTCCTGGT	This study
CCA1	CCA1_AS1F	TCCATTTCCGTAGCTTCTGGT	This study
	CCA1_AS1R	TCTGGACAGCAGTTTTTGTTGC	This study

	CCA1_MybF	ATCTGGTTATTAAGACTCGGAAGCC	(Pruneda-Paz et al., 2009)
	CCA1_MybR	GCCTCTTTCTCTACCTTGGAGAAAA	(Pruneda-Paz et al., 2009)
	CCA1_Int4_M13F	TGTAAAACGACGGCCAGTGGCAGAA-GATTGAAGAACATGTAGC	This study
	CCA1_Int4F	GGCAGAAGATTGAAGAACATGTAGC	(Pruneda-Paz et al., 2009)
	CCA1_Int4R	CGATCTTCATTGGCCATCTCAGGAT	(Pruneda-Paz et al., 2009)
	CCA1_AS2.2F	TAGTGGCAAAAGTGGGCTGA	This study
	CCA1_AS2.2R	TAATCCATTGGCAGCCCACC	This study
	CCA1_3'F	TCTGTGTCTGACGAGGGTCGAATT	(Pruneda-Paz et al., 2009)
	CCA1_3'R	ACTTTGCGGCAATACCTCTCTGG	(Pruneda-Paz et al., 2009)
LHY	LHY_Int1_M13F	TGTAAAACGACGGCCAGTGGCTACTCTCAAGGGTATAACAGTT	This study
	LHY_Int1F	GGCTACTCTCAAGGGTATAACAGTT	(Pruneda-Paz et al., 2009)
	LHY_Int1R	GCCACTTACCTGTTTCGTTGGTAAGG	(Pruneda-Paz et al., 2009)
	LHY_MybF	GAATTATTAGCTAAGGCAAGAAAGCC	(Pruneda-Paz et al., 2009)
	LHY_MybR	GCCTCTTTCTCCAACTTTGTGAAGA	(Pruneda-Paz et al., 2009)
	LHY_L-intF	AACGAATTGAAGAACATATTGGGAC	(Pruneda-Paz et al., 2009)
	LHY_L-intR	CCAGTTGATGTTTTCTCAGAGAACG	(Pruneda-Paz et al., 2009)
	LHY_E5aF	GTTCTTCACAAAGCCATCCATATGG	(Pruneda-Paz et al., 2009)
	LHY_E5aR	CCTCTTTCTCCAACCGAATTTTTCG	(Pruneda-Paz et al., 2009)
	LHY_3'F	ACGAAACAGGTAAGTGGCGACATT	(Pruneda-Paz et al., 2009)
	LHY_3'R	TGGGAACATCTTGAACCGCGTT	(Pruneda-Paz et al., 2009)
GI	GI_AS1F	TCCGATATTCGTCGATCTCTCA	This study
	GI_AS1R	AGAAGAAGTGCAGGGACAGC	This study
	GI_I5F	GCCCAAGTAGTGAGAATGAC	This study
	GI_I5R	ACCCATTACACCACTACACC	This study
	GI_AS2F	GCTGTCCCTGCACTTCTTCT	This study
	GI_AS2R	CAACCCAGACAATGCATGCG	This study

	GI_AS3F	TGGCACAAACTGATTGCTGC	This study
	GI_AS3R	AGACTAAACACCAGACGCACA	This study
ELF3	ELF3_Int1F	TGCCCCTTTCCTCTCTCTGT	This study
	ELF3_Int1R	TTGGCAATGGCTTTCCTTGC	This study
	ELF3_Int2_M13F	TGTAAAACGACGGCCAGTATCTA GTCAGCCTTGTGGTG	This study
	ELF3_Int2F	ATCTAGTCAGCCTTGTGGTG	This study
	ELF3_Int2R	GTGAACTTACTTGGCAATGGC	This study
	ELF3_Int3R	GACTTTACCTTAATCAGTCTGTGC	This study
TOC1	TOC1_ASaF	GTGTTCTTATCAAGTGACTGCAGT G	(Pruneda-Paz et al., 2009)
	TOC1_ASaR	CAAGTCCTAGCATGCGTCTTCTTC	(Pruneda-Paz et al., 2009)
	TOC1_ASbF	GAAGAAGACGCATGCTAGGACTT GC	(Pruneda-Paz et al., 2009)
	TOC1_ASbR	GCTGAAGAAAATTGCGCTGGATTA C	(Pruneda-Paz et al., 2009)
PRR9	PRR9_ASF	TGTTGAAGTGTATGCTGAGAGGT GC	This study
	PRR9_ASR	ATCATCACGCAAAGTCAGTCTTCT C	This study
	PRR9_I3F	TGTTGAAGTGTATGCTGAGAGGT GC	This study
	PRR9_I3R	ATCATCACGCAAAGTCAGTCTTCT C	This study
	PRR9_AS2F	GCTTCACAGCACAACCTTGA	This study
	PRR9_AS2R	AGCTTAGCCTGATCATTTGCAG	This study
PRR7	PRR7_AS1F	GGTGGCTTATGGTCGTCTCC	This study
	PRR7_AS1R	CCGCTCTCACTTCCACTACC	This study
	PRR7_ASaF	GTCTTTAAGTGCTTATCGAAAGGA GC	(Pruneda-Paz et al., 2009)
	PRR7_ASaR	CACTACCACTAGAACTTTGGCATC T	(Pruneda-Paz et al., 2009)
	PRR7_Int4_M13F	TGTAAAACGACGGCCAGTGATGC CAAAGTGTATGTCCTTGC	This study
	PRR7_Int4F	GATGCCAAAGTGTATGTCCTTGC	This study
	PRR7_Int4R	CGCTCTCACTTCCACTACCA	This study
	PRR7_AS3F	GAGGCGCCTAACACTCACTT	This study
	PRR7_AS3R	CGTACTATGCACGGCTCCAA	This study

PRR5	PRR5F	GTAGTTACAGAGTTGCTGCAGTACC	(Pruneda-Paz et al., 2009)
	PRR5R	TCAGGAGCAAGTGAAGTTTGTCTTCT	(Pruneda-Paz et al., 2009)
	PRR5cF	GTAGTTACAGAGTTGCTGCAGTACC	(Pruneda-Paz et al., 2009)
	PRR5cR	GCCTGTGAACGTACCGTGTGAAAGC	(Pruneda-Paz et al., 2009)
PRR3	PRR3F	TGGTTCTGGTCTTTAAGTGTTTGTCG	(Pruneda-Paz et al., 2009)
	PRR3R	CTTCCGCTAGAGCTGTGACATCTTC	(James et al., 2012)
UBC21	UBC21_M13F	TGTAAAACGACGGCCAGTTTAGAGATGCAGGCATCAAGAGCGC	This study
	UBC2F	TTAGAGATGCAGGCATCAAGAGCGC	(James et al., 2012)
	UBC2R	CATATTTCTCCTGTCTTGAAATGAA	(James et al., 2012)
SIC	AT4G24500.1-134caccF	CACCGAGACTTTTGTTTTCGGGCAATTGAG	This study
	AT4G24500.1+2112R	ACGGTACCTAACAAGTCACATTAATCTATAAAAA	This study
	SIC_CDS_caccF	CACCATGGAAGATTCAGAGAAAAGAAAACA	This study
	SIC_CDS_nostopR	ATTGCTTGACTCATCGCAT	This study
	SIC_CDS_wstopR	TTAATTGCTTGACTCATCGCAT	This study
	SIC_5UTR_caccF	CACCGAGAAAGTGAAGATTTTTGGAGAAAC	This study
SIC	SICalleleF	GCACAGGTCACCTCTCCAAT	This study
	SICalleleR	AAAATCGTTCTGCCCCTGGT	This study
	SIC-YFP_F	CAACCACGTCTTCAAAGCAA	This study
	SIC-YFP_R	TTAATTGCTTGACTCATCGCAT	This study
PRMT4a	PRMT4A_CDS_F	CACCATGGAGATTCCTTCTCTGAATAAGCAGC	This study
	PRMT4A_CDS+stp_R	CTAGAGCTGAGCGTTTGCGTTTTG	This study
	PRMT4A_CDS-nostp_R	GAGCTGAGCGTTTGCGTTTTGT	This study
PRMT4b	PRMT4B_CDS_F	CACCATGGAGGTATCTTCTGTGAAAAAGC	This study
	PRMT4B_CDS+stp_R	TTAGAGCTGGGCACTTGGGTT	This study
	PRMT4B_CDS-nostp_R	GAGCTGGGCACTTGGGTTCT	This study

XAB2	XAB2_CDS_F	CACCATGGCGATTTCCAAAGATC	This study
	XAB20_CDS-nostp_R	CTGATTAAGCTTCTGTCTCTTGATCC	This study
AQR	AQR_CDS_F	CACCATGACGAAGGTCTATGGAAC	This study
	AQR_CDS-nostp_R	ATTCTTCTCATCAGCCTTTCCATTC	This study
DBR1	DBR1_CDS_F	CACCATGAAGATTGCAATTGAAGG	This study
	DBR1_CDS-nostp_R	TGCATCGTCTCTTGTATGA	This study
	DBR1_CDS+stp_R	TCATGCATCGTCTCTTGTATGA	This study

Bibliography

Abe, M, M Fujiwara, K Kurotani, S Yokoi, and K Shimamoto. 2008. “Identification of Dynamin as an Interactor of Rice GIGANTEA by Tandem Affinity Purification (TAP).” *Plant and Cell Physiology* 49 (3): 420–32. doi:pcn019 [pii]10.1093/pcp/pcn019.

Altschul, S F, W Gish, W Miller, E W Myers, and D J Lipman. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215 (3): 403–10. doi:10.1016/S0022-2836(05)80360-2S0022-2836(05)80360-2 [pii].

Austin, R S, D Vidaurre, G Stamatiou, R Breit, N J Provart, D Bonetta, J Zhang, et al. 2011. “Next-Generation Mapping of Arabidopsis Genes.” *Plant Journal* 67 (4): 715–25. doi:10.1111/j.1365-313X.2011.04619.x.

Bedford, M T, and S Richard. 2005. “Arginine Methylation an Emerging Regulator of Protein Function.” *Molecular Cell* 18 (3): 263–72. doi:S1097-2765(05)01247-5 [pii]10.1016/j.molcel.2005.04.003.

Ben-Yehuda, Sigal, Ian Dix, Caroline S Russell, Margaret Mcgarvey, Jean D Beggs, and Martin Kupiec. 2000. “Genetic and Physical Interactions Between Factors Involved in Both Cell Cycle Progression and Pre-mRNA Splicing in Saccharomyces Cerevisiae.” *Genetics* 156: 1503–17.

Bendix, Claire, Carine M. Marshall, and Frank G. Harmon. 2015. “Circadian Clock Genes Universally Control Key Agricultural Traits.” *Molecular Plant* 8 (8). Cell Press: 1135–52.

Bentley, David L. 2005. “Rules of Engagement : Co-Transcriptional Recruitment of Pre-mRNA Processing Factors.” *Current Opinion in Cell Biology*, no. 17: 251–256. doi:10.1016/j.ceb.2005.04.006.

Bhardwaj, V, S Meier, L N Petersen, R A Ingle, and L C Roden. 2011. “Defence Responses of Arabidopsis Thaliana to Infection by Pseudomonas Syringae Are Regulated by the Circadian Clock.” *PLoS ONE* 6 (10): e26968. doi:10.1371/journal.pone.0026968PONE-D-11-15640 [pii].

Burgess, Sean M, and Christine Guthrie. 1993. “A Mechanism to Enhance mRNA Splicing Fidelity : The RNA-Dependent ATPase Prpl6 Governs Usage of a Discard Pathway for Aberrant Lariat Intermediates.” *Cell* 73: 1377–91.

Chanarat, Sittinan, and Katja Sträßer. 2013. “Splicing and beyond : The Many Faces of the Prp19 Complex.” *Biochimica et Biophysica Acta* 1833 (10): 2126–34. doi:10.1016/j.bbamcr.2013.05.023.

Chaudhury, A, K Okada, N V Raikhel, K Shinozaki, and V V Sundaresan. 1999. “A Weed Reaches New Heights down under.” *Plant Cell* 11 (10): 1817–26. http://www.ncbi.nlm.nih.gov/pubmed/10521514.

Chen, W F, K H Low, C Lim, and I Edery. 2007. “Thermosensitive Splicing of a Clock Gene and Seasonal Adaptation.” *Cold Spring Harbor Symposia on Quantitative Biology* 72: 599–606. doi:10.1101/sqb.2007.72.021.

Cheng, D, J Cote, S Shaaban, and M T Bedford. 2007. “The Arginine Methyltransferase

CARM1 Regulates the Coupling of Transcription and mRNA Processing." *Molecular Cell* 25 (1): 71–83. doi:S1097-2765(06)00790-8 [pii]10.1016/j.molcel.2006.11.019.
Chow, B Y, A Helfer, D A Nusinow, and S A Kay. 2012. "ELF3 Recruitment to the PRR9 Promoter Requires Other Evening Complex Members in the Arabidopsis Circadian Clock." *Plant Signaling & Behavior* 7 (2): 170–73. doi:10.4161/psb.18766.
Consortium, UniProt. 2015. "UniProt: A Hub for Protein Information." *Nucleic Acids Research* 43 (Database issue): D204-12. doi:10.1093/nar/gku989gku989 [pii].
Covington, M F, J N Maloof, M Straume, S A Kay, and S L Harmer. 2008. "Global Transcriptome Analysis Reveals Circadian Regulation of Key Pathways in Plant Growth and Development." *Genome Biology* 9 (8): R130. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18710561.
Cui, Z, Q Xu, and X Wang. 2014. "Regulation of the Circadian Clock through Pre-mRNA Splicing in Arabidopsis." *Journal of Experimental Botany* 65 (8): 1973–80. doi:10.1093/jxb/eru085eru085 [pii].
Dai, S, X Wei, L Pei, R L Thompson, Y Liu, J E Heard, T G Ruff, and R N Beachy. 2011. "BROTHER OF LUX ARRHYTHMO Is a Component of the Arabidopsis Circadian Clock." *Plant Cell* 23 (3): 961–72. doi:10.1105/tpc.111.084293.
Daniel, X, S Sugano, and E M Tobin. 2004. "CK2 Phosphorylation of CCA1 Is Necessary for Its Circadian Oscillator Function in Arabidopsis." *Proceedings of the National Academy of Sciences of the United States of America* 101 (9): 3292–97.
Demarsy, E, and C Fankhauser. 2009. "Higher Plants Use LOV to Perceive Blue Light." *Current Opinion in Plant Biology* 12 (1): 69–74. doi:10.1016/j.pbi.2008.09.002S1369-5266(08)00159-3 [pii].
Diernfellner, A, H V Colot, O Dintsis, J J Loros, J C Dunlap, and M Brunner. 2007. "Long and Short Isoforms of Neurospora Clock Protein FRQ Support Temperature-Compensated Circadian Rhythms." *FEBS Letters* 581 (30): 5759–64. doi:S0014-5793(07)01187-8 [pii]10.1016/j.febslet.2007.11.043.
Diernfellner, Axel C R, Tobias Schafmeier, Martha W Merrow, and Michael Brunner. 2005. "Molecular Mechanism of Temperature Sensing by the Circadian Clock of Neurospora Crassa." *Genes and Development* 19 (17): 1968–73. doi:gad.345905 [pii]10.1101/gad.345905.
Dixon, L E, K Knox, L Kozma-Bognar, M M Southern, A Pokhilko, and A J Millar. 2011. "Temporal Repression of Core Circadian Genes Is Mediated through EARLY FLOWERING 3 in Arabidopsis." *Current Biology* 21 (2): 120–25. doi:S0960-9822(10)01599-X [pii]10.1016/j.cub.2010.12.013.
Dodd, Antony N, Neeraj Salathia, Anthony Hall, Eva Kévei, Réka Tóth, Ferenc Nagy, Julian M Hibberd, Andrew J Millar, and Alex a R Webb. 2005. "Plant Circadian Clocks Increase Photosynthesis, Growth, Survival, and Competitive Advantage." *Science (New York, N.Y.)* 309 (5734): 630–33. doi:10.1126/science.1115581.
Dong, Malia A, Eva M Farré, and Michael F Thomashow. 2011. "CIRCADIAN CLOCK-ASSOCIATED 1 and LATE ELONGATED HYPOCOTYL Regulate Expression of the C-REPEAT BINDING FACTOR (CBF) Pathway in Arabidopsis." *Proceedings of*

the National Academy of Sciences 108 (17): 7241–46. doi:10.1073/pnas.1103741108.
Earley, Keith W., Jeremy R. Haag, Olga Pontes, Kristen Opper, Tom Juehne, Keming Song, and Craig S. Pikaard. 2006. “Gateway-Compatible Vectors for Plant Functional Genomics and Proteomics.” *Plant Journal* 45 (4): 616–29. doi:10.1111/j.1365-313X.2005.02617.x.
Edgar, R C. 2004. “MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput.” *Nucleic Acids Research* 32 (5): 1792–97. doi:10.1093/nar/gkh34032/5/1792 [pii].
Edwards, K D, P E Anderson, A Hall, N S Salathia, J C Locke, J R Lynn, M Straume, J Q Smith, and A J Millar. 2006. “FLOWERING LOCUS C Mediates Natural Variation in the High-Temperature Response of the Arabidopsis Circadian Clock.” *Plant Cell* 18 (3): 639–50.
Edwards, Kieron D., James R. Lynn, Péter Gyula, Ferenc Nagy, and Andrew J. Millar. 2005. “Natural Allelic Variation in the Temperature-Compensation Mechanisms of the Arabidopsis Thaliana Circadian Clock.” *Genetics* 170 (1): 387–400. doi:10.1534/genetics.104.035238.
Fankhauser, Christian, and Dorothee Staiger. 2002. “Photoreceptors in Arabidopsis Thaliana: Light Perception, Signal Transduction and Entrainment of the Endogenous Clock.” *Planta* 216 (1): 1–16. doi:10.1007/s00425-002-0831-4.
Farinas, B, and P Mas. 2011. “Histone Acetylation and the Circadian Clock: A Role for the MYB Transcription Factor RVE8/LCL5.” *Plant Signaling & Behavior* 6 (4): 541–43. doi:14837 [pii].
Farré, Eva M., Steve A. Kay, E M Farre, and Steve A. Kay. 2007. “PRR7 Protein Levels Are Regulated by Light and the Circadian Clock in Arabidopsis.” *Plant Journal* 52 (3): 548–60. doi:TPJ3258 [pii]10.1111/j.1365-313X.2007.03258.x.
Filichkin, Sergei A., Jason S. Cumbie, Palitha Dharmawardhana, Pankaj Jaiswal, Jeff H. Chang, Saiprasad G. Palusa, A. S N Reddy, Molly Megraw, and Todd C. Mockler. 2015. “Environmental Stresses Modulate Abundance and Timing of Alternatively Spliced Circadian Transcripts in Arabidopsis.” *Molecular Plant* 8 (2): 207–27. doi:10.1016/j.molp.2014.10.011.
Filichkin, Sergei A, and Todd C Mockler. 2012. “Unproductive Alternative Splicing and Nonsense mRNAs: A Widespread Phenomenon among Plant Circadian Clock Genes.” *Biology Direct* 7: 20. doi:10.1186/1745-6150-7-201745-6150-7-20 [pii].
Filichkin, Sergei, Henry D Priest, Molly Megraw, and Todd C Mockler. 2015. “Alternative Splicing in Plants: Directing Traffic at the Crossroads of Adaptation and Environmental Stress.” *Current Opinion in Plant Biology* 24: 125–35. doi:10.1016/j.pbi.2015.02.008S1369-5266(15)00028-X [pii].
Findlay, Gregory M., Evan a. Boyle, Ronald J. Hause, Jason C. Klein, and Jay Shendure. 2014. “Saturation Editing of Genomic Regions by Multiplex Homology-Directed Repair.” *Nature* 513 (7516): 120–23. doi:10.1038/nature13695.
Finn, R D, A Bateman, J Clements, P Coggill, R Y Eberhardt, S R Eddy, A Heger, et al. 2014. “Pfam: The Protein Families Database.” *Nucleic Acids Research* 42

(Database issue): D222-30. doi:10.1093/nar/gkt1223gkt1223 [pii].
Fogelmark, K, and C Troein. 2014. “Rethinking Transcriptional Activation in the Arabidopsis Circadian Clock.” *PLoS Computational Biology* 10 (7): e1003705. doi:10.1371/journal.pcbi.1003705PCOMPBIOL-D-14-00197 [pii].
Garcia-Viloca, M, J Gao, M Karplus, and D G Truhlar. 2004. “How Enzymes Work: Analysis by Modern Rate Theory and Computer Simulations.” *Science* 303 (5655): 186–95. doi:10.1126/science.1088172303/5655/186 [pii].
Gendron, J M, J L Pruneda-Paz, C J Doherty, A M Gross, S E Kang, and S A Kay. 2012. “Arabidopsis Circadian Clock Protein, TOC1, Is a DNA-Binding Transcription Factor.” *Proceedings of the National Academy of Sciences of the United States of America* 109 (8): 3167–72. doi:1200355109 [pii]10.1073/pnas.1200355109.
Goodstein, D M, S Shu, R Howson, R Neupane, R D Hayes, J Fazo, T Mitros, et al. 2012. “Phytozome: A Comparative Platform for Green Plant Genomics.” *Nucleic Acids Research* 40 (Database issue): D1178-86. doi:10.1093/nar/gkr944gkr944 [pii].
Gould, Peter D, James C W Locke, Camille Larue, Megan M Southern, Seth J Davis, Shigeru Hanano, Richard Moyle, et al. 2006. “The Molecular Basis of Temperature Compensation in the Arabidopsis Circadian Clock.” *Plant Cell* 18 (5): 1177–87. doi:tpc.105.039990 [pii]10.1105/tpc.105.039990.
Green, Rachel M, Sonia Tingay, Zhi-Yong Wang, and Elaine M Tobin. 2002. “Circadian Rhythms Confer a Higher Level of Fitness to Arabidopsis Plants.” *Plant Physiology* 129 (2): 576–84. doi:10.1104/pp.004374.
Green, R M, and E M Tobin. 1999. “Loss of the Circadian Clock-Associated Protein 1 in Arabidopsis Results in Altered Clock-Regulated Gene Expression.” *Proceedings of the National Academy of Sciences of the United States of America* 96 (7): 4176–79. doi:10.1073/pnas.96.7.4176.
Greenham, Kathleen, and C. Robertson McClung. 2015. “Integrating Circadian Dynamics with Physiological Processes in Plants.” *Nature Reviews Genetics* 16 (10). Nature Publishing Group: 598–610. doi:10.1038/nrg3976.
Guilgur, Leonardo Gastón, Pedro Prudêncio, and Daniel Sobral. 2014. “Requirement for Highly Efficient Pre-mRNA Splicing during Drosophila Early Embryonic Development.” *eLIFE* 3: e02181. doi:10.7554/eLife.02181.
Guindon, S, J F Dufayard, V Lefort, M Anisimova, W Hordijk, and O Gascuel. 2010. “New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0.” *Systems Biology* 59 (3): 307–21. doi:10.1093/sysbio/syq010syq010 [pii].
Gutierrez, R A, T L Stokes, K Thum, X Xu, M Obertello, M S Katari, M Tanurdzic, et al. 2008. “Systems Approach Identifies an Organic Nitrogen-Responsive Gene Network That Is Regulated by the Master Clock Control Gene CCA1.” *Proceedings of the National Academy of Sciences of the United States of America* 105 (12): 4939–44. doi:10.1073/pnas.0800211105 [pii].
Hanawalt, Philip C, and Graciela Spivak. 2008. “Transcription-Coupled DNA Repair : Two Decades of Progress and Surprises” 9 (DECEmbER): 958–70.

doi:10.1038/nrm2549.
Harmer, S L, J B Hogenesch, M Straume, H S Chang, B Han, T Zhu, X Wang, J a Kreps, and S a Kay. 2000. “Orchestrated Transcription of Key Pathways in Arabidopsis by the Circadian Clock.” *Science* 290 (5499): 2110–13. doi:10.1126/science.290.5499.2110.
Harmer, Stacey L. 2009. “The Circadian System in Higher Plants.” *Annual Review of Plant Biology* 60 (1): 357–77. doi:10.1146/annurev.arplant.043008.092054.
He, F, and A Jacobson. 2015. “Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story.” *Annual Review of Genetics* 49: 339–66. doi:10.1146/annurev-genet-112414-054639.
Helfer, A, D A Nusinow, B Y Chow, A R Gehrke, M L Bulyk, and S A Kay. 2011. “LUX ARRHYTHMO Encodes a Nighttime Repressor of Circadian Gene Expression in the Arabidopsis Core Clock.” *Current Biology* 21 (2): 126–33. doi:S0960-9822(10)01652-0 [pii]10.1016/j.cub.2010.12.021.
Hernando, Carlos E, Sabrina E Sanchez, Estefanía Mancini, and Marcelo J Yanovsky. 2015. “Genome Wide Comparative Analysis of the Effects of PRMT5 and PRMT4/CARM1 Arginine Methyltransferases on the Arabidopsis Thaliana Transcriptome.” *BMC Genomics* 16: 192. doi:10.1186/s12864-015-1399-210.1186/s12864-015-1399-2 [pii].
Herrero, E, E Kolmos, N Bujdoso, Y Yuan, M Wang, M C Berns, H Uhlworm, et al. 2012. “EARLY FLOWERING4 Recruitment of EARLY FLOWERING3 in the Nucleus Sustains the Arabidopsis Circadian Clock.” *Plant Cell* 24 (2): 428–43. doi:10.1105/tpc.111.093807.
Hogg, Rebecca, Joanne C Mcgrail, and Raymond T O Keefe. 2010. “The Function of the NineTeen Complex (NTC) in Regulating Spliceosome Conformations and Fidelity during Pre-mRNA Splicing.” *Biochemical Society Transa* 38: 1110–15. doi:10.1042/BST0381110.
Hong, Sunghyun, Sun A Kim, Mary Lou Guerinot, and C Robertson McClung. 2013. “Reciprocal Interaction of the Circadian Clock with the Iron Homeostasis Network in Arabidopsis.” *Plant Physiology* 161 (2): 893 LP-903. http://www.plantphysiol.org/content/161/2/893.abstract.
Hong, Sunghyun, Hae-Ryong R Song, Kerry Lutz, Randall A Kerstetter, Todd P Michael, and C Robertson McClung. 2010. “Type II Protein Arginine Methyltransferase 5 (PRMT5) Is Required for Circadian Period Determination in Arabidopsis Thaliana.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (49): 21211–16. doi:10.1073/pnas.1011987071011987107 [pii].
Hsu, Polly Yingshan, Upendra K Devisetty, and Stacey L Harmer. 2013. “Accurate Timekeeping Is Controlled by a Cycling Activator in Arabidopsis.” *Elife* 2: e00473. doi:10.7554/eLife.0047300473 [pii].
Hsu, Polly Yingshan, and Stacey L. Harmer. 2014. “Wheels within Wheels: The Plant Circadian System.” *Trends in Plant Science* 19 (4): 240–49. doi:10.1016/j.tplants.2013.11.007.

Huang, He, Sophie Alvarez, Rebecca Bindbeutel, Zhouxin Shen, Michael J Naldrett, Bradley S Evans, Steven P Briggs, et al. 2016. “Identification of Evening Complex Associated Proteins in Arabidopsis by Affinity Purification and Mass Spectrometry * □,” no. 3: 201–17. doi:10.6019/PXD002606.Each.

Huang, W., P. Perez-Garcia, a. Pokhilko, a. J. Millar, I. Antoshechkin, J. L. Riechmann, and P. Mas. 2012. “Mapping the Core of the Arabidopsis Circadian Clock Defines the Network Structure of the Oscillator.” *Science* 336 (6077): 75–79. doi:science.1219075 [pii]10.1126/science.1219075.

Ito, S, Y H Song, and T Imaizumi. 2012. “LOV Domain-Containing F-Box Proteins: Light-Dependent Protein Degradation Modules in Arabidopsis.” *Molecular Plant* 5 (3): 573–82. doi:10.1093/mp/sss013sss013 [pii].

James, A. B., N. H. Syed, S. Bordage, J. Marshall, G. A. Nimmo, G. I. Jenkins, P. Herzyk, J. W. S. Brown, and H. G. Nimmo. 2012. “Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes.” *Plant Cell* 24 (3): 961–81. doi:10.1105/tpc.111.093948tpc.111.093948 [pii].

James, Allan B., Naeem Hasan Syed, John W. S. Brown, and Hugh G. Nimmo. 2012. “Thermoplasticity in the Plant Circadian Clock.” *Plant Signaling & Behavior* 7 (10): 1219–23. doi:10.4161/psb.21491.

Johnson, C H, J Elliott, R Foster, K Honma, and R Kronauer. 2004. “Fundamental Properties of Circadian Rhythms.” In *Chronobiology: Biological Timekeeping*, edited by J Dunlap, J J Loros, and P J DeCoursey, 67–105. Sunderland, MA: Sinauer Associates.

Jones, Matthew A., Brian A. Williams, Jim McNicol, Craig G. Simpson, John W.S. Brown, and Stacey L. Harmer. 2012. “Mutation of Arabidopsis Spliceosomal Timekeeper locus1 Causes Circadian Clock Defects.” *Plant Cell* 24 (10): 4066–82. doi:10.1105/tpc.112.104828tpc.112.104828 [pii].

Joon, Pil, and Paloma Mas. 2014. “Multiple Layers of Posttranslational Regulation Re Fi Ne Circadian Clock Activity in Arabidopsis” 26 (January): 79–87. doi:10.1105/tpc.113.119842.

Kalyna, M, C G Simpson, N H Syed, D Lewandowska, Y Marquez, B Kusenda, J Marshall, et al. 2012. “Alternative Splicing and Nonsense-Mediated Decay Modulate Expression of Important Regulatory Genes in Arabidopsis.” *Nucleic Acids Research* 40 (6): 2454–69. doi:10.1093/nar/gkr932gkr932 [pii].

Karampelias, Michael, Pia Neyt, Steven De Groeve, Stijn Aesaert, Griet Coussens, Jakub Rolčík, Leonardo Bruno, et al. 2016. “ROTUNDA3 Function in Plant Development by Phosphatase 2A-Mediated Regulation of Auxin Transporter Recycling.” *Proceedings of the National Academy of Sciences of the United States of America* 113 (10): 2768–73. doi:10.1073/pnas.1501343112 1501343112 [pii].

Kataoka, Naoyuki. 2015. “Human RNA Lariat Debranching Enzyme Protein 1 – A Surveillant for Branch RNAs for Degradation” 1: 1–5. doi:10.14800/rd.963.

Kearse, M, R Moir, A Wilson, S Stones-Havas, M Cheung, S Sturrock, S Buxton, et al. 2012. “Geneious Basic: An Integrated and Extendable Desktop Software Platform for the Organization and Analysis of Sequence Data.” *Bioinformatics* 28 (12): 1647–

49. doi:10.1093/bioinformatics/bts199bts199 [pii].
Keller, Walter. 1984. “The RNA Lariat : A New Ring to the Splicing of mRNA Precursors.” *Cell* 39: 423–25.
Kersey, P J, J E Allen, I Armean, S Boddu, B J Bolt, D Carvalho-Silva, M Christensen, et al. 2016. “Ensembl Genomes 2016: More Genomes, More Complexity.” *Nucleic Acids Research* 44 (D1): D574-80. doi:10.1093/nar/gkv1209gkv1209 [pii].
Kiba, T, R Henriques, H Sakakibara, and N H Chua. 2007. “Targeted Degradation of PSEUDO-RESPONSE REGULATOR5 by an SCFZTL Complex Regulates Clock Function and Photomorphogenesis in Arabidopsis Thaliana.” *Plant Cell* 19 (8): 2516–30. doi:tpc.107.053033 [pii]10.1105/tpc.107.053033.
Kim, Woe-Yeon Y, Sumire Fujiwara, Sung-Suk S Suh, Jeongsik Kim, Yumi Kim, Linqu Han, Karine David, Joanna Putterill, Hong Gil Nam, and David E Somers. 2007. “ZEITLUPE Is a Circadian Photoreceptor Stabilized by GIGANTEA in Blue Light.” *Nature* 449 (7160): 356–60. doi:10.1038/nature06132.
Koncz, Csaba, Femke DeJong, Nicolas Villacorta, Dõra Szakonyi, and Zsuzsa Koncz. 2012. “The Spliceosome-Activating Complex : Molecular Mechanisms Underlying the Function of a Pleiotropic Regulator.” *Frontiers in Plant Science* 3: 1–12. doi:10.3389/fpls.2012.00009.
Kuhn, P, R Chumanov, Y Wang, Y Ge, R R Burgess, and W Xu. 2011. “Automethylation of CARM1 Allows Coupling of Transcription and mRNA Splicing.” *Nucleic Acids Research* 39 (7): 2717–26. doi:10.1093/nar/gkq1246gkq1246 [pii].
Kuraoka, Isao, Shinsuke Ito, Tadashi Wada, Mika Hayashida, Lily Lee, Masafumi Sijo, Yoshimichi Nakatsu, et al. 2008. “Isolation of XAB2 Complex Involved in Pre-mRNA Splicing, Transcription, and Transcription-Coupled Repair.” *Journal of Biological Chemistry* 283 (2): 940–50. doi:10.1074/jbc.M706647200.
Kwon, Young-Ju, Mi-Jeong Park, Sang-Gyu Kim, Ian T Baldwin, and Chung-Mo Park. 2014. “Alternative Splicing and Nonsense-Mediated Decay of Circadian Clock Genes under Environmental Stress Conditions in Arabidopsis.” *BMC Plant Biology* 14 (1): 136. doi:10.1186/1471-2229-14-136.
Lai, a. G., C. J. Doherty, B. Mueller-Roeber, S. a. Kay, J. H. M. Schippers, and P. P. Dijkwel. 2012. “CIRCADIAN CLOCK-ASSOCIATED 1 Regulates ROS Homeostasis and Oxidative Stress Responses.” *Proceedings of the National Academy of Sciences* 109 (42): 17129–34. doi:10.1073/pnas.1209148109.
Lareau, L F, M Inada, R E Green, J C Wengrod, and S E Brenner. 2007. “Unproductive Splicing of SR Genes Associated with Highly Conserved and Ultraconserved DNA Elements.” *Nature* 446 (7138): 926–29. doi:nature05676 [pii]10.1038/nature05676.
Lewis, B P, R E Green, and S E Brenner. 2003. “Evidence for the Widespread Coupling of Alternative Splicing and Nonsense-Mediated mRNA Decay in Humans.” *Proceedings of the National Academy of Sciences of the United States of America* 100 (1): 189–92. doi:10.1073/pnas.0136770100136770100 [pii].
Li, Gang, Hamad Siddiqui, Yibo Teng, Rongcheng Lin, Xiang-yuan Wan, Jigang Li, On-Sun Lau, et al. 2011. “Coordinated Transcriptional Regulation Underlying the Circadian Clock in Arabidopsis.” *Nat Cell Biol* 13 (5). Nature Publishing Group, a

division of Macmillan Publishers Limited. All Rights Reserved.: 616–22. http://dx.doi.org/10.1038/ncb2219.
Lim, C, and R Allada. 2013. “Emerging Roles for Post-Transcriptional Regulation in Circadian Clocks.” *Nature Neuroscience* 16 (11): 1544–50. doi:10.1038/nn.3543nn.3543 [pii].
Liu, X L, M F Covington, C Fankhauser, J Chory, and D R Wagner. 2001. “ELF3 Encodes a Circadian Clock-Regulated Nuclear Protein That Functions in an Arabidopsis PHYB Signal Transduction Pathway.” *Plant Cell* 13 (6): 1293–1304. doi:10.1105/tpc.13.6.1293.
Lu, Sheen X, Stephen M Knowles, Christos Andronis, May S Ong, and Elaine M Tobin. 2009. “CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL Function Synergistically in the Circadian Clock of Arabidopsis.” *Plant Physiology* 150 (2): 834 LP-843. http://www.plantphysiol.org/content/150/2/834.abstract.
Marshall, Carine M, Virginia Tartaglio, Maritza Duarte, and Frank G Harmon. 2016. “The Arabidopsis Sickle Mutant Exhibits Altered Circadian Clock Tesponses to Cool Temperatures and Temperature-Dependent Alternative Splicing.” *The Plant Cell* 28 (10). United States: 2560–75. doi:10.1105/tpc.16.00223.
Mas, P, W Y Kim, D E Somers, and S A Kay. 2003. “Targeted Degradation of TOC1 by ZTL Modulates Circadian Function in Arabidopsis Thaliana.” *Nature* 426 (6966): 567–70.
Más, Paloma, and Marcelo J. Yanovsky. 2009. “Time for Circadian Rhythms: Plants Get Synchronized.” *Current Opinion in Plant Biology* 12 (5): 574–79. doi:10.1016/j.pbi.2009.07.010.
Masaki, So, Rei Yoshimoto, Daisuke Kaida, Asuka Hata, and Takayuki Satoh. 2015. “Identification of the Specific Interactors of the Human Lariat RNA Debranching Enzyme 1 Protein.” *International Journal of Molecular Sciences* 16: 3705–21. doi:10.3390/ijms16023705.
Matera, A Gregory, and Zefeng Wang. 2014. “A Day in the Life of the Spliceosome.” *Nature Reviews Molecular Cell Biology* 15 (2): 108–21. doi:10.1038/nrm3742nrm3742 [pii].
Matsushika, A, S Makino, M Kojima, and T Mizuno. 2000. “Circadian Waves of Expression of the APRR1/TOC1 Family of Pseudo-Response Regulators in Arabidopsis Thaliana: Insight into the Plant Circadian Clock.” *Plant and Cell Physiology* 41 (9): 1002–12. http://www.ncbi.nlm.nih.gov/pubmed/11100772.
McGlincy, N J, and C W Smith. 2008. “Alternative Splicing Resulting in Nonsense-Mediated mRNA Decay: What Is the Meaning of Nonsense?” *Trends in Biochemical Sciences* 33 (8): 385–93. doi:10.1016/j.tibs.2008.06.001S0968-0004(08)00143-6 [pii].
McWatters, H G, R M Bastow, A Hall, and a J Millar. 2000. “The ELF3 Zeitnehmer Regulates Light Signalling to the Circadian Clock.” *Nature* 408 (6813): 716–20. doi:10.1038/35047079.
Michael, T P, and C R McClung. 2002. “Phase-Specific Circadian Clock Regulatory

Elements in Arabidopsis." *Plant Physiology* 130 (2): 627–38.
Michael, Todd P, Patrice A Salome, and C Robertson McClung. 2003. "Two Arabidopsis Circadian Oscillators Can Be Distinguished by Differential Temperature Sensitivity." *Proceedings of the National Academy of Sciences of the United States of America* 100 (11): 6878–83. doi:10.1073/pnas.1131995100.
Mizoguchi, Tsuyoshi, Kay Wheatley, Yoshie Hanzawa, Louisa Wright, Mutsuko Mizoguchi, Hae Ryong Song, I A Carre, George Coupland, Isabelle A. Carré, and George Coupland. 2002. "LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis." *Developmental Cell* 2 (5): 629–41. doi:S1534580702001703 [pii].
Mizuno, T, Y Nomoto, H Oka, M Kitayama, A Takeuchi, M Tsubouchi, and T Yamashino. 2014. "Ambient Temperature Signal Feeds into the Circadian Clock Transcriptional Circuitry through the EC Night-Time Repressor in Arabidopsis Thaliana." *Plant & Cell Physiology* 55 (5): 958–76. doi:10.1093/pcp/pcu030.
Moore, Melissa J. 2002. "Nuclear RNA Turnover." *Cell* 108: 431–34.
Nagel, D H, and S A Kay. 2012. "Complexity in the Wiring and Regulation of Plant Circadian Networks." *Current Biology* 22 (16): R648-57. doi:10.1016/j.cub.2012.07.025S0960-9822(12)00816-0 [pii].
Nakamichi, N, T Kiba, M Kamioka, T Suzuki, T Yamashino, T Higashiyama, H Sakakibara, and T Mizuno. 2012. "Transcriptional Repressor PRR5 Directly Regulates Clock-Output Pathways." *Proceedings of the National Academy of Sciences of the United States of America* 109 (42): 17123–28. doi:10.1073/pnas.12051561091205156109 [pii].
Niu, Lifang, Yong Zhang, Yanxi Pei, Chunyan Liu, and Xiaofeng Cao. 2008. "Redundant Requirement for a Pair of PROTEIN ARGININE METHYLTRANSFERASE4 Homologs for the Proper Regulation of Arabidopsis Flowering Time." *Plant Physiology* 148 (1): 490–503. doi:10.1104/pp.108.124727.
Nohales, Maria A, and Steve A Kay. 2016. "Molecular Mechanisms at the Core of the Plant Circadian Oscillator." *Nature Publishing Group* 23 (12). Nature Publishing Group: 1061–69. doi:10.1038/nsmb.3327.
Nusinow, D A, A Helfer, E E Hamilton, J J King, T Imaizumi, T F Schultz, E M Farre, and S A Kay. 2011. "The ELF4-ELF3-LUX Complex Links the Circadian Clock to Diurnal Control of Hypocotyl Growth." *Nature* 475 (7356): 398–402. doi:10.1038/nature10182nature10182 [pii].
Ooi, S. L., C. Dann, K. Nam, D. J. Leahy, M. J. Damha, and J. D. Boeke. 2001. "RNA Lariat Debranching Enzyme." In *Methods in Enzymology*, 342:233–48. doi:10.1016/S0076-6879(01)42548-1.
Pittendrigh, C S. 1954. "On Temperature Independence in the Clock System Controlling Emergence Time in Drosophila." *Proceedings of the National Academy of Sciences of the United States of America* 40 (10): 1018–29. http://www.ncbi.nlm.nih.gov/pubmed/16589583.
Plautz, J D, M Straume, R Stanewsky, C F Jamison, C Brandes, H B Dowse, J C Hall, and S A Kay. 1997. "Quantitative Analysis of Drosophila Period Gene Transcription

in Living Animals." *Journal of Biological Rhythms* 12 (3): 204–17.
Portolés, Sergi, and Paloma Más. 2010. "The Functional Interplay between Protein Kinase CK2 and cca1 Transcriptional Activity Is Essential for Clock Temperature Compensation in Arabidopsis." *PLoS Genetics* 6 (11). doi:10.1371/journal.pgen.1001201.
Rawat, R, N Takahashi, P Y Hsu, M A Jones, J Schwartz, M R Salemi, B S Phinney, and S L Harmer. 2011. "REVEILLE8 and PSEUDO-REPONSE REGULATOR5 Form a Negative Feedback Loop within the Arabidopsis Circadian Clock." *PLoS Genetics* 7 (3): e1001350. doi:10.1371/journal.pgen.1001350.
Roe, Jung-Hye, Richard R Burgess, and M.Thomas Record. 1985. "Temperature Dependence of the Rate Constants of the Escherichia Coli RNA Polymerase-λPR Promoter Interaction." *Journal of Molecular Biology* 184 (3): 441–53. doi:http://dx.doi.org/10.1016/0022-2836(85)90293-1.
Salome, P A, D Weigel, and C R McClung. 2010. "The Role of the Arabidopsis Morning Loop Components CCA1, LHY, PRR7, and PRR9 in Temperature Compensation." *Plant Cell* 22 (11): 3650–61. doi:10.1105/tpc.110.079087tpc.110.079087 [pii].
Salomé, Patrice A, C Robertson McClung, P A Salome, and C Robertson McClung. 2005. "PSEUDO-RESPONSE REGULATOR 7 and 9 Are Partially Redundant Genes Essential for the Temperature Responsiveness of the Arabidopsis Circadian Clock." *Plant Cell* 17 (3): 791–803. doi:10.1105/tpc.104.029504.In.
Salomé, Patrice A, Michele Oliva, Detlef Weigel, and Ute Krämer. 2013. "Circadian Clock Adjustment to Plant Iron Status Depends on Chloroplast and Phytochrome Function." *The EMBO Journal* 32 (4): 511–23. doi:10.1038/emboj.2012.330.
Samach, Alon, H Onouchi, S E Gold, G S Ditta, Z Schwarz-Sommer, M F Yanofsky, and G Coupland. 2000. "Distinct Roles of CONSTANS Target Genes in Reproductive Development of Arabidopsis." *Science* 288 (5471): 1613–16. doi:10.1126/science.288.5471.1613.
Sanchez, Sabrina E, Ezequiel Petrillo, Esteban J Beckwith, Xu Zhang, Matias L Rugnone, C Esteban Hernando, Juan C Cuevas, et al. 2010. "A Methyl Transferase Links the Circadian Clock to the Regulation of Alternative Splicing." *Nature* 468 (7320): 112–16. doi:10.1038/nature09470nature09470 [pii].
Schaffer, Robert, Nicola Ramsay, Alon Samach, Sally Corden, Joanna Putterill, Isabelle A Carre, George Coupland, Isabelle A. Carré, and George Coupland. 1998. "The Late Elongated Hypocotyl Mutation of Arabidopsis Disrupts Circadian Rhythms and the Photoperiodic Control of Flowering." *Cell* 93 (7): 1219–29. doi:10.1016/S0092-8674(00)81465-8.
Schlaen, R G, E Mancini, S E Sanchez, S Perez-Santangelo, M L Rugnone, C G Simpson, J W Brown, X Zhang, A Chernomoretz, and M J Yanovsky. 2015. "The Spliceosome Assembly Factor GEMIN2 Attenuates the Effects of Temperature on Alternative Splicing and Circadian Rhythms." *Proceedings of the National Academy of Sciences of the United States of America* 112 (30): 9382–87. doi:10.1073/pnas.15045411121504541112 [pii].
Seo, P J, M J Park, M H Lim, S G Kim, M Lee, I T Baldwin, and C M Park. 2012. "A

Self-Regulatory Circuit of CIRCADIAN CLOCK-ASSOCIATED1 Underlies the Circadian Clock Regulation of Temperature Responses in Arabidopsis." *Plant Cell* 24 (6): 2427–42. doi:10.1105/tpc.112.098723tpc.112.098723 [pii].

Shaw, Peter J., and John Ws S Brown. 2004. "Plant Nuclear Bodies." *Current Opinion in Plant Biology* 7 (6): 614–20. doi:S1369-5266(04)00132-3 [pii]10.1016/j.pbi.2004.09.011.

Song, L, M H Han, J Lesicka, and N Fedoroff. 2007. "Arabidopsis Primary microRNA Processing Proteins HYL1 and DCL1 Define a Nuclear Body Distinct from the Cajal Body." *Proceedings of the National Academy of Sciences of the United States of America* 104 (13): 5437–42. doi:0701061104 [pii]10.1073/pnas.0701061104.

Staiger, D, and R Green. 2011. "RNA-Based Regulation in the Plant Circadian Clock." *Trends in Plant Science* 16 (10): 517–23. doi:10.1016/j.tplants.2011.06.002S1360-1385(11)00128-2 [pii].

Stein, Paula, Petr Svoboda, Martin Anger, and Richard M Schultz. 2003. "RNAi: Mammalian Oocytes Do It without RNA-Dependent RNA Polymerase." *Rna* 9 (2): 187–92. doi:10.1261/rna.2860603.1.

Syed, N H, M Kalyna, Y Marquez, A Barta, and J W Brown. 2012. "Alternative Splicing in Plants--Coming of Age." *Trends in Plant Science* 17 (10): 616–23. doi:10.1016/j.tplants.2012.06.001S1360-1385(12)00127-6 [pii].

Thines, Bryan, and Frank G Harmon. 2010. "Ambient Temperature Response Establishes ELF3 as a Required Component of the Core Arabidopsis Circadian Clock." *Proceedings of the National Academy of Sciences of the United States of America* 107 (7): 3257–62.

Walter, M, C Chaban, K Schutze, O Batistic, K Weckermann, C Nake, D Blazevic, et al. 2004. "Visualization of Protein Interactions in Living Plant Cells Using Bimolecular Fluorescence Complementation." *Plant Journal* 40 (3): 428–38.

Wang, Guirong, Xiaoxuan Guo, and Joanna Floros. 2005. "Differences in the Translation Efficiency and mRNA Stability Mediated by 5′-UTR Splice Variants of Human SP-A1 and SP-A2 Genes." *American Journal of Physiology - Lung Cellular and Molecular Physiology* 289 (3). http://ajplung.physiology.org/content/289/3/L497.long.

Wang, Huai, Kristine Hill, and Sharyn E Perry. 2004. "An Arabidopsis RNA Lariat Debranching Enzyme Is Essential for Embryogenesis." *Journal of Biological Chemistry* 279 (2): 1468–73. doi:10.1074/jbc.M309106200.

Wang, Wei, Jinyoung Yang Barnaby, Yasuomi Tada, Hairi Li, M Tor, Daniela Caldelari, Dae-un U Lee, et al. 2011. "Timing of Plant Immune Responses by a Central Circadian Regulator." *Nature* 470 (7332): 110–14. doi:10.1038/nature09766nature09766 [pii].

Wang, X, F Wu, Q Xie, H Wang, Y Wang, Y Yue, O Gahura, et al. 2012. "SKIP Is a Component of the Spliceosome Linking Alternative Splicing and the Circadian Clock in Arabidopsis." *Plant Cell* 24 (8): 3278–95. doi:tpc.112.100081 [pii]10.1105/tpc.112.100081.

Wang, Zhi-Yong Y, and Elaine M Tobin. 1998. "Constitutive Expression of the

CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) Gene Disrupts Circadian Rhythms and Suppresses Its Own Expression." *Cell* 93 (7): 1207–17. doi:10.1038/sj.onc.1201613.

Wenden, Bénédicte, Lászlõ Kozma-Bognár, Kieron D. Edwards, Anthony J W Hall, James C W Locke, and Andrew J. Millar. 2011. "Light Inputs Shape the Arabidopsis Circadian System." *Plant Journal* 66 (3): 480–91. doi:10.1111/j.1365-313X.2011.04505.x.

Yachdav, G, E Kloppmann, L Kajan, M Hecht, T Goldberg, T Hamp, P Honigschmid, et al. 2014. "PredictProtein--an Open Resource for Online Prediction of Protein Structural and Functional Features." *Nucleic Acids Research* 42 (Web Server issue): W337-43. doi:10.1093/nar/gku366gku366 [pii].

Yeom, Miji, Hyunmin Kim, Junhyun Lim, Ah-Young Shin, Sunghyun Hong, Jeong-Il Kim, and Hong Gil Nam. 2017. "How Do Phytochromes Transmit the Light Quality Information to the Circadian Clock in Arabidopsis?" *Molecular Plant* 7 (11). Elsevier: 1701–4. doi:10.1093/mp/ssu086.

Yonemasu, Rie, Mitsuyoshi Minami, Yoshimichi Nakatsu, Masayo Takeuchi, Isao Kuraoka, Yoichi Matsuda, Yujiro Higashi, Hisato Kondoh, and Kiyoji Tanaka. 2005. "Disruption of Mouse XAB2 Gene Involved in Pre-mRNA Splicing, Transcription and Transcription-Coupled DNA Repair Results in Preimplantation Lethality." *DNA Repair* 4 (4): 479–91. doi:10.1016/j.dnarep.2004.12.004.

Zhan, Xiangqiang, Bangshing Wang, Hongjiang Li, Renyi Liu, Rajwant K. Kalia, J.-K. Jian-Kang K Zhu, and Viswanathan Chinnusamy. 2012. "Arabidopsis Proline-Rich Protein Important for Development and Abiotic Stress Tolerance Is Involved in microRNA Biogenesis." *Proceedings of the National Academy of Sciences of the United States of America* 109 (44): 18198–203. doi:10.1073/pnas.12161991091216199109 [pii].

Zhang, Y, Y Tian, Q Chen, D Chen, Z Zhai, and H B Shu. 2007. "TTDN1 Is a Plk1-Interacting Protein Involved in Maintenance of Cell Cycle Integrity." *Cellular and Molecular Life Sciences* 64 (5): 632–40. doi:10.1007/s00018-007-6501-8.