Coordinated transcriptional regulation between a reactive oxygen species-responsive gene network and the circadian clock in *Arabidopsis thaliana*

A thesis presented in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Plant Biology

at Massey University, Palmerston North,

New Zealand.

Alvina Grace Lai

2012

Abstract

Most organisms have evolved endogenous biological clocks as internal timekeepers to fine-tune physiological processes to the external environment. Energetic cycles such as photosynthesis and glycolytic cycles are physiological processes that have been shown to be under clock control. This work sought to understand the mechanism of the synchrony between the circadian oscillator and products of energetic cycles. The fact that plants rely on photosynthesis for survival, and that photosynthesis relies on the sun, this would have meant that oxygen levels would have fluctuated across the day. A common by-product of oxygen metabolism and photosynthesis is the Reactive Oxygen Species (ROS). Evidence has proposed ROS as regulators of cellular signaling and plant development. However, if ROS levels are left unmanaged, it may cause oxidative stress in organisms, which could damage cellular components and disrupt normal mechanisms of cellular signaling. Therefore, it is advantageous for plants to be able to anticipate such periodic burst in ROS. My research investigates the role of the circadian clock in regulating ROS homeostasis in the model plant *Arabidopsis thaliana*. I found that ROS production and scavenging wax and wane in a periodic manner under diurnal and circadian conditions. Not only that, at the transcriptional level, ROS-responsive genes exhibited time-of-day specific phases under diurnal and circadian conditions, suggesting the role of the circadian clock in ROS signaling. Mutations in the core-clock regulator, *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*), affect both the transcriptional regulation of ROS genes and ROS homeostasis. Furthermore, mis-expressions of other clock genes such as *EARLY FLOWERING 3* (*ELF3*), *LUX ARRHYTHMO* (*LUX*) and *TIMING OF CAB EXPRESSION 1* (*TOC1*) also have profound effects on ROS signaling and homeostasis, thus suggesting a global clock effect on ROS networks. Taken together, *CCA1* is proposed as a master regulator of ROS signaling where the response to oxidative stress is dependent on the time of *CCA1* expression. Plants exhibit the strongest response at dawn, the time when *CCA1* peaks. Moreover, CCA1 can associate to the Evening Element or CCA1-Binding Site on promoters of ROS genes *in vivo* to coordinate transcription. A common feature of circadian clocks is the presence of multiple interlocked transcriptional feedback loops. It is shown here that

the oscillator incorporates ROS as a component of the loop where ROS signals could feed back to affect circadian behavior by changing *CCA1* and *TOC1* transcription. The clock regulates a plethora of output pathways; particularly the transcription of an output gene *FLAVIN BINDING KELCH REPEAT F-BOX 1* (*FKF1*) is affected by ROS signals. Temporal coordination of ROS signaling by *CCA1* and the reciprocal control of circadian behavior by ROS revealed a mechanistic link of which plants match their physiology to the environment to confer fitness.

Preface

Additional published work undertaken during the course of study, as stated below, is included in section 7. Publication.

Lai, A. G., Denton-Giles, M., Mueller-Roeber, B., Schippers, J. H. M., & Dijkwel, P. P. 2011. Positional information resolves structural variations and uncovers an evolutionarily divergent genetic locus in accessions of *Arabidopsis thaliana. Genome Biology and Evolution* **3**: 627–640.

Authors' contributions:

AGL carried out sample preparations, LR-PCR, sequence analysis, data interpretation and validation experiments. MD-G participated in sequence analysis. JHMS and BM-R participated in research design and data interpretation. AGL and PPD conceived and designed the project. AGL drafted the manuscript. PPD critically reviewed the manuscript. All authors contributed to the draft, read and approved the final manuscript.

Table of contents

Abbreviations

3-AT	3-aminotriazole
ABA	Abscisic acid
ACT2	*ACTIN 2*
AGI	Arabidopsis Gene Identifier
AOX	Alternative oxidase
APX	Ascorbate peroxidase
APX4	*ASCORBATE PEROXIDASE 4*
AsA	Ascorbate
ASN1	*ASPARAGINE SYNTHASE 1*
AtCP1	*Ca2+ BINDING PROTEIN 1*
ATPase	Adenosine triphosphatase
bHLH128	*Basic helix-loop-helix DNA-binding superfamily*
BMAL1	*Brain and muscle ARNT-like 1*
BME	β-mercaptoethanol
BOA	*BROTHER OF LUX ARRHYTHMO*
°C	Degrees celcius
CAB2	*CHLOROPHYLL A/B BINDING PROTEIN 2*
cADPR	Cyclic adenosine diphosphate ribose
CAT	*CATALASE*
CBF	*CRT/DRE BINDING FACTORS*
CBS	CCA1-binding site
CCA1	*CIRCADIAN CLOCK ASSOCIATED 1*
CCD	Charged couple device
CCR2	*COLD CIRCADIAN RHYTHM RNA-BINDING 2*
CCT	CONSTANS (CO), CO-like, TOC1
cDNA	Complementary DNA
CHE	*CCA1 HIKING EXPEDITION*
ChIP	Chromatin immunoprecipitation
CLOCK	*Circadian locomotor output cycles kaput*

cm	Centimeter
CRY	*CRYPTOCHROME*
CSD2	*Cd/Zn superoxide dismutase*
Cytb6f	Cytochrome b6f
DAB	3,3-diaminobenzidine
DD	Constant darkness
DHA	Dehydroascorbate
DHAR	DHA reductase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribose nucleic acid
DNase	Deoxyribonuclease
DPI	Diphenylene iodonium
EC	Evening complex
EE	Evening element
ELF	*EARLY FLOWERING*
ERF2	*ETHYLENE RESPONSE FACTOR 2*
Fd	Ferredoxin
FeSOD	*Fe superoxide dismutases*
FFT-NLLS	Fast fourier transform-nonlinear least squares
FIO1	*FIONA 1*
FKF1	*FLAVIN BINDING KELCH REPEAT F-BOX 1*
FNR	Ferredoxin NADPH reductase
FW	Fresh weight
g	Gram
GA	Gibberellin
GCL	Glutamate cysteine ligase
GAST1	GIBBERELLIC ACID STIMULATED TRANSCRIPT1
GDH1	*GLUTAMINE DEHYDROGENASE*
GFP	Green fluorescent protein
GI	*GIGANTEA*
GLN 1.3	*GLUTAMINE SUNTHASE*

GLR	Glutaredoxin
Glu	Glutamate
GO	Gene ontology
GPX	Glutathione peroxidase
GR	Glutathione reductase
GSH	Glutathione
GSSG	Oxidized glutathione
h	Hour(s)
H_2O	Water
H_2O_2	Hydrogen peroxide
HRP	Horseradish peroxidase
HSFA4A	*HEAT SHOCK TRANSCRIPTION FACTOR A4A*
IPP2	*ISOPENTENYL PYROPHOSPHATE 2*
JMJD5	*JUMONJI DOMAIN PROTEIN 5*
KI	Potassium iodide
krpm	Kilo-revolutions per minute
LD	12 h light 12 h dark
LHCB	*LIGHT-HARVESTING CHLOROPHYLL A/B-BINDING PROTEIN*
LHY	*LATE ELONGATED HYPOCOTYL*
LKP2	LOV KELCH PROTEIN 2
LL	Constant light
LUC	Luciferase
LUX	*LUX ARRHYTHMO*
LWD	*LIGHT-REGULATED WD 1*
M	Molar
MAPK	Mitogen-activated protein kinase
MDA	Monodehydroascorbate
MDAR	Monodehydroascorbate reductase
MES18	*METHYL ESTERASE 18*
µE	Micro-Einstein
µg	Micro-gram

µl	Micro-litre
µM	Micro-molar
mg	Milli-gram
mL	Milli-liter
min	Minute(s)
mM	Milli-molar
mRNA	Messenger ribonucleic acid
MS	Murashige and Skoog
MV	Methyl viologen
MYB59	*MYB DOMAIN PROTEIN 59*
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced form)
ng	Nano-gram
nm	Nano-meter
NOX	NADPH oxidase
NTR	NADPH thioredoxin reductase
NTRX	NADP-linked thioredoxin
$^1O_2$	Singlet oxygen
O_2	Oxygen
O^{2-}	Superoxide anion
OH^-	Hydroxyl radical
PAL1	*PHENYLALANINE AMMONIA-LYASE 1*
PC	Plastocyanin
PCR	Polymerase chain reaction
PDX1	*PYRIDOXINE BIOSYNTHESIS 1*
PER	*PERIOD*
PHY	*PHYTOCHROME*
PIF7	*PHYTOCHROME INTERACTING FACTOR 7*
PIN	PIN-FORMED
PQ	Plastoquinone

PRMT5	*PROTEIN ARGININE METHYL TRANSFERASE 5*
PRR	*PSEUDO-RESPONSE REGULATOR*
PRX	Peroxiredoxin
PrxR	Peroxireductase
PSI	Photosystem I
PSII	Photosystem II
PUP1	*PURINE PERMEASE 1*
qPCR	Quantitative polymerase chain reaction
RAV1	*RAV family transcription factor*
RBOH	Respiratory burst oxidase homologue
RBOHC	*RESPIRATORY BURST OXIDASE HOMOLOG C*
RNA	Ribonucleic acid
RNase	Ribonuclease
ROS	Reactive oxygen species
RuBisCO	Ribulose-1,5-bisphosphate carboxylase oxygenase
RVE	*REVEILLE*
s.d.	Standard deviation
s.e.m	Standard error of mean
SA	Salicylic acid
SAND	*GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C2*
SCF	Skp1-Cullin-F-box
sec	Second(s)
SHAM	Salicylhydroxamic acid
SOD	Superoxide dismutase
TCA	Tricarboxylic acid
TCP	TB1, CYC and PCF
TE	Tris-ethylene diamine tetra-acetic acid
TF	Transcription factor
TIC	*TIME FOR COFFEE*
TOC1	*TIMING OF CAB EXPRESSION 1*
Trx	Thioredoxin

TTFL	Transcriptional-translational feedback loop
TUB2	*TUBULIN BETA-2*
U	Unit(s)
UPL7	*UBIQUITIN-PROTEIN LIGASE 7*
VIM	Video-intensified microscopy
VTC2	*VITAMIN C 2*
WT	Wild-type
XCT	*XAP5 CIRCADIAN TIME-KEEPER*
ZAT	Zinc-finger transcription factor
ZAT10	*SALT TOLERANCE ZINC FINGER*
ZT	Zeitgeber time
ZTL	*ZEITLUPE*

List of figures

List of tables

1. Introduction

1.1 The clock paradigm

1.1.1 Biological clocks and circadian rhythms

The earth's rotation on its axial tilt around the Sun gives rise to diel cycles of a 24 h period. Two contrasting environmental conditions, a result of the rising and setting of the sun, bring about inherently regular day and night cycles. It is therefore likely that organisms have evolved an autonomous internal timekeeper known as the circadian/biological clock to enable temporal coordination of physiological events and the synchronization of these events to successions of day and night (Pittendrigh, 1993).

Back in the 17th century, the first evidence of circadian rhythms is discovered where some organisms possess internal mechanisms to measure the progression of time. The first measurement of rhythmic changes is by Jean de Mairan (1729) whom observed leaf movements of a heliotrope plant that followed a periodic rhythm. He transferred the plants to a dark cellar and found that even in the absence of diurnal light signals, leaf movement persisted. It has only been recently that circadian rhythmicity is considered an essential regulatory process that controls molecular and biochemical functions in organisms, albeit the pioneering research of de Mairan (1729) and Bünning (1973) earlier on. Rhythmic leaf movement in plants actually reflect the rhythmic changes in plasmalemma of pulvinus cells, where such changes are related to periodic alteration in membrane permeability (Satter et al., 1988), the periodic synthesis and metabolism of membrane proteins and the periodic phosphorylation and dephosphorylation of transport proteins (Schweiger and Schweiger, 1977).

Several examples of pioneering eukaryotic research involve the discovery of circadian rhythms in carbon dioxide metabolism in *Bryophyllum fedtschenkoi* (Wilkins 1959, 1989), in conidiation of *Neurospora crassa* (Sargent et al., 1966), in photosynthetic capacity in *Acetabularia mediterranaea* (Sweeney and Haxo, 1961) and in luminescence in a marine dinoflagellate *Gonyaulax polyedra* (Sweeney, 1971). Circadian rhythms also prevail in the animal kingdom. In the 19th century, researchers found that animals could maintain a 24 h activity pattern in the absence of external

stimuli such as light and temperature. The first clock mutant was isolated in *Drosophila* where the *PERIOD* gene was discovered (Konopka & Benzer, 1971). Later in 1994, Takahashi discovered the first mammalian *CLOCK* gene in mice (Vitaterna et al., 1994). Many other physiological behaviors were discovered to be circadian controlled. For example in mice, the circadian clock regulates feeding, wheel running and body temperature (van der Horst et al., 1999). Also, egg laying and eclosion in *Drosophila* are behavioral outputs regulated by the circadian clock (Sehgal et al., 1992). Notably, the existence of a timekeeping mechanism that is independent of cues from Earth's rotation was discovered in *Neurospora crassa* grown in space. Despite the removal from earth's orbital cues, *Neurospora* could maintain a 23 h rhythmic growth under complete darkness (Sulzman et al., 1984).

Circadian rhythm denotes a periodic oscillation in biochemical or behavioral events. The circadian clock generates self-sustaining circadian rhythms that is the 24 h temporal oscillations in biological and metabolic processes to allow organisms to anticipate diurnal changes and coordinate their physiology according to such changes. Circadian rhythms are outputs of the clock if they fulfill the following criteria (Mas, 2008; Harmer, 2009; Mas & Yanovsky, 2009). Firstly, circadian rhythms are generated under entraining conditions (an environment that varies according to the time of the day). The sinusoidal waves of circadian rhythms are made up of three components, the period, phase and amplitude (McClung, 2006; Harmer, 2009; Fig. 1.1). Rhythms persist with approximately 24 h periodicity when the organism is transferred to free-running conditions (an environment that is unchanged or constant). Indeed, the word circadian was coined from Latin terms; *circa*: approximately and *dies*: day. Although circadian rhythms can persist in the absence of external cues, the clock does not run in isolation from the environment. Secondly, circadian systems receive input signals such as light and temperature to reset the time of onset of the rhythm. Zeitgeber, the German term for 'time-giver' coined by chronobiologist Jurgen Aschoff (1960), is any external stimulus that entrains the endogenous timekeeper according to earth's 24 h diel cycles. The time of onset of an input signal that resets the oscillator is defined as zeitgeber time 0 (ZT0). If the organism is entrained under 12 h light and 12 h dark cycles, under free-running conditions ZT0-ZT12 denotes the subjective day whereas

ZT12-ZT24 denotes the subjective night. Temporal fluctuations in light and temperature can synchronize the rhythmic expression of clock genes. Thirdly, circadian rhythms are temperature compensated and retain 24 h periodicity across a wide range of physiological temperatures, e.g., 12°C to 27°C (Edwards et al., 2005). Because of this feature, circadian systems are robust timekeepers that act as buffering systems against ambient temperatures changes. Indeed, temperature compensatory mechanisms are crucial for the survival of many organisms to enable persistence in various geographical zones. Biogeographical analysis of the model plant *Arabidopsis thaliana* revealed that almost all accessions never experience average monthly temperatures higher than 16°C (Hoffmann, 2002). Nevertheless, daily temperature maxima may exceed 16°C and therefore it is essential that the clock is buffered against such daily temperature variations so that a 24 h periodicity can be maintained for each cycle (Salomé et al., 2010).

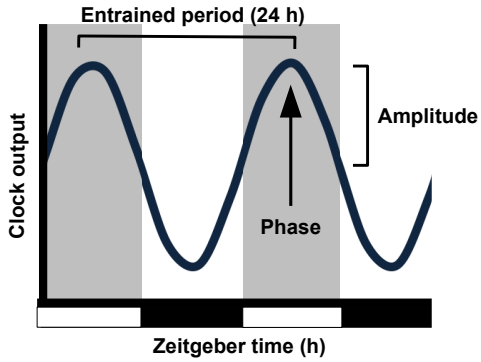


Figure 1.1: An idealized circadian rhythm
The entrained period, time taken to complete one cycle, is 24 h in light-dark cycles. Period can be measured from peak to peak, trough to trough or from any specific phase position. The amplitude is calculated as half of the peak-to-trough distance. The phase is the time of day for a particular event; if the rhythm peaked at dusk, the phase of the peak would be 12 h and so on. Zeitgeber is the term for 'time giver' that includes any stimulus that conveys time information to the clock such as light or temperature.

The circadian system essentially consists of three interconnected components (Fig. 1.2); 1) an environmental stimulus that provides input to allow the entrainment of the oscillator according to local environmental conditions, 2) the core oscillator that typically consists of negative feedback loops, which generates rhythms based on the

inputs and 3) output pathways that coordinate physiological processes based on cues from the core oscillator (Jones, 2009; McWatters & Devlin, 2011). The clock has evolved to incorporate multiple, partially redundant, interlocking components to allow greater flexibility in the modulation of clock inputs, i.e., altering light sensitivities, in order to increase the accuracy of the timing mechanism (Rand et al., 2004; Dunlap et al., 2007; Harmer, 2009).

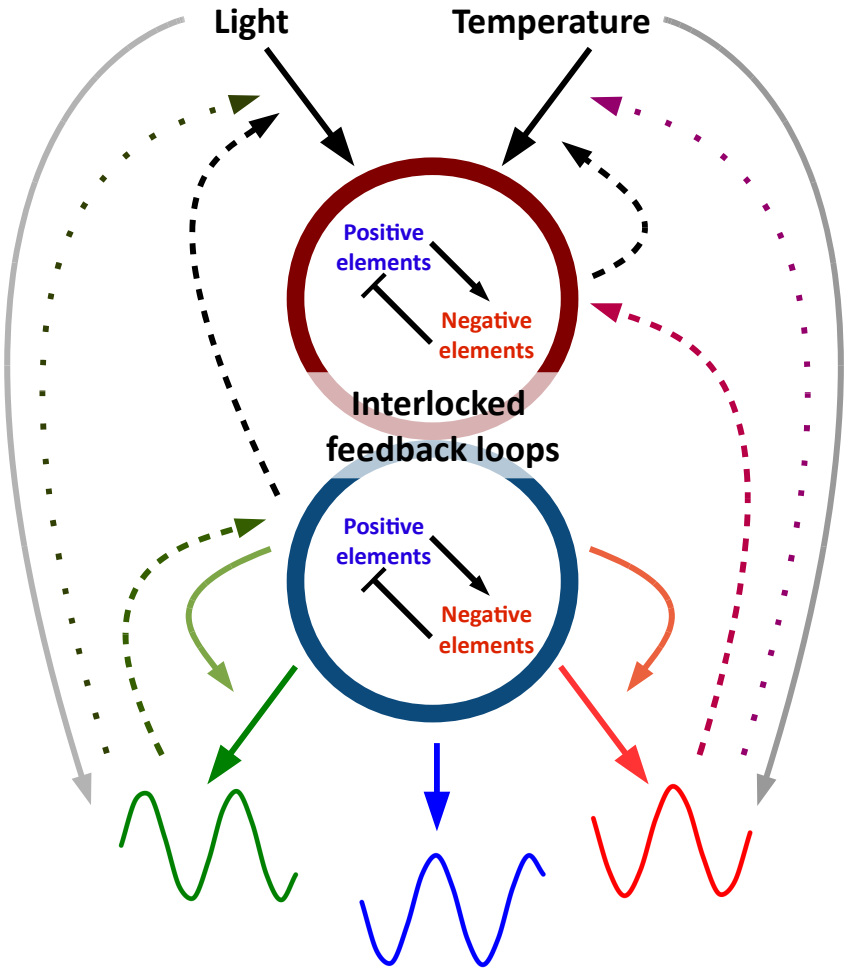


Figure 1.2: Model of circadian timekeeping
The circadian system consists of three components. The core oscillator is made up of interlocked negative feedback loops that can be entrained to environmental synchronizers (zeitgebers) through input pathways (solid black lines) to drive overt rhythms in output pathways (solid colored lines). Oscillator components may also act within input pathways (black dashed lines). Clock outputs may feed back to regulate the oscillator (colored dashed lines) and input pathways (colored dotted lines). Input pathways may also directly affect clock output pathways (solid grey lines).

Based on the observation that different taxonomic groups have unrelated clock proteins, multiple evolutionary origins for biological clocks have been proposed (Dunlap & Loros, 2004; Hardin, 2005; Iwasaki & Kondo, 2004; Lowrey & Takahashi, 2004; Mas, 2005; Bell-Pedersen et al., 2005; McClung, 2006; Panda & Hogenesch, 2004; O'Neill et al., 2011). The molecular circuitry of biological clocks, however, retains similar regulatory architecture and properties, whereby the clocks' autonomous oscillation is driven by multiple interlocking positive and negative feedback loops (Wijnen & Young, 2006). Plants, being non-motile organisms, rely on their internal timekeepers to integrate external cues in order to coordinate developmental processes according to a circadian schedule (Yanovsky & Kay, 2001). Indeed, effects of the clock are widespread in almost all aspects of plant development (Covington et al., 2008). Circadian rhythms in plants have been found to regulate rhythmic leaf movement (Millar et al., 1995), hypocotyl growth (Dowson-Day & Millar, 1999; Nozue et al., 2007; Nusinow et al., 2011), response to phytohormones (Covington & Harmer, 2007; Michael et al., 2008a; Legnaioli et al., 2009), day-length dependent (photoperiodic) flowering time (Fowler et al., 1999; Yanovsky & Kay, 2002; Imaizumi, 2010), pathogen defense (Wang et al., 2011), ultraviolet-B light responses (Frohnmeyer & Staiger, 2003), herbivory (Kerwin et al., 2011) and metabolism (Blasing et al., 2005; Fukushima et al., 2009; Graf et al., 2010). Many temporal physiological responses in plants are also dependent on the action of phytohormones; it has been shown that abscisic acid (ABA) activity is phased to mid-day when transpiration rate is the highest to ensure proper adjustments of stomatal closure (Legnaioli et al., 2009). It has also been shown that plants experience reduced growth in days longer than 24 h because productivity depends on the rate of carbohydrate utilization where it is adjusted according to time and starch content information (Graf et al., 2010). Indeed, one possible ecological reason for circadian timing of growth, where rapid growth occurred by the end of the night, is to allow temporal matching of this process to maximum water availability (Nozue et al., 2007). Collectively, the interplay between input pathways, the core oscillator and output pathways allows organisms to use this timekeeping mechanism to match internal processes according to the approximated conditions in the real world.

1.1.2 Circadian gating of environmental responses

Another inherent consequence of circadian timekeeping is that the extent of response to a stimulus of the same intensity varies according to the time at which the stimulus is perceived. This phenomenon, known as 'gating' (Carré, 2002), which adds yet another level of complexity to the circadian control of physiology as organisms can then better modulate their reactions to different environmental signals to further improve the synchronization of internal processes with rhythmic surroundings. Gating of a stimulus allows plants to respond only when it is advantageous and this may occur either through direct or indirect mechanisms (Fig. 1.3; Hotta et al., 2007) by regulating the availability or the abundance of metabolites and signaling molecules (Harmer et al., 2000). This is because, without gated responses, the clock may be vulnerable to constant resetting by fluctuations in light and temperature that prevents the clock from progressing. Furthermore, circadian gating ensures that an optimal amount of energy is expended for a particular response, which in turn will confer fitness to the organism. When clock outputs are part of the signaling pathway, gating occurs through direct mechanisms. On the other hand, if the clock regulates a single gating pathway that pathway in turn regulates other signaling pathway(s), gating occurs indirectly (Fig. 1.3). Processes that are gated by the clock include light signaling, growth, temperature responses, hormonal signaling and stomatal regulation (Fowler et al., 2005; Covington & Harmer, 2007; Legnaioli et al., 2009). As an example, the abundance of photoreceptors phytochromes (PHYs) and cryptochromes (CRYs) are clock controlled and this may result in gated light input to the oscillator (Harmer et al., 2000; Tóth et al., 2001). Also, light-induction of *CHLOROPHYLL A/B BINDING PROTEIN 2* (*CAB2*) expression coincided with the maximal and minimal levels of *CAB2* transcription in constant darkness (DD; Millar & Kay, 1996). Previously, it has been proposed that *EARLY FLOWERING* (*ELF*) *3* mediates the circadian gating of light responses by regulating light input to the clock because the null mutant *elf3* is arrhythmic for circadian outputs [*CAB2* and and *COLD CIRCADIAN RHYTHM RNA-BINDING 2* (*CCR2*)] in constant light (LL) but not in DD (Hicks et al., 1996; Covington et al., 2001). In *elf3*, *CAB2* transcription is constitutively activated in constant light and the phase of the *elf3* oscillator is always clamped at dusk, suggesting that the oscillator is arrested in the light and can be restarted upon light-to-dark transition (McWatters et al., 2000;

Covington et al., 2001). Circadian rhythms may have arisen as a result of evolutionary pressure; therefore, mechanisms to restrict responses to certain times of the day must be in place to ensure efficient energy utilization and to maximize carbon uptake for growth. Yet, it is possible that during extreme stresses, circadian gating mechanisms may be overridden by cellular demands as survival then becomes of paramount importance to the plant.

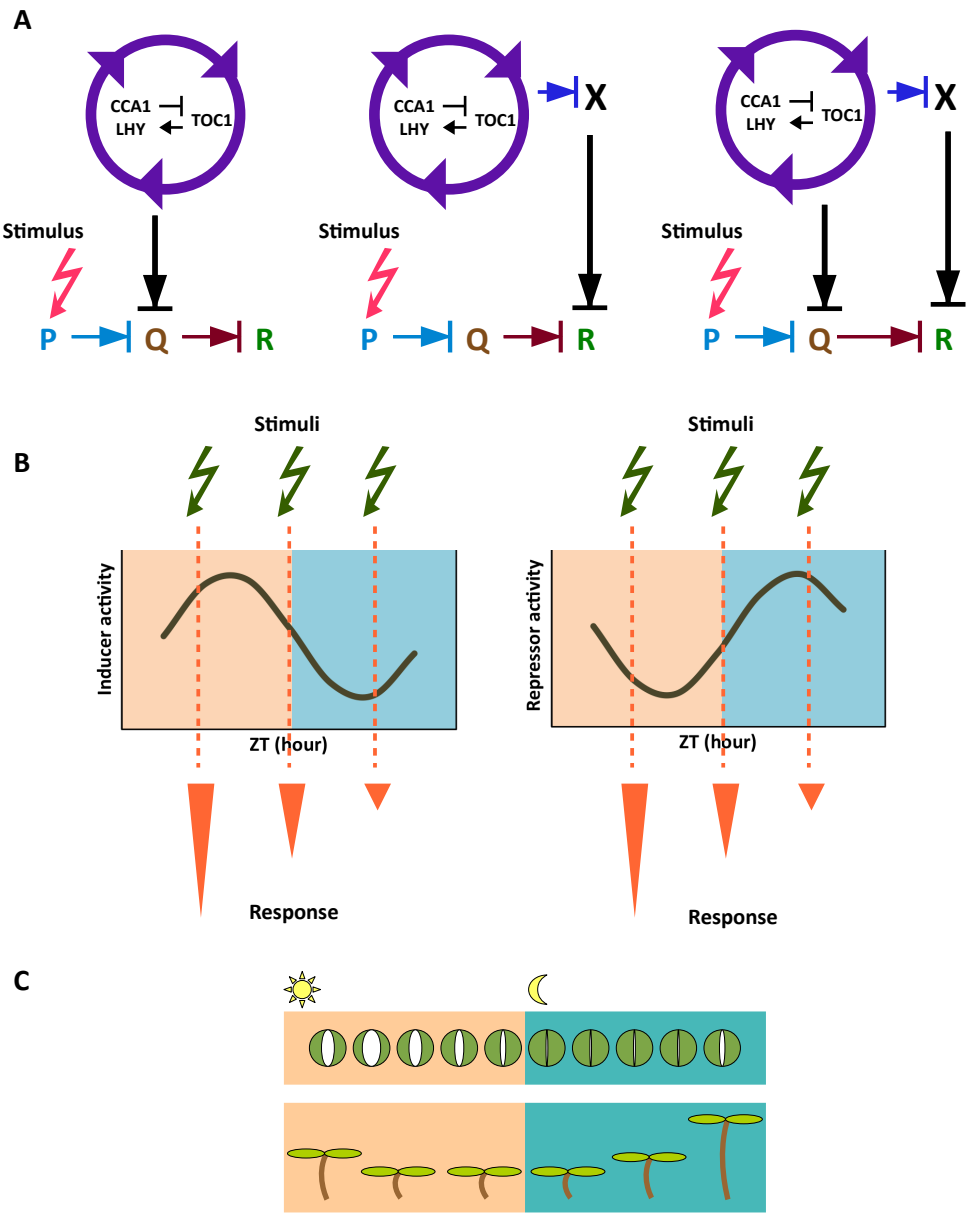


Figure 1.3: Circadian gating of environmental responses
(A) The oscillator gates responses to environmental stimuli through three plausible mechanisms. The model on the left shows that circadian gating occurs through the regulation of a pathway that is involved in transducing the stimulus. The model in the middle shows that a separate gating pathway regulates the stimulus-transducing pathway and the model on the right shows that both gating pathway and stimulus-transducing pathway act together to gate a response. The arrows with T-bars indicate that interactions between components could be positive or negative. (B) In addition, rhythmic regulation of any regulatory component within the pathways may be sufficient to confer gating even when stimulus of the same intensity is applied at different times of the day. This regulatory component may either be an activator (model on the leftl) or a repressor (model on the right) component of the pathway. When the level of the activator is the highest, the response is the greatest and vice versa. When the level of the repressor is the highest, the response is the smallest and vice versa. Indeed, similar patterns of gated responses are generated when the levels of activator and repressor are antiphasic to one another. (C) Examples of gated responses in stomata and hypocotyl elongation. Stimuli that promotes the opening and closing of stomatal aperture are gated by the clock. Rhythmic hypocotyl elongation represents another example of circadian gating of growth. Hypocotyl elongation is maximum during late night, suggesting that responses to acute changes in light is buffered and that growth occurs only in response to extended periods of darkness.

Modified from Hotta et al., 2007; Nozue et al., 2007.

1.1.3 Growing with time

With the biological clock, organisms have a built-in mechanism in place to enable the approximation of external time upon entrainment: a process that occurs when rhythmic physiological events match the environmental oscillation. The clock uses a combination of endogenous and exogenous signals to communicate time information to its outputs and this information is used to correctly phase the clock's targets (Dodd & Love, 2005). Time-keeping mechanisms allow predictions and adaptations to upcoming time-dependent environmental changes, e.g., changes in light and temperature intensities, where the periodicity in physiological and behavioral attributes of most organisms mirrors the periodicity in these environmental variables. Such coordination may increase the organism's fitness as energy is expended only on processes that are important at certain times of the day (Michael et al., 2003). It has been shown that organisms gain advantage when their endogenous period length matches the period of exogenous light-dark cycle. This phenomenon is known as 'circadian resonance' and may confer advantage to the organism by optimizing phase relations between external diel cycles and internal clock-controlled processes (Woelfle et al., 2004; Dodd et al., 2005; Hellweger, 2010). In plants, circadian resonance have been found to increase chlorophyll content, growth vigor, net photosynthetic carbon fixation, which can ultimately improve competitive advantage (Dodd et al., 2005). Likewise, mutant plants with dissonant clocks where their endogenous period lengths are either longer or shorter than 24 h were inferior to wild-type (WT) plants that have period length that matches the 24 h environmental period length (Dodd et al., 2005). Taken together, the ubiquitous nature of biological clocks at multiple levels of organization emphasizes the importance of having temporal resolution in metabolism and physiology to confer adaptive advantage (Green et al., 2002).

1.2 The clock's architecture

The initial proposal that the circadian network comprised of a single feedback loop has now been replaced by the current view that the clock of most organisms is built on multiple interlocked regulatory loops (Bell-Pedersen et al., 2005; Hamilton & Kay, 2006; Wijnen & Young, 2006). Although clock proteins have been shown to be structurally dissimilar among different taxonomic groups, molecular studies have deciphered conserved regulatory mechanisms that underlie clock function (Harmer et al., 2001). The common clock circuitry, with possible exception of the cyanobacteria clock that relies on oscillations in ATPase activity and protein phosphorylation (Tomita et al., 2005; Terauchi et al., 2007), is composed of interlocked autoregulatory transcriptional-translational feedback loops (TTFLs). TTFLs incorporate positive and negative components that regulate their own transcription as well as the expression of other oscillator components (Locke et al., 2005; Locke et al., 2006; Zeilinger et al., 2006).

To serve as accurate timekeepers, the clock's function is compartmentalized into three different modules; 1) the input module, where in plants, it is comprised of photoreceptors that perceive environmental stimuli, 2) the central oscillator that is comprised of transcriptional regulators and regulators of protein degradation that integrate environmental stimuli perceived by the input module to maintain the phase and period of the oscillator and 3) the output module that is responsible for transmitting temporal information to multiple output pathways (Harmer et al., 2009). Nevertheless, the actual molecular circuitry is actually far more complex than the idealized lineal pathway. This is because the oscillator may also regulate input modules and output modules may feed back to fine tune the sensitivity of the oscillator to environmental cues (Mas, 2008).

1.2.1 The *Arabidopsis* oscillator

Recently, intensive efforts to identify regulators of the *Arabidopsis* circadian circuitry began with large-scale genetic screens that employed luciferase (LUC) reporter lines of clock-controlled gene promoter fragments (Millar et al., 1992). Because of the short

half-life of the firefly LUC, they have been useful tools for the real-time monitoring of clock-driven transcriptional changes (Welsh et al., 2005). Indeed, promoter elements of the *CAB2* and *CCR2* genes were routinely used to drive LUC expression in genetic screens for clock regulators to identify mutants with altered period length under LL conditions (Millar et al., 1992; Strayer et al., 2000).

The input module of the *Arabidopsis* oscillator consists of two groups of photoreceptors, the PHYs and CRYs (Franklin et al., 2005). *Arabidopsis* has five *PHY* genes, *PHYA* to *PHYE* (Nagy et al., 2002; Quail, 2002) and two *CRY* genes, *CRY1* and *CRY2* (Lin, 2002). Apart from light, temperature can also entrain the clock (Gould et al., 2006), although the exact molecular mechanism of this remains to be fully investigated.

1.2.2 Myb transcription factors and a pseudo-response regulator form the core loop

The central oscillator consists of a core feedback mechanism that connects the morning- and evening-phase circuits (Fig. 1.4). The left arm of the core loop is made up of two morning-expressed, partially redundant Myb transcription factors (TF), the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*) that inhibit the expression of an evening-expressed pseudo-response regulator *TIMING OF CAB EXPRESSION 1* (*TOC1*) on the right arm of the loop through the association of CCA1 to the Evening Element (EE; AAAATATCT) motif in *TOC1* promoter (Wang & Tobin, 1998; Schaffer et al., 1998; Strayer et al., 2000; Alabadi et al., 2001, Harmer, 2009). In addition to the EE motif, DNA-binding activity of CCA1 to another motif known as the CCA1-binding site (CBS; AAAAATCT) has been characterized (Wang et al., 1997). Although the amino acid sequence of *LHY* suggests homology to the MYB family of TFs, *LHY* encodes an unusual member because its DNA-binding domain only comprised of a single MYB repeat instead of two or three repeats (Jin & Martin, 1999). Nevertheless, both LHY and CCA1 have essentially identical DNA-binding domains (Schaffer et al., 1998).

On the right arm of the core loop, *TOC1* functions as a transcriptional activator of *CCA1* and *LHY* (Alabadi et al., 2001; Makino et al., 2002). The expression of *TOC1* is antiphasic to that of *CCA1* and *LHY* where nocturnal TOC1 accumulation could induce the expression of both Myb TFs (Alabadi et al., 2001). The repression of *TOC1* expression at dawn by *CCA1* and *LHY* resulted in the decline in *TOC1* abundance during the day and consequently reducing the promotion of *CCA1* and *LHY*. By dusk, *CCA1* and *LHY* levels would have diminish and the repression on *TOC1* expression was then lifted. This results in TOC1 accumulation at night as the cycle repeats itself. Previously, there has been little biochemical data supporting the mechanism for TOC1 in *CCA1* and *LHY* reactivation as no DNA-binding domains could be found in TOC1 (Strayer et al., 2000). However, In a recent study, it has been proposed that TOC1 actually occupies specific genomic regions in the promoters of *CCA1* and *LHY* and that TOC1 binds DNA through a proposed DNA-binding domain known as the [CONSTANS (CO), CO-like, TOC1] CCT domain (Gendron et al., 2012). Furthermore, an additional factor is also involved in the co-regulation of *CCA1* and *LHY* by TOC1. This factor has been identified as *CCA1 HIKING EXPEDITION* (*CHE*), a member of the TB1, CYC and PCF (TCP) transcription factor family (Cubas et al., 1999; Pruneda-Paz et al., 2009). The CHE protein interacts with TOC1 and this complex binds to the promoter of *CCA1* but not *LHY*, to repress the transcription of *CCA1* and *CHE* itself (Pruneda-Paz et al., 2009).

1.2.3 *PSEUDO-RESPONSE REGULATORS (PRR) 7, PRR9* and *PRR5* take the morning shift

The oscillator's circuitry involves two additional phase-specific feedback loops, the morning and the evening loop (Farre et al., 2005; Zeilinger et al., 2006; Locke et al., 2006). The morning loop consists of *TOC1* homologs, the *PSEUDO-RESPONSE REGULATOR* (*PRR*) *7* and *PRR9*, where both are partially redundant in repressing the transcription of *CCA1* and *LHY* (Farre et al., 2005). *PRR7* and *PRR9* are in turn, activated by *CCA1* and *LHY* (Farre et al., 2005; Zeilinger et al., 2006). The mechanism of this regulation has been proposed to occur through the interaction of PRR7 and PRR9 with transcription factors bound to promoters *CCA1* and *LHY* (Pruneda-Paz & Kay, 2009). In addition, *PRR5* has been added to the morning loop and along with *PRR7* and *PRR9*, functions as transcriptional repressor of *CCA1* and *LHY* (Nakamichi et al., 2010).

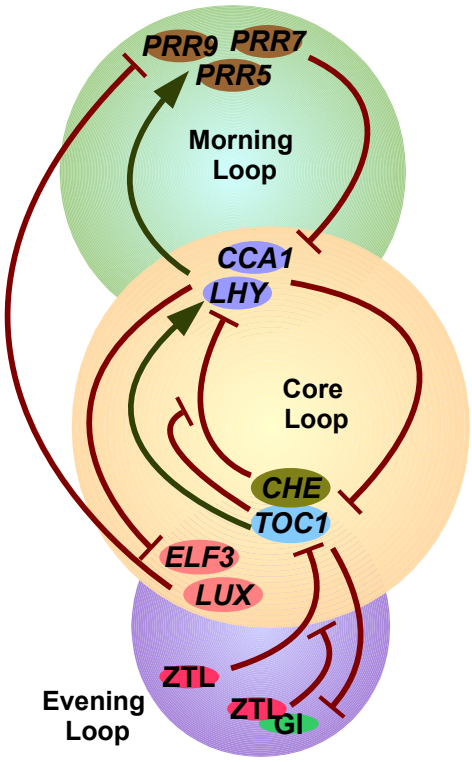


Figure 1.4: Schematic representation of the *Arabidopsis* genetic circuit
The core feedback loop contains Myb transcription factors *CCA1* and *LHY*, which peak in the morning and negatively regulate the expression of *TOC1*, which peaks in the evening. TOC1 subsequently promotes the expression of *CCA1* and *LHY* through an interaction with CHE. CHE represses *CCA1* expression and *CHE* expression is in turn inhibited by *CCA1*. TOC1 antagonizes the binding of CHE to the *CCA1* promoter. In the morning loop, *CCA1* and *LHY* promote the expression of *PRR9*, *PRR7* and *PRR5*, which reciprocally repress *CCA1* and *LHY* expression. The evening loop contains *TOC1* and *GI*. *TOC1* represses *GI* expression and GI in turn activates *TOC1* expression. TOC1 is degraded by ZTL and this degradation is inhibited by GI through its association with ZTL. Other components of the evening loop include *ELF3* and *LUX,* in which their expressions are inhibited by *CCA1* and *LHY*. *LUX* represses *PRR9* and *ELF3* represses *PRR7* and *PRR9*.

1.2.4 GIGANTEA and three F-box proteins take the evening shift

The evening loop involves *TOC1* as a negative regulator of *GIGANTEA* (*GI*) and another putative component 'Y' that feeds back to activate *TOC1* transcription (Locke et al., 2006). The role of *GI* in temperature compensation has been implicated, where *GI* is critical for increasing the temperature range that is permissive for accurate rhythmicity, i.e., the range at which *CAB2* rhythms could be maintained (Gould et al., 2006). Indeed, the dynamic balance between *GI* and *LHY* is needed for effective temperature compensation at high temperatures (28°C), whereas at low temperatures

(12°C) *CCA1* replaces *LHY* in this buffering mechanism (Gould et al., 2006). In addition to its role in the circuitry, *GI* has been found to regulate flowering time (Fowler et al., 1999; Martin-Tyron et al., 2007; Sawa & Kay, 2011) and responses to oxidative stress (Kurepa et al., 1998a). Post-translational regulation is also part of the clock's circuitry in addition to transcriptional feedback regulation. The role of photoreceptor F-box proteins, FLAVIN BINDING KELCH REPEAT F-BOX 1 (FKF1), LOV KELCH PROTEIN 2 (LKP2) and ZEITLUPE (ZTL) have been implicated to regulate TOC1 and PRR5 stability (Baudry et al., 2010). ZTL, a blue light photoreceptor with E3 ubiquitin ligase activity, is part of the Skp1-Cullin-F-box (SCF) complex that is involved in the proteasome-mediated protein degradation pathway (Han et al., 2004). ZTL targets the degradation of TOC1 and PRR5 at night (Kiba et al., 2007). GI participates in this regulation by binding to the F-box proteins. Evening-expressed GI associates with ZTL in a light-dependent manner and confer rhythms in ZTL protein abundance by post-transcriptional regulation (Kim et al., 2007; Sawa et al., 2007). GI also interacts with ZTL to regulate the proteasomal degradation of PRR5 (Kiba et al., 2007). Furthermore, it has been shown that GI and PRR3 regulate the ZTL-mediated degradation of TOC1 and PRR5 (Fujiwara et al., 2008); where in the evening, PRR3 interacts with TOC1 to protect TOC1 from ZTL-mediated degradation (Para et al., 2007).

1.2.5 Additional indispensable clock regulators

The reciprocal regulation of *CCA1-LHY-TOC1* alone cannot account for all rhythmicity in *Arabidopsis*. Other members of the *CCA1*, *LHY* Myb TF family includes the *REVEILLE* genes (*RVE1* to *RVE8*; Kuno et al., 2003; Yanhui et al., 2006). With the exception of *RVE5*, all other *RVEs* have been shown to bind the EE (Gong et al., 2008), which suggests their partially redundant function to *CCA1* and *LHY*. In addition, *RVEs* are also involved in various aspects of plant development. *RVE1* has been found to control growth by regulating the auxin signaling pathway (Rawat et al., 2009). *RVE2* and *RVE7* are directly involved in the oscillator's function. Constitutive expression of *RVE2* reduced the amplitude and shortened the period of *CCA1* and *LHY* (Zhang et al., 2007) while constitutive expression of *RVE7* repressed the *LIGHT-HARVESTING CHLOROPHYLL A/B-BINDING PROTEIN* (*LHCB*) expression without affecting *CCA1* or *LHY* rhythms (Kuno et al., 2003). On the other hand, *RVE8* has been proposed to be part of a

negative transcriptional feedback loop within the circuitry that functions to set the pace of the clock, affects temperature compensation and light signaling (Rawat et al., 2011).

Among the other regulators involved in proper clock function are the *ELF3*, *ELF4*, *LUX ARRHYTHMO* (*LUX*), *BROTHER OF LUX ARRHYTHMO* (*BOA*), *FIONA 1* (*FIO1*), *TIME FOR COFFEE* (*TIC*), *PROTEIN ARGININE METHYL TRANSFERASE 5* (*PRMT5*), *JUMONJI DOMAIN PROTEIN 5* (*JMJD5*), *XAP5 CIRCADIAN TIME-KEEPER* (*XCT*), *LIGHT-REGULATED WD 1* (*LWD1*) and *LWD2* (Hicks et al., 1996; Doyle et al., 2002; Hall et al., 2003; Hazen et al., 2005; Kim et al., 2008; Martin-Tryon et al., 2008; Wu et al., 2008; Hong et al., 2010; Jones et al., 2011; Dai et al., 2011). ELF4-ELF3-LUX forms an evening complex (EC) that links the clock to diurnal regulation of growth (Nusinow et al., 2011). The EC is required for accurate *CCA1* and *LHY* expression; *elf3*, *elf4* and *lux* mutants have decreased expression of *CCA1* and *LHY* (Doyle et al., 2002; Hazen et al., 2005; Onai & Ishiura, 2005; Helfer et al., 2011). Moreover, the EC is essential for maintaining rhythms under free-running conditions since mutations in each of these genes resulted in circadian arrhythmia (Hicks et al., 1996; Doyle et al., 2002, Hazen et al., 2005). The functions of *ELF3*, *ELF4* and *LUX* extend into the morning loop, where *ELF3* and *ELF4* act as transcriptional repressors of *PRR7* and *PRR9* while *LUX* acts as a nighttime repressor of *PRR9* (Dixon et al., 2011; Helfer et al., 2011; Kolmos et al., 2011).

The presence of multiple redundant regulators in the clock's circuitry suggests the robustness of the system. In *Arabidopsis,* rhythmicity is retained in single knockouts of any clock genes and many double knockouts are still rhythmic albeit with altered period and phase. Because of this genetic redundancy, clocks of most organisms are robust to perturbations. Indeed, to buffer against the loss of any single clock component and environmental noise, oscillations are sustained (accurate period, phase and amplitude) upon perturbations but yet amenable to re-entrainment by environmental signals (Hogenesch & Herzog, 2011; McWatters & Devlin, 2011).

1.3 The oxygen paradigm

1.3.1 Reactive oxygen species (ROS) and the oxygen paradox

Synchronization between external conditions and internal metabolism to occur at specific times exists to allow temporal separation of incompatible metabolic events (Mas et al., 2005; Hotta et al., 2007). As plants undergo aerobic metabolism, e.g., photosynthesis and respiration, they constantly face challenges from the by-products of molecular oxygen (O_2), collectively known as ROS. Plant organelles that have high oxidizing metabolic activity e.g., chloroplast, mitochondria and peroxisomes, are major sites of ROS production. In addition, detoxifying reactions by cytochromes in the endoplasmic reticulum and the cytoplasm may also generate ROS (Urban et al., 1997). Generation of ROS during photosynthesis occurs either through direct photoreduction of O_2 by electron transport components of PSI or through photorespiratory reactions involving RuBisCO in the chloroplast (Apel & Hirt, 2004). The photoreduction of O_2 to H_2O gives rise to singlet oxygen ($^1O_2$), superoxide anion (O^{2-}), hydrogen peroxide (H_2O_2) and hydroxyl radical (OH^-; Fig. 1.5; Foyer & Noctor, 2005). Chloroplasts produce ROS through the Mehler reaction in the antenna pigments (Asada and Takahashi, 1987) especially under conditions of limiting carbon dioxide fixation, which also activates the photorespiratory pathway. During photorespiration, H_2O_2 is generated in peroxisomes by glycolate oxidase (Foyer, 2002). The over-reduction of the mitochondria electron transport chain generates the main source of O^{2-} (Møller, 2001). Perhaps the most-studied enzymatic complex involved in the generation of ROS is the NADPH oxidase (NOX; Sagi & Fluhr, 2006). NOXs in plants are the respiratory burst oxidase homologue (RBOH) proteins, which constitute a multigenic family of ten *RBOH* genes in *Arabidopsis* (Torres & Dangl, 2005).

Under steady state conditions, ROS are efficiently scavenged by various antioxidant components (Alscher et al., 1997). Due to their reactive nature as lethal oxidants, ROS can potentially be dangerous when overproduced in organelles such as the mitochondria, chloroplasts and peroxisomes (D'Autréaux & Toledano, 2007; Rosenwasser et al., 2011). However, environmental stressors may perturb the equilibrium between ROS production and scavenging (Malan et al., 1990). The rapid

increase in ROS levels is known as 'oxidative burst' (Apostol et al., 1989). If ROS levels are left unmanaged, plants undergo oxidative stress when the overproduction of ROS causes an imbalance in cellular redox states that may eventually lead to cell death through damage inflicted on lipids, nucleic acids and cellular proteins (Apel & Hirt, 2004; Gechev et al., 2006; Møller et al., 2007; D'Autréaux & Toledano, 2007).

Prasad et al. (1994) have shown that chilling could impose oxidative stress in maize seedlings through the elevation of H_2O_2 during both the acclimation and chilling of non-acclimated seedlings. As a consequence, plants have evolved enzymatic and non-enzymatic scavenging machineries to maintain redox homeostasis to keep ROS at physiologically permissive levels (Mullineaux & Karpinski, 2002; Mittler, 2002; Overmyer et al., 2003; Mittler et al., 2004; Halliwell, 2006). Furthermore, to protect photosynthetic apparatus against ROS-induced photoinhibition, plants rely on photochemical and non-photochemical quenching mechanisms (Ort & Baker, 2002). Chilling tolerance is achieved by pre-treatment of maize seedlings with H_2O_2 or menadione at 27°C, possibly by the induction of antioxidant enzymes such as guaiacol peroxidases and catalases (CAT; Prasad et al., 1994). ROS is also generated by the activation of peroxidases and oxidases that produce ROS in response to environmental perturbations (Doke, 1985; Allan and Fluhr, 1997).

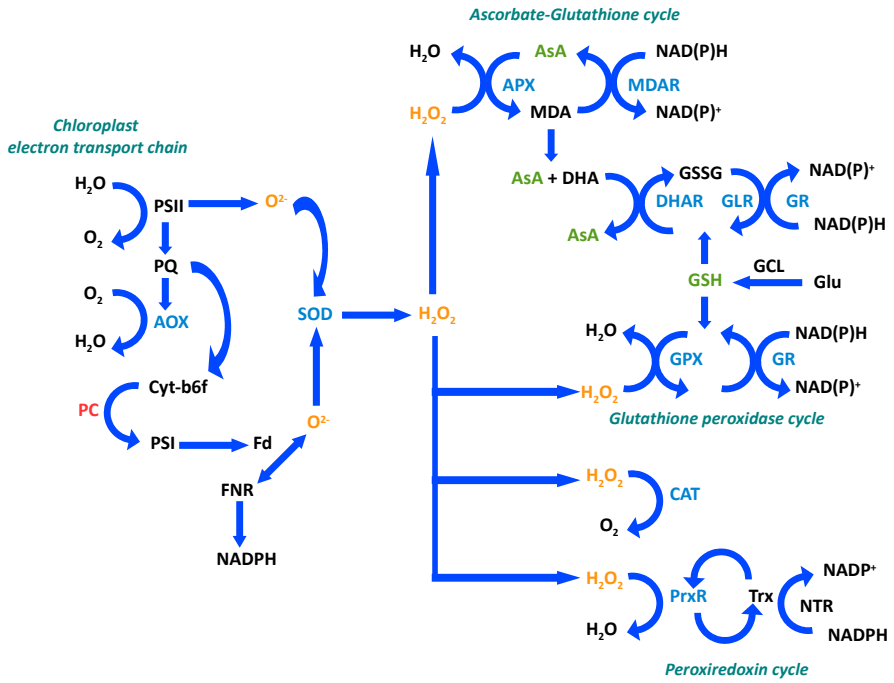


Figure 1.5: Schematic diagram depicting modes of ROS generation and scavenging in plants
AOX is involved in ROS detoxification in the photosynthetic electron-transport chain. Electrons from the photosynthetic apparatus are used by AOX to reduce O_2 to H_2O. In the first line of defense, SOD converts O^{2-} into H_2O_2, which is followed by the detoxification of H_2O_2 by APX, GPX, CAT and PrxR. In the ascorbate-glutathione cycle, H_2O_2 is converted into water. Ascorbate is oxidized into MDA by APX and MDA is then reduced into ascorbate by MDAR in the presence of NAD(P)H. DHA is produced by MDA and is reduced by DHAR into ascorbate in the presence of GSH that is oxidized to GSSG. GSSG is then converted back to GSH by GR in the presence of NAD(P)H. In the glutathione peroxidase cycle, H_2O_2 is converted into water by reducing equivalents from GSH. Oxidized GSSG is subsequently converted back to GSH by GR in the presence of NAD(P)H. H_2O_2 is also detoxified into water by CAT and PrxR. ROS are indicated in orange, ROS-scavenging enzymes in blue and antioxidants in green. Abbreviations: AOX: alternative oxidase, APX: ascorbate peroxidase, AsA: ascorbate, CAT: catalase, Cytb6f: cytochrome b6f, DHA: dehydroascorbate, DHAR: DHA reductase, Fd: ferredoxin, FNR: ferredoxin NADPH reductase, GCL: glutamate cysteine ligase, GLR: glutaredoxin, Glu: glutamate, GPX: glutathione peroxidase, GR: glutathione reductase, GSH: glutathione, GSSG: oxidized glutathione, MDA: monodehydroascorbate, MDAR: monodehydroascorbate reductase, NTR: NADPH thioredoxin reductase, PC: plastocyanin, PQ: plastoquinone, PrxR: peroxireductase, PSI: photosystem I, PSII: photosystem II, SOD: superoxide dismutase and Trx: thioredoxin.

Modified from Mittler, 2002; Apel & Hirt, 2004.

1.3.2 Enzymatic and non-enzymatic detoxification of ROS

ROS detoxification occurs through enzymatic and non-enzymatic scavenging mechanisms. Enzymatic scavengers of ROS include ascorbate peroxidase (APX), CAT, glutathione peroxidase (GPX) and superoxide dismutase (SOD; Fig. 1.5). The equilibrium in ROS homeostasis (production and scavenging) is determined by balance in CAT, APX and SOD activities, which will in turn determine the levels of cellular H_2O_2,

O^{2-} and OH^{-}. Perturbations in the equilibrium of enzymatic scavengers may induce compensatory mechanisms. For instance, reducing CAT activity may upregulate the activity of APX and GPX (Rizhsky et al., 2002). Catalase is encoded by a multigene family in *Arabidopsis* that forms at least six isozymes. Individual catalase isozymes display distinct patterns of organ specific expression, where six isozymes are found in leaves and flowers and two are detected in roots (Frugoli et al., 1996).

On the other hand, the non-enzymatic scavengers of ROS are glutathione (GSH), ascorbate, tocopherol, carotenoids and flavonoids. It has been shown that mutants with diminished levels of reduced ascorbate and GSH were hypersensitive to stress (Conklin & Williams, 1996; Creissen et al., 1999). Similar to that of the enzymatic scavengers, balance among different non-enzymatic scavengers must also be tightly regulated to buffer against changes in cellular redox states. Perturbations in ROS homeostasis due to enhanced glutathione biosynthesis in the chloroplast have been found to cause cellular oxidative damage (Creissen et al., 1999). In addition, the overexpression of the antioxidant ß-carotene hydroxylase has been found to increase oxidative stress tolerance under high light conditions (Davison et al., 2002).

1.3.3 Non-toxic levels of ROS is essential for plant development

At low non-toxic levels, however, ROS are not exclusively deleterious and may have signaling functions. ROS play key roles in multiple aspects of plant development such as growth (Gapper & Dolan, 2006), stomatal closure, cell proliferation (Dunand et al., 2007; Tsukagoshi et al., 2010), programmed cell death (Torres et al., 2006), senescence (Miao et al., 2004; Park et al., 2004; Schippers et al., 2008) and pathogen defense (Grant & Loake, 2000; Apel & Hirt, 2004). Initially regarded as toxic byproducts, ROS have now been recognized as players in cell signaling networks (Mittler et al., 2011). Because of the link between metabolic processes and ROS homeostasis, it is advantageous for plants to utilize ROS as signals to control different physiological processes. Since external conditions could alter the equilibrium between production and scavenging rates, it could alter ROS levels and the intensity of the generated signals. A burst of ROS triggers a cascade of events through the propagation of ROS signals over long distances, which affects many downstream processes (Nishimura &

Dangl, 2010; Mittler et al., 2011). Moreover, ROS signals can also be blocked by local application of scavengers (catalase or NADPH oxidase inhibitor) at distances of 5 to 8 cm away from the signal initiation site (Miller et al., 2009) and this suggests the versatility of ROS in signaling networks.

1.3.4 The dynamics of ROS signaling network

ROS signal transduction network is evolutionarily conserved in aerobic organisms (Mittler et al, 2011). Cells utilize this network to maintain non-toxic steady state levels of ROS while allowing ROS to act as signaling molecules when ROS transiently accumulate in subcellular spaces (Mittler et al., 2004). Key components in plants' ROS signal transduction pathway have been identified, although the actual receptors for ROS are presently unknown. Plant cells may engage in ROS sensing through three distinct mechanisms: 1) redox transcription factors, 2) receptor proteins or 3) inhibition of phosphatases (Rhee et al., 2000; Orozco-Cardenas et al., 2001). Other downstream ROS signaling events involve the activation of G-proteins (Baxter-Burrell et al., 2002), the activation of phospholipid signaling (Anthony et al., 2004) and the role of calcium-binding proteins (Coelho et al., 2002). A serine/threonine protein kinase has been found to participate in ROS sensing through the activation of mitogen-activated protein kinases (MAPK) by the calcium ion (Rentel et al., 2004). This MAPK cascade controls the activation of defense machineries during ROS-induced stress (Kovtun et al., 2000).

Taken together, there are several advantages of utilizing ROS as signaling molecules (Mittler et al., 2011); 1) Each individual cell autonomously activates its own ROS generation and rapidly propagates ROS signals over long distances to different plant organs. Propagation rate of 8.4cm/min has been reported in *Arabidopsis* (Miller et al., 2009), 2) cells can induce rapid dynamic changes in ROS homeostasis by altering the equilibrium between production and scavenging rates of different ROS forms, 3) since plants exert tight control over subcellular localization of ROS signals, limiting ROS production to specific cellular locations such as the organelle or cell membrane could therefore account for spatial-specific signaling functions (Monshausen et al., 2007; Takeda et al., 2008), 4) signaling versatility with regards to ROS mobility within cells

could be enhanced since different ROS forms have different molecular properties. O^{2-}, in its native charged form, could not passively cross the cell membrane. However, once converted to H_2O_2, it can undergo passive transport through water channels (Miller et al., 2010) and lastly 5) ROS signaling is tightly linked to cellular metabolism. Alterations in metabolism could cause changes in ROS homeostasis, which in turn allow plants to use ROS signals to monitor such changes and to fine tune biological processes. Indeed, increase in photorespiration could enhance ROS production in peroxisomes (Vanderauwera et al., 2011).

From the phylogenetic study of the *Arabidopsis* ROS gene network, the evolutionary pathway of ROS signaling could be inferred by comparative genomics with other members of the plant lineage, e.g., poplar, rice, grapevine, millet, green algae and moss, using the PLAZA tool (Proost et al., 2009). This method allows the reconstruction of ancestral ROS genes and the tracing of the origin of the genes. Reverse genetic screens have identified knockout lines that revealed the linkage between ROS signaling with growth and development. The knockdown of thylakoid-attached copper-zinc SOD (KD-SOD) has suppressed expression of water-water cycle enzymes that are required for maintaining electron flow through photosynthetic apparatus. Defects in photoprotection of the chloroplast from oxidative stress are manifested by reduced chloroplast size, photosynthetic activity and chlorophyll content in the plants (Rizhsky et al., 2003). Deficiency in H_2O_2 scavenging APX1 resulted in suppressed growth, altered stomatal responses and elevated induction of heat shock proteins during light stress (Pnueli et al., 2003). Furthermore, manipulation of the expression of programmed cell death genes could alter the superoxide-dependent cell death phenotype and normal hypersensitive response in plants (Epple et al., 2003). Although each knockdown mutant exhibits distinct phenotypes associated with the gene functions, the aforementioned lines are nevertheless viable, which indicates the redundancy of the ROS network.

The most well studied effects of ROS signals are on hormonal signaling networks. ROS affects auxin homeostasis through the regulation of auxin catabolism, transport and redistribution by altering transcription and cellular location of PIN-FORMED (PIN)

proteins (Gazarian et al., 1998; Jansen et al., 2001; Santelia et al., 2008; Grunewald & Friml, 2010). In turn, auxin can alter antioxidant levels by triggering cell-specific ROS generation (Joo et al., 2001; Pignocchi et al., 2006) and affecting the transcription of ROS-responsive genes (Huang et al., 2008; Tognetti et al., 2010). Indeed, the regulators of redox homeostasis such as NADP-linked thioredoxin (NTRX) and glutathione have been shown to alter auxin metabolism and transport (Bashandy et al., 2010). In addition, elevated ROS levels could result in salicylic acid (SA) accumulation (Chamnongpol et al., 1998) and affects SA-induced stomatal closure (Khokon et al., 2011). Reciprocally, inhibition of SA biosynthesis inhibits ROS-induced defense responses (Chaouch & Noctor, 2010). Gibberellin (GA) is also involved in the induction of genes encoding the redox-regulator GIBBERELLIC ACID STIMULATED TRANSCRIPT1 (GAST1)-like protein (Rubinovich & Weiss, 2010). Similarly, DELLA proteins of the GA signaling pathway is also proposed to regulate the transcription of antioxidant enzymes (Achard et al., 2008). In addition, ROS signaling is also integrated with many other signaling networks in plants, which include the protein phosphorylation, calcium and cellular redox networks through associations with peroxiredoxins, thioredoxins and glutaredoxins (Dietz et al., 2010; Rouhier, 2011; O'Neill & Reddy, 2011); these collectively demonstrate the intricate interactions between ROS networks, plant growth and the environment.

1.3.5 ROS as signals for transcriptional coordination

The role of ROS in transcriptional regulation has also been implicated (Mittler et al., 2004); ROS signals generated in organelles can diffuse into the nucleus to affect the expression of transcription (Gadjev et al., 2006). A dynamic gene network consisting of ROS-generating and ROS-scavenging proteins have been proposed where this network receives signals released from perturbations in ROS homeostasis. Intensities of different ROS signals at a specific time can affect the network through the induction of signaling cascades that affect the transcription of ROS-responsive genes (Mittler et al., 2004; Apel & Hirt, 2004). Transcriptomic analyses have revealed the response specificity against different ROS signals (Davletova et al., 2005; Gadjev et al., 2006; Scarpeci et al., 2008). For example, it is shown that H_2O_2 and O^{2-} responsive transcripts largely overlap with one another, which suggests that both ROS forms participate in

the same signaling cascade because most O^{2-} are catalytically or spontaneously dismutased to H_2O_2 (Davletova et al., 2005; Scarpeci et al., 2008).

1.4 Circadian regulation of other stress-responsive pathways and the ROS network?

The equilibrium in ROS homeostasis is not only affected by factors such as temperature and light intensity but also abiotic stresses (Malan et al., 1990; Prasad et al., 1994; Tsugane et al., 1999). Evidence for circadian gating of stress-responsive pathways has been discovered (Legnaioli et al., 2009). For example, in response to abiotic or biotic stresses, rapid wound-responsive genes are induced and these genes are also co-regulated by the circadian clock (Walley et al., 2007). In addition, the clock also controls the expression of temperature-responsive genes. The components of temperature-induced stress pathway, either from extreme heat or freezing, are integrated with clock function (Harmer et al., 2000; Bieniawska et al., 2008). Furthermore, the expression patterns of freezing tolerance transcription factors *CRT/DRE BINDING FACTORS* (*CBF1, 2* and *3*) and cold-induced genes such as *COLD REGULATED GENE 27* (*COR27*), zinc-finger transcription factor (*ZAT12*) and *RAV* family transcription factors (*RAV1*) are rhythmically expressed and are gated by the circadian clock (Bieniawska et al., 2008; Mikkelsen & Thomashow, 2009). Diurnal sensitivity to cold is, in part, mediated by the interaction of *TOC1* with *PHYTOCHROME INTERACTING FACTOR 7* (*PIF7*) to repress *CBF3* expression (Kidokoro et al., 2009). The role of *PRR5*, *PRR7* and *PRR9* has also been implicated in freezing-tolerance where the arrhythmic triple mutant *prr5/prr7/prr9* displayed upregulation of cold-responsive genes, elevated levels of antioxidants and tricarboxylic acid (TCA) cycle intermediates and enhanced freezing-tolerance (Nakamichi et al., 2009; Fukushima et al., 2009). Regulation of stress-responsive pathways is, therefore, likely to occur through signals transduced from the oscillator by core clock components to stress pathways (de Montaigu et al., 2010). Moreover, functional clustering of clock-regulated transcripts has revealed clock-regulated genes as overrepresented in stress-responsive pathways (Covington et al., 2008). This corroborated the observation that many metabolic and stress-responsive pathways are under circadian control.

As plants can utilize ROS as signaling molecules (Mittler, 2002; Mittler et al., 2004; Foyer & Noctor, 2005), it may therefore be essential for ROS homeostasis to be in tune with plants' photosynthetic activities and the daily light dark cycles in order to enhance

productivity and fitness. However, the underlying mechanisms and the biological importance of restricting stress responses to certain times of the day have not been fully elucidated. Conceivably, as the continuous signaling of stress-responsive components is metabolically demanding and may be deleterious to growth, gated stress responses and clock-regulated anticipation of stressful events may confer maximal tolerance to stress while minimizing the use of plants' resources. It is plausible that two mechanisms of ROS signaling exist in plant; 1) ROS signals are spatially compartmentalized (Miller et al., 2009) and 2) the circadian clock conveys temporal information to ROS networks and utilize such information to modulate ROS homeostasis as hypothesized in this work. Taken together, it is intuitive to suggest that ROS sensing is anticipatory and that the circadian clock coordinates ROS homeostasis according to the environment's circadian schedule.

1.5 Aims of this research:

Circadian clock research can be categorized into three complementary fields: 1) the core clock mechanism or architecture, 2) processes that are controlled by the clock (outputs) and 3) processes that control the clock (inputs). Time information is communicated to clock outputs at the molecular and physiological level where this information is converted to the correct phase for the output target based on endogenous and exogenous signals. This thesis aims to investigate and establish hypothetical mechanisms for signaling of circadian time to a proposed clock-controlled output: the ROS signaling network. Two hypotheses can be put forward in this research: 1) Temporal coordination in ROS signaling, homeostasis and the response to oxidative stress occurs and 2) The mechanistic link between ROS signaling and the circadian clock exists and is fine-tuned.

To test these hypotheses, the model plant *Arabidopsis thaliana* was used in this study to address the seven key objectives stated below:

1) Determine whether ROS homeostasis and signaling is controlled by diurnal day-night transitions or by the circadian oscillator. To address this question, temporal changes in ROS production (H_2O_2) and scavenging (catalase) were quantified and the temporal expression profiles of ROS-responsive genes were obtained.

2) Determine whether a functional clock is required for regulated ROS homeostasis. To address this question, ROS hypersensitivity assays were performed on plants with mutated version of clock genes. In addition, H_2O_2, catalase levels and temporal expression profiles of *CATALASE* (*CAT*) genes in plants having mutated core clock genes (*CCA1* and *LHY*) were obtained.

3) Determine the mode of clock-imposed transcriptional regulation of ROS-responsive genes. To address this question, bioinformatics approaches were used. Promoters of ROS-responsive genes were queried for the enrichments of the circadian-regulated EE and CBS. All *Arabidopsis* genes categorized under different ROS-

related Gene Ontology (GO) categories (referred to as ROS-GO genes hereafter) were queried for time-of-day specific phasing by the determination of phase enrichments.

4) Determine whether the clock regulates an anticipatory response to oxidative stress. To address this question, temporal expression profiles of ROS-responsive genes were compared between WT and clock mutant plants grown under non-stressed conditions.

5) Determine whether *CCA1* is a master regulator of plants' response to oxidative stress. To address this question, plants were treated with ROS-inducing agent at different times of the day to investigate whether the response is gated by diurnal cycles and whether the response is dependent on the time of *CCA1* expression. In addition, the effect of *CCA1* over-expression on the temporal response to ROS was also studied.

6) Determine whether transcriptional coordination of ROS-responsive genes occurs through the *in vivo* association of CCA1 to promoters of ROS genes. To address this question, Chromatin immunoprecipitation quantitative real-time PCR (ChIP-qPCR) was performed to identify enrichments in CCA1-bound promoter fragments of ROS-responsive genes.

7) Determine whether ROS signals could feed back to affect the oscillator and other clock-regulated output pathways. To address this question, luminescence imaging was performed using transgenic reporter lines. These reporter lines harbor the firefly LUC gene fused to promoters of clock-regulated genes, which allow the real-time monitoring of clock-driven expression. Dose-dependent ROS treatments were administered on the reporter lines to investigate the effects that altering ROS homeostasis has on transcription of oscillator components.

2. Materials and Methods

2.1 Plant materials and growth conditions

Plants were germinated and grown for 14-16 days on soil (Dalton Seed Raising Mix), unless otherwise stated, in chambers (Contherm, New Zealand) with cool white fluorescent lights set at 22°C, 65% relative humidity at a light intensity of 100-150 µE under long-day (16 h light, 8 h dark) or day neutral (LD; 12 h light, 12 h dark) or LL photocycles. WT controls and mutant lines (Appendix 5.1) were from the same genetic background in each experiment. Mutant lines in the Col-0 background were *CCA1-ox* (Wang & Tobin, 1998)*, elf3-1* (Hicks et al., 1996)*, elf4-101* (Kikis et al., 2005)*, prr5-1* (Eriksson et al., 2003)*, prr7-3, prr9-1* (Zeilinger et al., 2006)*, prr7-3/prr9-1* (Farré et al., 2005) and *prr5-1/prr9-1* (Eriksson et al., 2003). Mutant lines in the Ler-0 background were *cca1-1, lhy-11* and *cca1-1/lhy-11* (Mizoguchi et al., 2002). Mutant lines in C24 background were *lux-1* (Hazen et al., 2005)*, tic-1* (Hall et al., 2003)*, ztl-1* (Somers et al., 2000) and *toc1-1* (Millar et al., 1995). Mutant lines were gifts from C. R. McClung (Department of Biological Sciences, 6044 Gilman Laboratories, Dartmouth College, Hanover, NH 03755-3576, USA), E. Tobin (University of California; Department of Molecular, Cell and Developmental Biology; Los Angeles, CA USA) and D. Hincha (Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany).

2.2 H_2O_2 and catalase assays

Plants were grown under LD photocycles for 15 days. A subset of plants was transferred to LL on day 15 for LL samples and a subset remained in LD for LD samples. Samples were harvested on day 16 every 4 h in LD or LL and snap frozen in liquid nitrogen. H_2O_2 and catalase measurements (Standard curves in Appendix 5.7) were performed using the Amplex® Red Hydrogen Peroxide/Peroxidase Assay Kit (Invitrogen) and Amplex® Red Catalase Kit (Invitrogen) respectively.

Amplex® Red Hydrogen Peroxide/Peroxidase Assay

<u>Preparation of stock solutions:</u>

10 mM Amplex® Red reagent stock solution

- One vial of Amplex® Red reagent (Component A, blue cap) and DMSO (Component B, green cap) was thawed to room temperature.

- Just prior to use, the contents of the vial of Amplex® Red reagent was dissolved in 60 µL of DMSO.

1X Reaction Buffer

- 4 mL of 5X Reaction Buffer (Component C, white cap) was added to 16 mL of deionized water.

10U/mL Horseradish Peroxidase (HRP) stock solution

- The content of the vial of HRP (Component D, yellow cap) was dissolved in 1.0 mL of 1X Reaction Buffer.

20mM H_2O_2 working solution

- The 3.0% (v/v) H_2O_2 (Component E, red cap) was diluted into the appropriate volume of 1X Reaction Buffer. A 20 mM H_2O_2 working solution was prepared from a 3.0% (v/v) (0.88 M) H_2O_2 stock solution by diluting 22.7 µL of 3.0% (v/v) H_2O_2 into 977 µL of 1X Reaction Buffer.

Preparation of an H_2O_2 standard curve

The appropriate amount of 20 mM H_2O_2 working solution (prepared in step 1.5) was diluted into 1X Reaction Buffer (prepared in step 1.3) to produce H_2O_2 concentrations of 0 to 10 µM, each in a volume of 50 µL.

Amplex® red H_2O_2 assay

- 200 µL of 1X reaction buffer was added to ground frozen leaves and mixture was vortexed and kept on ice for 5 min.
- Mixture was centrifuged at 13,3000 x g for 13 min at 4°C.
- 50 µl of samples were added into individual wells of a 96-well microplate.
- A working solution of 100 µM Amplex® Red reagent and 0.2 U/mL HRP was prepared by mixing the following:

  50 µL of 10 mM Amplex® Red reagent stock solution

  100 µL of 10 U/mL HRP stock solution

  4.85 mL of 1X Reaction Buffer

- The reaction was started by adding 50 µl of the Amplex® Red reagent/HRP working solution to each microplate well containing the samples.
- The reaction was incubated at room temperature for 30 min, protected from light and absorbance was measured for 30 min at each 1 min interval at 560 nm.

Amplex® Red Catalase Assay

Preparation of stock solutions:

10 mM Amplex® Red reagent stock solution

- One vial of Amplex® Red reagent (Component A, blue cap) and DMSO (Component B, green cap) was thawed to room temperature.
- Just prior to use, the contents of the vial of Amplex® Red reagent was dissolved in 60 µL of DMSO.

1X Reaction Buffer

- 4 mL of 5X Reaction Buffer (Component C, white cap) was added to 16 mL of deionized water.

100 U/mL Horseradish Peroxidase (HRP) stock solution

- The content of the vial of HRP (Component D, yellow cap) was dissolved in 200 µL of 1X Reaction Buffer.

20 mM H_2O_2 working solution

- The 3.0% (v/v) H_2O_2 (Component E, red cap) was diluted into the appropriate volume of 1X Reaction Buffer. A 20 mM H_2O_2 working solution was prepared from a 3.0% (v/v) (0.88 M) H_2O_2 stock solution by diluting 22.7 µL of 3.0% (v/v) H_2O_2 into 977 µL of 1X Reaction Buffer.

1000 U/mL catalase solution

- The content of the vial of catalase was dissolved in 100 µL of dH2O

Preparation of a catalase standard curve

The appropriate amount of the 1000 U/mL catalase solution was diluted into 1X Reaction Buffer to produce catalase concentrations of 0 to 4.0 U/mL.

Amplex® red catalase assay

- 200 µL of 1X reaction buffer was added to ground frozen leaves and mixture was vortexed and kept on ice for 5 min.
- Mixture was centrifuged at 13,3000 x g for 13 min at 4°C.
- 25 µL of the samples were pipetted into individual wells of a 96-well microplate.
- 40 µM H_2O_2 solution was prepared by adding 10 µL of the 20 mM H_2O_2 solution to 4.99 mL 1X Reaction Buffer
- 25 µL of the 40 µM H_2O_2 solution was pipetted to each microplate well containing the samples.
- The reaction was incubated for 30 minutes at room temperature.
- A working solution of 100 µM Amplex Red reagent containing 0.4 U/mL HRP was prepared by adding 50 µL of the Amplex Red reagent stock solution and 20 µL of the HRP stock solution to 4.93 mL 1X Reaction Buffer.
- The second phase of the reaction was started by adding 50 µL of the Amplex Red/HRP working solution to each microplate well containing the samples.
- The reaction was incubated for 30 min at 37°C, protected from light and absorbance was measured for 30 min at each 1 min interval at 560 nm.

DAB staining

For *in planta* H_2O_2 detection, 3,3-diaminobenzidine (DAB) was used to stain H_2O_2 according to the method modified from Thordal-Christensen et al., (1997). Leaves were infiltrated with 1 mg/mL DAB dissolved in 50mM Tris-acetate (pH 5.0) for 6 h. Chlorophyll was removed by boiling in lactic acid: glycerol: ethanol (1:1:4) solution for 10 mins.

DAB staining solution

50 mM Tris-acetate pH 5

1 mg/mL DAB

Destaining solution (960mL)

Lactic acid	160mL
Glycerol	160mL
Ethanol	640mL

2.3 ROS hypersensitivity assay

WT and mutant lines were grown under LD photocycles for 14 days according to the aforementioned conditions (Section 2.1) and were transferred to LL on day 15 prior to the experiment on day 16. Trays of plants were sprayed with 5 mL of 5 µM methyl viologen (MV) at ZT3 to induce superoxide burst (Bowler et al., 1992) and were scored for hyper- or hyposensitivity (measuring number of wilted leaves) 24 h later.

2.4 ROS treatments for quantitative PCR (qPCR) analysis

WT and mutant lines were grown under LD photocycles for 15 days according to the aforementioned conditions (Section 2.1) and were transferred to LL on day 15 prior to the experiment. Plants were sprayed with the stated dose of MV at the ZT3, ZT11 and ZT19 on day 16 and were harvested 4 h after treatment for expression studies.

2.5 Transcript analysis by qPCR

WT (Col-0, Ler-0 and C24) and mutant lines (*CCA1-ox, cca1-11/lhy-21, elf3-1, lux-1, toc1-1*) were grown on soil under LD photocycles according to the aforementioned conditions (Section 2.1). On day 15, plants were transferred to LL and samples were collected on day 16 every 4 h across two days and were snap frozen in liquid nitrogen. RNA was isolated using the ZR Plant RNA Miniprep Kit (Zymo Research) followed by DNAse I (Roche Applied Sciences) treatment. The digestion of genomic DNA (Table 2.1) in RNA samples was performed.

RNA extraction using the ZR Plant RNA Miniprep Kit

1. Ground frozen leaves were transferred into a ZR Bashing Bead™ Lysis Tube and 800 µl of RNA Lysis Buffer was added to the sample.
2. The ZR Bashing Bead™ Lysis Tube was centrifuged at 13,3000 x g for 1 min.
3. 400 µl of the supernatant was transferred into a Zymo-Spin™ IIIC Column in a Collection Tube and was centrifuged at 7,000 x g for 30 sec.
4. 320 µl 95% (v/v) ethanol was added to the flow-through in the Collection Tube.
5. Mixture was transferred to a Zymo-Spin™ IIC Column in a Collection Tube and centrifuged at 13,3000 x g for 30 sec. Flow-through was discarded.
6. 400 µl RNA Prep Buffer was added to the column and centrifuged at 13,3000 x g for 1 min. Flow-through was discarded and the Zymo-Spin™ IIC Column was placed back into the Collection Tube.
7. 800 µl RNA Wash Buffer was added to the column and centrifuge at 13,3000 x g for 30 sec. Flow-through was discarded and the Zymo-Spin™ IIC Column was placed back into the Collection Tube. The wash step was repeated with 400 µl RNA Wash Buffer.
8. The Zymo-Spin™ IIC Column was centrifuged at 13,3000 x g for 2 mins in the emptied Collection Tube to ensure complete removal of the wash buffer.
9. The Zymo-Spin™ IIC Column was removed from the Collection Tube and placed into a DNase/RNase-Free Tube. 25 µl DNase/RNase-Free Water was added directly to the column matrix and left for 1 min.
10. The tube was centrifuged at 13,3000 x g for 30 sec to elute the RNA from the column. RNA was stored at -80°C until future use.

Table 2.1: Genomic DNA digestion in RNA samples

Component	Final concentration
Total RNA	50 µg
10X incubation buffer	5 µl
DNase I recombinant, RNase-free	10 units

Protector RNase Inhibitor	10 units
RNase free water	Up to 50 µl

The above reaction was incubated at 37°C for 20 min. The reaction was then stopped by adding 2 µl of 0.2M EDTA (pH 8.0) to a final concentration of 8 mM and heated to 75°C for 10 min. First-strand cDNA synthesis was performed with 2 µg of RNA using the Transcriptor First Strand cDNA Synthesis kit (Roche Applied Sciences).

Table 2.2: Template-primer mixture

Component A	**Final volume or concentration**
Total RNA	2 µg
Anchored-oligo(dT) primer	1 µl (2.5 µM)
PCR-grade water	Up to 13 µl

Table 2.3: Reverse transcription reaction mixture

Component B	**Final volume or concentration**
5X reaction buffer	4 µl (8 mM $MgCl^2$)
Protector RNase Inhibitor	0.5 µl (20 units)
Deoxynucleotide mix	2 µl (1 mM each)
Trancriptor reverse transcriptase	0.5 µl (10 units)

Component A (Table 2.2) was heat denatured at 65°C for 10 min and transferred to ice. Component B (Table 2.3) was added to component A and mixture was incubated at 50°C for 60 min. The reverse transcription reaction was stopped by heating the mixture at 85°C for 5 min. The cDNA was stored at -20°C until future use. The qPCR reaction mix was prepared with the Power SYBR® Green Master Mix (Table 2.4; Applied Biosystems) and thermal cycling (Applied Biosystems, Germany) was performed (Table 2.5).

Table 2.4: QPCR reaction mixture

Component	Final concentration
Power SYBR Green PCR master mix	0.2X
Reverse primer	50 nM
Forward primer	50 nM
cDNA	10 ng
Nuclease-free water	Up to 5 µL

Table 2.5: QPCR setup

Step	Hold	40 cycles PCR	
		Denature	Anneal/Extend
Time	10 min	15 sec	60 sec
Temp (°C)	95	95	60

As normalization controls, the *ACTIN 2* (*ACT2*), *UBIQUITIN-PROTEIN LIGASE 7* (*UPL7*), *GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE C2* (*SAND*), *ISOPENTENYL PYROPHOSPHATE 2* (*IPP2*), and *TUBULIN BETA-2* (*TUB2*) genes were used as reference genes. Expression data were analyzed using the comparative C_T method (Schmittgen & Livak, 2008), where C_T is the average threshold cycle for three biological replicates. For time series expression profiles (Fig. 3.1*D*; 3.2*D*; 3.5*D*; 3.6*D*; 3.8; 3.9; 3.10; 3.11*C*; 3.13), the relative expression values were calculated using the equation 2^{dCT}, where dCT = C_T (target gene) – C_T (reference gene). For MV treated samples (Fig. 3.11*A* and *B*; Fig. 3.12), normalized expression values were calculated using the equation 2^{ddCT}, where ddCT = [C_T target gene (control – treated)] – [C_T reference gene (control – treated)]; control: samples treated with water at each corresponding time point. Primers were designed with the QuantPrime tool (Arvidsson et al., 2008) and primer sequences were listed in Appendix 5.2 to 5.6.

2.6 Chromatin Immunoprecipitation-quantitative PCR (ChIP-qPCR) assay

CCA1::CCA1-GFP transgenic lines (Pruneda-Paz et al., 2009) used for ChIP were grown under LD photocycles according to the aforementioned conditions (Section 2.1). Leaves were harvested at ZT1 on day 16 of growth. To determine *in vivo* binding of CCA1 to promoters of ROS genes, ChIP was performed using the EpiQuik^{TM} Plant ChIP Kit (Epigentek Group Inc.). For the immunopurification of the CCA1-GFP-DNA complexes, the anti-GFP antibody was used (Roche Applied Science). Cells were cross-linked with 1% (v/v) formaldehyde and samples were sonicated followed by performing the subsequent steps below.

Chromatin Immunnoprecipitation with the EpiQuik™ Plant ChIP Kit

Components:

CP1 (Wash Buffer)

CP2 (Antibody Buffer)

CP3C (5X Lysis Buffer I)

CP3D (Lysis Buffer II)

CP3E (Lysis Buffer III)

CP3F (Lysis Buffer IV)

CP4 (ChIP Dilution Buffer)

CP5 (DNA Release Buffer)

CP6 (Reverse Buffer)

CP7 (Binding Buffer)

CP8 (Elution Buffer)

Protease Inhibitor Cocktail (100X)

Normal Mouse IgG (1 mg/mL)

Anti-Dimethyl H3-K9 (1 mg/mL)

Proteinase K (10 mg/mL)

8-Well Assay Strips (with frame) 8-Well Strip Caps

F-Spin Column

F-Collection Tube

Formaldehyde

Glycine

2-mercaptoethanol

Antibody of interest

TE buffer (pH 8.0)

Ethanol (95%, v/v)

Protocol:

The following required solutions were prepared: 90% (v/v) Ethanol; 70% (v/v) Ethanol; 37% (v/v) Formaldehyde; 2 M Glycine Solution; 14.3 M 2-mercaptoethanol (BME); 1X TE Buffer (pH 8.0).

Antibody Binding to the Assay Plate

1. Strip wells were washed once with 150 µl of CP1.
2. 100 µl of CP2 was added to each strip well the antibodies were then added: 1 µl of Normal Mouse IgG as the negative control, 1 µl of Anti-Dimethyl H3-K9 as the positive control, and 2 µg of an GFP antibody.
3. The strip wells were covered with Parafilm M and incubated at room temperature for 90 min. After incubation, the incubated antibody solution was removed and washed three times with 150 µl of CP2 by pipetting in and out.

Tissue Collection and In Vivo Cross-Link

1. 0.8-1 g of leaf tissue was harvested in a 50 mL Falcon tube.
2. Tissue was rinsed gently with 20 mL of deionized water two times. As much water as possible was removed from the tissue and 20 mL of 1.0% (v/v) formaldehyde was added.
3. The top of the 50 mL conical tube (containing the formaldehyde soaked tissue) was stuffed with nylon mesh to keep the tissue immersed during vacuum infiltration (and aid later rinse steps).
4. The tissue was vacuum-infiltrated for 10 min in a desiccator attached to a vacuum pump. The formaldehyde solution should boil.

Tissue Lysis and DNA Shearing

1. 1.25 mL of 2 M Glycine solution (final concentration 0.125 M) was added to quench cross-linking and vacuum infiltration was continued for an additional 5 min.
2. The formaldehyde was removed and the tissue was rinsed two times with 20 mL of deionized water. As much water as possible was removed (at this stage the cross-linked tissue can either be frozen in liquid nitrogen and stored at –80°C, or used directly for chromatin extraction).
3. CP3C was diluted with distilled water at a 1:5 ratio (1X CP3C). 3.5 µl of BME was added to each 10 mL of 1X CP3C. The tissue was ground in liquid nitrogen to a fine powder. The powder was added to 20 mL of cold 1X CP3C in a 50 mL conical tube, then vortexed, and placed on ice.
4. The solution was filtered through two layers of Miracloth into a 50 mL tube and centrifuged at 4,000 x g for 20 min.
5. 1 µl of BME was added into each 1 mL of CP3D. Supernatant was removed and pellet was resuspended in 1 mL of CP3D containing BME. The resuspended pellet was transferred to a 1.5 mL vial and centrifuged at 13,3000 x g for 10 min at 4°C to pellet nuclei (white pellet should be seen at this stage).
6. 1 µl of BME was added into each 1 mL of CP3E. Supernatant was removed and pellet was resuspended in 300 µl of CP3E containing BME.
7. 300 µl of CP3E containing BME was added into a new 1.5 mL vial. The resuspended pellet from step 6 was layered on top of this 300 µl cushion and centrifuged at 13,3000 x g for 45 min at 4°C.
8. Supernatant was removed and chromatin pellet was resuspended in 500 µl of CP3F containing Protease Inhibitor Cocktail. DNA was sheared by sonication.
9. Sample was placed on ice for 1 min between each sonication treatment. The length of sheared DNA should be between 200-1000 bp.
10. Cell debris were pelleted by centrifuging at 13,3000 x g for 10 min at 4°C.

Protein/DNA Immunoprecipitation

1. Clear supernatant was transferred to a new 1.5 mL vial. The required volume of supernatant was diluted with CP4 at a 1:1 ratio.

2. 5 µl of the diluted supernatant was removed to a 0.5 mL vial. The vial was labeled as “input DNA” and then placed on ice.
3. 100 µl of the diluted supernatant was transferred to each antibody-bound strip well. The strip wells were covered with Parafilm M and incubated at room temperature (25°C) for 90 min on an orbital shaker (100 rpm).
4. Supernatant was removed. The wells were washed six times with 150 µl of CP1. 2 min on a rocking platform (100 rpm) was allowed for each wash. The wells were washed once (for 2 min) with 150 µl of 1X TE Buffer.

Cross-Linked DNA Reversal/DNA Purification

1. 1 µl of Proteinase K was added to each 40 µl of CP5 and mix. 40 µl of CP5 containing Proteinase K was added to the samples (including the “input DNA” vial). The sample wells were covered with strip caps and incubated at 65°C in a waterbath for 15 min.
2. 40 µl of CP6 was added to the samples; mixed, and the wells were re-covered with strip caps and incubated at 65°C in a waterbath for 90 min. 40 µl of CP6 was also added to the vial containing supernatant, labeled as “input DNA”, mixed and incubated at 65°C for 90 min.
3. A spin column was placed into a 2 mL collection tube. 150µl of CP7 was added to the samples and mixed solution was transferred to the column, centrifuged at 13,3000 x g for 20 sec.
4. 200 µl of 70% (v/v) ethanol was added to the column, centrifuged at 13,3000 x g for 15 sec. The column was removed from the collection tube and flow-through was discarded.
5. The column was replaced to the collection tube. 200 µl of 90% ethanol was added to the column and centrifuged at 13,3000 x g for 20 sec.
6. The column was removed and flow-through was discarded. The column was replaced to the collection tube and the column was washed again with 200 µl of 90% (v/v) ethanol at 13,3000 x g for 35 sec.
7. The column was placed in a new 1.5 mL vial. 20 µl of CP8 was added directly to the filter in the column and centrifuged at 13,3000 x g for 20 sec to elute purified DNA.
8. DNA was stored at –20°C until future use.

For each qPCR reaction, 1 μl of the purified DNA was used to determine enrichment of potential binding sites of CCA1 to promoters of the selected ROS genes. Fold enrichment values were calculated as 2^{dCT}, where dCT = C_T (sample) – C_T (input). Primer sequences were listed in Appendix 5.5.

2.7 Luminescence assay

2.7.1 Seed surface sterilization

Seed aliquots were placed in a sterile 1.5 mL microcentrifuge tube and rinsed with 1 mL of 95% (v/v) ethanol for 5 min. Ethanol was then decanted and 1 mL of bleach solution (3% v/v sodium hypochlorite and 0.01% Tween 20) was added to the tube to rinse the seeds for 2 min. Bleach solution was decanted and the seeds were rinsed with 1 mL sterile water for 3 times. Seeds were resuspended in 0.01% (w/v) sterile agarose and stored in 4°C before plating onto agar plates. Sterilized seeds of reporter lines CCA1::LUC (Pruneda-Paz et al., 2009), TOC1::LUC (Alabadi et al., 2001), FKF1::LUC (Schultz et al., 2001), CAB2::LUC (Millar et al., 1995) and CAT3::LUC (Michael & McClung, 2002) were plated on Murashige and Skoog (MS) media with 0.8% phytoagar, stratified for 3 days at 4°C and grown under 12 h light 12 h dark photocycles at 22°C.

MS media

4.4 g/L MS

0.5 g/L 2-(N-morpholino)ethanesulfonic acid

0.8% Phytoagar pH 5.7

2.7.2 Video-intensified microscopy (VIM) imaging of luminescence

On day 10, seedlings were then transferred to new media plates (Fig. 2.1*A*). D-luciferin (Biosynth L-8820) was used as the substrate for luciferase. Stock solutions of 50 mM luciferin was made up and stored in aliquots at -20°C in the dark. A working solution of 5 mM luciferin in 0.01% Triton X-100 was prepared fresh, filter sterilized and stored in

an amber spray bottle. Seedlings were then sprayed with the luciferin working solution and transferred to LL conditions.

50 mM D-luciferin stock

1 g Firefly D-Luciferin

71.3 mL 1M Triphosphate buffer (Na_2HPO_4) pH 8

5 mM D-luciferin working solution

7.5 mL 50 mM D-Luciferin stock 67.5 mL 0.01% (w/v) Triton X-100

On day 11, seedlings were sprayed with ROS inducers at ZT3. ROS inducers used were H_2O_2 (2, 5, 10, 20 mM), MV (0.5, 1, 2, 5, 10 µM), 3-aminotriazole (3-AT; 0.5, 1 mM), salicylhydroxamic acid (SHAM; 1, 2, 5 mM), diphenylene iodonium (DPI; 10, 20, 50 µM) and potassium iodide (KI; 1, 2 mM). Real-time monitoring of luciferase activity was performed using a VIM-intensified charge-couple device (CCD) camera mounted with an ARGUS-50 photon-counting system (Hamamatsu). A wide-angle Nikon Nikkor 35mm photographic lens was used for image acquisition. Imaging started on day 12. Plants were temporarily moved to a light-tight black box for the imaging session (Fig. 2.1*B*) and luminescence from individual seedlings were captured in each 25 min exposure across five consecutive days.

H_2O_2 solutions

30% (v/v) Stock (9.79 M, density 1.1 g/cm3)

Working solution (mM)	**Stock (µl)**	**0.01% (w/v) Triton X-100 (mL)**
2	4	20
5	10	20
10	20	20
20	40	20

MV solutions

10 mM stock

25.7 mg in 10 mL 0.01% (w/v) Triton X-100

Working solution (µM)	Stock (µl)	0.01% (w/v) Triton X-100 (mL)
0.5	1	20
1.0	2	20
2.0	4	20
5.0	10	20
10.0	20	20

3-AT solutions

10 mM stock

8.4 mg in 10 mL 0.01% (w/v) Triton X-100

Working solution (mM)	Stock (mL)	0.01% (w/v) Triton X-100 (mL)
0.5	1	19
1.0	2	18

SHAM solutions

100 mM stock

153 mg in 10 mL DMSO

Working solution (mM)	Stock (µl)	0.01% (w/v) Triton X-100 (mL)
1	200	19.8
2	400	19.6
5	1000	19.0

DPI solutions

10 mM stock

31.5 mg in 10 mL DMSO

Working solution (µM)	Stock (µl)	0.01% (w/v) Triton X-100 (mL)
10	20	20
20	40	20
50	100	20

KI solutions

100 mM stock

166 mg in 10 mL 0.01% (w/v) Triton X-100

Working solution (mM)	Stock (µl)	0.01% (w/v) Triton X-100 (mL)
1	200	19.8
2	400	19.6

2.7.3 Luminescence data processing with Metamorph

Images were processed using the MetaMorph imaging software (Molecular Devices). A set of images (time series) of the same experiment was loaded as an image stack in MetaMorph. The region tool was used to draw circular regions of interest around each seedling to measure the average luminescence intensity of each seedling within the region of interest (Fig. 2.1*C*). The size of each region was kept constant across the time series and the positions of the regions were adjusted, if needed, to account for seedling movements during the VIM. To log MetaMorph data into Microsoft Excel, the dynamic data exchange tool was used. The data were acquired using the graph intensities function and the parameters were configured to ensure that only the required information ('image name', 'image plane', 'region name' and 'integrated') was saved in the data log (Hall and Brown, 2007).

2.7.4 Data analysis with BioDare

Data processing with BioDare involved two stages: 1) Experiment conditions and the data were firstly described and 2) Data were processed and analyzed with the web-based interface provided by the BioDare data repository (www.biodare.ed.ac.uk).

In step 1, all details required for reproducing the experiment, e.g., genotypes of measured plants and biological description of the numerical time series, were described using a stand-alone application called PedroApp (Fig. 2.2*A*). The experimental description was indexed in the metadata generated using PedroApp. The following is a step-by-step guide for PedroApp and BioDare.

Step 1: Experiment description and metadata generation with PedroApp

1. The PedroApp.jar application is started.
2. Experiment was described by filling all the relevant fields:
    - Bold font in field labels marks the mandatory fields.
    - Do not use the simple forms for the temperature and light settings. Look for the 'Fill' button to choose appropriate settings from the new dialog window.
    - The growth and experimental conditions have to be described before describing biological samples.
    - Do not use the simple forms to describe samples. Look for the 'FillUp' button and use the wizard that opens in the new dialog window.
3. Descriptions can be saved at any time using 'File/Save' option in the menu.
4. Once finished, go to 'View/Show Errors' to check if all the necessary information were provided.
5. Once error free, a new file with metadata for the experiment is then generated using the 'File/Export to final submission format' option. Two files are generated by PedroApp, one called 'Pedro file' which stored the description form and another called the 'xml file' which contained proper metadata.

Step 2: Creating new experiment and data processing in BioDare

1. Choose *Add experiment*. The metadata (.xml) generated by PedroApp is uploaded (Fig. 2.2*B*).
2. Choose *Add raw data* to attach the raw luminescence data files containing numeric time series for this experiment.
3. Go to *Data graphs* to see the time series. Click on *Change display criteria* to choose the data type (raw, normalized, etc) and the biological samples to display. The data can

be filtered according to genotype and experimental setting by selecting options on the *Change display criteria* panel.

4. For period and phase estimation of rhythmic LUC activity, the Fast Fourier Transform-Nonlinear Least Squares (FFT-NLLS) method was used to process the time-series data (Plautz et al., 1997). Data were fitted to a series of linearly damped cosine curves and the relative amplitude (amplitude of the residuals relative to the primary fitted amplitude) was used to measure rhythm significance; values vary from 0 (perfect cosine fit) to 1 (residual amplitude equals fitted amplitude).

5. Go to *Start FFT* to start new FFT analysis. Data window used for analysis is selected. For example for data recorded under 2 days LD + 5 days LL, the most likely data window will be form ZT50-168. If one of the boundaries is left as blank or 0, the min or max values will be used respectively.

6. Go to *Show FFT* to review the results and click on *Expand the FFT results* to see data for individual time series. Summary statistics is shown for each genotype based on the results that the system could determine period. If period could not be determined, it is excluded from the summary statistic.

7. Phase shifts were calculated by subtracting the mean phase of the WT traces from the phase of each reporter line.

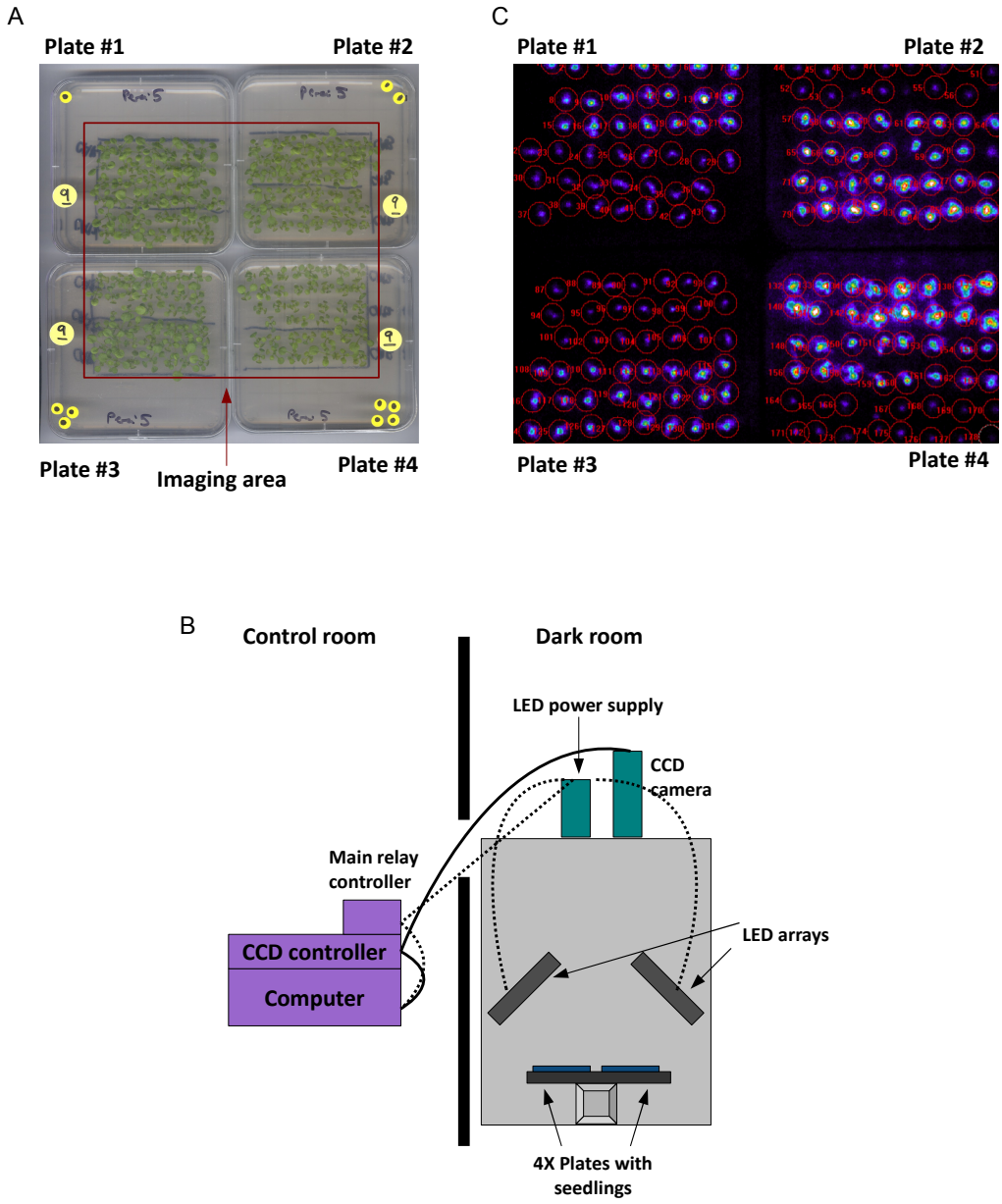
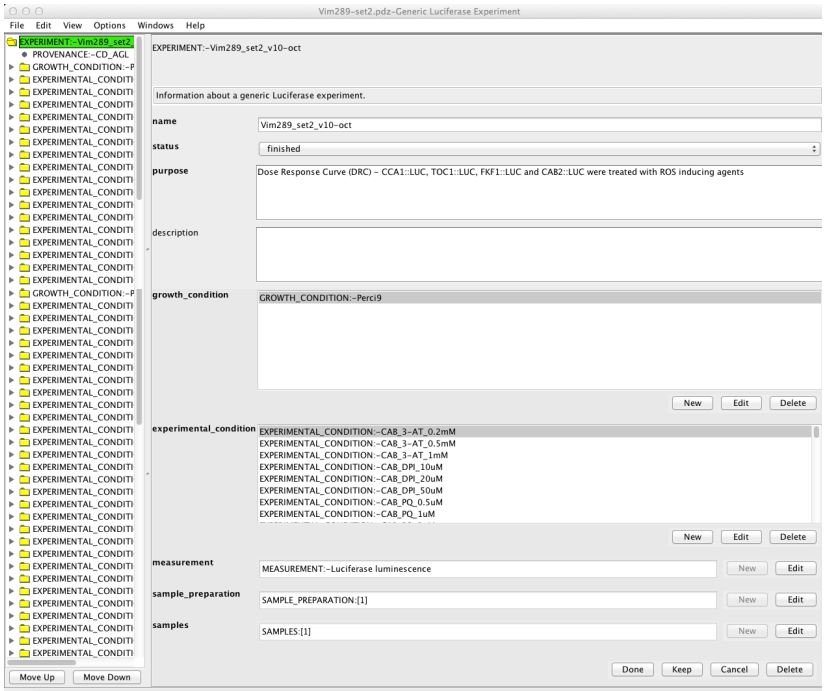
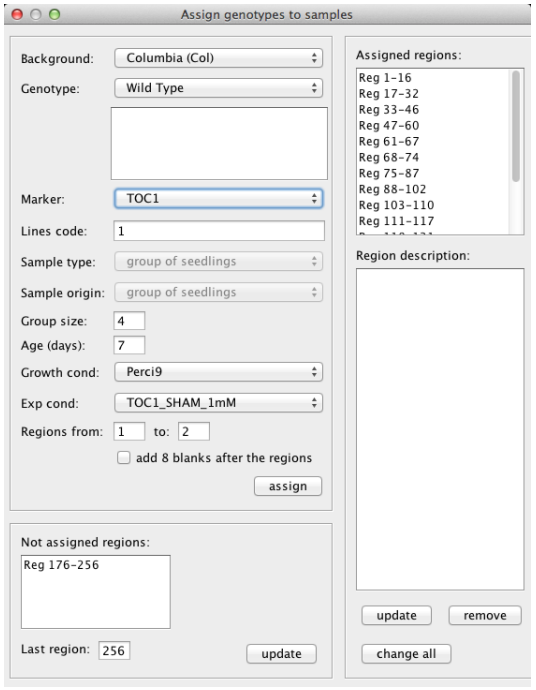


Figure 2.1: Real-time imaging of circadian-regulated gene expression with the VIM system
(A) Transgenic *Arabidopsis* reporter seedlings harboring the CCA1::LUC, TOC1::LUC, FKF1::LUC, CAT3::LUC and CAB2::LUC constructs were transferred to square MS-agar plates (9.5cm x 9.5cm) two days prior to imaging. Approximately 36 seedlings can be fitted onto each plate (6 seedlings in 6 rows). The imaging area (6cm x 6cm) included a total of approximately 144 seedlings. (B) Diagram illustrates the setup of the VIM equipped with a CCD system for image acquisition. The CCD camera is connected to the CCD controller and the computer. The LED power supply is connected to the CCD controller, computer and the LED arrays. The entire camera setup is placed in a dark room (Modified from Hall and Brown, 2007). (C) The acquired image was processed using the MetaMorph program. Circular regions were drawn around each seedling to quantify the average luminescence intensity.

A





B

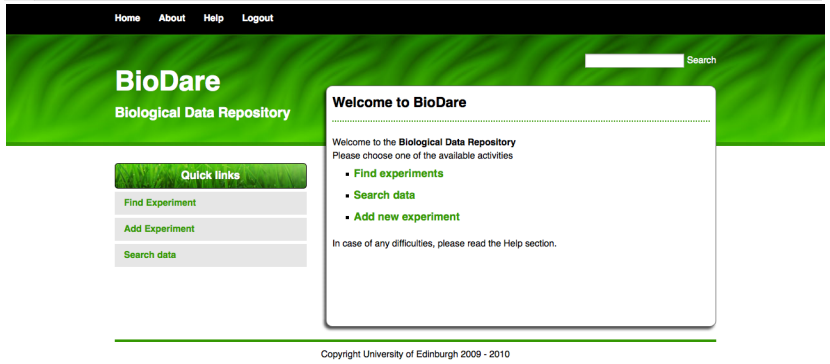


Figure 2.2: Luminescence data analysis with PedroApp and BioDare

(A) The PedroApp was used to generate the metadata file (.xml) containing descriptions of experimental and growth conditions of the VIM. (B) Data processing, trace graphs display and FFT-NLLS analysis were done using the BioDare web-based tool.

2.8 Bioinformatics analyses

GO overrepresentations of the selected ROS genes and *Arabidopsis* genes under different ROS GO categories were obtained using the ATCOECIS tool (http://bioinformatics.psb.ugent.be/ATCOECIS/; Vandepoele et al., 2009). The DIURNAL tool (http://diurnal.cgrb.oregonstate.edu/; Michael et al., 2008b) was used to identify cycling genes. The time of peak in gene expression for rhythmic genes were obtained using the PHASER tool (http://phaser.cgrb.oregonstate.edu; Michael et al., 2008b) by querying the five diurnal time series: LDHH_ST (Blasing et al., 2005), LDHH_SM (Smith et al., 2004), Col_SD, Ler_SD and LHYOX_SD (Michael et al., 2008b). Promoter analysis for EE and CBS motifs in ROS-responsive genes was performed using the ATHENA tool (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/home.pl; O'Connor et al., 2005).

2.9 Statistical analysis

Parametric statistics (one-way ANOVA and Student's t-test) were performed using the Graphpad Prism v5 software. One-way ANOVA was used to test for the effect of time.

3. Results

3.1 ROS homeostasis and signaling are regulated by diurnal cycles

In this section, the first hypothesis on whether ROS homeostasis is controlled by diurnal day-night transitions and the circadian oscillator was investigated. Given the widespread effects of ROS on plant growth (Mathur, 2004), it is likely that ROS homeostasis may be regulated by diurnal rhythms. It has been shown previously that genes associated with the light-harvesting apparatus peaked approximately at ZT4 and coincide with the peak in photosynthesis and light-harvesting capacity (Harmer et al., 2000). Such metabolic fluctuations at a particular time may tilt the balance of ROS production and scavenging. It is therefore likely for ROS levels to peak shortly after or concomitant with the peak in photosynthetic capacity.

To investigate whether periodic fluctuations in ROS production and scavenging exist under entraining and free-running conditions, H_2O_2 quantifications and catalase activity measurements were performed. H_2O_2 levels were quantified from plants grown under 12 h light, 12 h dark (LD) photocycles. Indeed, it was observed that H_2O_2 production peaked at noon, ZT7 (0.117 ± 0.009 µmol/mgFW) and reached trough levels (0.015 ± 0.008 µmol/mgFW) at midnight, ZT19 (Fig. 3.1*A* and *B*). The scavenger of H_2O_2 is catalase. The scavenging capacity at a specific time may mirror the pattern of ROS production. To determine the time-of-day scavenging of H_2O_2, catalase activity was quantified across two diurnal cycles in LD grown plants. Consistent with the prediction, catalase activity was found to peak at ZT7 (0.611 ± 0.021 U/mL) and dip at ZT19 (0.156 ± 0.039 U/mL; Fig. 3.1*C*), coinciding with the peak and trough of H_2O_2.

Because of the observed diurnal pattern in ROS homeostasis, it is likely that at the transcriptional level, ROS signaling may be regulated in a similar way. To investigate the interplay between diurnal rhythms and ROS signaling, time-course expression profiles of 167 ROS-responsive genes (Appendix 5.8) were obtained by qPCR. These genes were selected from a group of general oxidative stress response markers that consists of 3925 genes where the genes displayed at least 5-fold change in transcript abundance in *Arabidopsis* treated with different ROS agents (Gadjev et al., 2006). Samples were collected every 4 h or 6 h across one diurnal cycle from 16-day-old

plants entrained under 16 h light, 8 h dark (long-day) photocycles. Out of the 167 genes, 140 genes displayed time-of-day specific phases (one-way ANOVA, P<0.0001) under long-day photocycles (Fig. 3.1*D*). As expected, the largest gene cluster peaked at noon (Fig. 3.2*E*) that is about the same time when ROS levels were the highest (Fig. 3.1*A* and *B*). To further ascertain other molecular functions and biological processes that might be co-regulated by these 167 ROS genes, the ATCOECIS resource (Vandepoele et al., 2009) was used to determine enrichments in GO categories. In addition to the role of these genes in ROS signaling, categories associated with stresses and abiotic stimuli were enriched (Table 3.1). These results suggest that ROS production and scavenging peak at noon under diurnal cycles. ROS-responsive genes also display similar phase specific expression. The next section investigates whether such phase relations exist under free-running LL conditions.

Table 3.1: GO overrepresentation of the 167 ROS-responsive genes

GO enrichment	Enrichment fold	P-value	Description
GO:0009408	39.39	2.41E-25	Response to heat
GO:0009266	17.65	2.84E-21	Response to temperature stimulus
GO:0009628	5.59	3.29E-21	Response to abiotic stimulus
GO:0042221	6.54	2.75E-19	Response to chemical stimulus
GO:0050896	4.24	2.91E-19	Response to stimulus
GO:0006950	6.7	4.73E-19	Response to stress
GO:0042542	58.9	2.50E-14	Response to H_2O_2
GO:0009642	42.74	3.05E-14	Response to light intensity
GO:0000302	42.83	6.02E-13	Response to reactive oxygen species
GO:0009644	45.28	7.49E-12	Response to high light intensity
GO:0006979	13.54	1.62E-11	Response to oxidative stress
GO:0006800	12.43	4.75E-11	Oxygen and ROS metabolism
GO:0051707	6.89	5.34E-09	Response to other organism
GO:0009611	10.17	6.00E-08	Response to wounding
GO:0009416	7.13	4.60E-07	Response to light stimulus
GO:0009314	7	5.52E-07	Response to radiation
GO:0009719	4.1	8.90E-07	Response to endogenous stimulus

GO:0003700	2.98	1.35E-06	Transcription factor activity
GO:0009607	4.13	1.72E-06	Response to biotic stimulus
GO:0009605	6.75	2.54E-06	Response to external stimulus
GO:0009723	11.1	3.46E-06	Response to ethylene stimulus
GO:0009613	7.3	4.36E-06	Response to pest, pathogen or parasite
GO:0030528	2.7	7.52E-06	Transcription regulator activity
GO:0009725	4.4	8.69E-06	Response to hormone stimulus
GO:0006952	4.21	1.41E-05	Defense response
GO:0009407	18.21	6.94E-05	Toxin catabolism
GO:0009404	18.21	6.94E-05	Toxin metabolism
GO:0009861	8.43	8.30E-05	Jasmonic acid and ethylene-dependent systemic resistance
GO:0004364	17.09	8.91E-05	Glutathione transferase activity
GO:0009814	6.49	1.10E-04	Defense response to pathogen, incompatible interaction
GO:0042828	6.49	1.10E-04	Response to pathogen
GO:0042829	6.49	1.10E-04	Defense response to pathogen
GO:0003677	2.28	1.16E-04	DNA binding
GO:0009409	7.71	1.36E-04	Response to cold
GO:0009636	13.3	2.38E-04	Response to toxin
GO:0009617	8.95	2.51E-04	Response to bacteria
GO:0050832	13.09	2.53E-04	Defense response to fungi
GO:0042742	12.32	3.19E-04	Defense response to bacteria
GO:0051242	12.14	3.38E-04	Positive regulation of cellular physiological process
GO:0048522	11.97	3.57E-04	Positive regulation of cellular process
GO:0009753	8.18	3.81E-04	Response to jasmonic acid stimulus
GO:0009737	5.87	5.84E-04	Response to ABA stimulus
GO:0043119	10.21	6.51E-04	Positive regulation of physiological process
GO:0009620	9.63	8.14E-04	Response to fungi

Go enrichments were obtained from the ATCOECIS resource (http://bioinformatics.psb.ugent.be/ATCOECIS/; Vandepoele et al., 2009).

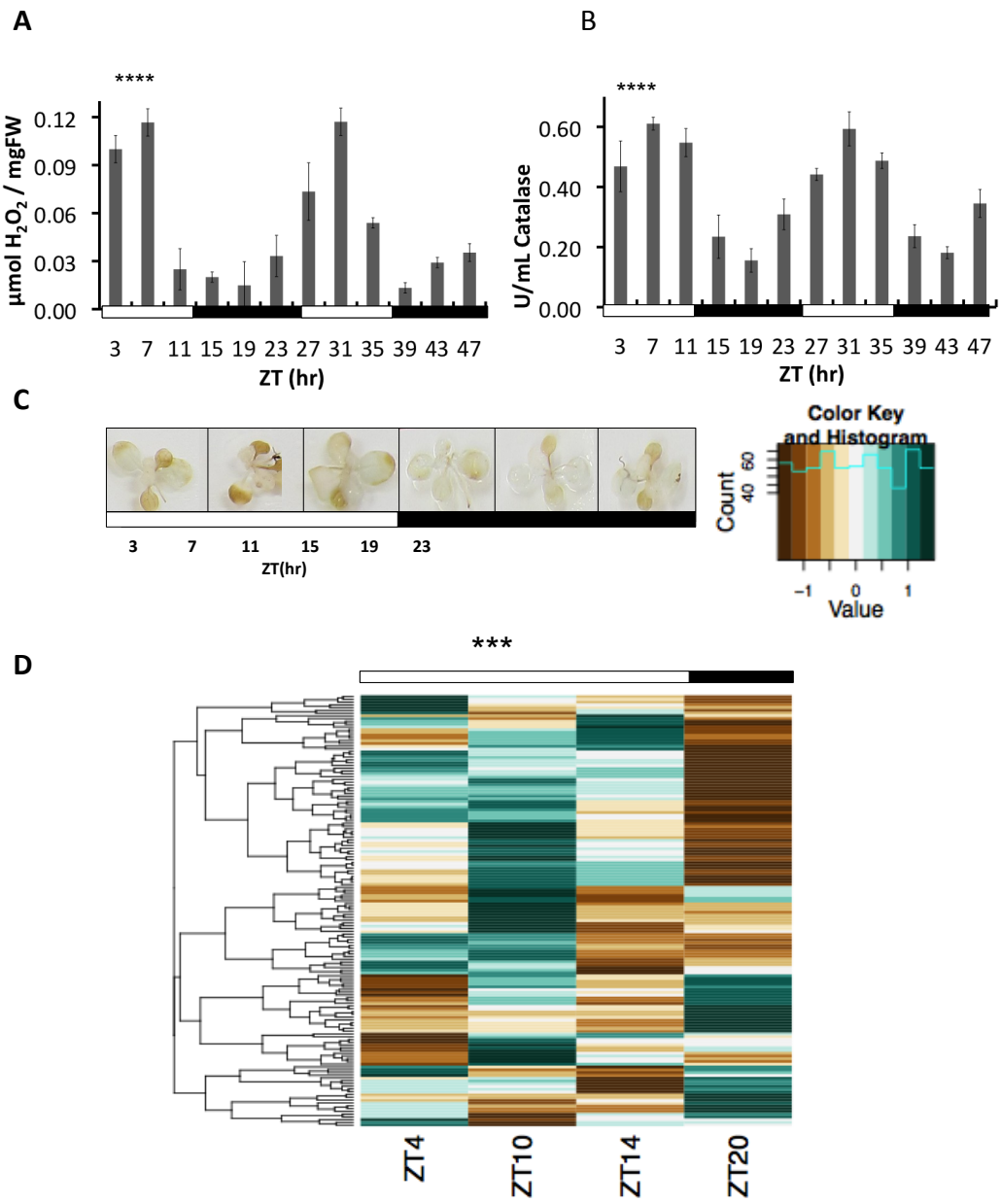


Figure 3.1: ROS homeostasis and signaling are regulated by diurnal cycles

Solid light grey bars represent (A) H_2O_2 levels and (B) catalase enzyme activity quantified from 16 day old Col-0 plants entrained in 12 h light 12 h dark photocycles sampled every 4 h across two days. Solid bars represent mean values of 3 biological replicates (N=15), with error bars indicating s.e.m. One-way ANOVA (effect of time) for LD profiles of H_2O_2 and catalase were significant (****P<0.0001). (C) Representative images of H_2O_2 accumulation in plants stained with DAB. (D) Expression profiles of 167 ROS transcripts obtained by qPCR from 16 day old Col-0 plants entrained in 16 h light 8 h dark photocycles from samples collected every 4 to 6 h. Up-regulation and down-regulation of gene expressions were depicted in red and green respectively. Gene expression data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression in long-day were significant (***P<0.0001). White bars: day; black bars: night.

3.2 The circadian clock regulates ROS homeostasis and signaling

Previously, the observed phase relationship between ROS homeostasis and ROS-driven transcription suggests coordinated regulation of this network within the diel cycle. The next question is whether a clock-regulated effect exists in LL conditions. Similar experiments were performed; ROS homeostasis and gene expression were this time investigated under LL conditions. It can be expected that H_2O_2 levels would remain high in LL due to constant light signaling of photoreceptors. Although LL condition is essential for studying relationships between the oscillator and its outputs, constant energy supply to photoreceptors have been proposed to generate ROS through carbohydrate accumulation and over-reduction of electron acceptors (Velez-Ramirez et al., 2011). H_2O_2 levels were elevated in both subjective day and night in LL (Fig. 3.2*A* and *C*). Nevertheless, a time-of-day specific peak at mid-day (ZT7; one-way ANOVA, $P<0.001$) for H_2O_2 could still be observed in LL albeit with a lower peak-trough ratio of 1.49 (Fig. 3.2*A*); peak-trough ratio in LD was 7.86. In the same way, catalase activity was also elevated in LL (Fig. 3.2*B*) with a similarly lowered peak-trough ratio in LL (1.59) than in LD (3.92).

The finding that ROS homeostasis was time-of-day specific suggests that the circadian clock may also transcriptionally coordinate ROS signaling. To test this, expression profiling of the same 167 genes were performed on plants transferred to free-running LL conditions. In agreement with the previous diurnal expression profile (Fig. 3.1*D*), ROS genes were phased to the subjective mid-day (ZT10; one-way ANOVA, $P<0.0001$) under LL conditions (Fig. 3.2*D*) and this implied a clock-regulated effect. Interestingly, a transient up-regulation of transcripts (Student's t-test, $P<0.01$) in the subjective mid-day was observed (Fig. 3.2*D*). Under LD conditions, only 39% of the tested ROS genes exhibited a noon phase (ZT10) whereas in LL, over 75% of those genes peaked at the subjective mid-day (Fig. 3.2*E*). It has been proposed that a metaphorical 'gate' regulates the differential sensitivity of the clock to signals presented at different times. Here, the transient up-regulation of ROS transcripts levels (Fig. 3.2*D*) is likely to be regulated by a clock-controlled gating mechanism, where the 'gate' opens at mid-day to allow the induction of these genes. How the clock controls ROS production and

scavenging remains to be fully investigated. The mechanistic role of the circadian clock in ROS homeostasis is studied next.

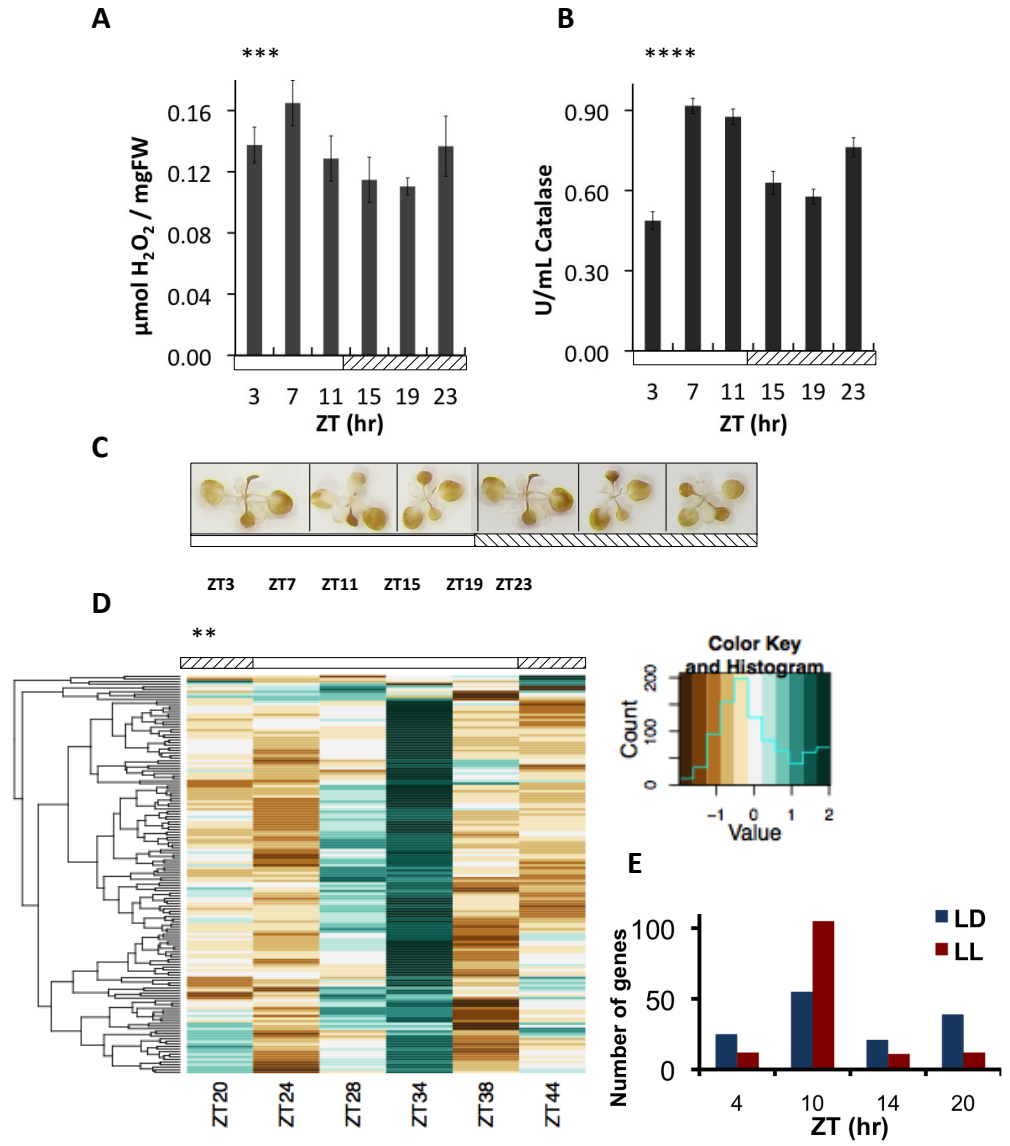


Figure 3.2: The circadian clock regulates ROS homeostasis and signaling

Solid dark grey bars represent (A) H_2O_2 levels and (B) catalase enzyme activity quantified from 16 day old Col-0 plants entrained for 14 days in 12 h light 12 h dark and transferred to constant light (LL) for a day prior to sample collection at every 4 h. Solid bars represent mean values of 3 biological replicates (N=15), with error bars indicating s.e.m. One-way ANOVA (effect of time) for H_2O_2 and catalase data under LL were significant (****P<0.0001, ***P<0.001). (C) Representative images of H_2O_2 accumulation in plants stained with DAB. (D) Expression profiles of 167 ROS transcripts obtained by qPCR from 16-day-old Col-0 plants entrained for 14 days in 16 h light 8 h dark photocycles and transferred to LL for a day prior to sample collection at every 4 to 6 h. Up-regulation and down-regulation of gene expressions were depicted in red and green respectively. Gene expression data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression in long-day were significant (***P<0.0001). Student t-test for expression levels at noon, ZT10 (long-day samples) and subjective mid-day ZT32 (LL samples) were significant (**P<0.01). (E) Genes exhibiting time-of-day specific phases were separated into phase clusters. White bars: day; hatched bars: subjective night.

3.3 A functional clock is required to maintain ROS homeostasis

Results have suggested that the circadian clock regulates ROS homeostasis. In this section, the effects circadian clock genes mutations on ROS homeostasis was investigated. If a functional clock is required for ROS homeostasis, it is then expected that circadian clock mutants would have altered ROS homeostasis. Perturbations in ROS homeostasis can be tested using hypersensitivity assays by administering ROS-generating agents to clock mutants followed by hyper- or hyposensitivity scoring.

Hypersensitivity assays were performed by administering 5µM MV to induce a superoxide burst (Bowler et al., 1992) in plants mis-expressing well-studied clock genes such as *TOC1*, *ELF3*, *ELF4*, *LUX*, *TIC*, *ZTL*, *PRR5, PRR7* and *PRR9* in addition to *CCA1* and *LHY*. Plants carrying mutations in the core clock transcription factors, *cca1-1*, *lhy-11* and *cca1-1/lhy-11* (Mizoguchi et al., 2002), were all hypersensitive to MV treatments (Fig. 3.3). In a reciprocal manner, over-expression of *CCA1* (*CCA1-ox;* (Wang & Tobin, 1998) resulted in MV hyposensitivity (Fig. 3.3). The *elf3-1, elf4-101. lux-1*, *tic-1*, *prr5-1, prr9-1, prr7-3* and *prr5-1/prr9-1* mutants exhibited MV hypersensitivity (Fig. 3.4). On the contrary, *prr7-3/prr9-1* mutant was hyposensitive to MV treatment (Fig. 3.4). Interestingly, the C24 wild-type (WT) accession is completely hyposensitive to 5µM MV treatment (Fig. 3.4). Therefore, the observed MV hyposensitivity of *toc1-1* and *ztl-1* mutants remained inconclusive as both genotypes were of the C24 background (Fig. 3.4).

Hypersensitivity of *cca1-1*, *lhy-11* and *cca1-1/lhy-11* mutants to ROS-generating agents could be a result of altered ROS homeostasis in these genotypes. Indeed, under both LD and LL conditions, *cca1-1/lhy-11* mutants exhibited high H_2O_2 and low catalase levels (Fig. 3.5). Interestingly, it was observed that H_2O_2 levels were clamped low in *CCA1-ox* under both LD and LL (Fig. 3.5*A* and *C*). However, catalase activity in *CCA1-ox* plants was found to be lower than WT levels only during the day (in LD) and subjective day (in LL; Fig. 3.5*B*). Both *cca1-1* and *lhy-11* single mutants exhibited high H_2O_2 levels (Fig. 3.6*C*) in the evening and night under diurnal conditions (Fig. 3.6*A*). In LL conditions, however, H_2O_2 levels were similar to that of the WT in both single mutants (Fig. 3.6*A*). Similar to the profile of *cca1-1/lhy-11* mutants, catalase levels were also

dampened in *cca1-1* and *lhy-11* single mutants during the day in LD and the subjective day LL (Fig. 3.6*B*). *Arabidopsis* has three genes that encode the catalase enzyme, i.e., *CAT1, CAT2* and *CAT3* (Frugoli et al., 1996). All three catalase genes displayed time-of-day specific phases in LL (Fig. 3.5*D*). Under the present growth conditions, *CAT1* and *CAT3* peaked at noon (ZT7) and *CAT2* peaked at dawn (ZT3; Fig. 3.5*D*). This was similar to the expression patterns previously reported by Zhong and McClung (1996). *CCA1* over-expression, loss of *CCA1* and *LHY* functions all resulted in *CAT1*, *CAT2* and *CAT3* arrhythmia (Fig. 3.5*D*).

Here, it was shown that loss of *CCA1* and *LHY* functions resulted in increased ROS production and reduced scavenging (Fig. 3.5*A*, *B* and *C*). Reduction in ROS production was observed upon *CCA1* over-expression (Fig. 3.5*A* and *C*). Taken together, these results suggest that the observed hypersensitivity of *cca1-1*, *lhy-11* and *cca1-1*/*lhy-11* plants to MV (Fig. 3.3) could be attributed to higher basal ROS levels in these genotypes (Fig. 3.5*A* and *C*; Fig. 3.6*A* and *C*). Likewise, the hyposensitivity of *CCA1-ox* to ROS treatments may be due to lower basal ROS levels in this genotype (Fig. 3.5*A* and *C*). The reduction in catalase activity in *CCA1-ox* and *cca1-1/lhy-11* mutants (Fig. 3.5*B*) could also be the result of altered ROS signaling at the transcriptional level, i.e., catalase genes no longer displayed time-of-day specific expression (Fig. 3.5*D*).

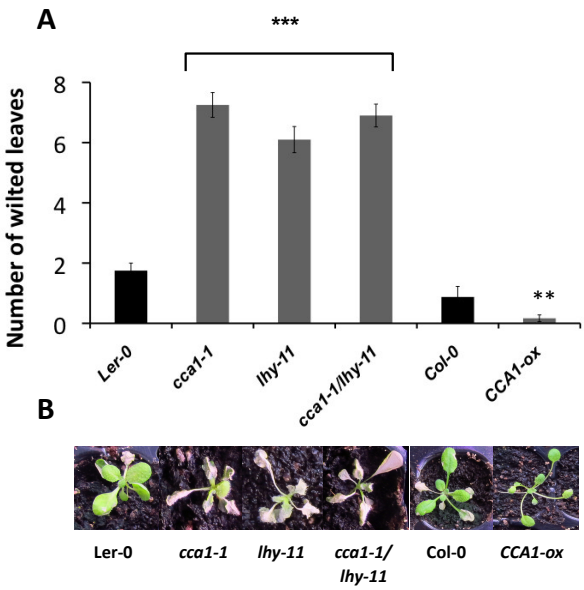


Figure 3.3: Mutations in *CCA1* and *LHY* resulted in ROS hypersensitivity
Plants were entrained for 14 days under 12 h light 12 h dark photocycles and transferred to LL for a day prior to MV treatments administered at ZT3 on day 16. (A) Solid bars represent mean number of wilted leaves of 3 biological replicates (N=15) after 24 h of treatment with 5µM MV, with error bars indicating s.e.m. Student t-test for number of wilted leaves were significant (***P<0.0001 and **P<0.05). (B) Phenotypes of WT and mutant plants after 24 h treatment with 5µM MV. The *cca1-1, lhy-11,* and *cca1-1/lhy-11* mutants are of the Ler-0 background. The *CCA1-ox* mutant is of the Col-0 background.

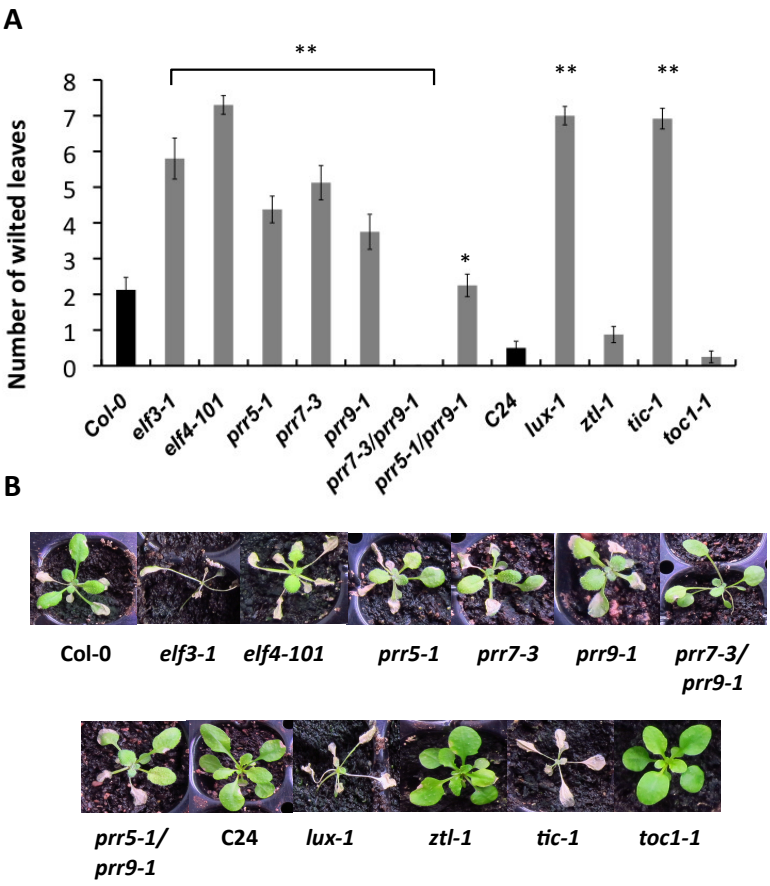
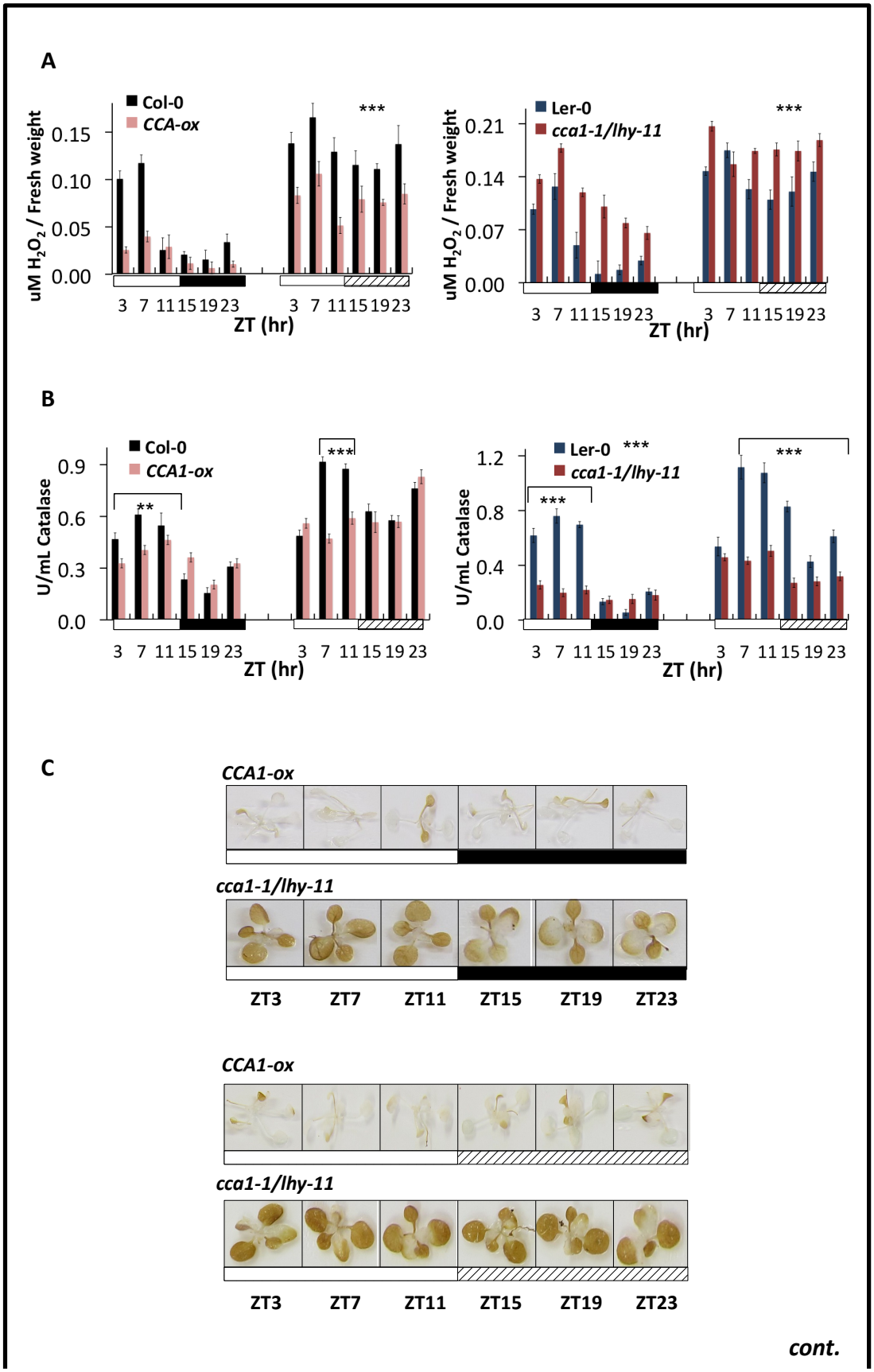


Figure 3.4: A functional clock is required for ROS homeostasis
Plants were entrained for 14 days under 12 h light 12 h dark photocycles and transferred to LL for a day prior to MV treatments administered at ZT3 on day 16. (A) Solid bars represent mean number of wilted leaves of 3 biological replicates (N=15) after 24 h of treatment with 5µM MV, with error bars indicating s.e.m. Student t-test for number of wilted leaves in WT versus mutant plants were significant (**P<0.001 and *P<0.05). (B) Phenotypes of WT and mutant plants after 24 h treatment with 5µM MV. The *elf3-1, elf4-101, prr5-1, prr7-3, prr9-1, prr7-3/pr9-1* and *prr5-1/prr9-1* mutants are of the Col-0 background. The *lux-1, ztl-1, tic-1* and *toc1-1* mutants are of the C24 background.

To further investigate the global clock effect on ROS homeostasis, H_2O_2 levels were determined in other clock mutants. Loss-of-function in clock genes have been known to cause circadian arrhythmia (Dixon et al., 2011; Helfer et al., 2011). In addition, perturbations in ROS signaling could also be the outcome of an arrested circadian oscillator. Indeed, H_2O_2 levels were high in *elf3-1, elf4-101*, and *tic-1* in comparison to WT levels (Fig. 3.6*C*). Low H_2O_2 levels were detected in *prr5-1, prr9-1, prr7-3, prr5-1/prr9-1*, *prr7-3/prr9-1, toc1-1* and *ztl-1* (Fig. 3.6*C*). To test for global clock effects on ROS signaling, temporal expression profiles of *CAT1*, *CAT2* and *CAT3* were obtained from *elf3-1*, *lux-1* and *toc1-1* mutants. It is noteworthy that both *elf3-1* and *lux-1* exhibited hypersensitivity to exogenous ROS (Fig. 3.4). Indeed, the expressions of *CAT1* and *CAT2* were arrhythmic in *elf3-1* and *lux-1* (Fig 3.6*D*). *CAT3* was arrhythmic in *elf3-1* in LL conditions (Fig. 3.6*D*). However, the expression of *CAT3* in *lux-1* is apparently still phased to mid-day (ZT7) in LL (Fig. 3.6*D*). The altered *CAT* expression in *elf3-1* and *lux-1* may be the result of altered ROS homeostasis (Fig. 3.6*C*). In *toc1-1*, however, no significant alteration in *CAT* expression profiles could be observed (Fig. 3.6*D*). Taken together, the results suggested that a link between the circadian clock and ROS signaling exist and is likely to occur through transcriptional mechanisms. The potential mechanisms by which the circadian clock could mediate these responses were next investigated.


A
Col-0
CCA-ox
uM H2O2 / Fresh weight
0.15
0.10
0.05
0.00
***
3 7 11 15 19 23 3 7 11 15 19 23
ZT (hr)
Ler-0
cca1-1/lhy-11
0.21
0.14
0.07
0.00
***
3 7 11 15 19 23 3 7 11 15 19 23
ZT (hr)
B
Col-0
CCA1-ox
U/mL Catalase
0.9
0.6
0.3
0.0
**
***
3 7 11 15 19 23 3 7 11 15 19 23
ZT (hr)
Ler-0
cca1-1/lhy-11
1.2
0.8
0.4
0.0
***
***
***
3 7 11 15 19 23 3 7 11 15 19 23
ZT (hr)
C
CCA1-ox
cca1-1/lhy-11
ZT3 ZT7 ZT11 ZT15 ZT19 ZT23
CCA1-ox
cca1-1/lhy-11
ZT3 ZT7 ZT11 ZT15 ZT19 ZT23

cont.

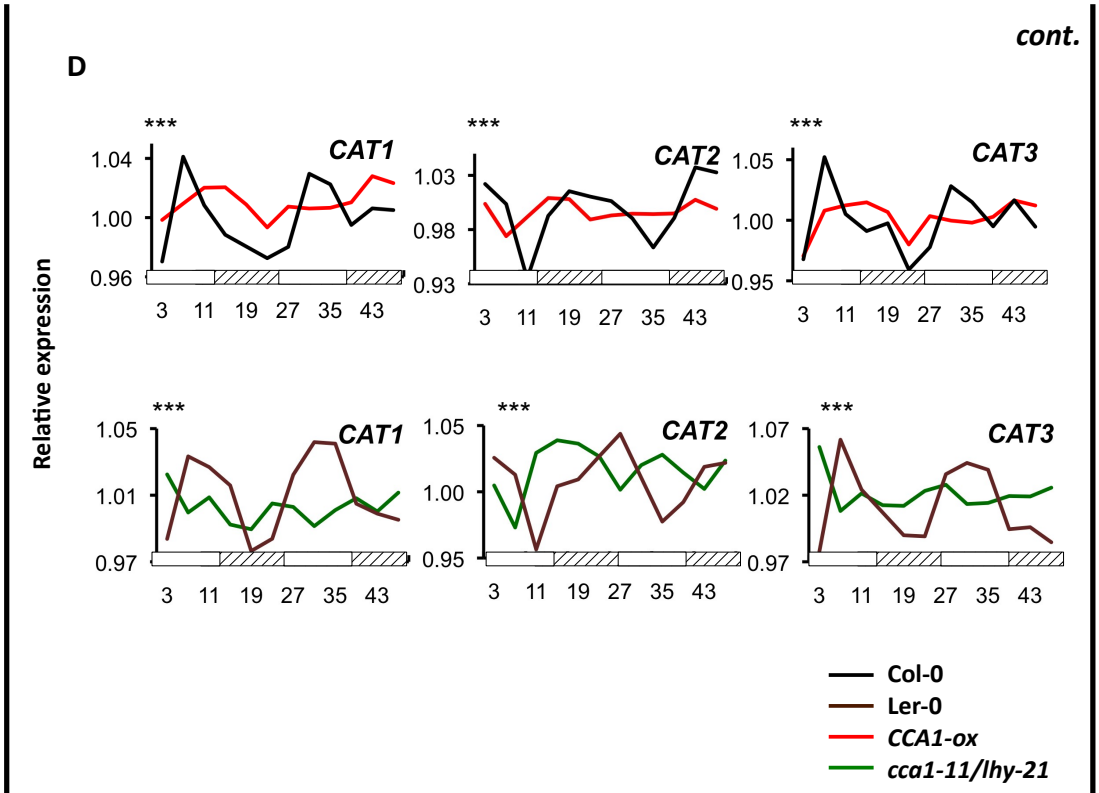
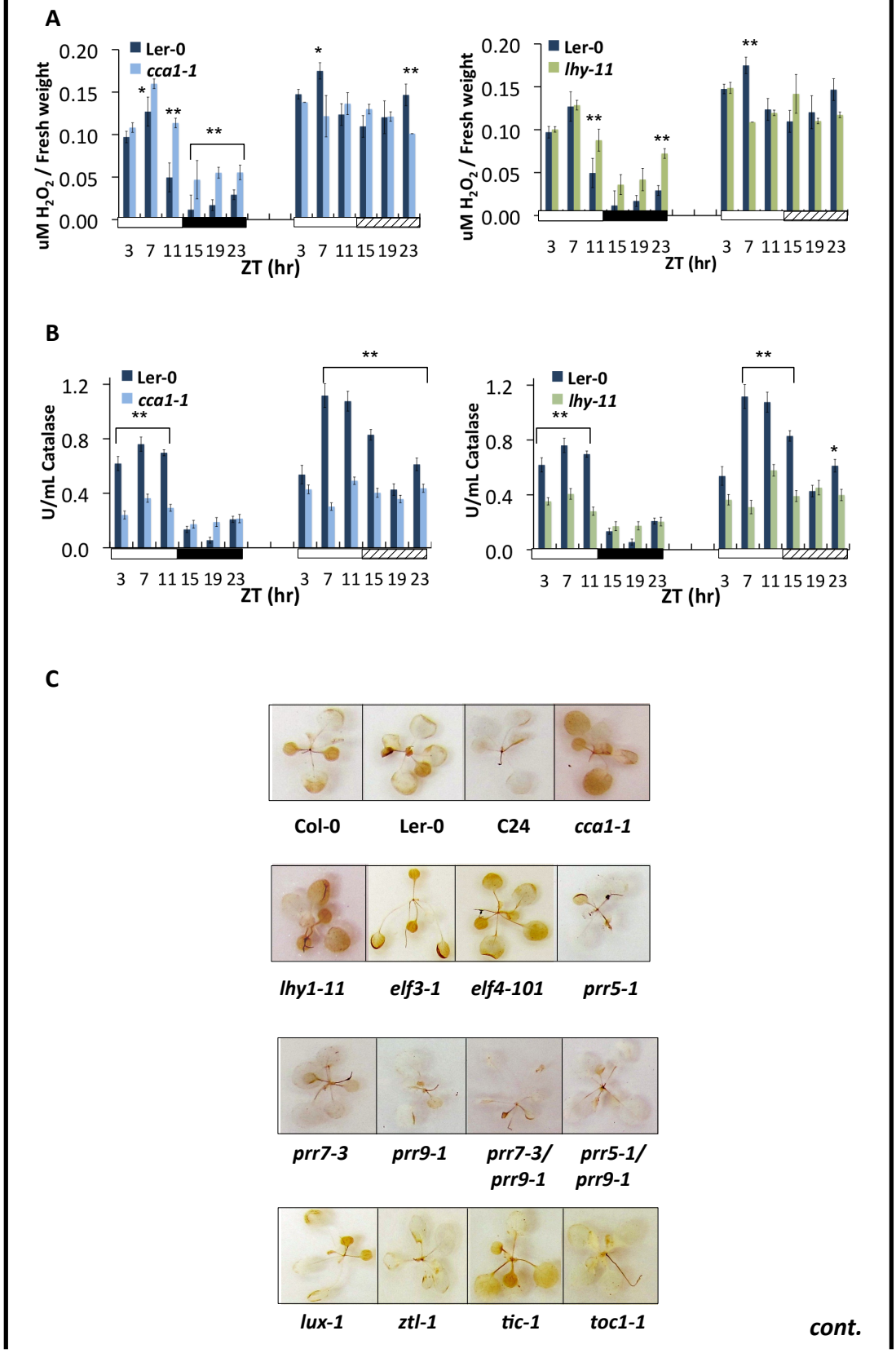


Figure 3.5: H_2O_2 and catalase rhythms are regulated by *CCA1*

Solid bars represent (A) H_2O_2 levels and (B) catalase enzyme activity quantified from 16 day old WT (black bars: Col-0, blue bars: Ler-0), *CCA1-ox* (light red bars, Col-0 background) and *cca1-1/lhy-11* (dark red bars; Ler-0 background) plants entrained in 12 h light 12 h dark photocycles. Solid bars represent mean values of 3 biological replicates (N=12), with error bars indicating s.e.m. One-way ANOVA (effect of group; mutant versus WT) for H_2O_2 and catalase measurements were significant (***P<0.0001, **P<0.01). (C) Representative images of H_2O_2 accumulation in *CCA1-ox* and *cca1-1/lhy-11* plants stained with DAB. (D) Expression profiles of *CAT1*, *CAT2* and *CAT3* from 16 day old Col-0, Ler-0, *CCA1-ox* and *cca1-1/lhy-11* plants that were entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. Gene expression data represent mean values of 3 biological replicates (N=12). One-way ANOVA (effect of time in Col-0 and Ler-0) for *CAT1*, *CAT2* and *CAT3* expressions were significant (***P<0.001). White bars: day; black bars: night, hatched bars: subjective night.


A
Ler-0
cca1-1
lhy-11
uM H2O2 / Fresh weight
0.20
0.15
0.10
0.05
0.00
3 7 11 15 19 23 3 7 11 15 19 23
ZT (hr)
B
U/mL Catalase
1.2
0.8
0.4
0.0
C
Col-0
Ler-0
C24
cca1-1
lhy1-11
elf3-1
elf4-101
prr5-1
prr7-3
prr9-1
prr7-3/ prr9-1
prr5-1/ prr9-1
lux-1
ztl-1
tic-1
toc1-1

cont.

cont.

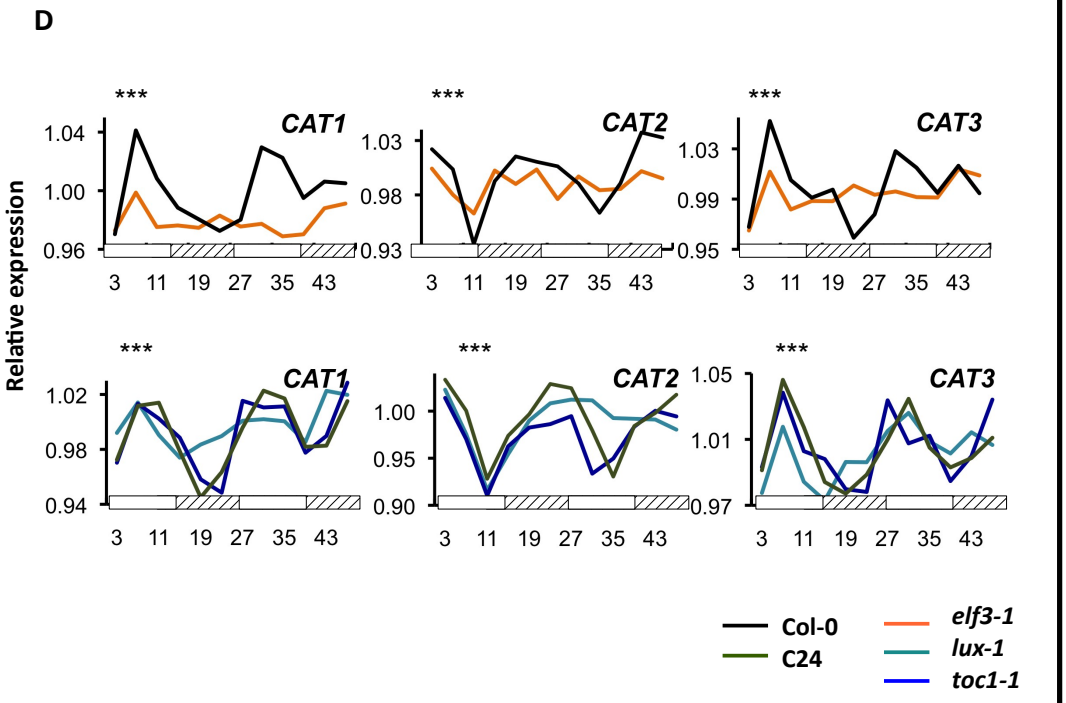


Figure 3.6: ROS homeostasis is altered in clock mutants

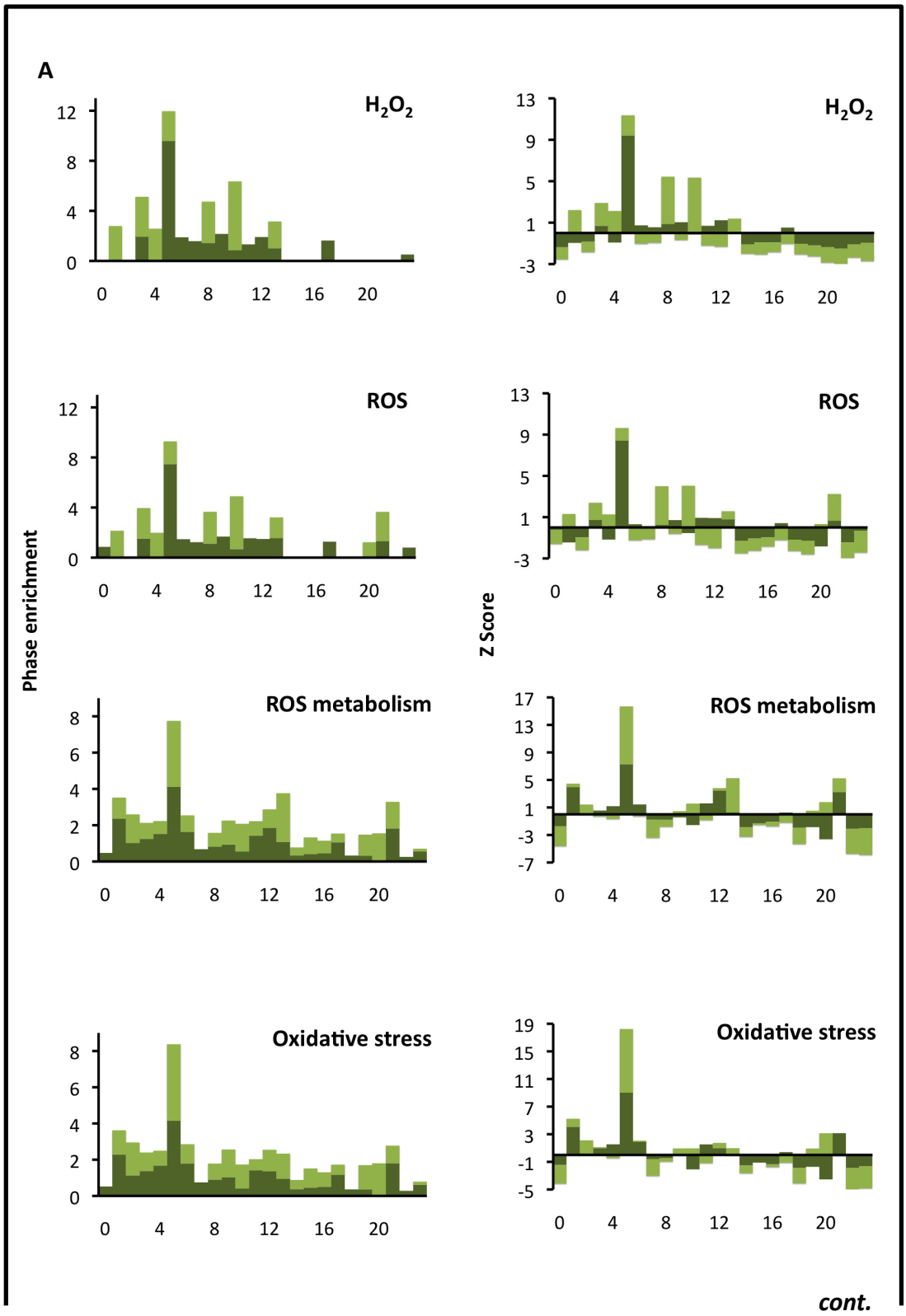
Solid bars represent (A) H_2O_2 levels and (B) catalase enzyme activity quantified from 16 day old WT Ler-0 (dark blue bars), *cca1-1* (light blue bars, Ler-0 background) and *lhy-11* (light green bars; Ler-0 background) plants entrained in 12 h light 12 h dark photocycles. Solid bars represent mean values of 3 biological replicates (N=12), with error bars indicating s.e.m. One-way ANOVA (effect of group; mutant versus WT) for H_2O_2 and catalase measurements were significant (**P<0.001, *P<0.05). (C) Representative images of H_2O_2 accumulation in clock mutants stained with DAB. (D) Expression profiles of *CAT1*, *CAT2* and *CAT3* from 16 day old Col-0, C24, *elf3-1, lux-1*and *toc-1* plants that were entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. Gene expression data represent mean values of 3 biological replicates (N=12). One-way ANOVA (effect of time in WT) for *CAT1*, *CAT2* and *CAT3* expressions were significant (***P<0.001). White bars: day; black bars: night, hatched bars: subjective night. The *elf3-1, elf4-101, prr5-1, prr7-3, prr9-1, prr7-3/pr9-1* and *prr5-1/prr9-1* mutants are of the Col-0 background. The *lux-1, ztl-1, tic-1* and *toc1-1* mutants are of the C24 background.

3.4 The Evening Element is enriched in ROS-responsive genes

In the previous section, mutations in *CCA1*, *LHY*, *TOC1*, *ELF3* and *LUX* appear to affect the transcription of catalase genes. The clock exerts a global effect on ROS signaling at the transcriptional level possibly to time ROS production and scavenging to a specific time of the day in order to minimize energy consumption and maximize productivity. The *cca1-1*/*lhy-11* mutant appears to accumulate H_2O_2, possibly due to the abolishment of rhythmic expression of catalase genes. Interestingly, overexpression of *CCA1* produced the opposite phenotype; that is being hyposensitive to ROS-agents and having reduced levels of H_2O_2. Since mutations in *CCA1* produced the most dramatic phenotype, it is predicted that *CCA1* would act as a master regulator of the ROS response. To investigate the role of *CCA1* in transcriptional coordination of ROS genes, promoter analysis for motifs recognized by CCA1 (EE and CBS) was performed.

Promoter regions (1000bp upstream of the ATG site) of ROS-responsive genes obtained from publicly available microarray datasets were analyzed using the ATHENA tool (O'Connor et al., 2005). Enrichments of the circadian-regulated EE and/or CBS (Table 3.2) were found in the promoters, which suggested the regulation of these genes by CCA1. Of the previously selected 167 ROS genes used for expression profiling, 32 genes contained the CBS and 15 genes contained the EE in their promoters (Table 3.3). The ATCOECIS tool (Vandepoele et al., 2009) was used to obtain a list of all *Arabidopsis* genes grouped under various ROS-related GO categories (referred to as ROS GO genes hereafter) to investigate whether these genes exhibit circadian rhythmicity. With the DIURNAL tool (Michael et al., 2008b) it was observed that out of the 517 ROS GO genes, on average 73% and 39% of the genes were rhythmic in two or more diurnal and circadian conditions respectively (Table 3.4). Of these rhythmic genes, the PHASER tool (Michael et al., 2008b) was used to investigate the time-of-day at which they peak. As expected, the mid-day phase (ZT5) was found to be overrepresented in ROS GO genes (Z-score > 1.96: 95% prediction interval; Fig. 3.7*A*) when cycling calls were made on two 12 h light 12 h dark diurnal datasets (LDHH_SM and LDHH_ST; Smith et al., 2004; Blasing et al., 2005). Interestingly, cycling calls on three short-day (8 h light 16 h dark) datasets (COL_SD, LER_SD and LHY_OX) revealed that the ZT4 phase was enriched in ROS GO genes (Z-score > 1.96), with the LHYOX_SD

dataset (obtained from plants overexpressing *LHY*) showing the highest enrichment the first two GO categories (Fig. 3.7*B*). These results were further supported by an independent study where 34% of ROS-responsive genes were found to be clock-regulated with a majority of them peaking in the subjective day (Covington et al., 2008). Thus, it appears that the clock controls ROS signaling at the transcriptional level possibly through the association of CCA1 with the EE and/or CBS in promoters of ROS genes.


A
H2O2
H2O2
ROS
ROS
ROS metabolism
ROS metabolism
Oxidative stress
Oxidative stress
Phase enrichment
Z Score
cont.

cont.

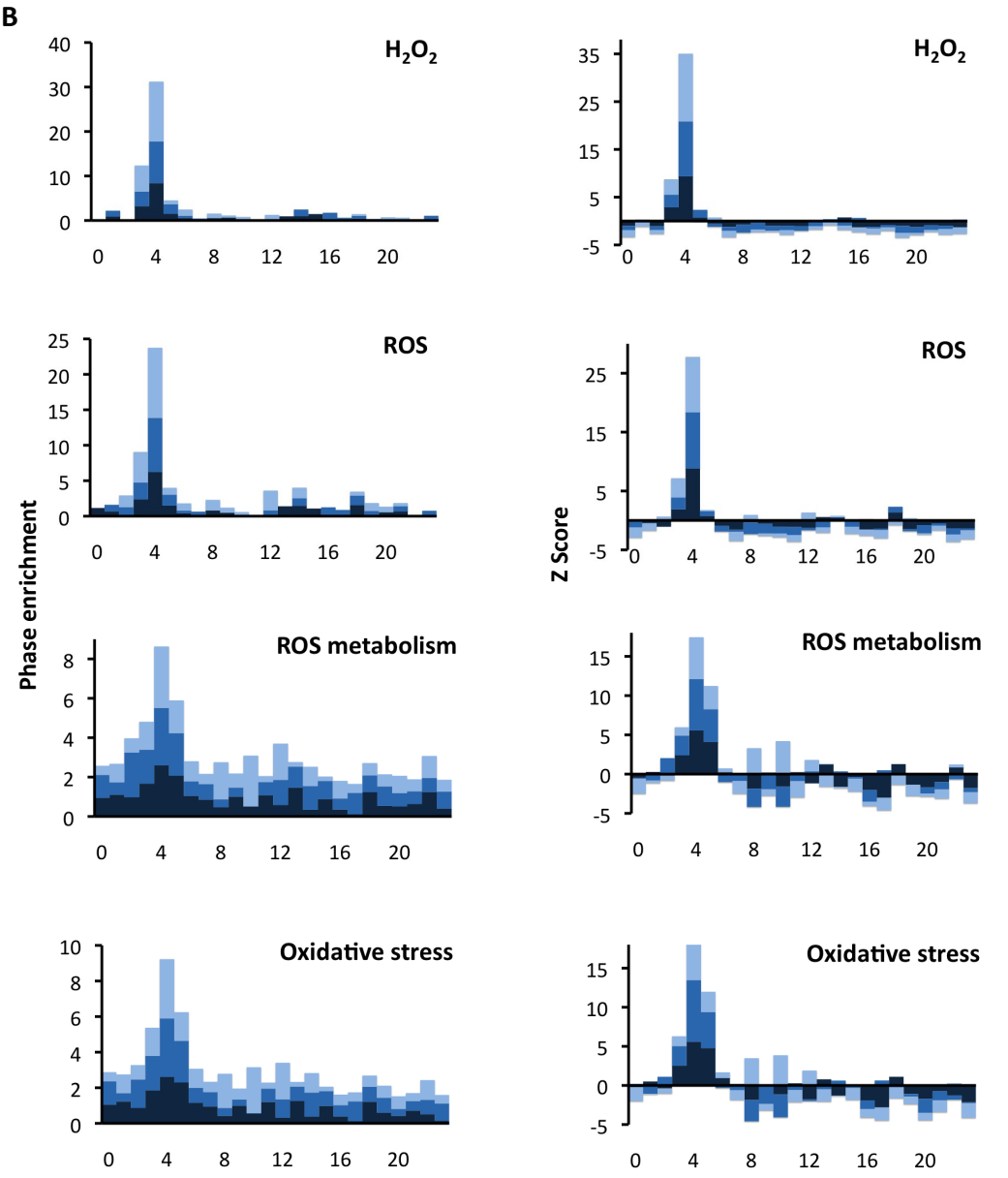


Figure 3.7: Phase distribution of ROS genes for the four GO categories (ROS GO genes) and the datasets that were used to call cycling genes (LDHH_SM, LDHH_ST, COL_SD, LER_SD and LHYOX_SD).

(A) PHASER revealed that the 517 genes (GO H_2O_2: 32 genes; GO ROS: 45 genes; GO ROS metabolism: 232 genes; GO oxidative stress: 208 genes) exhibited the highest Z-score (significance of overrepresentation at any given time of day) at ZT5 in plants grown under 12 h light 12 h dark photocycles (LDHH_SM and LDHH_ST). (B) ROS GO genes exhibited the highest Z-score at ZT4 in plants grown under 8 h light 16 h dark photocycles (COL_SD, LER_SD and LHYOX_SD). LDHH_SM: light green, LDHH_ST: dark green, LHYOX_SD: light blue, LER_SD: blue, COL_SD: dark blue.

Table 3.2: Enrichments of EE and CBS in ROS-responsive genes obtained from publicly available microarray datasets

ROS conditions	CBS (p-value)	EE (p-value)	References
Up-regulated in the ex1/flu, ex2/flu and ex1/ex2/flu	**0.0437**	**0.0016**	Lee et al., 2007
Up & down-regulated in gr1 (glutathione reductase 1) mutant	**0.0132**	**0.0078**	Mhamdi et al., 2010
Transcripts up-regulated specifically in flu, O_3 and PQ experiments	**0.0227**	**0.0108**	Gadjev et al., 2006
Up-regulated in flu (fluorescent) mutant	**0.0889**	**0.0345**	Gadjev et al., 2006
Up-regulated in KD-SOD (Cu/Zn SOD knock-down) mutant	0.4882	**<10-5**	Gadjev et al., 2006
Up-regulated in CAT2HP1 (catalase deficient) mutant	0.2058	**<10-4**	Gadjev et al., 2006
Down-regulated in KO-APX (ascorbate peroxidase knock-out) mutant	0.393	**<10-3**	Gadjev et al., 2006

Cis-element enrichment values (in bold) were obtained using the ATHENA tool (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/visualize_select.pl; O'Connor et al., 2005). Numerical values in bold represent $p<0.05$.

Table 3.3: Positions of putative EE and CBS in promoters of ROS-responsive genes obtained using the ATHENA tool

Gene name	AGI	EE (AAAATATCT)	CBS (AAMAATCT)
NIMIN1	AT1G02450		(+)-161
DUF676	AT1G10040		(+)-441; (+)-565
VPS4	AT1G13340		(-)-961
MYB51	AT1G18570		(+)-690
IGMT2	AT1G21120		(+)-727
FAD	AT1G26380		(+)-271
Unknown protein	AT1G28190		(+)-296
GRX480	AT1G28480		(-)-653
Unknown protein	AT1G35210		(+)-122
SHM7	AT1G36370		(-)-744
Unknown protein	AT1G61340	(-)-135	(-)-368
Hydroxyproline-rich glycoprotein family protein	AT1G63720		(+)-457
Chaperone DnaJ-domain superfamily protein	AT1G71000	(-)-161; (+)-234	(+)-120
PLAT/LH2	AT1G72520		(-)-413
Leucine-rich repeat protein kinase family protein	AT1G74360		(+)-917
NAC032	AT1G77450		(+)-353
ATCCR2	AT1G80820		(+)-201
WRKY40	AT1G80840		(+)-588
RNS1	AT2G02990		(-)-464
HSP20	AT2G29500		(+)-148
NAT	AT2G32030		(-)-425
BAG6	AT2G46240		(-)-46
Unknown protein	AT3G10020		(-)-48
HSP70	AT3G12580		(+)-559
NAI2	AT3G15950		(-)-733
ERF/AP2	AT3G23230	(-)-485	(-)-462
AGC2	AT3G25250		(+)-573; (-)-942
CYP81D11	AT3G28740		(+)-869
HSP17.4	AT3G46230		(+)-541

LSU1	AT3G49580		(+)-843
CDC48	AT3G53230		(+)-227
DREB2A	AT5G05410	(-)-265; (+)-476	(+)-42
DUF1645	AT5G14730	(+)-804	(+)-876; (-)-913
ChaC-like family protein	AT5G26220		(-)-450
Glycine-rich protein	AT5G28630		(+)-620
Protein kinase superfamily protein	AT5G46080		(-)-34; (+)-115
ERF5	AT5G47230		(+)-372
ROF2	AT5G48570		(+)-187
ERF/AP2	AT5G51190		(+)-828
PBP1	AT5G54490		(-)-108; (+)-598
ZAT12	AT5G59820		(+)-303
Concanavalin A-like lectin protein kinase family protein;	AT5G65600		(-)-141

Cis-element sites were obtained using the ATHENA tool (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/visualize_select.pl; O'Connor et al., 2005). +' denotes sense strand and '-' denotes antisense strand.

Table 3.4: *Arabidopsis* genes under five ROS GO categories that are called rhythmic by the DIURNAL tool under at least two diurnal and circadian conditions

GO category	GO term	% Diurnal in ≥ 2 conditions*	% Circadian in ≥ 2 conditions*
GO:0042542	Response to H_2O_2	81	47
GO:0000302	Response to ROS	82	39
GO:0006800	ROS metabolism	64	42
GO:0006979	Response to oxidative stress	63	27

ROS genes under different ROS-related GO categories were obtained using the ATCOECIS (http://bioinformatics.psb.ugent.be/ATCOECIS/) resource (Vandepoele et al., 2009). Rhythmicity of ROS GO genes was determined using the DIURNAL (http://diurnal.cgrb.oregonstate.edu/; Michael et al., 2008b) tool. *Appendix 5.9.

Table 3.5: GO overrepresentation of genes involved in ROS signaling

GO enrichment	Enrichment fold	P-value	Description
GO:0006979	45.49	2.48E-09	Response to oxidative stress
GO:0006800	41.75	4.14E-09	Oxygen and reactive oxygen species metabolism
GO:0042221	12.25	1.06E-08	Response to chemical stimulus
GO:0006950	11.48	1.52E-07	Response to stress
GO:0009628	8.51	2.45E-07	Response to abiotic stimulus
GO:0009408	60.36	5.09E-07	Response to heat
GO:0042542	142.87	1.16E-06	Response to H_2O_2
GO:0000302	103.91	3.09E-06	Response to reactive oxygen species
GO:0003700	11.62	1.65E-05	Transcription factor activity
GO:0030528	10.52	2.69E-05	Transcription regulator activity
GO:0003677	8.87	6.15E-05	DNA binding
GO:0003676	6.54	2.69E-04	Nucleic acid binding
GO:0045449	7.35	5.79E-03	Regulation of transcription
GO:0019219	7.28	5.95E-03	Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
GO:0006350	6.96	6.75E-03	Transcription
GO:0045941	130.63	7.63E-03	Positive regulation of transcription
GO:0045935	126.41	7.88E-03	Positive regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
GO:0015931	122.46	8.14E-03	Nucleobase, nucleoside, nucleotide and nucleic acid transport
GO:0016564	118.75	8.39E-03	Transcriptional repressor activity
GO:0015932	105.92	9.40E-03	Nucleobase, nucleoside, nucleotide and nucleic acid transporter activity

Go enrichments were obtained from the ATCOECIS resource (http://bioinformatics.psb.ugent.be/ATCOECIS/; Vandepoele et al., 2009).

3.5 ROS-responsive genes display time-of-day specific expression phase in anticipation of oxidative stress under regular growth conditions

Enrichments for EE and/or CBS in promoters of ROS genes were observed previously. Along with the altered sensitivity and the arrhythmic expression of ROS genes in *CCA1* mutants, it is likely that *CCA1* is required for the ROS response. To investigate the mechanism of mis-regulated ROS signaling in clock mutants, expression profiling of 25 ROS genes were obtained *in WT, CCA1-ox, cca1-1/lhy-11, elf3-1, lux-1* and *toc1-1* plants in a 48h time course under LL conditions.

These genes were selected based on two criteria: 1) Genes involved in ROS signaling (Table 3.5) and 2) ROS genes that have EE and/or CBS in their promoter regions (Table 3.6). Two ROS genes that do not have the EE or CBS, the *RESPIRATORY BURST OXIDASE HOMOLOG C* (*RBOHC*) and *HEAT SHOCK TRANSCRIPTION FACTOR A4A* (*HSFA4A*) were also selected for this experiment. *RBOHC* is involved in ROS signaling in roots (Foreman et al., 2003; Monshausen et al., 2007) and *HSFA4A* is involved in H_2O_2 sensing (Davletova et al., 2005). Of the 25 genes, 18 genes exhibited time-of-day specific phases in WT plants (one-way ANOVA, P<0.001; Fig. 3.8*A*, Table 3.6) while no overt phases could be detected in the remaining seven genes (Fig. 3.10).

Table 3.6: Positions of putative EE and CBS in promoters of selected ROS-responsive genes used for expression analysis and ChIP-qPCR

Gene name	AGI	EE (AAAATATCT)	CBS (AAMAATCT)	Phase
CCT	AT1G07050	(-)-880	(-)-491	ND
BCAT	AT1G10060	(-)-54		Noon
BBOX	AT1G28050	(+)552; (+)-531; (-)-268		ND
ADOF1	AT1G51700		(+)-41; (+)-792	Evening
PLATZ	AT1G76590	(+)-316	(+)-862	Midnight
ATIPS2	AT2G22240	(+)-657		Midnight
PEROXIDASE	AT2G22420		(-)606; (+)-710	Noon
PAL1	AT2G37040		(+)-364	Noon
INVC	AT3G06500	(+)-576		Noon

JMJD5	AT3G20810	(+)-11		Noon
ZAT7	AT3G46090		(+)-389	ND
APX4	AT4G09010		(+)-117	Noon
ISA3	AT4G09020	(-)-198	(+)-492	Midnight
GSH1	AT4G23100		(-)-387	ND
VTC2	AT4G26850		(-)-278	Midnight
PDX1	AT5G01410		(-)-360; (-)-680	ND
DREB2A	AT5G05410	(-)265; (+)-476	(+)-42	Midnight
COR27	AT5G42900	(+)-129; (-)-148		Noon
PLATZ	AT5G46710	(-)-119		ND
MES18	AT5G58310		(+)-537	Evening
LTP3	AT5G59320		(+)-245	Midnight
HSP18.2	AT5G59720		(+)-153; (-)-458; (-)-565	Midnight
MYB59	AT5G59780		(+)-640	Evening
RBOHC	AT5G51060	NA	NA	ND
HSFA4A	AT4G18880	NA	NA	Noon
CAT1	AT1G20630	(-)-50		Noon
CAT2	AT4G35090	NA		Dawn
CAT3	AT1G20620	(+)-182		Noon

Cis-element sites were obtained using the ATHENA tool (http://www.bioinformatics2.wsu.edu/cgi-bin/Athena/cgi/visualize_select.pl; O'Connor et al., 2005). ND: none detected. +' denotes sense strand and '-' denotes antisense strand.

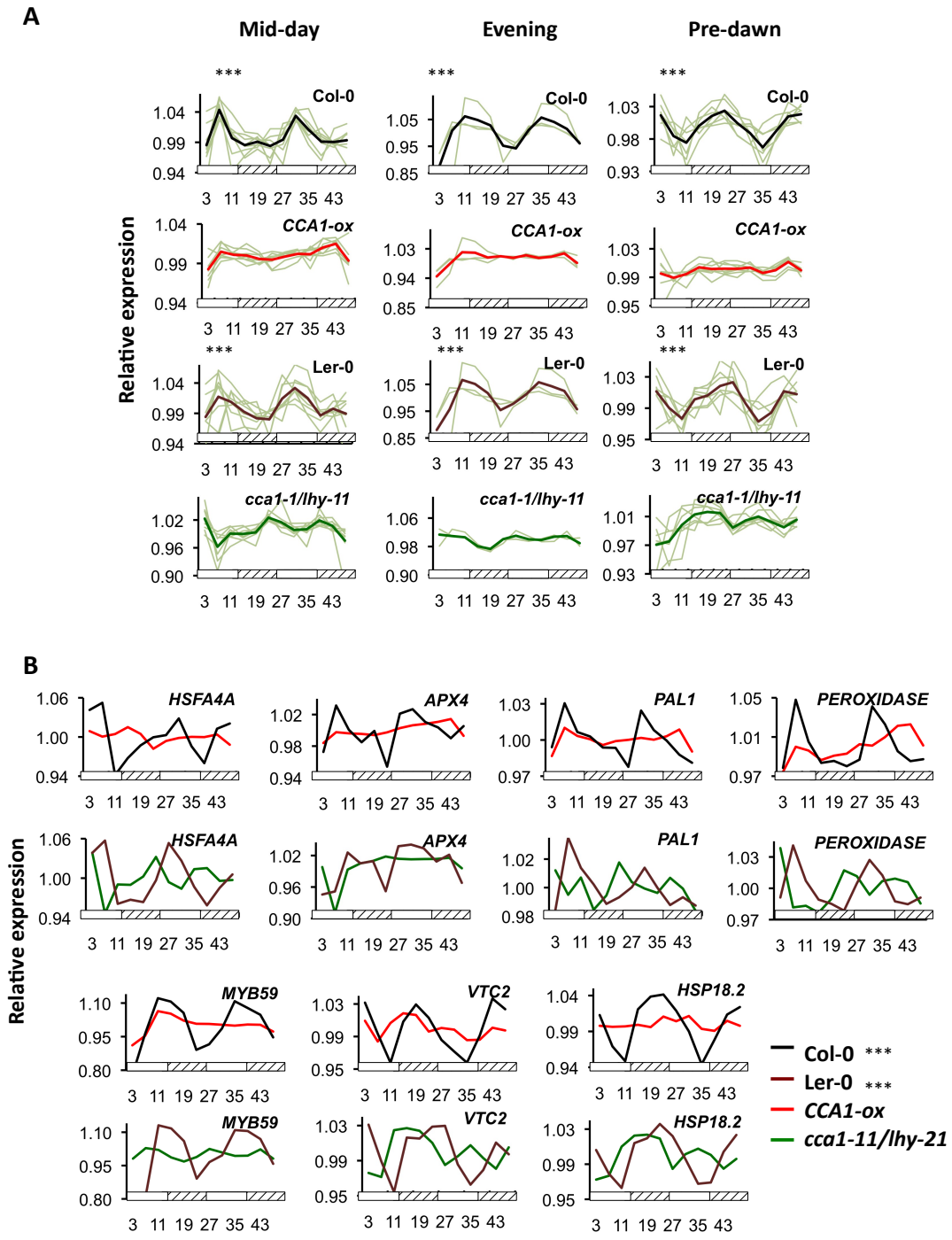


Figure 3.8: *CCA1* and *LHY* are involved in transcriptional coordination of ROS gene expression
(A) Expression profiles of three ROS gene clusters were obtained from 16 day old plants entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. *CCA1-ox* is of the Col-0 background and *cca1-1/lhy-11* is of the Ler-0 background. Bolded lines represent mean expression values of ROS genes in four genotypes; Col-0, *CCA1-ox,* Ler-0 and *cca1-1/lhy-11*. (B) Expression profiles of representative genes from three ROS gene clusters in Col-0, *CCA1-ox,* Ler-0 and *cca1-1/lhy-11.* Data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression was significant in WT samples (***P<0.001). White bars: day; hatched bars: subjective night.

These 18 genes can be grouped into three clusters based on the time when they peak in expression, i.e., mid-day (ZT7), evening (ZT11) and pre-dawn (ZT23; Fig. 3.8*A*). Previous studies have reported that genes encoding enzymes of the flavanoid biosynthetic pathway, which act as 'sunscreens' to absorb UV light, peaked at pre-dawn (Shirley, 1996; Harmer et al., 2000). Indeed, it was observed that the gene that encodes a flavanoid ROS scavenger (Foyer & Noctor, 2005), *VITAMIN C 2* (*VTC2*), had a pre-dawn phase (Fig. 3.8*B*). Phase-specific expression of *VTC2* was abolished altogether in *CCA1-ox* and *cca1-1/lhy-11* (Fig. 3.8*B*). In *lux-1*, *VTC2* displayed a broadened peak while in *elf3-1* gene expression amplitude was altered (Fig. 3.9*B*). Another well-studied H_2O_2 scavenger is the ascorbate peroxidase enzyme where deficiency in this enzyme resulted in light-induced necrosis (Giacomelli et al., 2007). The arrhythmic expression profiles of *ASCORBATE PEROXIDASE 4* (*APX4*) in *CCA1-ox* and *cca1-1/lhy-11* were distinct from that of the WT where a mid-day phase (ZT7) could be observed in WT plants (Fig. 3.8*B*). Interestingly, the phase of *APX4* appeared to be shifted in *elf3-1* and *lux-1* (Fig. 3.9*B*). Notably, examples of other well-studied genes that exhibited phase-specific expression were the pre-dawn phased (ZT23) gene *HEAT SHOCK PROTEIN 18.2* (*HSP18.2*), the mid-day phased (ZT7) genes *PHENYLALANINE AMMONIA-LYASE 1* (*PAL1*) and *PEROXIDASE* and the evening-phased (ZT11) gene *MYB DOMAIN PROTEIN 59* (*MYB59*; Fig. 3.8*B*; Stracke et al., 2001; Rohde et al., 2004; Cao et al., 2006a; Nishizawa et al., 2006; Mu et al., 2009). All four genes were arrhythmic in *CCA1-ox* and *cca-1/lhy-11* (Fig. 3.8*B*) while only *PAL1*, *PEROXIDASE* and *MYB59* showed altered expression in *elf3-1, lux-1* and *toc1-1* (Fig. 3.9*B*). In general, phase-specific expressions of the remaining 16 genes from all three clusters were abolished altogether in *CCA1-ox* and *cca1-1/lhy-11*, hence providing affirmation that rhythmic oscillation in *CCA1* mRNA *per se* is essential for transcriptional coordination of ROS genes (Fig. 3.8*A*). In *elf3-1*, *lux-1* and *toc-1,* mid-day phased genes showed dramatically altered expression with many of them becoming arrhythmic (Fig. 3.9*A*). Interestingly, although no EE/CBS could be found in the promoter of *HSFA4A*, its expression was phased to mid-day (ZT7) in WT but became arrhythmic in *CCA1-ox* and *cca1-1/lhy-11* (Fig. 3.8*B*). Genes from the evening and pre-dawn cluster displayed altered amplitudes in *toc1-1* (Fig. 3.9*A*). In *elf3-1*, evening and pre-dawn genes were arrhythmic (Fig. 3.9*A*). Evening phased genes showed altered phase of expression in

lux-1, while the pre-dawn cluster still appeared be correctly phased although showing a broadened peak (Fig. 3.9*A*). The alteration in gene expression of circadian-regulated ROS transcripts in the clock mutants tested provided additional support that the circadian clock affects the transcriptional coordination of ROS network.

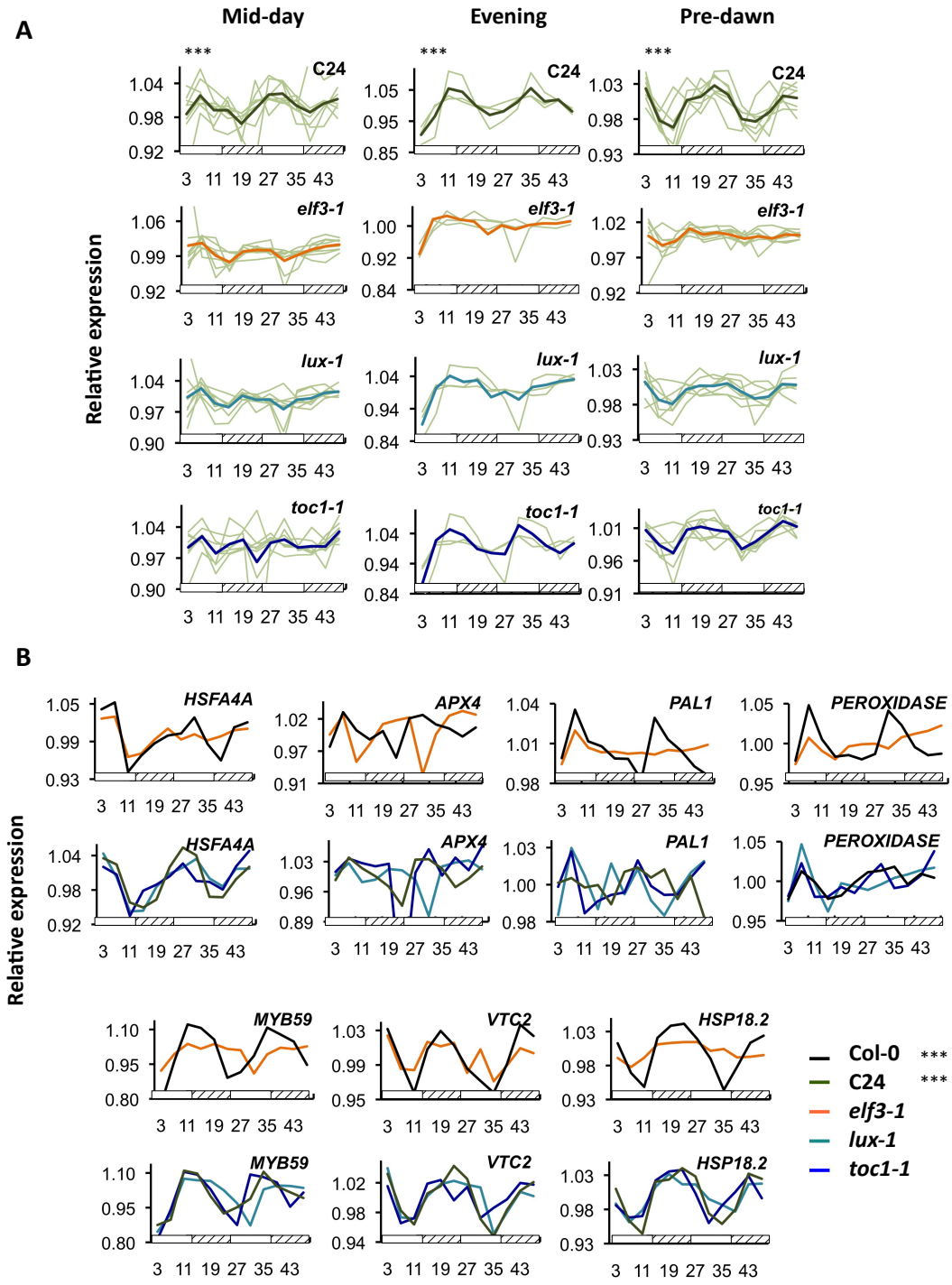


Figure 3.9: *ELF3*, *LUX* and *TOC1* are partly involved in transcriptional coordination of ROS genes expression

(A) Expression profiles of three ROS gene clusters were obtained from 16 day old plants entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. *Elf3-1* is of the Col-0 background and *lux-1* and *toc1-1* are of the C24 background. Bolded lines represent mean expression values of ROS genes in five genotypes; Col-0 (Fig. 3.8*A*), C24, *elf3-1*, *lux-1* and *toc1-1*. (B) Expression profiles of representative genes from three ROS gene clusters in Col-0, C24, *elf3-1*, *lux-1* and *toc1-1*. Data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression was significant in WT samples (***P<0.001). White bars: day; hatched bars: subjective night.

As mentioned previously, seven of the ROS genes tested did not display any specific time-of-day phase. Interestingly, all seven genes exhibited altered expression in *CCA1-ox, cca1-1/lhy-11, elf3-1, lux-1* and *toc1-1* plants (Fig. 3.10), which suggested potential for regulation of ROS signaling beyond circadian-regulated ROS genes. Of the seven genes, *PYRIDOXINE BIOSYNTHESIS 1* (*PDX1*) encodes a ROS quencher pyridoxine (vitamin B6) that increased in levels upon UV-B irradiation (Ristilä, 2011). Noticeably, the expression of *PDX1* was up-regulated in *cca1-1/lhy-11* (Fig. 3.10). Although *RBOHC*, another non-EE/CBS containing gene, did not show any time-of-day specific phase, it is interesting that this gene was up-regulated in *CCA1-ox* and *toc1-1* mutants (Fig. 3.10). Taken together, the results suggest that the clock also mediates the expression of non-circadian ROS transcriptome possibly through the recruitment of additional regulators or though the association of clock genes to other cognate cis-regulatory elements.

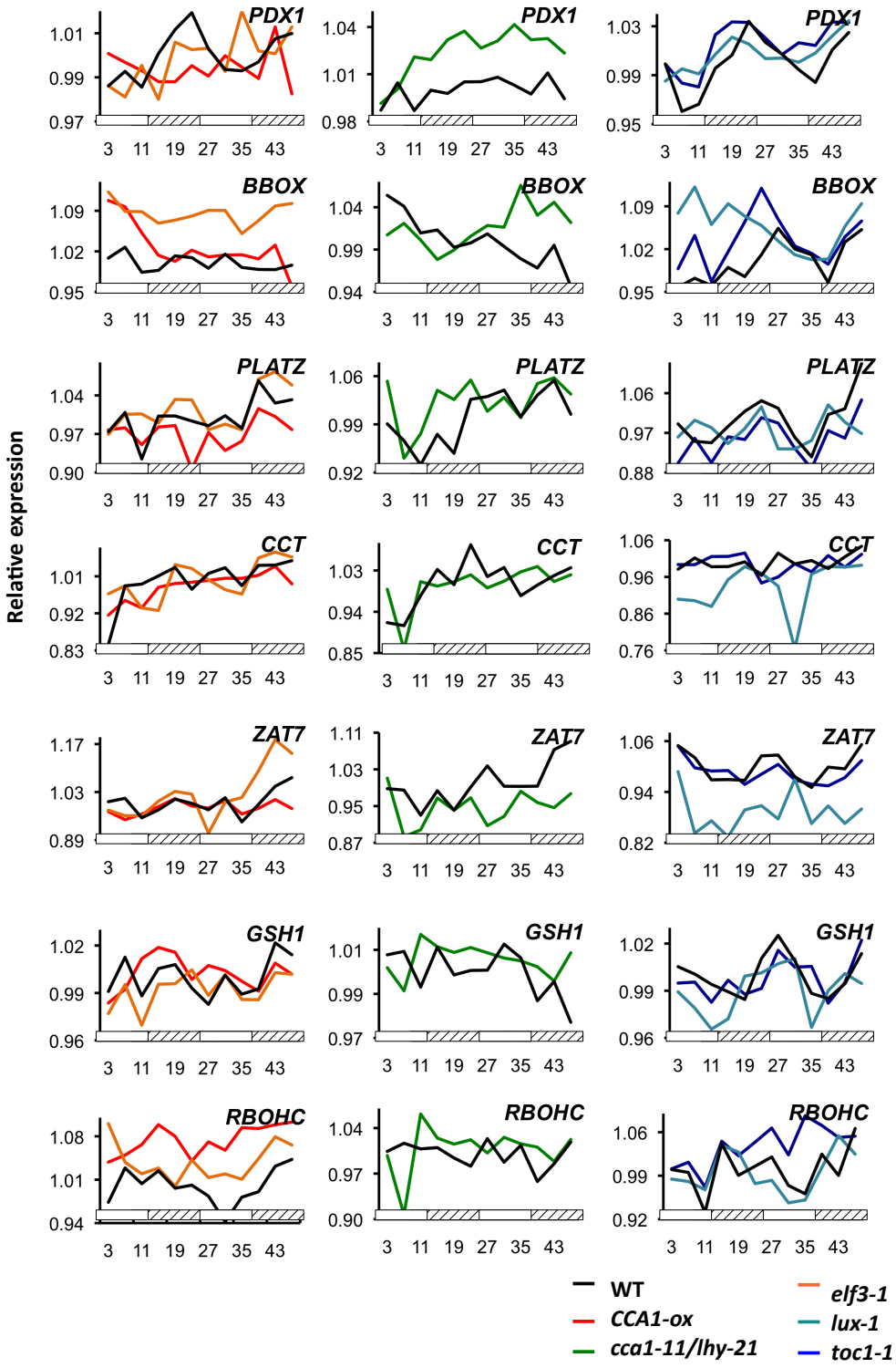


Figure 3.10: Clock regulation of non-circadian ROS transcripts
Expression profiles of seven non-circadian ROS genes were obtained from 16 day old plants entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. Mutant expression data were grouped according their WT background as represented in 3 columns: Col-0: 1st column; Ler-0: 2nd column; C24: 3rd column. *CCA1-ox* and *elf3-1* (1st column) are of the Col-0 background; *cca1-1/lhy-11* (2nd column) are of the Ler-0 background; *lux-1* and *toc1-1* (3rd column) are of the C24 background. Data represent mean values of 3 biological replicates (N=12). White bars: day; hatched bars: subjective night.

3.6 *CCA1* regulates plants' response to oxidative stress

The circadian clock in plants may regulate phase-specific expression of ROS genes under regular growth conditions to allow the anticipation of oxidative stress according to a diurnal schedule. Previous results suggest that this anticipatory response is mediated by *CCA1*. Therefore, plants would be responsive to ROS at dawn when *CCA1* is expressed and not at other times of the day. To test this, WT plants were treated with 2µM MV at three different times in a single diurnal cycle, i.e., morning (ZT3), evening (ZT11) and midnight (ZT19).

In order to determine the effect that *CCA1* has on ROS signaling, qPCR analysis were performed on five ROS-related transcription factors (TFs), i.e., *SALT TOLERANCE ZINC FINGER* (*ZAT10*), *ZAT12*, basic helix-loop-helix DNA-binding superfamily protein (*bHLH128*), *WRKY DNA-BINDING PROTEIN 11* (*WRKY11*) and *ETHYLENE RESPONSE FACTOR 2* (*ERF2*) along with two regulatory genes associated with ROS signaling, i.e., *PURINE PERMEASE 1* (*PUP1*) and Ca^{2+} *BINDING PROTEIN 1* (*AtCP1*) (Brenner et al., 2005; Davletova et al., 2005; Journot-Catalino et al., 2006; Mittler et al., 2006) on MV-treated plants at the three selected ZTs. In all the seven genes tested, treatments in the morning (ZT3) resulted in significant fold-inductions (Student's t-test, $P<0.0001$) (Fig. 3.11*A*). However, this induction was not observed in evening or night treated plants (Fig. 3.11*A*), consistent with the prediction that the system is most responsive in the morning when *CCA1* is expressed. Interestingly, in six genes, down-regulation (Student's t-test, $P<0.01$) of gene expression was observed in evening- (ZT11) treated plants (Fig. 3.11*A*). To determine if the response was mediated through *CCA1*, the response to ROS treatments was examined in the *CCA1-ox* background. Interestingly, the tested genes were able to respond to ROS treatments at all three tested times (Fig. 3.11*B*) presumably because in the *CCA1-ox* background, *CCA1* is continuously present all the time to permit the response. The diurnal response to ROS, observed in WT plants with the peak response at ZT3 (Fig. 3.11*A*), was abolished altogether in *CCA1-ox* (Fig. 3.11*B*). However, in comparison with WT plants, the fold-induction was reduced in *CCA1-ox* in a dose-dependent manner (Fig. 3.11*B*; Fig. 3.12). Here, plants' response to ROS was investigated by determining time-of-day specific fold-induction of known ROS TFs and ROS regulators. One can assume that if these genes display time-of-day

specific expression phase in the absence of exogenous ROS and if their temporal expression profiles are altered in plants mis-expressing *CCA1*, it would further imply that the clock mediates an anticipatory response to oxidative stress through *CCA1*. Indeed, with the exception of *ZAT10*, time-of-day specific phasing was observed in six out of the seven genes under LL conditions and all seven genes were either arrhythmic or displayed altered expression in *CCA1-ox* and *cca1-1/lhy-11* plants (Fig. 3.11*C*). To address the question of a global clock effect on the response to ROS, the expressions of all seven genes were tested *elf3-1*, *lux-1* and *toc1-1* (Fig. 3.13). In general, expression profiling revealed that these genes were also in part affected by mutations in *ELF3, LUX* and *TOC1* (Fig. 3.13). *WRKY11* and *PUP1* were both arrhythmic in *elf3-1, lux-1* and *toc1-1* while the expression profiles of *bHLH128*, *ZAT10, AtCP1, ZAT12* and *ERF2* were altered in all three mutant lines (Fig. 3.13), supporting the role of circadian regulation of these genes. Taken together, as part of an anticipatory response to oxidative stress, the expression patterns of ROS genes were found to be clock-controlled under regular growth conditions (Fig. 3.8; Fig. 3.9) and that plants were most responsive to oxidative stress during the day when *CCA1* is expressed but not at night (Fig. 3.11*A*).

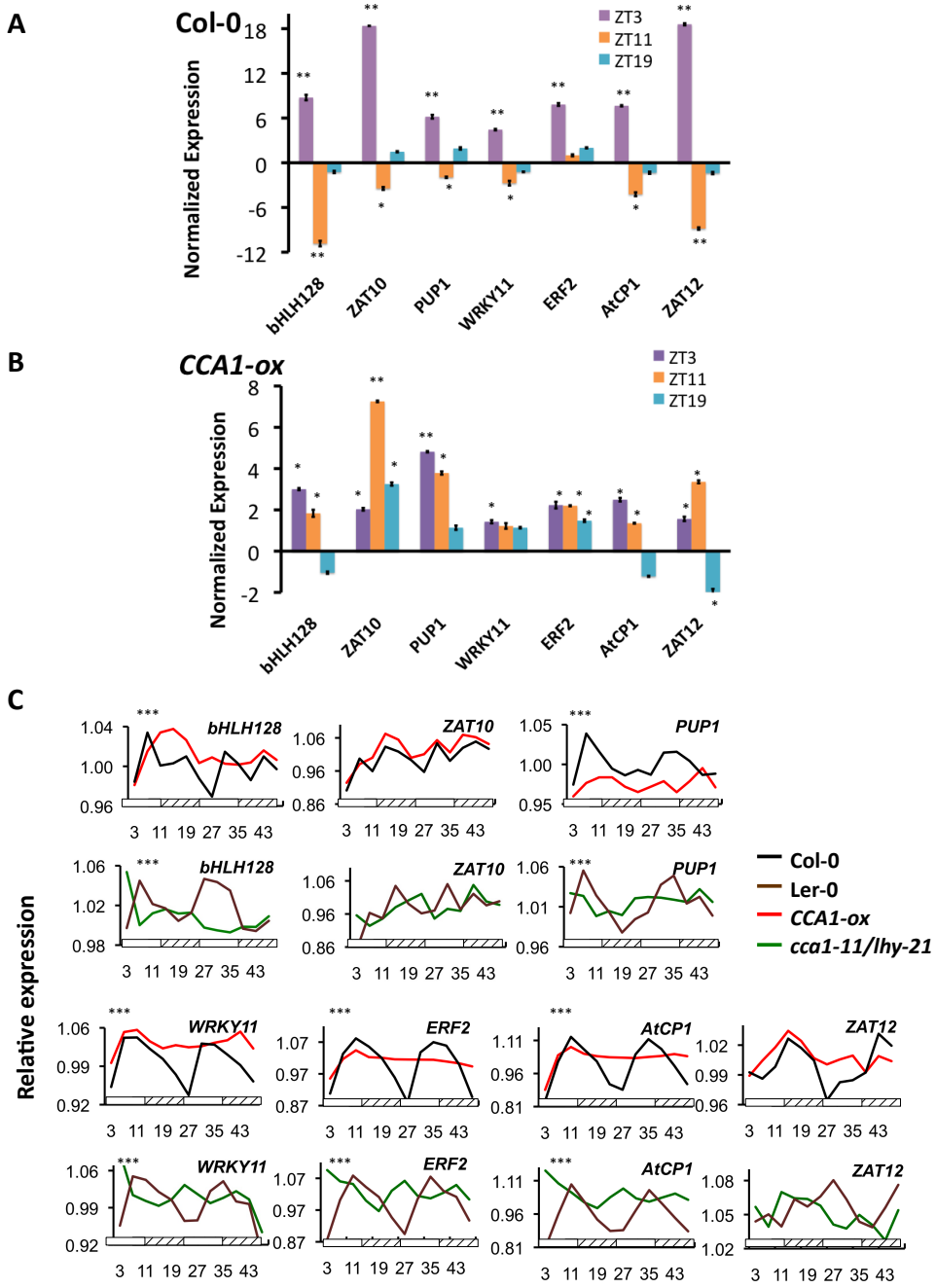


Figure 3.11: Response to ROS is regulated by diurnal cycles and is dependent on the time of *CCA1* expression

Solid bars represent fold-changes of ROS TFs and regulatory genes expression in (A) Col-0 and (B) *CCA1-ox* plants, entrained under 12 h light 12 h dark photocycles and treated with 2µM MV at ZT3 (purple bars), ZT11 (orange bars) and ZT19 (blue bars) on day 16 of growth. Solid bars represent mean values of 3 biological replicates (N=15), with error bars indicating s.e.m. In Col-0 samples (A), student t-test for the effects of time on normalized expression values were significant (**P<0.0001; *P<0.01). In *CCA1-ox* samples (B), student t-test for effects of genotypes (Col-0 versus *CCA1-ox*) on normalized expression values were significant (**P<0.0001; *P<0.01). (C) Expression profiles of seven ROS TFs and regulatory genes were obtained from plants entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. *CCA1-ox* is of the Col-0 background and *cca1-1/lhy-11* is of the Ler-0 background. Gene expression data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression was significant in WT samples (**P<0.001). White bars: day; hatched bars: subjective night.

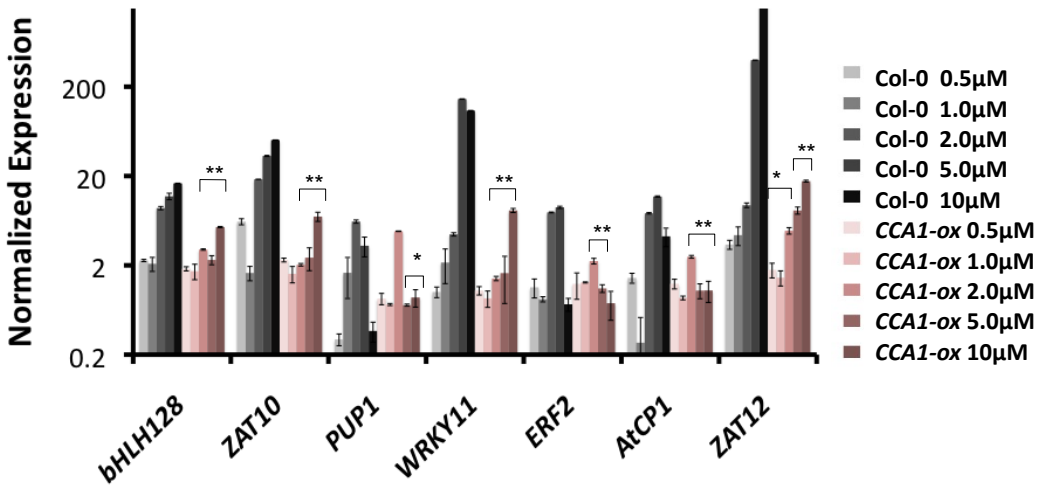


Figure 3.12: Response to ROS is attenuated in *CCA1-ox*
Solid bars represent normalized expression values of ROS signaling genes in Col-0 and *CCA1-ox* plants treated with 0.5, 1, 2, 5 and 10µM MV. Plants were grown under 12 h light 12 h dark photocycles and treatments were performed at ZT3. Data represent mean values of 3 biological replicates (N=12) and error bars represent s.d. Student t-test for normalized expression values of Col-0 versus *CCA1-ox* were significant (*P<0.05; **P<0.01).

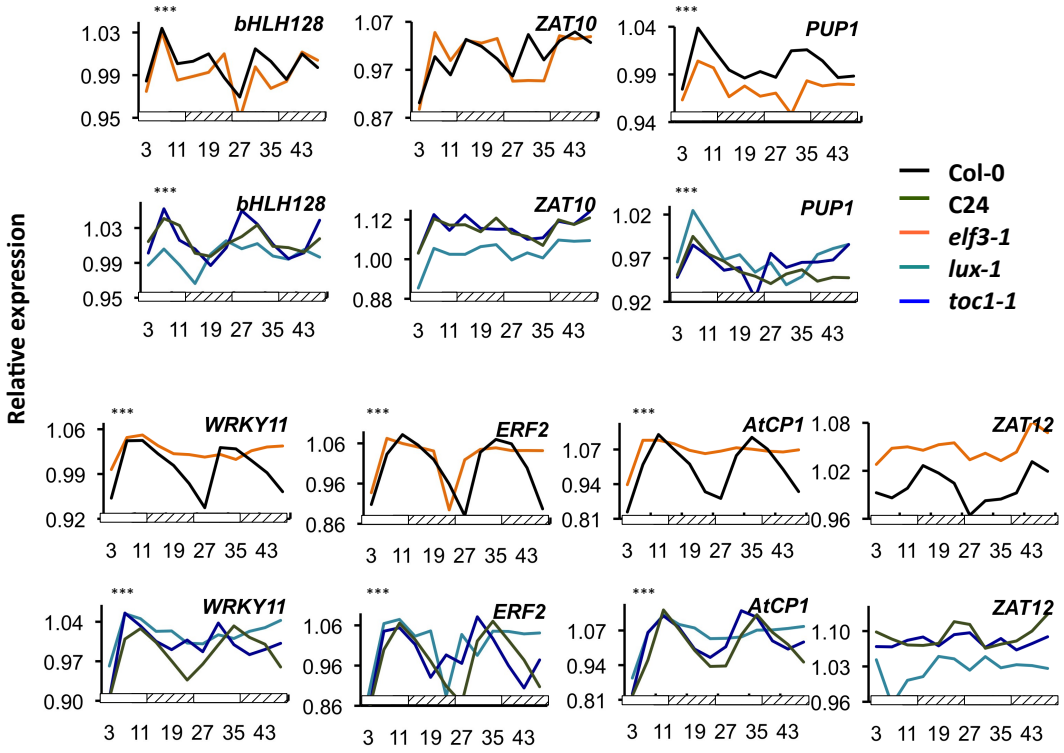


Figure 3.13: ROS signaling is altered at the transcriptional level in *elf3-1, lux-1* and *toc1-1* mutants

Expression profiles of seven ROS TFs and regulatory genes were obtained from plants entrained under 12 h light 12 h dark photocycles and transferred to LL prior to sampling. *Elf3-1* is of the Col-0 background while *lux-1* and *toc1-1* are of the C24 background. Data represent mean values of 3 biological replicates (N=12). One way ANOVA (effect of time) for gene expression was significant in WT samples (***P<0.001). White bars: day; hatched bars: subjective night.

3.7 *WRKY11, MYB59, PAL1* and *ZAT12* are direct targets of CCA1 *in vivo*

Response to ROS was strongest when *CCA1* is expressed in the morning. Furthermore, *CCA1* overexpression had permitted this response to occur throughout the day. Along with the observation of EE and/or CBS in ROS promoters, these suggest that CCA1 is a direct regulator of ROS signaling and could bind to the promoters to impart transcriptional coordination. To determine whether CCA1 associate with ROS promoters *in vivo*, ChIP-qPCR assay was performed using CCA1::CCA1-GFP transgenic lines in which *CAT3* was used as a positive control (Table 3.6) since enrichments for EE-containing promoter fragments of *CAT3* have previously been observed (Michael & McClung, 2002). ChIP was performed at the peak of *CCA1* expression (ZT1). Of the ten ROS genes tested, enrichments were detected for EE/CBS-containing promoter fragments of *COR27* (Mikkelsen & Thomashow, 2009), *JMJD5* (Lu et al., 2011), *ZAT12, MYB59, WRKY11* and *PAL1* but not in negative control sequences (Fig. 3.14). However, enrichment was not observed in *PUP1*, *PEROXIDASE* and *METHYL ESTERASE 18* (*MES18*) although all three genes contained putative CBS (Fig. 3.15). As shown previously, response to oxidative stress is regulated by *CCA1* where the induction of ROS genes was highest when *CCA1* is expressed (Fig. 3.11*A*). These results and the effect of altered *CCA1* expression on ROS signaling genes (Fig. 3.8) demonstrated that CCA1 can directly bind to the promoters of ROS genes to coordinate transcription and therefore affect ROS signaling and homeostasis accordingly.

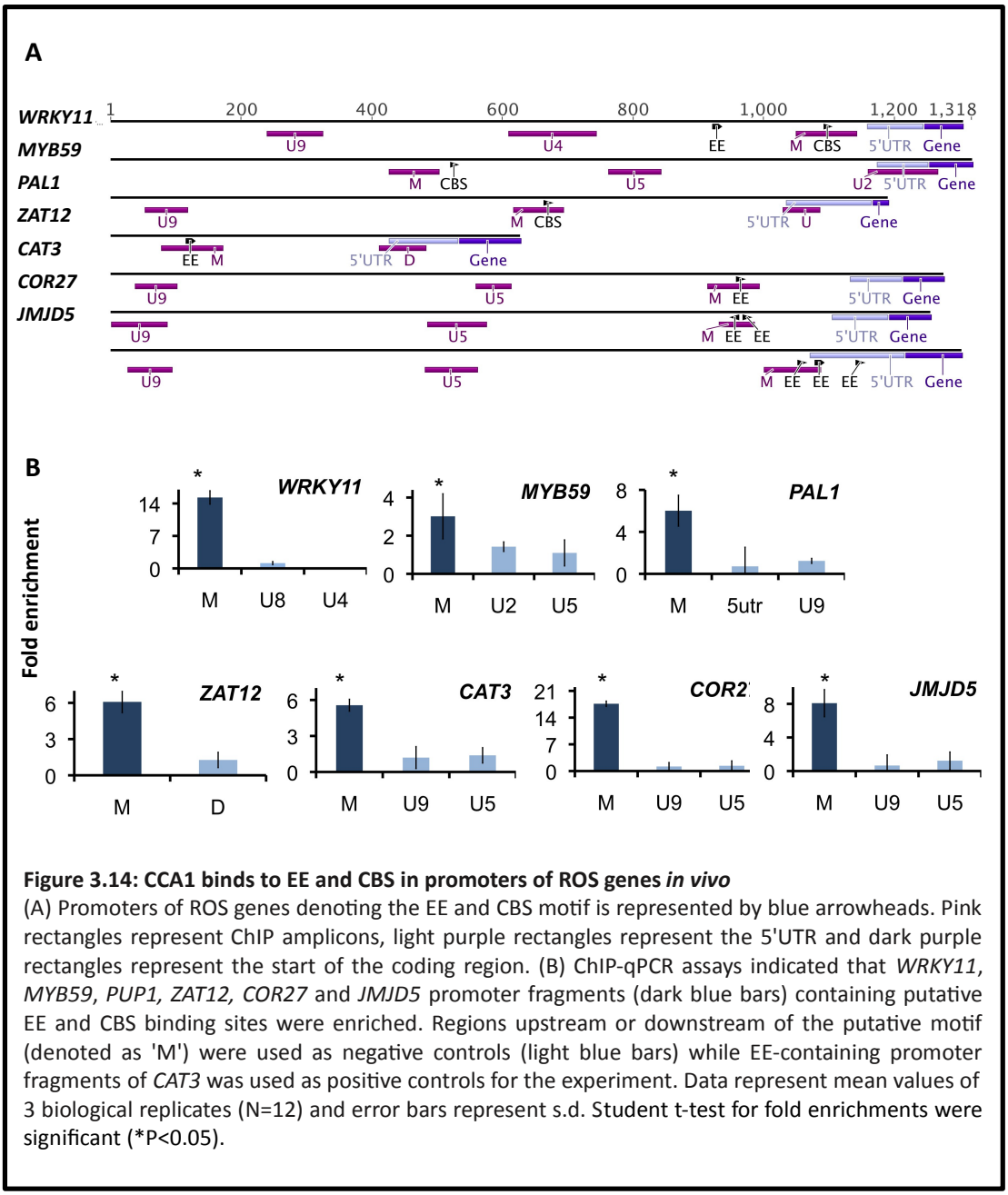


Figure 3.14: CCA1 binds to EE and CBS in promoters of ROS genes *in vivo*
(A) Promoters of ROS genes denoting the EE and CBS motif is represented by blue arrowheads. Pink rectangles represent ChIP amplicons, light purple rectangles represent the 5'UTR and dark purple rectangles represent the start of the coding region. (B) ChIP-qPCR assays indicated that *WRKY11*, *MYB59*, *PUP1, ZAT12, COR27* and *JMJD5* promoter fragments (dark blue bars) containing putative EE and CBS binding sites were enriched. Regions upstream or downstream of the putative motif (denoted as 'M') were used as negative controls (light blue bars) while EE-containing promoter fragments of *CAT3* was used as positive controls for the experiment. Data represent mean values of 3 biological replicates (N=12) and error bars represent s.d. Student t-test for fold enrichments were significant (*P<0.05).

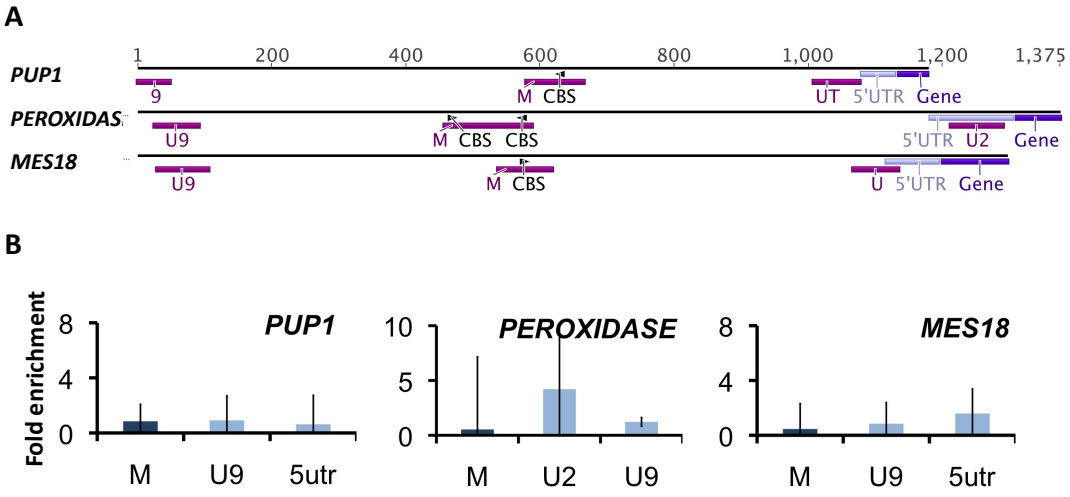


Figure 3.15: Not all ROS genes with the CBS are bound by CCA1 *in vivo*
(A) Promoters of ROS genes denoting the CBS motif is represented by blue arrowheads. Pink rectangles represent ChIP amplicons, light purple rectangles represent the 5'UTR and dark purple rectangles represent the start of the coding region. (B) ChIP-qPCR assays indicated that *PUP1*, *PEROXIDASE* and *MES18* promoter fragments (dark blue bars) containing putative CBS binding sites were not enriched. Regions upstream or downstream of the putative motif (denoted as 'M') were used as negative controls (light blue bars) while EE-containing promoter fragments of *CAT3* was used as positive controls for the experiment. Data represent mean values of 3 biological replicates (N=12) and error bars represent s.d.

3.8 ROS signals feed back to affect circadian behavior

Results have shown that the circadian clock not only controls ROS homeostasis, but also the transcription of ROS genes. *CCA1* is proposed to be a master regulator of this response because *CCA1* mutants have altered sensitivities to ROS agents, altered ROS homeostasis and ROS gene expressions were arrhythmic in these plants. Furthermore, the response to exogenous applications of ROS is dependent on the time of *CCA1* expression. Collectively, these results suggest the ROS homeostasis is an output of the clock. Nevertheless, it might also be possible for ROS to feed back to influence circadian behavior. Since interlocking feedback loops have been shown to be common in other metabolic processes (Dodd et al., 2007; Gutiérrez et al., 2008; O'Neill et al., 2011), it is hypothesized that ROS signals may feed back to affect processes controlled by the oscillator, possibly as a mechanism to reset ROS signaling cascade through crosstalk with other pathways and thereby allowing plants to continuously monitor the increase or decrease in ROS levels under various physiological conditions.

To investigate the effects of ROS on the oscillator, LUC reporter lines were used for the real-time monitoring of clock-regulated gene transcription. For markers of the clock, LUC reporter lines harboring promoter fragments of *CCA1* and *TOC1* fused to the luciferase gene were used. For markers of circadian output, LUC reporter lines of *FKF1*, *CAB2* and *CAT3* were used. These output genes were selected on the basis of both being likely targets of *CCA1* and having the EE present in their promoters (Millar et al., 1992; Nelson et al., 2000; Michael & McClung, 2002). Chronic ROS treatments were achieved by administering different doses of MV, H_2O_2, 3-AT, an inhibitor of catalases (Gechev et al., 2002) and SHAM, an inhibitor of peroxidases (Brouwer et al., 1986). Since ROS signaling is mediated by both ROS production and scavenging, inhibition of ROS production can be expected to have similar effects. DPI and KI were used to inhibit NOX activity and to scavenge H_2O_2 respectively. Luminescence imaging revealed that several chronic ROS treatments (MV, H_2O_2, KI and DPI) resulted in the dose-dependent lengthening of *TOC1* free-running periods and phase delays (peak was shifted to the right) of *TOC1* transcription (Fig. 3.16*A*). Similar responses were observed in CCA1::LUC lines that were treated with MV and H_2O_2 (Fig. 3.16*B*). Notably, the amplitudes of *TOC1* rhythms were increased by ROS induction and inhibition (Fig.

3.16*A*). Of all the three output genes tested, ROS treatments only affected the *FKF1* period and phase (Fig. 3.17*A*) in a similar manner as observed in *CCA1* and *TOC1*. The period and phase of *CAB2* (Fig. 3.17*B*) and *CAT3* (Fig. 3.17*C*) were not significantly affected by either ROS induction or inhibition, if anything only slight phase delays were observed in CAB2::LUC lines treated with 3-AT and DPI (Fig. 3.17*B*). Taken together, the results suggested that the clock is involved in coordinated transcriptional regulation of ROS genes with *CCA1* being a likely master regulator of ROS signaling. In addition, real-time monitoring of clock-driven gene transcription revealed a feed back mechanism at which ROS can affect components of the core oscillator and also an output pathway.

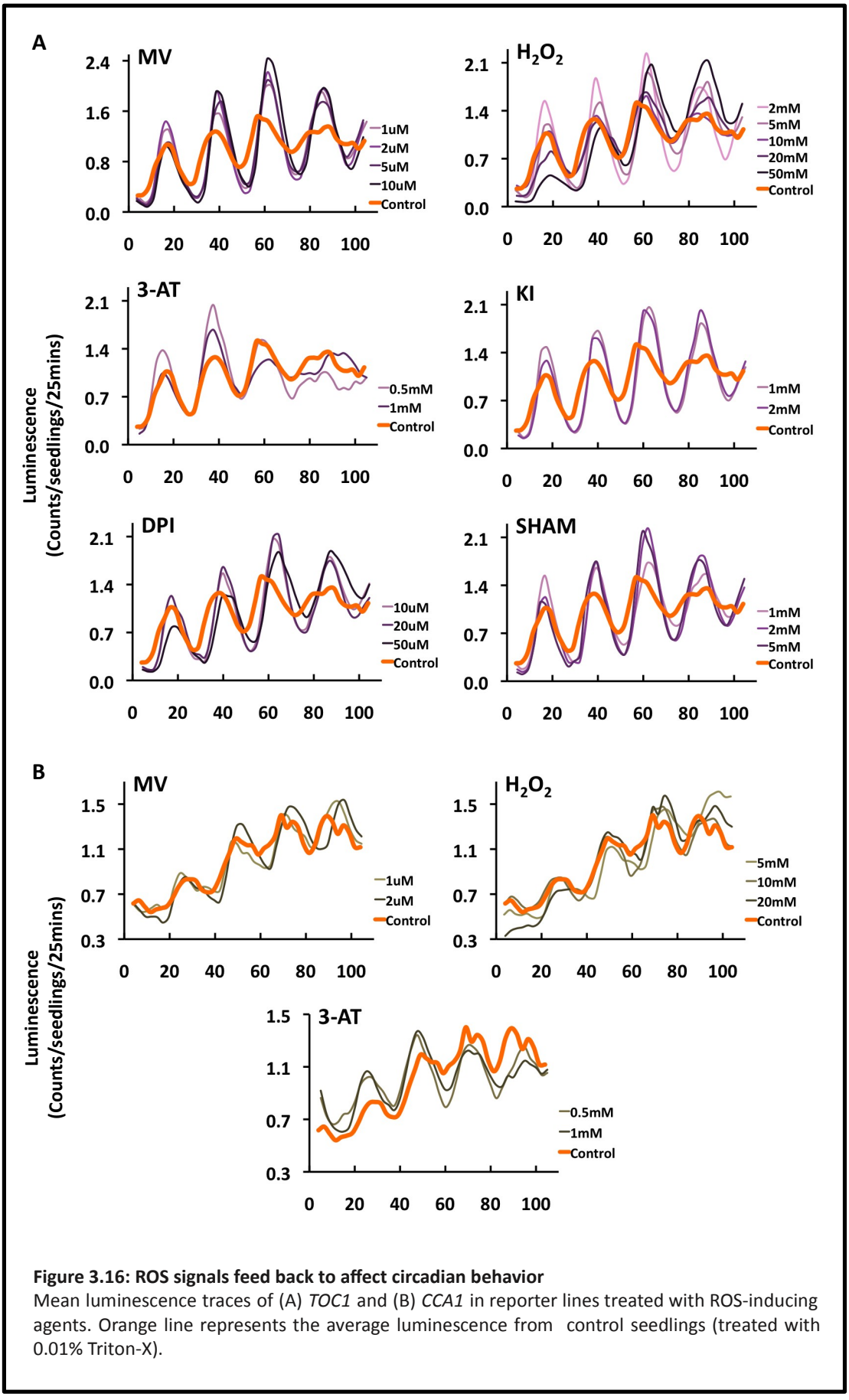
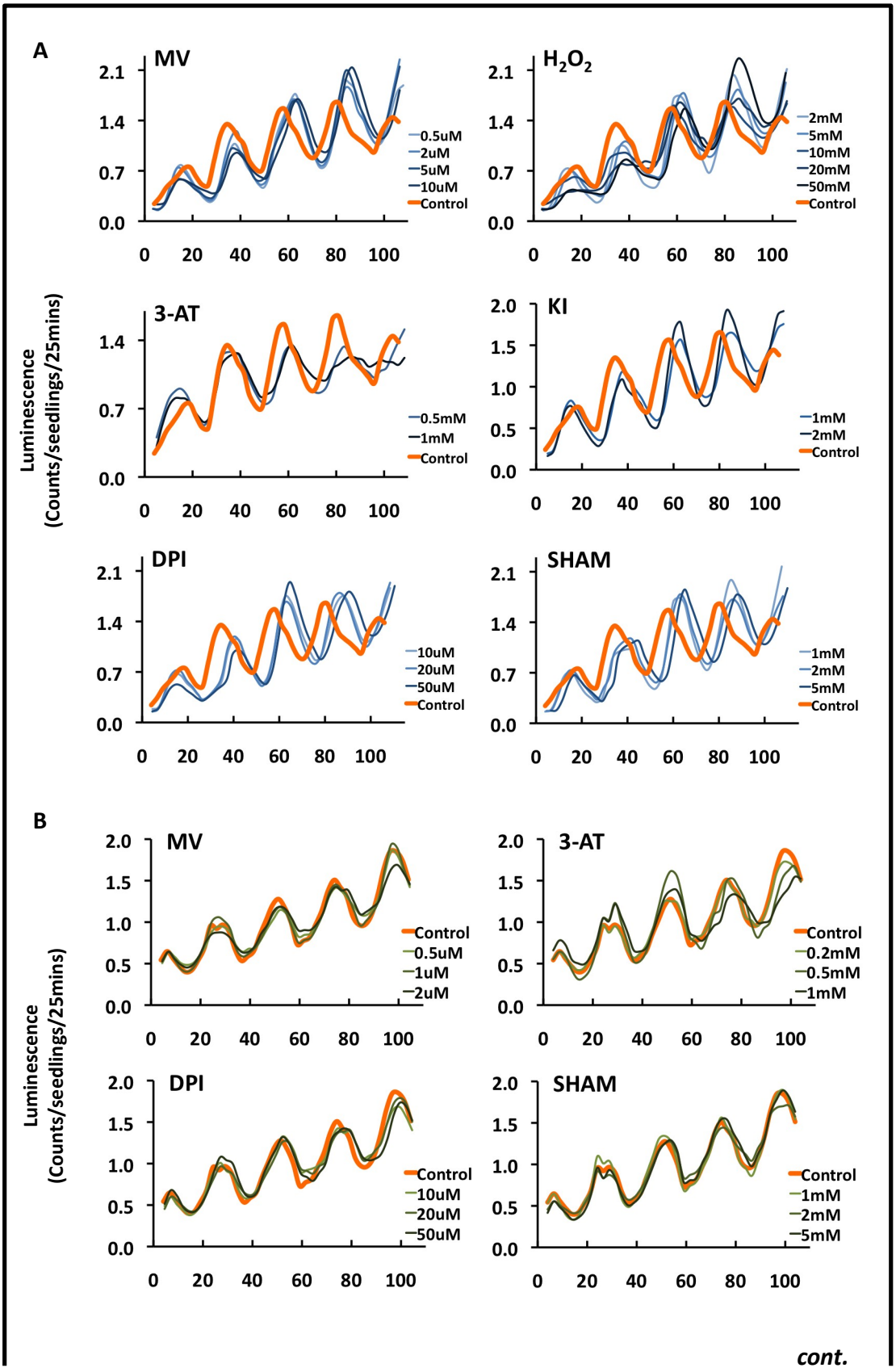


Figure 3.16: ROS signals feed back to affect circadian behavior
Mean luminescence traces of (A) *TOC1* and (B) *CCA1* in reporter lines treated with ROS-inducing agents. Orange line represents the average luminescence from control seedlings (treated with 0.01% Triton-X).


A
MV
0.5uM
2uM
5uM
10uM
Control
H_2O_2
2mM
5mM
10mM
20mM
50mM
Control
3-AT
0.5mM
1mM
Control
KI
1mM
2mM
Control
DPI
10uM
20uM
50uM
Control
SHAM
1mM
2mM
5mM
Control
Luminescence (Counts/seedlings/25mins)
B
MV
Control
0.5uM
1uM
2uM
3-AT
Control
0.2mM
0.5mM
1mM
DPI
Control
10uM
20uM
50uM
SHAM
Control
1mM
2mM
5mM
Luminescence (Counts/seedlings/25mins)
cont.

cont.

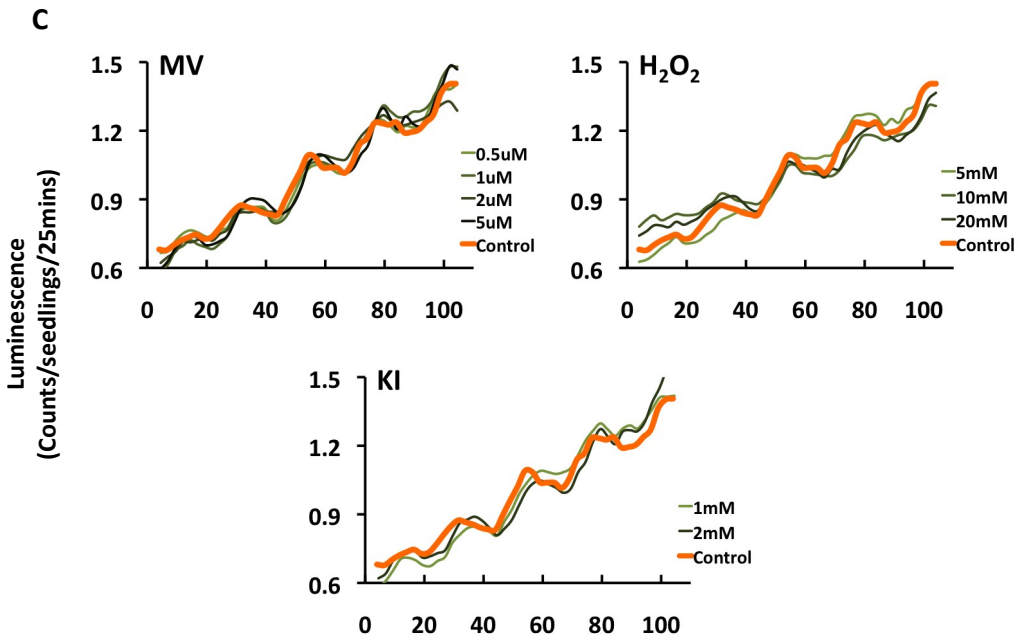


Figure 3.17: ROS signals feed back into clock-regulated genes
Mean luminescence traces of (A) *FKF1,* (B) *CAB2* and (C) *CAT3* in reporter lines treated with ROS-inducing agents. Orange line represents the average luminescence from control seedlings (treated with 0.01% Triton-X).

4. Discussion

4.1 Coupling of the circadian clock and metabolism

The exact molecular mechanism that links ROS signaling to the circadian clock has yet to be studied. In this work, an interaction between the circadian clock and ROS signaling is described. Seven key objectives were outlined and confirmed. ROS production and scavenging were found to exhibit a diurnal and circadian pattern. Not only that, at the transcriptional level, ROS genes display phase specific expression. ROS genes exhibit time-of-day specific phases in the absence of exogenous stresses (Fig. 3.1*D*; 3.2*D*; 3.8*A*). This may be part of a clock-mediated anticipatory response to maximize protection to the plant while minimizing the adverse effects of oxidative stress by effectively allocating energy for these processes. The crosstalk between the clock and ROS transcriptional networks suggests the existence of an anticipatory response that is in place to act a coping mechanism for plants to deal with toxic-by-products of oxygen. To confirm the involvement of the clock in ROS signaling, differential sensitivities to ROS applications and altered ROS homeostasis were observed in clock mutants. Furthermore, global clock effects on ROS homeostasis and signaling are also observed because loss-of-function of clock genes could tilt the balance in ROS production and scavenging (Fig. 3.5; Fig. 3.6). Since *CCA1* overexpression and null mutation produced the most dramatic phenotype, this suggests that *CCA1* might be a master regulator of the ROS response. The involvement of *CCA1* is confirmed by transcriptional analyses, where plants are found to be most responsive to ROS during the morning when *CCA1* is expressed. Therefore, at the transcriptional level, it appears that *CCA1* is a likely master regulator of ROS signaling (Fig. 3.8) and is required for the time-of-day response to oxidative stress (Fig. 3.11*B*). Nevertheless, the involvement of other clock regulators such as *ELF3*, *LUX* and *TOC1* could not be ruled out completely (Fig. 3.9; Fig. 3.13). Also, physical bindings of CCA1 to promoters of ROS genes further corroborate the involvement of CCA1 in ROS signaling.

It has been shown previously that *PRR5*, *PRR7* and *PRR9* were crucial for the regulation of mitochondrial homeostasis, respiration and photosynthesis (Fukushima et al., 2009). Notably, this study reveals that *prr5-1, prr9-1, prr7-3* and *prr5-1/prr9-1* mutants are hypersensitive to ROS treatments (Fig. 3.4). Interestingly, although the triple mutant

prr9/prr7/prr5 had defects in primary metabolism involving the TCA cycle and other biosynthetic pathways involving carotenoid, ABA, and chlorophyll, the authors found that this mutant was resistant to drought and had higher freezing tolerance (Fukushima et al., 2009; Nakamichi et al., 2009). Taken together, it is tempting to speculate that the Myb TFs *CCA1/LHY* and the *PRRs* foster a molecular link between primary metabolism and the circadian clock, possibly acting via distinct mechanisms such as through the coordination of ROS homeostasis (this study) or through the coordination of mitochondria respiration and photosynthesis (Fukushima et al., 2009). Another well-studied clock regulator *GI* is reported to be involved in metabolic coordination. Although the role of *GI* in circadian clock function and regulation of flowering time has been well researched (Fowler et al., 1999; Park et al., 1999), its role in oxidative stress (Kurepa et al., 1998a) and starch accumulation (Eimert et al., 1995; Messerli et al., 2007) has receive little attention to date. The role of *GI* in starch metabolism is evident in the *gi* mutant where it exhibited a starch-excess phenotype and had higher proportions of simple carbohydrates and carbon fixation rates (Eimert et al., 1995; Messerli et al., 2007). Interestingly, ROS scavenging genes, *APX1* and Cd/Zn and Fe superoxide dismutases (*CSD2* and *FeSOD*), were constitutively expressed, which could in part explain the MV-induced oxidative stress resistant phenotype of *gi* (Kurepa et al., 1998a; Kurepa et al., 1998b; Cao et al., 2006b).

Many metabolic pathways are incorporated as clock feedback loops possibly to increase the overall clock's robustness by providing multiple input and output points to the oscillator (Panda et al., 2002; Dodd et al., 2007; Yin et al., 2007). A study by Dodd et al. (2007) revealed that small signaling molecules were involved in the clock's feedback mechanism. The authors showed that cyclic adenosine diphosphate ribose (cADPR), a cytosolic ligand synthesized from nicotinamide adenine dinucleotide (NAD) that promotes Ca^{2+} release into cytosolic space (Allen et al., 1995), is a morning-phased clock-controlled output. Moreover, cADPR could affect the oscillator where decrease in cADPR concentrations by nicotinamide resulted in period lengthening of clock genes (Dodd et al., 2007). Thus, daily oscillations in Ca^{2+} may encode circadian clock signaling information to modulate other physiological responses in the Ca^{2+} signaling network (Love et al., 2004; Imaizumi et al., 2007).

The effects of the clock on metabolism are also well studied in animals. The TTFL of mammalian clocks involves a heterodimer of TFs *circadian locomotor output cycles kaput* (*CLOCK*) and *brain and muscle ARNT-like 1* (*BMAL1*) that by binding to E-box elements, repress their own transcription and drive the transcription of period (*PER1*, *PER2* and *PER3*), *CRY1* and *CRY2* (Green et al., 2008). Interestingly, metabolic cues such as rhythmically secreted hormones were found to convey temporal information to circadian clocks in mammalian peripheral organs (Damiola et a., 2000) and mice with mutated versions of the *CLOCK* gene developed metabolic syndromes (Turek et al., 2005). Furthermore, *CRY1* and *CRY2* could interact with the glucocorticoid receptor to alter transcriptional response to glucocorticoids; the regulators of glucose homeostasis in liver and pancreas (Marcheva et al., 2010; Sadacca et al., 2010; Lamia et al., 2011). The observation that metabolic cues set the timing of peripheral clocks suggests an adaptation process where organisms adjust to changes in metabolic needs over the course of each day (Asher et al., 2008; Lamia et al., 2009; Asher et al., 2010). Apart from the archetypal TTFL-based genetic oscillator, biochemical flux in NADP redox states have been shown to drive metabolic cycles (Rutter et al., 2002; Nakahata et al., 2008; Bass & Takahashi, 2011). An example of metabolic rhythms involves a class of protein found in mature red blood cells known as the peroxiredoxins. Peroxiredoxin (PRX) is involved in ROS inactivation where the cysteine residue in its active site undergoes oxidation when ROS levels are high. This results in circadian periodicity of PRX oligomerization pattern, i.e., the monomer-dimer transitions of the enzyme (O'Neill & Reddy, 2011; O'Neill et al., 2011). Indeed, the mammalian clock can also be driven by PRX rhythms and the rhythms *per se* mirror the changes in cellular redox states (O'Neill et al., 2011; Bass & Takahashi, 2011). Collectively, mechanisms for integrating redox or cellular signals to plant and animal clocks, represent yet another layer of feedback regulation within circadian systems to fine-tune metabolic homeostasis (Imaizumi et al., 2007).

4.2 The circadian clock communicates temporal information to regulate ROS network transcriptomes

This study reveals the interconnection between ROS transcriptomic and metabolic networks and the circadian clock. Here, the results suggest that *CCA1* and *LHY* are both

necessary for maintaining ROS homeostasis under normal physiological growth conditions (Fig. 3.5) and under oxidative stress conditions (Fig. 3.11*A*). The physiological relevance of this is apparent in the increased hypersensitivity of *cca1-1/lhy-11* to oxidative stress (Fig. 3.3) that could be a result of perturbed ROS homeostasis in this mutant (Fig. 3.5). Since ROS levels may reflect the metabolic state of the plant and transcriptional profiling of ROS genes reveals temporal gene clusters of coordinated expression (Fig. 3.8), this suggests that metabolic needs may be partitioned to different times of the diel cycle. Notably, transcriptional coordination of ROS signaling is driven in part by *CCA1* rhythms *per se* because ROS genes became arrhythmic upon mis-expression of *CCA1* (Fig. 3.8).

The hypothesis that CCA1 targets ROS genes *in vivo* to impart transcriptional coordination is validated with ChIP-qPCR assays using transgenic seedlings expressing GFP-tagged CCA1, CCA1::CCA1-GFP (Fig. 3.14). Although cognate cis-regulatory elements (EE or CBS) of CCA1 are present in all the ten genes tested, only seven genes show enrichments in EE/CBS-containing promoter fragments bound by CCA1 (Fig. 3.14; Fig. 3.15). One explanation for this is that since the GFP-tagged lines were generated using the native promoter of *CCA1* instead of a 35S-driven promoter, it is may be possible that other clock-regulated EE-binding proteins are recruited via distinct molecular mechanisms to regulate this subset of ROS genes. Indeed, three *CCA1/LHY* related clock-regulated Myb transcription factors that contained the Myb-like/SANT DNA binding domain; *RVE1, RVE4* and *RVE8*, have been shown to bind the EE (Rawat et al., 2009; Rawat et al., 2011). *RVE1* links the circadian clock to auxin signaling through the regulation of daily rhythms of auxin production (Rawat et al., 2009). *RVE8* acts as a transcriptional activator of dusk-phased-EE-containing clock genes, sets the pace of the clock and participates in temperature compensation mechanisms (Gong et al., 2008; Rawat et al., 2011). Questions remained to be answered now is whether these RVEs form physical associations with ROS genes and if such associations exist, what roles do these RVEs play in ROS signaling.

4.3 The circadian clock mediates the time-of-day sensitivity to oxidative stress signals

A metaphorical 'gate' governs the sensitivity of the clock to resetting signals presented at different times of the day (Fig. 1.3). For instance, in the middle of the day, light does not convey time information to the clock; hence the 'gate' is closed at mid-day to prevent clock resetting, which would otherwise be detrimental as the ongoing process of resetting prevents the clock from progressing (McWatters & Devlin, 2011). In addition, this process can partly explain the phenomenon where the clock either intensifies or attenuates physiological activities at different times of the day to generate permissive or non-permissive states for responses. In the event of oxidative stress, *CCA1* positively regulates the expression of ROS signaling genes only when it is expressed at dawn, but not at other times (Fig. 3.11*A*). When plants experience oxidative stress in the morning (permissive state), *CCA1* is present to coordinate ROS signaling genes expression (Fig. 3.11*A*). However, if oxidative stress is administered at night when *CCA1* levels are low (non-permissive state), the synergetic effect between the circadian oscillator and ROS signaling could not be observed (Fig. 3.11*A*). Furthermore, the response observed in the *CCA1-ox* background is likely to reflect the phase state of an arrested clock. In *CCA1-ox*, the clock may be locked in the day state when the system is in a permissive state for ROS responses (Fig. 3.11*B*) and could therefore account for the low basal ROS levels in this mutant (Fig. 3.5*A* and *C*). Consequently, lower basal ROS levels in *CCA1-ox,* presumably because of its constitutive day state, could also result in attenuated induction of ROS genes when plants are exposed to oxidative stress even in a dose-dependent manner (Fig. 3.11*B*; Fig. 3.12).

Interestingly, negative regulation of ROS genes is observed in the evening where gene expressions are down-regulated upon oxidative stress (Fig. 3.11*A*). This suggests that other factors may be involved in the regulation of these genes. Based on the results, it is observed that *CCA1* and *LHY* are not the only clock factors that control ROS signaling. The role of other clock genes in this regulation could not be ruled out completely because in most cases, phase-specific expressions of ROS genes are abolished in *elf3-1* and *lux-1* (Fig. 3.9). Nevertheless, this could also be interpreted as the outcome of an

arrested clock on ROS signaling in both genotypes since the clock is rendered arrhythmic by mutations in *ELF3* and *LUX* (Helfer et al., 2011; Dixon et al., 2011).

Clock-oriented experiments need to be performed in free-running conditions (LL or DD). Consequently, it is difficult to tease apart the interactions between the clock, ROS signaling, light signaling and LL-induced oxidative stress. Therefore, under LL conditions, circadian control alone cannot account for ROS homeostasis. Both ROS production and scavenging were up-regulated in LL albeit still exhibiting time-of-day specific phases (Fig. 3.2). Concomitant with the phase-specific expression of ROS-responsive genes in LL (Fig. 3.2*D*), this suggests that the clock still predominantly controls ROS homeostasis. The high ROS background levels in LL (Fig. 3.2*A* and *C*) could be due to the combined effect of constant energy supply for photosynthesis, constant photoreceptor signaling and the over-reduction of electron acceptors in which all may be the cause of LL-induced photo-oxidative damage (Bae & Choi, 2008; Velez-Ramirez et al., 2011). Since ROS homeostasis is an output of the clock, protection against photo-oxidative stress fluctuates (Fig. 3.2*B*) and LL-induced damage might occur at times when plants are not properly protected.

4.4 Biological importance of non-photic influences on the circadian oscillator

The circadian clock receives input from environmental stimuli and multiple metabolic pathways and uses such information to fine tune clock function. Rather than conferring large changes such as those observed in light responses, non-photic cues produce subtle changes not to reset the clock, but to improve clock-mediated coordination in metabolism and physiology based on the organisms' metabolic status (Stephan et al., 2002; Gibon et al., 2006; Feillet et al., 2006). During the day, photosynthesis fuels autotrophic plant growth. At night, plants derive their energy from starch reserves for nocturnal metabolism and growth. Indeed, starch degradation and mobilization is precisely timed to the anticipated night length and repression of growth-promoting genes ensued when the night was unexpectedly extended (Gibon et al., 2004; Smith & Stitt, 2007). Indeed, WT plants grown under periods longer than 24 h were inferior because of the mismatch between nighttime carbohydrate utilization and their internal oscillator (Graf et al., 2010). It was also discovered that *CCA1* and

LHY regulate the rate to starch degradation to ensure that enough carbohydrate remained until the next anticipated dawn (Graf et al., 2010). Another study addressing the effects of exogenous sucrose on the circadian clock revealed that the circadian system is particularly sensitive to sucrose in the dark when endogenous sucrose levels were low (Dalchau et al., 2011). Interestingly, continuous supply of sucrose in dark-adapted plants could sustain the oscillator and topical sucrose application could reestablish rhythms of an arrested clock, hence suggesting that the oscillator integrates long-term metabolic status and gates short-term metabolic signals to enhance growth (Dodd et al., 2005; Ni et al., 2009; Dalchau et al., 2011). In a comprehensive study by Gutiérrez et al. (2008), the authors discovered that changes in nitrogen metabolite status might feed back to affect the clock's oscillation. Indeed, it was proposed that *CCA1* affects the expression of target genes in the nitrogen assimilation network; *ASPARAGINE SYNTHASE 1* (*ASN1*), *GLUTAMINE SUNTHASE* (*GLN 1.3*) and *GLUTAMINE DEHYDROGENASE* (*GDH1*) and *CCA1* itself is regulated by glutamate-derived metabolites (Gutiérrez et al., 2008).

It is also noteworthy that the results suggest the existence of a ROS feedback loop within the circadian oscillator. The clock controls ROS homeostasis (Fig. 3.2) and ROS in turn affects the transcription of core clock genes (Fig. 3.16). The existence of a ROS feedback loop that affects the behavior of the oscillator may increase the adaptive value of circadian control on ROS homeostasis during periods of unanticipated environmental fluctuations. Indeed, both continuous supply and suppression of ROS could alter the free-running period and phase of core oscillator genes, *CCA1* and *TOC1* in a concentration-dependent manner (Fig. 3.16). Since the clock also controls a plethora of output pathways, it is possible for ROS signals to exert indirect effects on these output pathway(s) when the oscillator is perturbed. Notably, transcription of *FKF1* is altered by ROS treatments (Fig. 3.17*A*). FKF1 is found to be involved in the photoperiodic control of flowering (Imaizumi & Kay, 2006) and the proteasome-mediated degradation of TOC1 through the interaction with GI. Indeed, the role of *GI* in oxidative stress tolerance has been implicated (Kurepa et al., 1998a). In the case of *CAB2* and *CAT3*, ROS treatments did not have a pronounce effect on the transcription of both genes (Fig. 3.17*B* and *C*). It is likely that posttranscriptional regulation may

buffer against ROS-induced changes in gene expression and this could explain why the *CAB2* and *CAT3* rhythms were unaltered by ROS. The data presented in this work reveals an interface between ROS homeostasis and the circadian clock and a plausible molecular mechanism of how this occurs through *CCA1* is described. Moreover, ROS input to affect the oscillator and its output suggests that it may be advantageous to fine-tune oscillator functions based on changes in the environment.

4.5 Future research

This work represents an extensive study on the role of the circadian clock in regulating ROS homeostasis and ROS signaling. However, many areas remained underexplored. Although many key players of ROS signaling have emerged, this only adds to the complexity of the ROS pathway. One can speculate the existence of crosstalk between other clock output pathways and the ROS pathway since *FKF1* transcription is altered by ROS treatments. The nature of such crosstalk is far more complicated than direct transcriptional regulation. Since *CCA1* is involved in regulating *FKF1*, it may be possible that the alteration in *FKF1* transcription is from ROS affecting *CCA1*. To ascertain the definitive mechanism, it would be useful to determine whether ROS can affect *FKF1* in the absence of *CCA1* and *LHY*, since both *CCA1* and *LHY* have partially redundant functions. Also, it would be interesting to test the effects of ROS on other clock genes such as the evening complex *ELF3*, *ELF4* and *LUX*. This would require the use of more luciferase clock reporter lines. Apart from CCA1, LUX and CHE also functions as TFs and they may form physical associations with promoters of ROS genes. One could perform a promoter search on ROS genes for the LUX binding site and subsequently perform a ChIP experiment to confirm potential interactions. As mentioned previously, a closely related group of Myb TFs is the RVEs. Since RVEs bind the EE, they may be involved in co-regulating ROS homeostasis together with CCA1. Exploring the functions of RVEs in ROS signaling could yield potentially interesting results. Since I have identified several ROS TFs, e.g., *WRKY11*, *MYB59* and *ZAT12*, future work can be focused on isolating interacting clock protein partners from a Yeast Two Hybrid screen. Once a potential ROS-Clock protein interaction is discovered, it would be interesting to investigate how such protein complex works in concert to control ROS signaling.

4.6 Concluding remarks

In conclusion, the observed ROS phenotypes in loss-of-function clock mutants due to disruption in ROS homeostasis may occur through desynchronized transcriptional control of ROS signaling genes. Likewise, the function of the oscillator can be fine-tuned according to ROS levels; perturbations in the balance in ROS production and scavenging produce phase and period changes in the transcription of oscillator components that may in turn affect ROS network through other mechanisms (Fig. 4.1). Collectively, the results from this study not only draws attention to the adaptability and the plastic nature of the circadian clock that allows for flexible responses to the ever changing environment, but also contribute to a deeper understanding on the roles of clock components in regulating a plethora of processes that diverge from their role in generating the oscillation. This study opens up additional fields of investigation to 'solve' the clock by understanding how the oscillator controls plants' metabolic outputs to generate biological rhythms, how this occurs through reciprocal signaling between cellular metabolism and the environment and how the reciprocal communication of temporal information to the oscillator and its outputs could be used for the enhancement of growth and fitness.

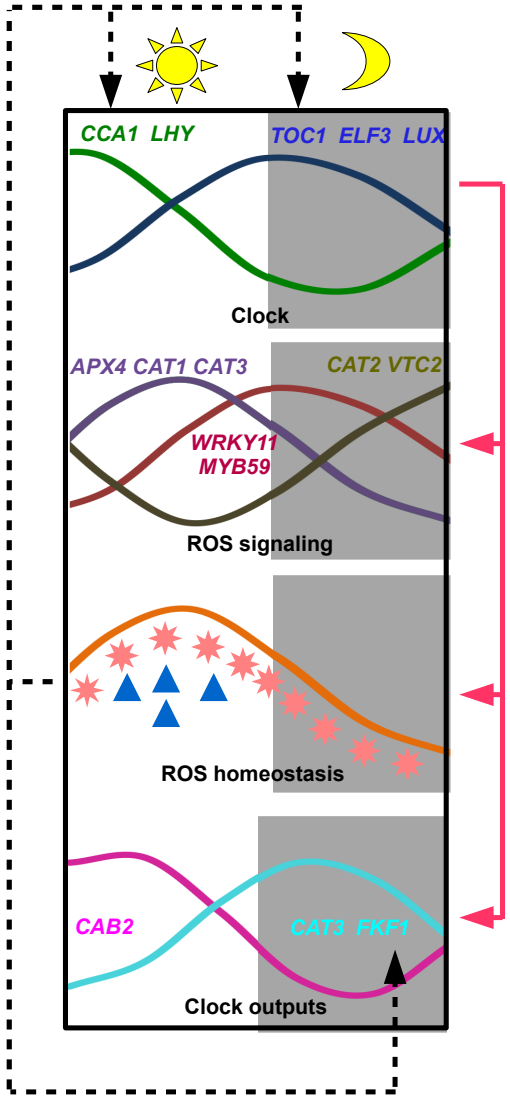


Figure 4.1: Schematic representation of a proposed model depicting the interactions between the *Arabidopsis* circadian clock, ROS networks and clock-controlled outputs
Green line represents the expression profile of morning-phased clock regulators (*CCA1* and *LHY*). Dark blue line represents the expression profile of evening-phased clock regulators (*TOC1, ELF3* and *LUX*). Purple, red and dark brown lines indicate phase-specific expression of ROS signaling gene clusters. Examples of genes from the three clusters are also indicated. Orange line indicates that ROS homeostasis exhibited a time-of-day specific phase. Light orange stars represent metabolic ROS and blue triangles represent environmental ROS. Clock outputs are evening-phased *CAT3* and *FKF1* (light blue line) and morning-phased *CAB2* (pink line). Red arrows indicate the role of the oscillator in the transcriptional coordination of ROS genes, the regulation of ROS production and scavenging and the transcriptional coordination of clock output genes. Dashed black lines indicate that ROS signals feed back to affect oscillator components, *CCA1* and *TOC1* and a clock output, *FKF1*.

5. Appendices

5.1 Genes of the *Arabidopsis* circadian clock their proposed molecular function

Gene name	AGI	Phenotype of mutant plants	Molecular function	References
CCA1	AT2G46830	Short period (redundant with *LHY*). CCA1-ox is arrhythmic	Transcription factor	Wang & Tobin, 1998; Mizoguchi et al., 2002
LHY	AT1G01060	Short period (redundant with *CCA1*)	Transcription factor	Mizoguchi et al., 2002
PRR9	AT2G46790	Long period (redundant with *PRR*7)	Transcription factor	Eriksson et al., 2003; Farré et al., 2005; Zeilinger et al., 2006
PRR7	AT5G02810	Long period (redundant with *PRR9* and *PRR5*)	Transcription factor	Zeilinger et al., 2006; Farré et al., 2005
PRR5	AT5G24470	Short period (redundant with *PRR7*)	Transcription factor	Eriksson et al., 2003
TOC1	AT5G61380	Short period	Transcription factor	Millar et al., 1995; Strayer et al., 2000
ZTL	AT5G57360	Long period (redundant with *FKF1* and *LKP2*)	F-box, blue-light receptor	Somers et al., 2000
LUX	AT3G46640	Arrhythmic in LL	Transcription factor	Hazen et al., 2005
ELF3	AT2G25930	Arrhythmic in LL	Transcription factor	Hicks et al., 1996
ELF4	AT2G40080	Arrhythmic in LL	Transcription factor	Kikis et al., 2005
TIC	AT3G22380	Short period	Unknown	Hall et al., 2003

5.2 Primer sequences of the 167 ROS-responsive genes used in time-course expression studies

AGI	Forward (5´-3´)	Reverse (5´-3´)
At1G01560	TGTTCCTCCTCTTCGTCCTATCG	CTCCGAATTCCATGCAGCACAG
AT1G02450	CGTCGTGAGGAGGAAATCTAACGG	AGAAATCCTCCGGCTGAAACGC
AT1G03940	TGGTGACTCTAGCTTTCATTTGGG	GTCTTCCTCGTTGGCCTTTGTTTC
AT1G05340	GATCACAACCAGTCTGCAGGAG	GGTGGTGGTGATGTACAGGTAGAC
AT1G05575	TGTTCTTCAAGACATGCCAGCAG	AGAGAAGGTCCCATGGCAGAAG
AT1G07160	TGGAGGATCGTTTCTCTGCCATC	ACTCCGAATATTGCCTGTTTGGG
AT1G10040	ACGCGAATACAAGTTTCGACCAC	AGCTCATTGTGACGCCGGATAG
AT1G10585	GGATCGAAGGATGCGCATGAAAC	AGGTGAGGCACTGGTAACTTGC
AT1G13340	TCTACGTCCACTGAGGGAACAAC	ACCATAACCTTCTCCTCTGCCATC
AT1G14040	CGCAGCTCCACTTTACAAGGTTAC	TGCTTCGAATCGCTTGAACCTG
AT1G14200	ATATGCCACGTGTCGTGATCGG	TCCAAACAAATCGCGCAAGAACC
AT1G15520	CCGCTTTCTACGGCATCTGGAATC	TTCCCACCACACAGGCATACTG
AT1G16030	ACGAAGGAGAACGTGCAAGGAC	TGCCCTTGAGCTCGAAAGTTCC
AT1G17170	TGCCAAATTCTGGGCCGACTTC	TCACCTTTGACCGCCCAAATCC
AT1G17180	TCAGCATTGAAGCCGAGTGTCC	TAGCCACACTCTCTCTCTCCACAC
AT1G18570	CCTTCACGGCAACAAATGGTCTG	TACCGGAGGTTATGCCCTTGTG
AT1G19020	ATGAACGGAACATCGTGGGCTGAC	TGTTTGACGTTGCTCCACTTCCG
AT1G21100	TCCATGGTCAAGTGTGGTAAGGC	AGGAAGACGCTGTCGAAATTGACG
AT1G21120	TCCCTTGCCTCTCAAGTCATTGTC	TCTAGCACCACGTCTTTCAGTTG
AT1G21310	ACCTCCTCCGGTTTACCATTCC	GGAGGAGGAGGAGATTTGTACACG
AT1G22400	GATCCGGCCAGATTTAGTAGCG	GCATACTGCGGTCTTTAGTCTCC
AT1G26380	ACCCAAACGCCACAGAGAGTAAC	GTTACTTGACACGTACGGTTCTGC
AT1G26420	AGCAGAACGTGGCGATAGAACC	ATCAGCTGGACCCAAGAACTGAG
AT1G27730	TCACAAGGCAAGCCACCGTAAG	TTGTCGCCGACGAGGTTGAATG
AT1G28190	CAACAGCAACCTCAGCCGTTTC	ACACGGATCCTTAAACGCTCTCC
AT1G28480	ACGGAGAGGATGTTGCATGTGTC	AATCTCAAGGACCGCCGGATTC
AT1G32170	ACCGGTTCTGGTTTCATATCATCG	ACGACGCCGGCAGTATAATCAG
AT1G35210	CAAAGCAGCTTCTTCACGTAGACC	TCGTCGGCGATCAACAATAGGAG
AT1G35230	GGCACAAATGTTCCGTCTGGAG	TGCGTTTGGAGAGCCACTTAGG
AT1G36370	TCTCCTGGAGGTGTAAGAATAGGG	TCCGCCATTGTCTCGAAATCAG
AT1G54050	CGGCGAAAGTAATAACGAGAGTCG	CCTCTCTTCCTCCACTGTAACCTG
AT1G59590	TTGTATCGCTTGGTACTTGGTTCC	TCGCGTAAGACGGTGATGATCC
AT1G61340	GATAGCAAAGCAGTCACATTTCGC	CCCAACCAAATCTGCCATGGTG

AT1G61820	GGTTAAGCTCCGAGATGCAGAAAG	CGCTGGTGGAAACAAACCACTTC
AT1G62300	TGTCCCGTTCGCAAACAAGTTC	CGGCAACGGATGGTTATGGTTTC
AT1G63720	AGAAACCTCGCTGTTCCCTCAC	ACACCAGCAGTCTTTGGACTCAAC
AT1G66060	AATGGAGCCACGCTTTGACCTG	TGTTTGCCTCATCGTCCCTGTC
AT1G69890	TCCGTTTAAGGAGCTGCTACGG	TCCGAGCAAGAACCGCTCATTC
AT1G71000	TGGAAGGTGGCATTAGGCAAACG	TCTGCTCCGCCGAAGAATCAAC
AT1G72520	GGAAGACCACATCATCGGTCAAC	AAACGGTTCGTCTCTAACGCTTG
AT1G73010	TGCCTTGGAACACACTCATGGAC	CAGATTTGATGGCAGGGACGAC
AT1G74310	CTCTTGCTGAGCAGCTGTTTGATG	ACCTTCCTCGTGACCAACATACCC
AT1G74360	GCCATGTGAGCACAGTGATTGC	TGGTGGCTTGCCAAGTTTGTCC
AT1G74590	ACCAACAGGTGTTTGAGGTCATGG	ACAGACTTTGCTTGAGCTTCACC
AT1G76070	TTCCACCGTTTAGTCCCGGAAG	CCGCCGTGTAATTATGGTGGTTG
AT1G76600	TCCGGTTGTGACGTTGAATCAGG	CGTTGCTTCGCCTCCAATTCTG
AT1G77450	CCTGCTCCGATTATCACCGAAC	ACAAAGCCATGTCGGGAAGGTC
AT1G80820	GCGGAGAGGTTGTTGAGATTCTGG	GGATTCTTCTCGGTCCGAACACTTG
AT1G80840	AGCTTCTGACACTACCCTCGTTG	TTGACAGAACAGCTTGGAGCAC
AT2G02990	AATCTTGCCTTCTGTCTTCTCTGC	GTCACAGTATGATCCTGGCCATTG
AT2G15480	CAACGACCAGCTGTTAGAGATCG	ACCACTCTTCATTGTCACCTTGG
AT2G18210	TTAAGCACGAGGACGAGGAGTC	TGCTCCAAGTAGGCTGAAGGTG
AT2G18690	TGAAGCAAGTTTGTTGCCTTCGAC	TCATAAGGCGGGCTGCGTATTC
AT2G20560	TGCTGATGTCAGTGGCAAGACG	GCCTTTCTTCCATCCTGGTTTCAC
AT2G21640	TGCCTCAATGGATGAACACAAGAC	GAGAAGCTCCCGAATATCTTGTCC
AT2G22470	AGTGCCGCATGGGCTAACAAAG	ACAGCGTATAAAGCTCCGGCAAC
AT2G22880	TCCGTTCGCGTTTGATCTTTCACC	AGCAGAGGGAAAGCTTCCCAGAAC
AT2G25735	AGGCTCTTGCCCAAATGGAAGG	TGGAAGAAGCACGAGAAGGCAAG
AT2G26150	GCAGCGTTGGATGTGAAAGTGG	TTGGCTGTCCCAATCCAAAGGC
AT2G26530	TTCATCAGCTTGCGTGTCTTGC	GCACTTCGGAACAGAAGAAAGTCC
AT2G28210	GCTTCATTGGCATTCTCCCTCTG	TGTGTAGCTCGAGAGCAAACCTTC
AT2G29460	GATGAACAGGTTGGACCAGTAGC	TGAGCCTCCTTGATTGCAACCTC
AT2G29470	CGACCAAACATGGACAAACAATCC	CCTGAGCCCTATCATTGTGACTTG
AT2G29490	TCGATCAGACGTGGAAGAACAGTC	AAATCGAGCCATGGCCTTCTC
AT2G29500	TGGATCAGGTTAAGGCTGCGATGG	TCAGCCTTAGGCACCGTAACAGTC
AT2G30770	GTGCTTCGGTTGCATCCTTCTC	CGCCCAAGCATTGATTATCACCTC
AT2G31945	TGGCGACGATATGTCTCCTTCATC	GTTGTTGCTGCCTCATCTTGCG
AT2G32030	TTGCCACACCCTTGGCTTAGAG	GTTACGGAGATCGAGCCAATCG
AT2G32120	TCCTTTCGGTGGTGACAGAAATC	GCTACTGTGTATGCTGCTTCTGC
AT2G33710	ACCGACGGAGACAACAACATGC	CCGGTCCGGTCAACACATTATCTC
AT2G37430	TAGGTCAAGCATGACGGTCGAG	ACCGGCATGGATGGAATCATCG

AT2G38250	AAACGCCGGAAATAGCCTCTCC	TTGTCTCTTCGACGCTCCACTG
AT2G38340	TGAGTCACCGTGGTGCAAACTC	CTGCAGTAGCAAACGTGCCAAG
AT2G38470	CTTCCACTTGTTTCAGTCCCTCTC	CTGTGGTTGGAGAAGCTAGAACG
AT2G40000	TGTAAGTACACGCTCGGACTCG	TCGCATCGTACCGGAGATTTGG
AT2G41380	ACGATCGCTAGCCGGAATCTAC	TCAAGTTGCTTTGAGCTCGTGTC
AT2G41640	GGAGGTGGTTGTTGCTACTCCAAG	GCGAAACGAATCTAGGAGGGAGAG
AT2G43000	TCTCCAGCTCAACAAGCAGAGG	CGGTTTGGTGGTAAGATGGTTGGG
AT2G43510	AGGCTATCGTTTCCATCTTCGTTG	TCTCCGGCACATCGGAAATGAC
AT2G43820	TGCATTAACTGGCTCGACACAAGG	AGCCATGCTTCCGAATGCTACG
AT2G46240	TTGGATGCTGTCGAGGGATTGC	CTCTGTGGCCAAAGCTTTCCTG
AT2G47000	TTCGAGCTGGAAAGACAGTTGC	CCTCTGCAACAATGCAATCACC
AT3G02550	GAAGCGCAAGCTAACGCAACTG	ATCCCAGGACGAAGGTGATTGG
AT3G02800	CATACATTGCAAACGCGGGAAGC	AACCCGCGTTCTTCTGGTACTC
AT3G02840	ACGCTTTGATCGTTCCTCTTCTGG	CGAACATTGTGTCGCGAGATCC
AT3G04070	AACTCTTCTAGGAGCGGTGGTAGC	GGCAAAGAACCCAATCATCCAGTC
AT3G08720	AAGGACAAGATCAAGCTTCCACAG	TTTGCAGCAGCCCTTTCAGC
AT3G09270	TTCCAAGAAGCTTCAGGCGTCAC	CCCACGAAATCTTCCGACCAAC
AT3G09350	AGGCTATGCAATCGCAGACTGTAG	GCAGCTCATCCAGCAAATCTTGG
AT3G09410	ACCACTTCCTCTGCTTGTGGAC	TACAGACTGGTCCCGGATTGAG
AT3G10020	GTGGCGGTTGAGCTTGTGAAAG	ACCTCAACCCTAGCAACACCAG
AT3G10320	ATTGCGCTAAGAGATCGAAGCC	AATCGATGAGAAACAGAGCAGAGG
AT3G10930	ACCCGGAGAGTTTGTTCCAGTC	ACCGCCGACTTAGAAACAGGAG
AT3G11840	TATCCAACACTGGTGCGTCGAG	TCCTAAGGCTTCTTGGCCGAAC
AT3G12580	CTACCAACACCGTCTTCGATGC	GACTCTTATCCGCTTGAACAGAGG
AT3G13790	ATTCAGTGGCCGGTTAGCGAAG	CAATACTTCTACATCCGCCTGTGC
AT3G15500	CGTCGAAATGGAAGCACCAAGC	TGTCGACGAACCATTGTTGCTG
AT3G15950	CTATTGAACGCGAGTTTGAAGCC	CGTTCGAATGCTTCCAACATACTG
AT3G16530	AGAGAGGTTCAAGGCTTGGGTTG	AGGCGCGAGTGTAACCGTAATC
AT3G17690	TACTACAACCGCCAAGCTGCTC	CACCATACAGAACAGTGCGAAACC
AT3G23230	TTGGCTCGGGACATTTGAGACC	AAAGGCTGCTCGGTCATAAGCC
AT3G24500	AAGACCTACGCGATCCGAAAGC	TAACCGTTTGAACCGCGACACC
AT3G25250	ACCGTCACTGTCTAAACCATCGC	TGCAGAAACTGGTGAAGCGGAAG
AT3G26440	ACCCGGAATCGAAGTTCACAAAG	AGGTAGGTGAGGGTCATTGCAG
AT3G26830	CTTTAAGCTCGTGGTCAAGGAGAC	TGGGAGCAAGAGTGGAGTTGTTG
AT3G28210	TTGCGACGGCTGCAAATTGGTG	AAACCGTTCTGCTCCCATGATCG
AT3G28580	TTCGAATACTCCGGGTCAGAGC	AACGTAGCTGGATGCTCAAAGG
AT3G28740	CAGCCTGATTACTACACGGATGTG	AGTCCCGGCAAGTATCATAACAAG
AT3G46080	AAACGGCGACGTCTCATCCTTG	TGTAACCAACGTGCCTGGTGAG

AT3G46230	TTGCCAGAGAATGCAAAGGTGGAG	TGAACTTTCGGCACCGTAACCG
AT3G48450	GCAACGGGAAATAGAGCGAGAC	CCAGGATTAGTCGCATCCCATTCC
AT3G48650	AGCAGTCTCACTTGTGGTAGCAG	GCAATACTCGGTGTTGCCACTTC
AT3G49580	AGCTGGAGGTCGAGTCTTTAGAAC	TGAGGAAGAGCATGCGATCGTG
AT3G49620	GCACGAAGCCATTGATTGTTACAG	TTGGCCCTTCCAAGACCTTTCC
AT3G53230	ACATGAAGCTTGCGGAAGATGTTG	TGCATAATGCTGCAAGATCAGCAC
AT3G54150	TTGGGATGTTGGAACTGGTAATGG	TTCTCGTAATGCTCCACAAGCC
AT3G55980	AAACGGCGGATCATGGCAGAAC	TCCACCATTGAGCTGCAAAGCC
AT4G11280	ACGGCGAGAATTCCTCTTATTTCG	ACGCATCAAATCTCCACAAAGCTG
AT4G12400	CGTATTTGACGAATCGTGCTGCTG	CTTCAATGCACTCCTCGTACTTCC
AT4G17490	TCGAATCCTCCTCGCGTTACTG	TTCGGTGGTGCGATCTTCAACG
AT4G18880	AGCAGGAGCGAATGATGGCTTC	AACTTCCCGTTGCTCGGTTGAG
AT4G21390	ACGGTTCGCGTTGTTGGAACTTAC	AAGTCCTTCCATCGCGTATTCAGG
AT4G21990	CAGAGAAGCTAGATGTGGTGGAAG	AACATCTTCAGCTCCACTAAAGGC
AT4G22530	TTATCCCGCCGATTGGTACTCC	TCTTTCGTAGTGTTCCGCGATTCC
AT4G24160	CGTCTTCTCTCCATCCTCAAGACG	ACCTGGTGGTCCTGAACCAATAC
AT4G24570	ACGCAATCAGGAGCATGGTTAAGG	TCGCTCGGTTAATCGTCAACGC
AT4G25200	ACGAGGAGCGTTAGTCAAGTGC	GCCACGAGTAGCTGATAACAGAGG
AT4G34410	TGGGCCAAGGGCTAAACTCAAC	CAGCAGCAACAGGAGATGAAACTG
AT4G37370	CCAACCGGATTACTTCACGGATCG	TCGGTCCCTGCTAGTATCAAAGCG
AT4G37990	ACCGTCTCTGCGACTCATTCAC	CTCAAAGATGAGAGGCATGACAGG
AT4G39670	CGTTGAAAGACGCTGCGACAAC	TCTAACGGCCCACGTATGGAAC
AT5G04340	TCTACAAGCCACGTCAGCAGTG	TTCCGGTATCGGCGGTATGTTG
AT5G05410	TGTCTGGAGAATGGTGCGGAAG	TCGCTCAGCCAATGCTTATCCG
AT5G05730	ATGCATATAAGCTCCACGGTGAC	GTACGTCCCAGCAAGTCAAACC
AT5G12020	AAGACCCGCAACAACCCTTCAC	GTGTAGCAGCCATTGCCTTAGC
AT5G12030	TGCGGCTTGTAATGACGGTGTG	TTGGCTCAGGAGGAGGAAGTTTCG
AT5G13080	AGTGGACCAAGAAGTGGTCGTG	TTCTCGATGGGATGCGAATGCAC
AT5G14700	AATCGTCGTTGATTGTCCAGAGG	TCAGCGTCAGCTTCCATCTCAC
AT5G14730	AATTTCTGAGGAGAAGCCACAGC	GAAACTGCACCGTTTAGTTGCG
AT5G14760	TGGTCGCTGGTGCTCATCTTTG	AGGCCCTTCAGTACACACAACTC
AT5G20230	TTCGCTGCCGTTGTTGTCTTCG	AATTCGAGCTCGTCGCCTACAC
AT5G24110	TCTCGGAGCCAAATTTCCAAGAGG	TCCTCGGTAACTGATCTCAAGGAG
AT5G25450	CTCCAAGACATGTTGGCTCTTGTG	AAGCTCCCAAAGCCTCTCTCTCAG
AT5G26220	CAATCCACCGGAGCTATTTGCTG	TCTTTCTCTGGTCCTCCACGAAC
AT5G27420	ACGTGACTTGTCCGGTTTGTCG	ATCGGTTTCCACCACCACTTCC
AT5G28630	CTGCAGCAATGGTTGGTCTTGG	TTCCACCACCTTTGTGCTTCTTC
AT5G35320	CCAGGAAGATTCGTCGTCTGGATG	ACCGGCTCAAGCTCGATTTCTTC

AT5G37670	TGAGTTGCCGGAGAATGTGAAAG	CCCGGTTTAGACATCAAAGCTTGC
AT5G38900	ATGTTTGCAAGAATGTCCGAGGTC	ACCAGCTGTGTCAAACTCTAAGCC
AT5G42380	TCCGGTGAAGAGCTACAAAGCTG	TCCTCCACCTCACGTGACGATAAC
AT5G46080	CAGCGACTGCGGTTGAATTTGG	TACTCCAATCCTCTCGCTGCTC
AT5G47230	TGACCCGAATAAACGCGGATCTC	TCTAGCCGCTTCAATCGCTGTATC
AT5G48570	AACGGTGACGAAGTTGAAGTGC	GTCACGGCTCGAATCAAACTTGG
AT5G48850	TCACAAAGTTCCCTGTGGAGACAC	CGGGTTCTTCTCTATCAACTGAGC
AT5G51060	CAAAGGAGTCTTCTCTGAGGTGTG	CCATGCAACATGTCGGCTCTAAG
AT5G51190	TCTAAGCCAACGCAAACCACCTC	TTCTCTTGAACCACCGGTGCAG
AT5G51440	ACTCGTGGAATGGGAGCTTCTG	GATGCAACGCGTCGTCTTTCTC
AT5G52640	AAGCTCGATGGACAGCCTGAAC	TCCCAAGTTGTTCACCAAATCTGC
AT5G52760	TCGAGCCATGACCGCAAAGAAC	AATCGAGAAACGGTGACAAGCG
AT5G54490	GCAAAGGGTTCGAGCTTCTTATGG	CGTCGATGCGTTTCTTCGTAAGC
AT5G57220	CATCATCAAAGGGCTCATGCTCAG	AATGTTACGGCCGCAGTATCCG
AT5G58390	TGGGTGCTTCTCTCCTTCGTTTG	AGTAAGGATCCGTCACACCCATTG
AT5G59820	TTGGTTACACGCGCTTTGTTGC	ACAAGCCACTCTCTTCCCACTG
AT5G63790	TTAACGTTCCGGTTATCGCAGAG	CGTACAACGCCATTTCTGGAAGC
AT5G64310	AGCCGCATTGACTCCAGAATCC	AGTCAGCGGTAGGAGAATCAGC
AT5G64510	AAGCGTCTTGCTCCAGGACTTG	CGTTGGCACCAATGGTTGAGAG
AT5G65600	TCTTGGTTTAGCCTCTGCTCTGC	ACTTGCCTTGATGTCTCTGTGC

5.3 Primer sequences of ROS-responsive genes selected for expression study in clock mutants

Gene name	AGI	Forward (5´-3´)	Reverse (5´-3´)
RBOHC	AT5G51060	CAAAGGAGTCTTCTCTGAGGTGTG	CCATGCAACATGTCGGCTCTAAG
CCT	AT1G07050	AAGACTTGGGTCTGAAGCTGAATC	TCTGGTGATCAGACCAAGCATCG
BCAT	AT1G10060	CACGCGCTTATCTTGGAGGAAC	TCTTCTCATCACTTCAAGCACTGG
BBOX	AT1G28050	GGTGGCCAGAGTTCTCAGATATGG	AGCTTCCACTCGACTCGTATCC
ADOF1	AT1G51700	ACGGCGAAACAGAACCAACCAG	AGTCACAGCGAGGACACTTTAGC
PLATZ	AT1G76590	GCGAGTGCAACATGTTCTGTCTTG	TGGTACGATGATCTCCGTATCTGG
ATIPS2	AT2G22240	AGCTGGCATTAAGCCTACTTCG	TGAGAGGTTCATCCCGTCGTTG
PEROXIDASE	AT2G22420	TGTCCAGCTACTGTTTCTTGTGC	TCCCAATCTGGTCCTCCTGTAAG
PAL1	AT2G37040	GCAGTGCTACCGAAAGAAGTGG	TGTTCGGGATAGCCGATGTTCC
INVC	AT3G06500	GTTAGCCCTGTTGACTCTGGTTTG	TGTAGTCACCCGTGAGCTTTCC
JMJD5	AT3G20810	GCTGGGACAGTTACTCCGTTACAC	ACTTCTTGCCAACAACCTGAGC
ZAT7	AT3G46090	TTTGGGAGGTCATCGTGCAAGC	AGCTTGTCCCATCGGAAACTTCAC
APX4	AT4G09010	GCATATGGTTCAGCTGGTCAGTGG	TCAGCCTCTGTTGCATCACTCC
ISA3	AT4G09020	CTTTCATGCCACTGTCACCCTTG	GGAATGGGATGTACCGGGATGATG
HSFA4A	AT4G18880	AGCAGGAGCGAATGATGGCTTC	AACTTCCCGTTGCTCGGTTGAG
GSH1	AT4G23100	TGTCCTGAAACTCGCAAAGGATG	CATCGACTGCGTTCAAGAAACCG
VTC2	AT4G26850	TCAGACTGCTGTGTTTGCCTTC	TGTTTCTCTGCGTAACACTGTGG
PDX1	AT5G01410	TTTCCGGATCCCGTTCGTTTGC	TAATCATCGCCGCACCTTCACG
DREB2A	AT5G05410	TGTCTGGAGAATGGTGCGGAAG	TCGCTCAGCCAATGCTTATCCG
COR27	AT5G42900	AGCTAAGCTCTCATTCGCGTGAC	TCTGATCCGATACCTCTGCTTCTC
PLATZ	AT5G46710	GCAAGGTTGCAGGAACATCTGG	TCTCCACTCCAGAACTGTTGTTTG
MES18	AT5G58310	AAACCTGAGTATGTTCGGGACAAG	GTGGTTGCCAGTGTGTAATCCTC
LTP3	AT5G59320	AGCTTGCAGATGCATCCAGTCC	TGTTGTTGCAGTTAGTGCTCATGG
HSP18.2	AT5G59720	TTCACGCCATCTTCTGCGTTGG	TGTAAACGCTGCCACATCACGAG
MYB59	AT5G59780	GGTCAACTTTGTCCACTTGTTCGG	TCCCTCCACCTTCAAACCTGAAAC
CAT3	AT1G20620	TCGGGAAGGAGAACAACTTCAAGC	TCACGAATCGTTCTTGCCTGTC
CAT1	AT1G20630	AGGTACAGATCATGGGCACCAG	AAGGATCGATCAGCCTGAGACC
CAT2	AT4G35090	AAGTATCCAACTCCGCCTGCTG	TGGATGAATCGTTCTTGCCTCTC

5.4 Primer sequences of ROS transcription factors and regulatory genes used for expression analysis in response to MV treatments

Gene name	AGI	Forward (5´-3´)	Reverse (5´-3´)
bHLH128	AT1G10585	GGATCGAAGGATGCGCATGAAAC	AGGTGAGGCACTGGTAACTTGC
ZAT10	AT1G27730	TCACAAGGCAAGCCACCGTAAG	TTGTCGCCGACGAGGTTGAATG
PUP1	AT1G57990	GTCGCAACGGTTGTTTGTCTCG	AAACCTGTGGCTATCGCCTTGC
WRKY11	AT4G31550	CCCACGTGGTTACTACAAGTGC	TGGATCATCTAATGCTCGTTCCAC
ERF2	AT5G47220	ACACGTCATCATCGGACTTGAGC	TCGCCGTAAAGTTCTCAGTTGGC
ATCP1	AT5G49480	TCGCGATTCCGATAACCACCAC	ACCGTCGTGATCTGTGTCGATG
ZAT12	AT5G59820	TTGGTTACACGCGCTTTGTTGC	ACAAGCCACTCTCTTCCCACTG
CCA1	AT2G46830	CCAGATAAGAAGTCACGCTCAGAA	GTCTAGCGCTTGACCCATAGCT
LHY	AT1G01060	GACTCAAACACTGCCCAGAAGA	CGTCACTCCCTGAAGGTGTATTT

5.5 Primer sequences of ROS genes used for ChIP-qPCR assay

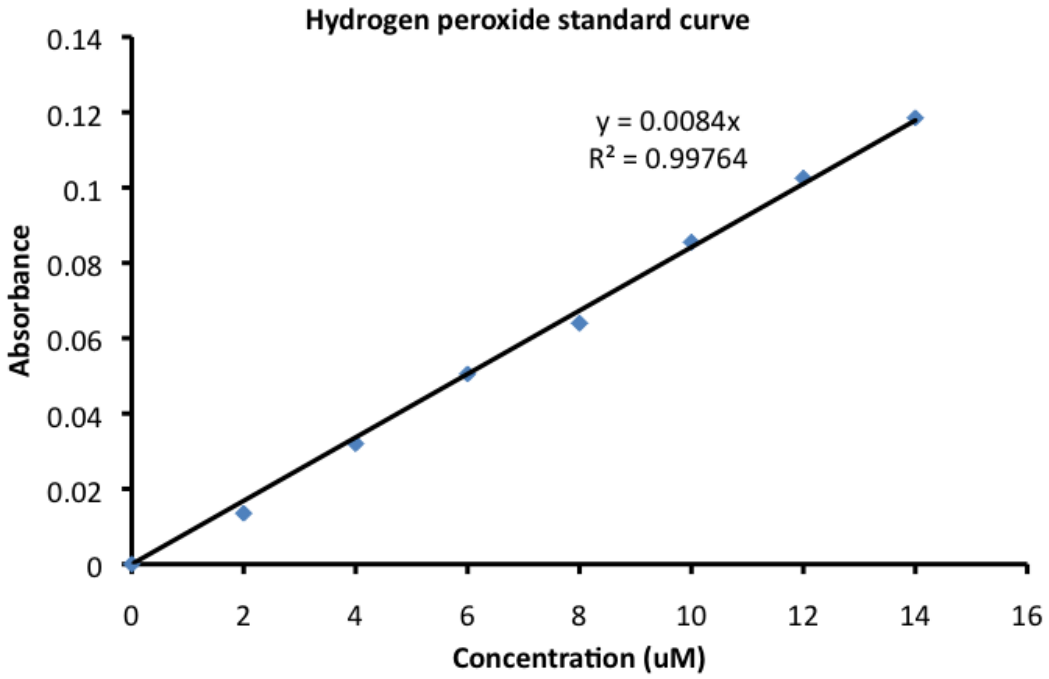
Gene name	AGI	Primer ID	Forward (5´-3´)	Reverse (5´-3´)
CAT3	AT1G20620	cat3M	TCACCGAACGAATTTTTCTTG	CAATAATGTAGTGCCCCCACTT
		cat3U5	ACAGTTCCACCTCGTGTCTGA	TGTTCGATCTTATCGCATAGAAA
		cat3U9	GACACTAAAGGGACCCCAAGT	TCGATGATGAACTGATCCACA
PAL1	AT2G37040	pal1U	TCCTCAGGAACAAATACAATTCC	AGGAGCTTATTTGTAATATTGTTGG
		pal1M	CACAATGAAACGCTTTACTCCA	CGGTTGGTGAGACTACGAGAT
		pal1U9	CGCCTCCTATTGGGAATCTTA	CGTGGAAATATCGCAAACTCT
JMJD5	AT3G20810	jmjM	ATTTTGGAGGGTTCGTTTTCA	TTTTGAAGTGATGATTAGGATCAGA
		jmjU5	GGGACTTGAATGTTTCATTGC	TGGGCTTACTGCAAACACACT
		jmjU9	TTTCTGCATATTACGTAAATTTTAACC	TCACAAAATGGCTGCAAAAAT
WRKY11	AT4G31550	wrkyM	GTTGGATCAAAACCTCCTTCC	ATGCGTGATATATGGGGAAAAG
		wrkU4	GAAGATTCAAATTGGGACAAACA	GAGGCAAGAGAAGGAAGAGTTG
		wrkyU9	GTCAGCTTTTCACCTTCCTTTG	GCAGTAATGGTTTGACTAATGATTTT
COR27	AT5G42900	cor27M	GAGAAGAGGACGGTGGAAGAT	ATCAAAGATATTTTCTTCAACTGCAA
		cor27U5	CATATGATGAAAAGAAGAATAATTTGA	GCTAGTTAGGGTGGAATTAATATGTGA
		cor27U9	TCAAACTCATCATTGGTGTAACCT	AAGAAACAGGAAAACTCCTTCAA
MYB59	AT5G59780	myb59U2	TTCCCATAAACCAGCTGAGAA	TTGCACAAGTTTCATCTCTCTCTC
		myb59U5	TGGTTTGATATTTGGATGCTG	TTGGCATGTCCCAATAGTTTC
		myb59M	TTCAATCTCAAGTGCTGTTGC	AAAAGGGGTCGAATTGTTTTG
ZAT12	AT5G59820	zat12d	AAACGGACACGAACACATCAT	TCCTCTTGTGGTTTGTGTTTTT

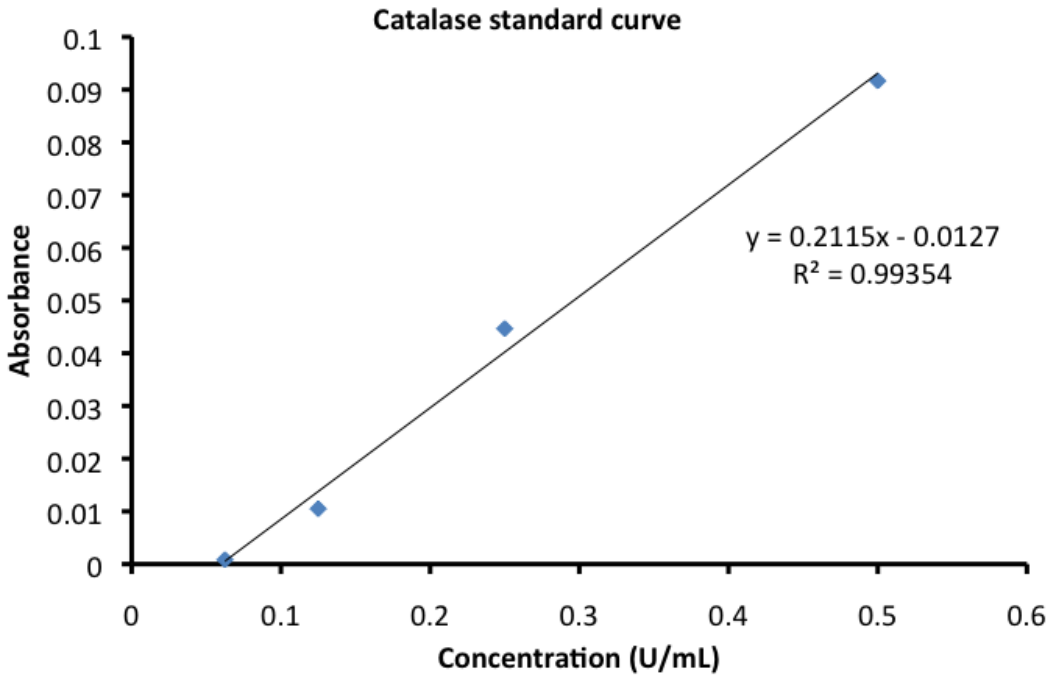
		zat12M	TCAAAAGAGTTTAGTTTTTGATGTGT	CGTGTTTTCTCGAAACCATTT
PUP1	AT1G57990	pupUT	TTCTCCTCCTCAACCATTTATCA	AAATTTCATTCTCTGTGGATTTTT
		pupM	TTTTGTTTGTTTATTTGGTTTTCG	ACGGATGAGACTGATCACACC
		pup9	TTCTGACGACTTTTGAAAACACA	TCTTGGTAATTATGTTGTAATTTGATG
PEROXIDASE	AT2G22420	perU2	CGAAAACCACCAAAAACCTCT	TGGTGTTGTGTTTCAGTGATTG
		perM	TTGATTATGAACATTGACCAGGA	TTTATGAAATGCGAATGATCTAGC
		perU9	TTGAGCACTACGACGACACTG	AAGCATGCATCTTTTATTCATTT
MES18	AT5G58310	mes18U	CTGCCGTTTACATCGTTCATT	TTTCTGTATGGCTGTAATGTTTTCA
		mes18M	CATGAAGTTGTCAACACCTGA	TGAACAATGTCTGTTTGGTGAA
		mes18u9	TAATGGCCGTCGTGACTATTT	CTTTTGACTGCATGAGCCTTC

5.6 Primer sequences of reference genes used for qPCR normalization control

Gene name	AGI	Forward (5´-3´)	Reverse (5´-3´)
UPL7	AT3G53090	TTCAAATACTTGCAGCCAACCTT	CCCAAAGAGAGGTATCACAAGAGACT
SAND	AT1G13440	AACTCTATGCAGCATTTGATCCACT	TGATTGCATATCTTTATCGCCATC
ACTIN2	AT3G18780	TCCCTCAGCACATTCCAGCAGAT	AACGATTCCTGGACCTGCCTCATC
IPP2	AT3G02780	AAAATATTGAAGCCAAGAATTTGC	AGTCCACACTAGAGGGAACGTAAC
TUB2	AT5G62690	GCCAATCCGGTGCTGGTAACA	CATACCAGATCCAGTTCCTCCTCCC

5.7 Standard curves





5.8 The 167 ROS-responsive genes used in expression profiling

AGI	Gene description
At1G01560	Member of MAP Kinase
AT1G02450	NIMIN1 modulates PR gene expression
AT1G03940	HXXXD-type acyl-transferase family protein
AT1G05340	Molecular functions uncharacterized
AT1G05575	Molecular functions uncharacterized, involved in anaerobic respiration
AT1G07160	Protein phosphatase 2C family protein, has protein serine/threonine phosphatase activity, catalytic activity
AT1G10040	Alpha/beta-Hydrolases superfamily protein
AT1G10585	Basic helix-loop-helix (bHLH) DNA-binding superfamily protein has sequence-specific DNA binding transcription factor activity
AT1G13340	Regulator of Vps4 activity in the MVB pathway protein
AT1G14040	EXS (ERD1/XPR1/SYG1) family protein
AT1G14200	RING/U-box superfamily protein, involves in zinc ion binding
AT1G15520	ABC transporter family involved in ABA transport and resistance to lead
AT1G16030	Heat shock protein 70B (Hsp70b), functions in protein folding, response to heat
AT1G17170	Encodes glutathione transferase belonging to the tau class of GSTs
AT1G17180	Encodes glutathione transferase belonging to the tau class of GSTs
AT1G18570	Encodes a member of the R2R3-MYB transcription family, involved in indole glucosinolate biosynthesis.
AT1G19020	Molecular functions uncharacterized
AT1G21100	O-methyltransferase family protein, has methyltransferase activity, O-methyltransferase activity, protein dimerization activity

AT1G21120	O-methyltransferase family protein, has methyltransferase activity, O-methyltransferase activity, protein dimerization activity
AT1G21310	Encodes extensin 3.
AT1G22400	UGT85A1, involved in metabolic process
AT1G26380	FAD-binding Berberine family protein, has electron carrier activity, oxidoreductase activity, FAD binding, catalytic activity
AT1G26420	FAD-binding Berberine family protein, has electron carrier activity, oxidoreductase activity, FAD binding, catalytic activity
AT1G27730	Related to Cys2/His2-type zinc-finger proteins found in higher plants. Compensated for a subset of calcineurin deficiency in yeast. Salt tolerance produced by ZAT10 appeared to be partially dependent on ENA1/PMR2, a P-type ATPase required for Li+ and Na+ efflux in yeast. The protein is localized to the nucleus, acts as a transcriptional repressor and is responsive to chitin oligomers. Also involved in response to photooxidative stress.
AT1G28190	Molecular functions uncharacterized
AT1G28480	Encodes GRX480, a member of the glutaredoxin family that regulates protein redox state. GRX480 interacts with TGA factors and suppresses JA-responsive PDF1.2 transcription. GRX480 transcription is SA-inducible and requires NPR1. Maybe involved in SA/JA cross-talk.
AT1G32170	Xyloglucan endotransglycosylase-related protein (XTR4)
AT1G35210	Molecular functions uncharacterized
AT1G35230	Encodes arabinogalactan-protein (AGP5).
AT1G36370	Encodes a putative serine hydroxymethyltransferase.
AT1G54050	HSP20-like chaperones superfamily protein
AT1G59590	ZCF37
AT1G61340	F-box family protein
AT1G61820	beta glucosidase 46 (*BGLU46*), involves in cation binding, hydrolase activity, hydrolyzing O-glycosyl compounds, catalytic activity,

	lignin biosynthetic process
AT1G62300	Encodes a transcription factor WRKY6. Regulates Phosphate1 (Pho1) expression in response to low phosphate (Pi) stress.
AT1G63720	Hydroxyproline-rich glycoprotein family protein
AT1G66060	Molecular functions uncharacterized
AT1G69890	Molecular functions uncharacterized
AT1G71000	Chaperone DnaJ-domain superfamily protein, involves in heat shock protein binding, protein folding
AT1G72520	PLAT/LH2 domain-containing lipoxygenase family protein, involves in oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen, lipoxygenase activity, iron ion binding, metal ion binding, growth, jasmonic acid biosynthetic process, response to wounding, defense response
AT1G73010	Encodes PPsPase1, a pyrophosphate-specific phosphatase catalyzing the specific cleavage of pyrophosphate (Km 38.8 uM) with an alkaline catalytic pH optimum
AT1G74310	Encodes ClpB1, which belongs to the Casein lytic proteinase/heat shock protein 100 (Clp/Hsp100) family. Involved in refolding of proteins, which form aggregates under heat stress. Also known as AtHsp101. AtHsp101 is a cytosolic heat shock protein required for acclimation to high temperature
AT1G74360	Leucine-rich repeat protein kinase family protein, involves in protein serine/threonine kinase activity, protein kinase activity, transmembrane receptor protein tyrosine kinase signaling pathway, protein amino acid phosphorylation
AT1G74590	Encodes glutathione transferase belonging to the tau class of GSTs
AT1G76070	Molecular functions uncharacterized
AT1G76600	Molecular functions uncharacterized
AT1G77450	NAC domain containing protein 32 (*NAC032*), involves in sequence-specific DNA binding transcription factor activity, multicellular organismal development, regulation of transcription

AT1G80820	Encodes an cinnamoyl CoA reductase isoform, involves in lignin biosynthesis
	Pathogen-induced transcription factor. Binds W-box sequences *in vitro*. Forms protein complexes with itself and with WRKY40 and WRKY60. Coexpression with WRKY18 or WRKY60 made plants more susceptible to both P. syringae and B. cinerea. WRKY18, WRKY40, and WRKY60 have partially redundant roles in response to the hemibiotrophic bacterial pathogen Pseudomonas syringae
AT1G80840	and the necrotrophic fungal pathogen Botrytis cinerea, with WRKY18 playing a more important role than the other two
	Member of the ribonuclease T2 family, responds to inorganic phosphate starvation, and inhibits production of anthocyanin. Also
AT2G02990	involved in wound-induced signaling independent of jasmonic acid
	UDP-glucosyl transferase 73B5 (*UGT73B5*), involves in quercetin 3-O-glucosyltransferase activity, UDP-glycosyltransferase activity,
AT2G15480	UDP-glucosyltransferase activity, transferase activity, transferring glycosyl groups, response to other organism
AT2G18210	Molecular functions uncharacterized
AT2G18690	Molecular functions uncharacterized
AT2G20560	DNAJ heat shock family protein, involves in unfolded protein binding, heat shock protein binding, protein folding
AT2G21640	Encodes a protein of unknown function that is a marker for oxidative stress response.
AT2G22470	Encodes arabinogalactan-protein (AGP2).
AT2G22880	VQ motif-containing protein
AT2G25735	Molecular functions uncharacterized
	Member of Heat Stress Transcription Factor (Hsf) family, involved in response to misfolded protein accumulation in the cytosol.
AT2G26150	Regulated by alternative splicing and non-sense-mediated decay
AT2G26530	Molecular functions uncharacterized
	Alpha carbonic anhydrase 2 (ACA2), involves in carbonate dehydratase activity, zinc ion binding, response to carbon dioxide, one-
AT2G28210	carbon metabolic process

AT2G29460	Encodes glutathione transferase belonging to the tau class of GSTs
AT2G29470	Encodes glutathione transferase belonging to the tau class of GSTs
AT2G29490	Encodes glutathione transferase belonging to the tau class of GSTs
AT2G29500	HSP20-like chaperones superfamily protein
AT2G30770	Putative cytochrome P450
AT2G31945	Molecular functions uncharacterized
AT2G32030	Acyl-CoA N-acyltransferases (NAT) superfamily protein, involves in N-acetyltransferase activity, metabolic process
AT2G32120	Heat-shock protein 70T-2 (HSP70T-2), involves in ATP binding, protein folding, response to high light intensity, response to hydrogen peroxide, response to heat
AT2G33710	Encodes a member of the ERF (ethylene response factor) subfamily B-4 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 7 members in this subfamily
AT2G37430	C2H2 and C2HC zinc fingers superfamily protein, involves in sequence-specific DNA binding transcription factor activity, zinc ion binding, nucleic acid binding, response to chitin, regulation of transcription
AT2G38250	Homeodomain-like superfamily protein, involves in sequence-specific DNA binding transcription factor activity
AT2G38340	Encodes a member of the DREB subfamily A-2 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are eight members in this subfamily including DREB2A AND DREB2B that are involved in response to drought
AT2G38470	Member of the plant WRKY transcription factor family. Regulates the antagonistic relationship between defense pathways mediating responses to P. syringae and necrotrophic fungal pathogens. Located in nucleus. Involved in response to various abiotic stresses especially salt stress
AT2G40000	Ortholog of sugar beet HS1 PRO-1 2 (HSPRO2)
AT2G41380	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein, involves in methyltransferase activity, response to

	cadmium ion
AT2G41640	Glycosyltransferase family 61 protein, involves in transferase activity, transferring glycosyl groups
AT2G43000	NAC domain containing protein 42 (NAC042), involves in sequence-specific DNA binding transcription factor activity
AT2G43510	Member of the defensin-like (DEFL) family. Encodes putative trypsin inhibitor protein, which may function in defense against herbivory
AT2G43820	Induced by Salicylic acid, virus, fungus and bacteria. Involved in the tryptophan synthesis pathway. Independent of NPR1 for their induction by salicylic acid. UGT74F1 transfers UDP:glucose to salicylic acid (forming a glucoside (SAG) and a glucose ester (SGE)), benzoic acid, and athranilate in vitro. UGT74F2 shows a weak ability to catalyze the formation of the p-aminobenzoate-glucose ester in vitro. But, UGT75B1 appears to be the dominant pABA acylglucosyltransferase in vivo based on assays in leaves, flowers, and siliques. The true biological substrate(s) of UGT74F2 are not known, but mutant plants lacking UGT74F2 have a decreased level of SAG and SGE
AT2G46240	A member of Arabidopsis BAG (Bcl-2-associated athanogene) proteins, plant homologs of mammalian regulators of apoptosis. Expression of BAG6 in leaves was strongly induced by heat stress. Knockout mutants exhibited enhanced susceptibility to fungal pathogen Botrytis cinerea. Plant BAG proteins are multi-functional and remarkably similar to their animal counterparts, as they regulate apoptotic-like processes ranging from pathogen attack, to abiotic stress, to plant development
AT2G47000	Multidrug resistance P-glycoprotein (MDR/PGP) subfamily of ABC transporters. Functions in the basipetal redirection of auxin from the root tip. Exhibits apolar plasma membrane localization in the root cap and polar localization in tissues above
AT3G02550	LOB domain-containing protein 41 (LBD41)
AT3G02800	Encodes an atypical dual-specificity phosphatase
AT3G02840	ARM repeat superfamily protein
AT3G04070	NAC domain containing protein 47 (NAC047)

AT3G08720	Encodes a ribosomal-protein S6 kinase. Gene expression is induced by cold and salt (NaCl). Activation of AtS6k is regulated by 1-naphthylacetic acid and kinetin, at least in part, via a lipid kinase-dependent pathway. Phosphorylates specifically mammalian and plant S6 at 25 degrees C but not at 37 degrees C. Involved in translational up-regulation of ribosomal proteins
AT3G09270	Encodes glutathione transferase belonging to the tau class of GSTs
AT3G09350	Encodes one of the Arabidopsis orthologs of the human Hsp70-binding protein 1 (HspBP-1) and yeast Fes1p: Fes1A (AT3G09350), Fes1B (AT3G53800), Fes1C (AT5G02150). Fes1A is cytosolic and associates with cytosolic Hsp70. Mutants showed increased heat-sensitive phenotype suggestion the involvement of Fes1A in acquired thermotolerance. Does not have nucleotide exchange factor activity *in vitro*
AT3G09410	Pectinacetylesterase family protein
AT3G10020	Molecular functions uncharacterized
AT3G10320	Glycosyltransferase family 61 protein
AT3G10930	Molecular functions uncharacterized
AT3G11840	Encodes a U-box-domain-containing E3 ubiquitin ligase that acts as a negative regulator of PAMP-triggered immunity.
AT3G12580	Heat shock protein 70 (HSP70)
AT3G13790	Encodes a protein with invertase activity.
AT3G15500	Encodes an ATAF-like NAC-domain transcription factor that doesn't contain C-terminal sequences shared by CUC1, CUC2 and NAM. Note: this protein (AtNAC3) is not to be confused with the protein encoded by locus AT3G29035, which, on occasion, has also been referred to as AtNAC3
AT3G15950	Similar to TSK-associating protein 1 (TSA1), contains 10 EFE repeats, a novel repeat sequence unique to plants. Expressed preferentially in the roots. Protein is localized to ER bodies- an endoplasmic reticulum derived structure. Loss of function mutations lack ER bodies

AT3G16530	Lectin like protein whose expression is induced upon treatment with chitin oligomers
AT3G17690	Member of Cyclic nucleotide gated channel family
AT3G23230	Encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, and ATERF-5
AT3G24500	One of three genes in A. thaliana encoding multiprotein bridging factor 1, a highly conserved transcriptional coactivator. May serve as a bridging factor between a bZIP factor and TBP. Its expression is specifically elevated in response to pathogen infection, salinity, drought, heat, hydrogen peroxide, and application of abscisic acid or salicylic acid. Constitutive expression enhances the tolerance of transgenic plants to various biotic and abiotic stresses
AT3G25250	Arabidopsis protein kinase
AT3G26440	Molecular functions uncharacterized
AT3G26830	Mutations in pad3 are defective in biosynthesis of the indole derived phytoalexin camalexin. Encodes a cytochrome P450 enzyme that catalyzes the conversion of dihydrocamalexic acid to camalexin
AT3G28210	Encodes a putative zinc finger protein (PMZ)
AT3G28580	P-loop containing nucleoside triphosphate hydrolases superfamily protein
AT3G28740	Encodes a member of the cytochrome p450 family. Expression is upregulated in response to cis-jasmonate treatment. Overexpression induces synthesis of volatile compounds that affect chemical ecology and insect interactions
AT3G46080	C2H2-type zinc finger family protein, involves in sequence-specific DNA binding transcription factor activity, zinc ion binding, nucleic acid binding, response to chitin, regulation of transcription
AT3G46230	Member of the class I small heat-shock protein (sHSP) family, which accounts for the majority of sHSPs in maturing seeds
AT3G48450	RPM1-interacting protein 4 (RIN4) family protein; CONTAINS InterPro DOMAIN/s: RPM1-interacting protein 4, defence response
AT3G48650	pseudogene, At14a-related protein, similar to At14a (GI:11994571 and GI:11994573) (Arabidopsis thaliana)

AT3G49580	*RESPONSE TO LOW SULFUR 1* (*LSU1*)
	Encodes a protein similar to 2-oxoacid-dependent dioxygenase. Expression is induced after 24 hours of dark treatment, in
	senescing leaves and treatment with exogenous photosynthesis inhibitor. Induction of gene expression was suppressed in excised
AT3G49620	leaves supplied with sugar
	ATPase, AAA-type, CDC48 protein, involves in hydrolase activity, nucleoside-triphosphatase activity, binding, nucleotide binding,
AT3G53230	ATP binding, response to cadmium ion
	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein, involves in methyltransferase activity, metabolic
AT3G54150	process
AT3G55980	Salt-inducible zinc finger 1 (*SZF1*)
	Encodes a a member of the 1-aminocyclopropane-1-carboxylate (ACC) synthase (S-adenosyl-L-methionine methylthioadenosine-
AT4G11280	lyase, EC 4.4.1.14) gene family
	Encodes one of the 36 carboxylate clamp (CC)-tetratricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with
AT4G12400	potential to interact with Hsp90/Hsp70 as co-chaperones
	Encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family (ATERF-6). The
AT4G17490	protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, and ATERF-5
AT4G18880	Member of Heat Stress Transcription Factor (Hsf) family
	B120, involves in protein serine/threonine kinase activity, sugar binding, protein kinase activity, ATP binding, protein amino acid
AT4G21390	phosphorylation, recognition of pollen
	Encodes a protein disulfide isomerase-like (PDIL) protein, a member of a multigene family within the thioredoxin (TRX)
AT4G21990	superfamily. This protein also belongs to the adenosine 5'-phosphosulfate reductase-like (APRL) group
AT4G22530	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein, involves in methyltransferase activity, metabolic

	process
AT4G24160	Encodes a soluble lysophosphatidic acid acyltransferase with additional triacylglycerol lipase and phosphatidylcholine hydrolyzing enzymatic activities. Plays a pivotal role in maintaining the lipid homeostasis by regulating both phospholipid and neutral lipid levels
AT4G24570	Encodes one of the mitochondrial dicarboxylate carriers
AT4G25200	AtHSP23.6-mito mRNA, nuclear gene encoding mitochondrial
AT4G34410	Encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, and ATERF-5
AT4G37370	Member of CYP81D
AT4G37990	Encodes an aromatic alcohol:NADP+ oxidoreductase whose mRNA levels are increased in response to treatment with a variety of phytopathogenic bacteria. Though similar to mannitol dehydrogenases, this enzyme does not have mannitol dehydrogenase activity
AT4G39670	Glycolipid transfer protein (GLTP) family protein, involves in glycolipid transporter activity, glycolipid binding, glycolipid transport
AT5G04340	Putative C2H2 zinc finger transcription factor mRNA
AT5G05410	Encodes a transcription factor that specifically binds to DRE/CRT cis elements (responsive to drought and low-temperature stress). Belongs to the DREB subfamily A-2 of ERF/AP2 transcription factor family (DREB2A). There are eight members in this subfamily including DREB2B. The protein contains one AP2 domain. Overexpression of transcriptional activation domain of DREB2A resulted in significant drought stress tolerance but only slight freezing tolerance in transgenic *Arabidopsis* plants. Microarray and RNA gel blot analyses revealed that DREB2A regulates expression of many water stress inducible genes
AT5G05730	ASA1 encodes the alpha subunit of anthranilate synthase, which catalyzes the rate-limiting step of tryptophan synthesis. ASA1 is induced by ethylene, and forms a link between ethylene signalling and auxin synthesis in roots

AT5G12020	17.6 kDa class II heat shock protein (HSP17.6II)
	Encodes a cytosolic small heat shock protein with chaperone activity that is induced by heat and osmotic stress and is also
AT5G12030	expressed late in seed development
	WRKY75 is one of several transcription factors induced during Pi deprivation. It is nuclear localized and regulated differentially during Pi starvation. RNAi mediated suppression of WRKY75 made the plants more susceptible to Pi stress as indicated by the
AT5G13080	higher accumulation of anthocyanin during Pi starvation
	NAD(P)-binding Rossmann-fold superfamily protein, involves in coenzyme binding, binding, cinnamoyl-CoA reductase activity,
AT5G14700	catalytic activity, lignin biosynthetic process, cellular metabolic process, metabolic process
AT5G14730	Molecular functions uncharacterized
AT5G14760	At5g14760 encodes for L-aspartate oxidase involved in the early steps of NAD biosynthesis
AT5G20230	Al-stress-induced gene
AT5G24110	Member of WRKY Transcription Factor; Group III
	Cytochrome bd ubiquinol oxidase, 14kDa subunit, involves in ubiquinol-cytochrome-c reductase activity, mitochondrial electron
AT5G25450	transport, ubiquinol to cytochrome c; LOCATED IN: mitochondrion, mitochondrial respiratory chain complex III
AT5G26220	ChaC-like family protein
	Encodes CNI1 (Carbon/Nitrogen Insensitive1) (also named as ATL31), a RING type ubiquitin ligase that functions in the
AT5G27420	Carbon/Nitrogen response for growth phase transition in Arabidopsis seedlings
AT5G28630	Glycine-rich protein
AT5G35320	Molecular functions uncharacterized
AT5G37670	HSP20-like chaperones superfamily protein
AT5G38900	Thioredoxin superfamily protein, involves in protein disulfide oxidoreductase activity, defense response to fungus, incompatible

	interaction
AT5G42380	Calmodulin like 37 (CML37), involves in calcium ion binding, response to ozone
AT5G46080	Protein kinase superfamily protein, involves in protein kinase activity, kinase activity, ATP binding, protein amino acid phosphorylation
AT5G47230	Encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family (ATERF-5). The protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, and ATERF-5
AT5G48570	Encodes one of the 36 carboxylate clamp (CC)-tetratricopeptide repeat (TPR) proteins (Prasad 2010, Pubmed ID: 20856808) with potential to interact with Hsp90/Hsp70 as co-chaperones
AT5G48850	Homologous to the wheat sulphate deficiency-induced gene sdi1. Expression in root and leaf is induced by sulfur starvation. Knockout mutants retained higher root and leaf sulfate concentrations, indicating a role in regulation of stored sulfate pools
AT5G51060	RHD2 (along with RHD3 and RHD4) is required for normal root hair elongation. Has NADPH oxidase activity. Gene is expressed in the elongation and differention zone in trichoblasts and elongating root hairs. RDH2 is localized to the growing tips of root hair cells. It is required for the production of reactive oxygen species in response to extracellular ATP stimulus. The increase in ROS production stimulates Ca^{2+} influx
AT5G51190	Encodes a member of the ERF (ethylene response factor) subfamily B-3 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 18 members in this subfamily including ATERF-1, ATERF-2, and ATERF-5.
AT5G51440	HSP20-like chaperones superfamily protein
AT5G52640	Encodes a cytosolic heat shock protein AtHSP90.1. AtHSP90.1 interacts with disease resistance signaling components SGT1b and RAR1 and is required for RPS2-mediated resistance
AT5G52760	Copper transport protein family
AT5G54490	Encodes a PINOID (PID)-binding protein containing putative EF-hand calcium-binding motifs. The interaction is dependent on the

	presence of calcium. mRNA expression is up-regulated by auxin. Not a phosphorylation target of PID, likely acts upstream of PID to regulate the activity of this protein in response to changes in calcium levels
AT5G57220	Member of CYP81F, involved in glucosinolate metabolism. Mutants had impaired resistance to fungi
AT5G58390	Peroxidase superfamily protein, involves in peroxidase activity, heme binding, response to oxidative stress, oxidation reduction
AT5G59820	Encodes a zinc finger protein involved in high light and cold acclimation. Overexpression of this putative transcription factor increases the expression level of 9 cold-responsive genes and represses the expression level of 15 cold-responsive genes, including CBF genes. Also, lines overexpressing this gene exhibits a small but reproducible increase in freeze tolerance. Because of the repression of the CBF genes by the overexpression of this gene, the authors speculate that this gene may be involved in negative regulatory circuit of the CBF pathway
AT5G63790	Encodes a member of the NAC family of transcription factors. ANAC102 appears to have a role in mediating response to low oxygen stress (hypoxia) in germinating seedlings
AT5G64310	Encodes arabinogalactan-protein (AGP1)
AT5G64510	Molecular functions uncharacterized
AT5G65600	Concanavalin A-like lectin protein kinase family protein, involves in kinase activity, protein amino acid phosphorylation

Gene descriptions were obtained from The Arabidopsis Information Resource (http://arabidopsis.org/).

5.9 Diurnal and circadian experimental conditions

Conditions	Plant age (days)	Accession	Media	Light intensity	Photoperiod	Temperature	Reference
Diurnal							
LLHC	7	Col-0	MS agar	100 uE	Constant Light (LL)	12 h 22 C /12 h 12 C	Michael et al., 2008b
LDHC	7	Col-0	MS agar	100 uE	12 h light /12 h dark	12 h 22 C /12 h 12 C	Michael et al., 2008b
LDHH-ST	35	Col-0	Soil	130 uE	12 h light /12 h dark	22 C	Blasing et al., 2005
LDHH-SM	29	Col-0	Soil	180 uE	12 h light /12 h dark	20 C	Blasing et al., 2005
Short Day	7	Ler	MS agar, 3% sucrose	180 uE	8 h light/ 16 h dark	22 C	Yanovsky & Kay, 2002
Long Day	7	Ler	MS agar, 3% sucrose	90 uE	16 h light/ 8 h dark	22 C	Yanovsky & Kay, 2002
Circadian							
LL(LLHC)	9	Col-0	MS agar	100 uE	Constant Light (LL)	22 C	Michael et al., 2008b
LL(LDHC)	9	Col-0	MS agar	100 uE	Constant Light (LL)	22 C	Michael et al., 2008b
LL12(LDHH)	7	Col-0	MS agar, 3% sucrose	100 uE	Constant Light (LL)	22 C	Harmer et al., 2000
LL23(LDHH)	8	Col-0	MS agar, 3% sucrose	60 uE	Constant Light (LL)	22 C	Edwards et al., 2006
DD(DDHH)	8	Col-0	MS agar, 3% sucrose	0 uE	Constant Dark (DD)	22 C	Hazen et al., 2005

6. Bibliography

Achard, P., Renou, J.-P., Berthomé, R., Harberd, N. P., & Genschik, P. 2008. Plant DELLAs restrain growth and promote survival of adversity by reducing the levels of reactive oxygen species. *Current Biology* **18**: 656–660.

Alabadí, D., Oyama, T., Yanovsky, M. J., Harmon, F. G., Más, P., & Kay, S. A. 2001. Reciprocal regulation between *TOC1* and *LHY/CCA1* within the *Arabidopsis* circadian clock. *Science* **293**: 880–883.

Allan, A. C., & Fluhr, R. 1997. Two distinct sources of elicited reactive oxygen species in tobacco epidermal cells. *Plant Cell* **9**: 1559–72.

Allen, G. J., Muir, S. R., & Sanders, D. 1995. Release of Ca^{2+} from individual plant vacuoles by both InsP3 and cyclic ADP-ribose. *Science* **268**: 735–737.

Alscher, R. G., Donahue, J. H., & Cramer, C.L. 1997. Reactive oxygen species and antioxidants: relationships in green cells. *Physiologia Plantarum* **100**: 224–33.

Anthony, R. G., Henriques, R., Helfer, A., Meszaros, T., Rios, G., Testerink, C., Munnik, T., Deak, M., Koncz, C., & Bogre, L. 2004. A protein kinase target of a PDK1 signalling pathway is involved in root hair growth in *Arabidopsis*. *EMBO Journal* **23**: 572–581.

Apel, K., & Hirt, H. 2004. Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. *Annual Review of Plant Biology* **55**: 373–399.

Apostol, I., Heinstein, P. F., & Low, P. S. 1989. Rapid stimulation of an oxidative burst during elicidation of cultured plant cells. Role in defense and signal transduction. *Plant Physiology* **90**: 106–16.

Asada, K., & Takahashi, M. 1987. Production and scavenging of active oxygen in photosynthesis. In Photoinhibition. Kyle, D.J. et al., eds: 227–287, Elsevier.

Aschoff, J. 1960. Exogenous and endogenous components in circadian rhythms. *Cold Spring Harbor Symposia on Quantitative Biology* **25**: 11–28.

Asher, G., Gatfield, D., Stratmann, M., Reinke, H., Dibner, C., Kreppel, F., Mostoslavsky, R., Alt, F. W., & Schibler, U. 2008. SIRT1 regulates circadian clock gene expression through PER2 deacetylation. *Cell* **134**: 317–328.

Asher, G., Reinke, H., Altmeyer, M., Gutierrez-Arcelus, M., Hottiger, M. O., & Schibler, U. 2010. Poly(ADP-ribose) polymerase 1 participates in the phase entrainment of circadian clocks to feeding. *Cell* **142**: 943–953.

Bae, G., & Choi, G. 2008. Decoding of light signals by plant phytochromes and their interacting proteins. *Annual Review of Plant Biology* **59**: 281–311.

Bashandy, T., Guilleminot, J., Vernoux, T., Caparros-Ruiz, D., Ljung, K., Meyer, Y., & Reichheld, J.-P. 2010. Interplay between the NADP-linked thioredoxin and glutathione systems in *Arabidopsis* auxin signaling. *Plant Cell* **22**: 376–391.

Bass, J., & Takahashi, J. S. 2011. Circadian rhythms: Redox redux. *Nature* **469**: 476–478.

Baudry, A., Ito, S., Song, Y. H., Strait, A. A., Kiba, T., Lu, S., Henriques, R., Pruneda-Paz, J. L., Chua, N.-H., Tobin, E. M., Kay, S. A., & Imaizumi, T. 2010. F-box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control *Arabidopsis* clock progression. *Plant Cell* **22**: 606–622.

Baxter-Burrell, A., Yang, Z., Springer, P. S., & Bailey-Serres, J. 2002. Rop GAP4-dependent Rop GTPase rheostat control of *Arabidopsis* oxygen deprivation tolerance. *Science* **296**: 2026–2028.

Bell-Pedersen, D., Cassone, V. M., Earnest, D. J., Golden, S. S., Hardin, P. E., Thomas, T. L., & Zoran, M. J. 2005. Circadian rhythms from multiple oscillators: lessons from diverse organisms. *Nature Reviews Genetics* **6**: 544–556.

Bieniawska, Z., Espinoza, C., Schlereth, A., Sulpice, R., Hincha, D., & Hannah, M. 2008. Disruption of the *Arabidopsis* circadian clock is responsible for extensive variation in the cold-responsive transcriptome. *Plant Physiology* **147**: 263.

BioDare Data Repository. Retrieved January 8, 2012, from http://www.biodare.ed.ac.uk

Bläsing, O. E., Gibon, Y., Günther, M., Höhne, M., Morcuende, R., Osuna, D., Thimm, O., Usadel, B., Scheible, W.-R., & Stitt, M. 2005. Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in *Arabidopsis. Plant Cell* **17**: 3257–3281.

Brenner, W. G., Romanov, G. A., Köllmer, I., Bürkle, L., & Schmülling, T. 2005. Immediate-early and delayed cytokinin response genes of *Arabidopsis* thaliana identified by genome wide expression profiling reveal novel cytokinin sensitive processes and suggest cytokinin action through transcriptional cascades. *Plant Journal* **44**: 314–333.

Brouwer, K. S., Van Valen, T., & Day, D. 1986. Hydroxamate-stimulated O_2 uptake in roots of *Pisum sativum* and *Zea mays*, mediated by a peroxidase. *Plant Physiology* **82**: 236–240.

Bünning, E. 1973. The Physiological Clock. The English Universities Press Limited., London and Springer-Verlag, New York, Heidelberg, Berlin.

Cao, D., Cheng, H., Wu, W., Soo, H. M., & Peng, J. 2006a. Gibberellin mobilizes distinct DELLA-dependent transcriptomes to regulate seed germination and floral development in *Arabid dopsis*. *Plant Physiology* **142**: 509–525.

Cao, S., Jiang, S., & Zhang, R. 2006b. The role of *GIGANTEA* gene in mediating the oxidative stress response and in *Arabidopsis*. *Plant Growth Regulation* **48**: 261–270.

Carré, I. 2002. *ELF3*: a circadian safeguard to buffer effects of light. *Trends in Plant Science* **7**: 4–6.

Chamnongpol, S., Willekens, H., Moeder, W., Langebartels, C., Sandermann, H., Van Montagu, M., InzÈ, D., & Van Camp, W. 1998. Defense activation and enhanced pathogen tolerance induced by H_2O_2 in transgenic tobacco. *Proceedings of the National Academy of Sciences U.S.A* **95**: 5818–5823.

Chaouch, S., & Noctor, G. 2010. Myo-inositol abolishes salicylic acid-dependent cell death and pathogen defence responses triggered by peroxisomal hydrogen peroxide. *New Phytologist* **188**: 711–718.

Coelho, S. M., Taylor, A. R., Ryan, K. P., Sousa-Pinto, I., Brown, M. T., & Brownlee, C. 2002. Spatiotemporal patterning of reactive oxygen production and Ca^{2+} wave propagation in Fucus rhizoid cells. *Plant Cell* **14**: 2369–2381.

Conklin, P. L., Williams, E. H., & Last, R. L. 1996. Environmental stress sensitivity of an ascorbic acid-deficient *Arabidopsis* mutant. *Proceedings of the National Academy of Sciences U.S.A* **93**: 9970–9974.

Covington, M.F., Panda, S., Liu, X.L., Strayer, C.A., Wagner, D.R. & Kay, S.A. 2001. ELF3 modulates resetting of the circadian clock in *Arabidopsis*. *Plant Cell* **13**: 1305–1315.

Covington, M., & Harmer, S. 2007. The circadian clock regulates auxin signaling and responses in *Arabidopsis*. *PLoS Biology* **5**: e222.

Covington, M. F., Maloof, J. N., Straume, M., Kay, S. A., & Harmer, S. L. 2008. Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development. *Genome Biology* **9**: R130.

Creissen, G., Firmin, J., Fryer, M., Kular, B., Leyland, N., Reynolds, H., Pastori, G., Wellburn, F., Baker, N., Wellburn, A., & Mullineaux, P. 1999. Elevated glutathione biosynthetic capacity in the chloroplasts of transgenic tobacco plants paradoxically causes increased oxidative stress. *Plant Cell* **11**: 1277–1292.

Cubas, P., Lauter, N., Doebley, J., & Coen, E. 1999. The TCP domain: a motif found in proteins regulating plant growth and development. *Plant Journal* **18**: 215–222.

D'Autréaux, B., & Toledano, M. B. 2007. ROS as signalling molecules: Mechanisms that generate specificity in ROS homeostasis. *Nature Reviews Molecular Cell Biology* **8**: 813–824.

Dai, S., Wei, X., Pei, L., Thompson, R. L., Liu, Y., Heard, J. E., Ruff, T. G., & Beachy, R. N. 2011. *BROTHER*

OF LUX ARRHYTHMO is a component of the *Arabidopsis* circadian clock. *Plant Cell* **23**: 961–972.

Dalchau, N., Baek, S. J., Briggs, H. M., Robertson, F. C., Dodd, A. N., Gardner, M. J., Stancombe, M. A., Haydon, M. J., Stan, G. B., & GonÁalves, J. M. 2011. The circadian oscillator gene *GIGANTEA* mediates a long-term response of the *Arabidopsis thaliana* circadian clock to sucrose. *Proceedings of the National Academy of Sciences U.S.A* **108**: 5104.

Damiola, F., Le Minh, N., Preitner, N., Kornmann, B., Fleury-Olela, F., & Schibler, U. 2000. Restricted feeding uncouples circadian oscillators in peripheral tissues from the central pacemaker in the suprachiasmatic nucleus. *Genes & Development* **14**: 2950–2961.

Davison, P. A., Hunter, C. N., & Horton, P. 2002. Overexpression of B-carotene hydroxylase enhances stress tolerance in *Arabidopsis*. *Nature* **418**: 203–206.

Davletova, S., Schlauch, K., Coutu, J., & Mittler, R. 2005. The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in *Arabidopsis*. *Plant Physiology* **139**: 847–856.

de Mairan, M. 1729. Observation botanique. Histoiredel Academie Royale des Sciences, Paris, 35.

de Montaigu, A., Tóth, R., & Coupland, G. 2010. Plant development goes like clockwork. *Trends in Genetics* **26**: 296–306.

Dietz, K.-J., Jacquot, J.-P., & Harris, G. 2010. Hubs and bottlenecks in plant molecular signalling networks. *New Phytologist* **188**: 919–938.

Dixon, L. E., Knox, K., Kozma-Bognár, L., Southern, M. M., Pokhilko, A., & Millar, A. J. 2011. Temporal repression of core circadian genes is mediated through *EARLY FLOWERING 3* in *Arabidopsis*. *Current Biology* **21**: 120–125.

Dodd, A. N., & Love, J. 2005. The plant clock shows its metal: circadian regulation of cytosolic free Ca^2. *Trends in Plant Science* **10**: 1360–1385.

Dodd, A., Gardner, M., Hotta, C., Hubbard, K., Dalchau, N., Love, J., Assie, J., Robertson, F., Jakobsen, M., & Goncalves, J. 2007. The *Arabidopsis* circadian clock incorporates a cADPR-based feedback loop. *Science* **318**: 1789–1792.

Dodd, A., Salathia, N., Hall, A., Kevei, E., & Tóth, R. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science* **309**: 630–633.

Doke, N. 1985. NADPH-dependent O2- generation in membrane fractions isolated from wounded potato tubers inoculated with *Phytophtora infestans*. *Physiological and Molecular Plant Pathology* **27**: 311–22.

Dowson-Day, M. J., & Millar, A. J. 1999. Circadian dysfunction causes aberrant hypocotyl elongation patterns in *Arabidopsis. Plant Journal* **17**: 63–71.

Doyle, M. R., Davis, S. J., Bastow, R. M., McWatters, H. G., Kozma-Bogn r, L., Nagy, F., Millar, A. J., & Amasino, R. M. 2002. The *ELF4* gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Nature* **419**: 74–77.

Dunand, C., Crèvecoeur, M., & Penel, C. 2007. Distribution of superoxide and hydrogen peroxide in *Arabidopsis* root and their influence on root development: Possible interaction with peroxidases. *New Phytologist* **174**: 332–341.

Dunlap, J. C., & Loros, J. J. 2004. The Neurospora circadian system. *Journal of Biological Rhythms* **19**: 414–424.

Dunlap, J. C., Loros, J. J., Colot, H. V., Mehra, A., Belden, W. J., Shi, M., Hong, C. I., Larrondo, L. F., Baker, C. L., Chen, C. H., Schwerdtfeger, C., Collopy, P. D., Gamsby, J. J., & Lambreghts, R. 2007. A circadian clock in Neurospora: How genes and proteins cooperate to produce a sustained, entrainable, and compensated biological oscillator with a period of about a day. *Cold Spring Harbor Symposia on Quantitative Biology* **72**: 57–68.

Edwards, K. D., Anderson, P. E., Hall, A., Salathia, N. S., Locke, J. C. W., Lynn, J. R., Straume, M., Smith, J. Q., & Millar, A. J. 2006. *FLOWERING LOCUS C* mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell* **18**: 639–650.

Edwards, K. D., Lynn, J. R., Gyula, P., Nagy, F., & Millar, A. J. 2005. Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. *Genetics* **170**: 387–400.

Eimert, K., Wang, S. M., Lue, W. I., & Chen, J. 1995. Monogenic recessive mutations causing both late floral initiation and excess starch accumulation in *Arabidopsis. Plant Cell* **7**: 1703–1712.

Epple, P., Mack, A. A., Morris, V. R. F., & Dangl, J. L. 2003. Antagonistic control of oxidative stress-induced cell death in Arabidopsis by two related, plant-specific zinc finger proteins. *Proceedings of the National Academy of Sciences U. S. A.* **100**: 6831–6836.

Eriksson, M. E., Hanano, S., Southern, M. M., Hall, A., & Millar, A. J. 2003. Response regulator homologues have complementary, light-dependent functions in the *Arabidopsis* circadian clock. *Planta* **218**: 159–162.

Farré, E. M., Harmer, S. L., Harmon, F. G., Yanovsky, M. J., & Kay, S. A. 2005. Overlapping and distinct roles of *PRR7* and *PRR9* in the *Arabidopsis* circadian clock. *Current Biology* **15**: 47–54.

Feillet, C., Ripperger, J., Magnone, M., & Dulloo, A. 2006. Lack of food anticipation in *Per2* mutant mice. *Current Biology*: 2016–2022.

Foreman, J., Demidchik, V., Bothwell, J. H. F., Mylona, P., Miedema, H., Torres, M. A., Linstead, P., Costa, S., Brownlee, C., Jones, J. D. G., Davies, J. M., & Dolan, L. 2003. Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature* **422**: 442–446.

Fowler, S., Lee, K., Onouchi, H., Samach, A., Richardson, K., Morris, B., Coupland, G., & Putterill, J. 1999. *GIGANTEA*: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *EMBO Journal* **18**: 4679–4688.

Fowler, S.G., Cook, D. & Thomashow, M.F. 2005. Low temperature induction of *Arabidopsis CBF1, 2,* and *3* is gated by the circadian clock. *Plant Physiology* **137**: 961–968.

Foyer, C. H. 2002. The contribution of photosynthetic oxygen metabolism to oxidative stress in plants. In Oxidative Stress in Plants. Inzé, D. & Van Montagu, M., eds: 33–68, Taylor and Francis.

Foyer, C. H., & Noctor, G. 2005. Oxidant and antioxidant signalling in plants: A re-evaluation of the concept of oxidative stress in a physiological context. *Plant, Cell & Environment* **28**: 1056–1071.

Franklin, K. A., Larner, V. S., & Whitelam, G. C. 2005. The signal transducing photoreceptors of plants. *International Journal of Developmental Biology* **49**: 653–664.

Frohnmeyer, H. & Staiger, D. 2003. Ultraviolet-B radiation-mediated responses in plants. balancing damage and protection. *Plant Physiology* **133**: 1420–1428.

Fujiwara, S., Wang, L., Han, L., Suh, S. S., SalomÈ, P. A., McClung, C. R., & Somers, D. E. 2008. Post-translational regulation of the *Arabidopsis* circadian clock through selective proteolysis and phosphorylation of Pseudo-response Regulator proteins. *Journal of Biological Chemistry* **283**: 23073–23083.

Fukushima, A., Kusano, M., Nakamichi, N., Kobayashi, M., Hayashi, N., Sakakibara, H., Mizuno, T., & Saito, K. 2009. Impact of clock-associated *Arabidopsis* pseudo-response regulators in metabolic coordination. *Proceedings of the National Academy of Sciences U.S.A* **106**: 7251–7256.

Gadjev, I., Vanderauwera, S., Gechev, T., Laloi, C., Minkov, I., Shulaev, V., Apel, K., InzÈ, D., Mittler, R., & Van Breusegem, F. 2006. Transcriptomic footprints disclose specificity of reactive oxygen species signaling in *Arabidopsis*. *Plant Physiology* **141**: 436–445.

Gapper, C., & Dolan, L. 2006. Control of plant development by reactive oxygen species. *Plant Physiology* **141**: 341–345.

Gazarian, I. G., Lagrimini, L. M., Mellon, F. A., Naldrett, M. J., Ashby, G. A., & Thorneley, R. N. 1998. Identification of skatolyl hydroperoxide and its role in the peroxidase-catalysed oxidation of indol-3-yl acetic acid. *Biochemical Journal* **333**: 223–232.

Gechev, T. S., Van Breusegem, F., Stone, J. M., Denev, I., & Laloi, C. 2006. Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. *Bioessays* **28**: 1091–1101.

Gechev, T., Gadjev, I., Van Breusegem, F., InzÈ, D., Dukiandjiev, S., Toneva, V., & Minkov, I. 2002. Hydrogen peroxide protects tobacco from oxidative stress by inducing a set of antioxidant enzymes. *Cellular and Molecular Life Sciences* **59**: 708–714.

Gendron, J. M., Pruneda-Paz, J. L., Doherty, C. J., Gross, A. M., Kang, S. E., & Kay, S. A. 2012. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proceedings of the National Academy of Sciences U.S.A*. 10.1073/pnas.1200355109.

Giacomelli, L., Masi, A., Ripoll, D. R., Lee, M. J., & van Wijk, K. J. 2007. *Arabidopsis thaliana* deficient in two chloroplast ascorbate peroxidases shows accelerated light-induced necrosis when levels of cellular ascorbate are low. *Plant Molecular Biology* **65**: 627–644.

Gibon, Y., Bläsing, O. E., Palacios-Rojas, N., Pankovic, D., Hendriks, J. H. M., Fisahn, J., Höhne, M., Günther, M., & Stitt, M. 2004. Adjustment of diurnal starch turnover to short days: Depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period. *Plant Journal* **39**: 847–862.

Gibon, Y., Usadel, Bjoern, Blaesing, O. E., Kamlage, B., Hoehne, M., Trethewey, R., & Stitt, M. 2006. Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. *Genome Biology* **7**: R76.

Gong, W., He, K., Covington, M., Dinesh-Kumar, S. P., Snyder, M., Harmer, S. L., Zhu, Y.-X., & Deng, X. W. 2008. The development of protein microarrays and their applications in DNA-protein and protein-protein interaction analyses of *Arabidopsis* transcription factors. *Molecular Plant* **1**: 27–41.

Gould, P. D., Locke, J. C. W., Larue, C., Southern, M. M., Davis, S. J., Hanano, S., Moyle, R., Milich, R., Putterill, J., Millar, A. J., & Hall, A. 2006. The molecular basis of temperature compensation in the *Arabidopsis* circadian clock. *Plant Cell* **18**: 1177–1187.

Graf, A., Schlereth, A., Stitt, M., & Smith, A. M. 2010. Circadian control of carbohydrate availability for growth in *Arabidopsis* plants at night. *Proceedings of the National Academy of Sciences U.S.A* **107**: 9458–9463.

Grant, J. J., & Loake, G. J. 2000. Role of reactive oxygen intermediates and cognate redox signaling in

disease resistance. *Plant Physiology* **124**: 21–29.

Green, C.B., Takahashi, J.S. & Bass, J. 2008. The meter of metabolism. *Cell* **134**: 728–742.

Green, R. M., Tingay, S., Wang, Z. Y., & Tobin, E. M. 2002. Circadian rhythms confer a higher level of fitness to *Arabidopsis* plants. *Plant Physiology* **129**: 576–584.

Grunewald, W., & Friml, J. 2010. The march of the PINs: developmental plasticity by dynamic polar targeting in plant cells. *EMBO Journal* **29**: 2700–2714.

Gutiérrez, R. A., Stokes, T. L., Thum, K., Xu, Xiaodong, Obertello, M., Katari, M. S., Tanurdzic, M., Dean, A., Nero, D. C., McClung, C. R., & Coruzzi, G. M. 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 U.S.A* **105**: 4939–4944.

Hall, A., Bastow, R. M., Davis, S. J., Hanano, S., McWatters, H. G., Hibberd, V., Doyle, M. R., Sung, S., Halliday, K. J., & Amasino, R. M. 2003. The *TIME FOR COFFEE* gene maintains the amplitude and timing of *Arabidopsis* circadian clocks. *Plant Cell* **15**: 2719–2729.

Hall, A., & Brown, P. 2007. Monitoring circadian rhythms in *Arabidopsis* thaliana using luciferase reporter genes. *Methods in Molecular Biology* **362**: 143–152.

Halliwell, B. 2006. Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life. *Plant Physiology* **141**: 312–322.

Hamilton, E. E., & Kay, S. A. 2008. SnapShot: Circadian clock proteins. *Cell* **135**: 368–368.

Han, L., Mason, M., Risseeuw, E. P., Crosby, W. L., & Somers, D. E. 2004. Formation of an SCF(ZTL) complex is required for proper regulation of circadian timing. *Plant Journal* **40**: 291–301.

Hardin, P. E. 2005. The circadian timekeeping system of *Drosophila*. *Current Biology* **15**: R714–22.

Harmer, S. L. 2009. The circadian system in higher plants. *Annual Review of Plant Biology* **60**: 357–377.

Harmer, S. L., Panda, S., & Kay, S. A. 2001. Molecular bases of circadian rhythms. *Annual Review of Cell and Developmental Biology* **17**: 215–253.

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

Hazen, S. P., Schultz, T. F., Pruneda-Paz, J. L., Borevitz, J. O., Ecker, J. R., & Kay, S. A. 2005. *LUX ARRHYTHMO* encodes a Myb domain protein essential for circadian rhythms. *Proceedings of the*

National Academy of Sciences U.S.A **102**: 10387–10392.

Helfer, A., Nusinow, D., Chow, B., Gehrke, A., Bulyk, M., & Kay, S. 2011. *LUX ARRHYTHMO* encodes a nighttime repressor of circadian gene expression in the *Arabidopsis* core clock. *Current Biology* **21**: 126–133.

Hellweger, F. L. 2010. Resonating circadian clocks enhance fitness in cyanobacteria in silico. *Ecological Modelling* **221**: 1620–1629.

Hicks, K. A., Millar, A. J., Carré, I. A., Somers, D. E., Straume, M., Meeks-Wagner, D. R., & Kay, S. A. 1996. Conditional circadian dysfunction of the *Arabidopsis early-flowering 3* mutant. *Science* **274**: 790–792.

Hoffmann, M. H. 2002. Biogeography of *Arabidopsis thaliana* (L.) Heynh. (Brassicaceae). *Journal of Biogeography* **29**: 125–134.

Hogenesch, J. B., & Herzog, E. D. 2011. Intracellular and intercellular processes determine robustness of the circadian clock. *FEBS letters* **585**: 1427–1434.

Hong, S., Song, H. R., Lutz, K., Kerstetter, R. A., Michael, T. P., & McClung, C. R. 2010. Type II protein arginine methyltransferase 5 (PRMT5) is required for circadian period determination in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences U.S.A* **107**: 21211–21216.

Hotta, C. T., Gardner, M. J., Hubbard, K. E., Baek, S. J., Dalchau, N., Suhita, D., Dodd, A. N., & Webb, A. A. R. 2007. Modulation of environmental responses of plants by circadian clocks. *Plant, Cell & Environment* **30**: 333–349.

Huang, Y.-C., Chang, Y.-L., Hsu, J.-J., & Chuang, H.-W. 2008. Transcriptome analysis of auxin-regulated genes of *Arabidopsis thaliana. Gene* **420**: 118–124.

Imaizumi, T. 2010. *Arabidopsis* circadian clock and photoperiodism: time to think about location. *Current Opinion in Plant Biology* **13**: 83–89.

Imaizumi, T. & Kay, S.A. 2006. Photoperiodic control of flowering: not only by coincidence. *Trends in Plant Science* **11**: 550–558.

Imaizumi, T., Kay, S. A., & Schroeder, J. I. 2007. Circadian rhythms. Daily watch on metabolism. *Science* **318**: 1730–1731.

Iwasaki, H., & Kondo, T. 2004. Circadian timing mechanism in the prokaryotic clock system of cyanobacteria. *Journal of Biological Rhythms* **19**: 436–444.

Jansen, M. A., van den Noort, R. E., Tan, M. Y., Prinsen, E., Lagrimini, L. M., & Thorneley, R. N. 2001.

Phenol-oxidizing peroxidases contribute to the protection of plants from ultraviolet radiation stress. *Plant Physiology* **126**: 1012–1023.

Jin, H. & Martin, C. 1999. Multifunctionality and diversity within the plant MYB-gene family. *Plant Molecular Biology* **41**: 577–585.

Jones, M. A. 2009. Entrainment of the *Arabidopsis* circadian clock. *Journal of Plant Biology* **52**: 202–209.

Jones, M. A., Covington, M. F., DiTacchio, L., Vollmers, C., Panda, S., & Harmer, S. L. 2010. Jumonji domain protein JMJD5 functions in both the plant and human circadian systems. *Proceedings of the National Academy of Sciences U.S.A* **107**: 21623–21628.

Joo, J. H., Bae, Y. S., & Lee, J. S. 2001. Role of auxin-induced reactive oxygen species in root gravitropism. *Plant Physiology* **126**: 1055–1060.

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

Kerwin, R.E., Jimenez-Gomez, J.M., Fulop, D., Harmer, S.L., Maloof, J.N. & Kliebenstein, D.J. 2011. Network quantitative trait loci mapping of circadian clock outputs Identifies metabolic pathway-to-clock linkages in *Arabidopsis*. *Plant Cell* **23**: 471–485.

Khokon, A. R., Okuma, E., Hossain, M. A., Munemasa, S., Uraji, M., Nakamura, Y., Mori, I. C., & Murata, Y. 2011. Involvement of extracellular oxidative burst in salicylic acid-induced stomatal closure in *Arabidopsis. Plant, Cell & Environment* **34**: 434–443.

Kiba, T., Henriques, R., Sakakibara, H., & Chua, N.-H. 2007. Targeted degradation of PSEUDO-RESPONSE REGULATOR5 by an SCFZTL complex regulates clock function and photomorphogenesis in *Arabidopsis thaliana. Plant Cell* **19**: 2516–2530.

Kidokoro, S., Maruyama, K., Nakashima, K., Imura, Y., Narusaka, Y., Shinwari, Z. K., Osakabe, Y., Fujita, Y., Mizoi, J., & Shinozaki, K. 2009. The phytochrome-interacting factor PIF7 negatively regulates *DREB1* expression under circadian control in *Arabidopsis*. *Plant Physiology* **151**: 2046–2057.

Kikis, E., & Khanna, R. 2005. *ELF4* is a phytochrome-regulated component of a negative-feedback loop involving the central oscillator components *CCA1* and *LHY*. *Plant Journal* **44**: 300–313.

Kim, J, Kim, Y., Yeom, M., Kim, JH, & Nam, H. 2008. *FIONA1* is essential for regulating period length in the *Arabidopsis* circadian clock. *Plant Cell* **20**: 307–319.

Kim, W.-Y., Fujiwara, S., Suh, S.-S., Kim, Jeongsik, Kim, Y., Han, L., David, K., Putterill, J., Nam, H.-G., &

Somers, D. E. 2007. ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature* **449**: 356–360.

Kolmos, E., Herrero, E., Bujdoso, N., Millar, A. J., Tóth, R., Gyula, P., Nagy, F., & Davis, S. J. 2011. A reduced-function allele reveals that *EARLY FLOWERING3* repressive action on the circadian clock is modulated by phytochrome signals in *Arabidopsis. Plant Cell* **23**: 3230–3246.

Kovtun, Y., Chiu, W-L., Tena, G., & Sheen, J. 2000. Functional analysis of oxidative stress- activated mitogen-activated protein kinase cascade in plants. *Proceedings of the National Academy of Sciences U. S. A.* **97**: 2940–2945.

Kuno, N., Møller, S., Shinomura, T., & Xu, XM. 2003. The novel MYB protein EARLY-PHYTOCHROME-RESPONSIVE1 is a component of a slave circadian oscillator in *Arabidopsis*. *Plant Cell* **15**: 2476–2488.

Kurepa, J., Smalle, J., Va, M., Montagu, N., & InzÈ, D. 1998a. Oxidative stress tolerance and longevity in *Arabidopsis*: the late flowering mutant gigantea is tolerant to paraquat. *Plant Journal* **14**: 759–764.

Kurepa, J., Smalle, J., Van Montagu, M., & InzÈ, D. 1998b. Polyamines and paraquat toxicity in *Arabidopsis thaliana. Plant & Cell Physiology* **39**: 987–992.

Lamesch, P., Berardini, T., & Li, D. 2012. The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Research* **40**: D1202–D1210.

Lamia, K. A., Papp, S. J., Yu, R. T., Barish, G. D., Uhlenhaut, N. H., Jonker, J. W., Downes, M., & Evans, R. M. 2011. Cryptochromes mediate rhythmic repression of the glucocorticoid receptor. *Nature* **0**: 1–6.

Lamia, K. A., Sachdeva, U. M., DiTacchio, L., Williams, E. C., Alvarez, J. G., Egan, D. F., Vasquez, D. S., Juguilon, H., Panda, S., Shaw, R. J., Thompson, C. B., & Evans, R. M. 2009. AMPK regulates the circadian clock by cryptochrome phosphorylation and degradation. *Science* **326**: 437–440.

Lee, K. P., Kim, C., Landgraf, F., & Apel, K. 2007. EXECUTER1-and EXECUTER2-dependent transfer of stress-related signals from the plastid to the nucleus of *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences U.S.A* **104**: 10270–10275.

Legnaioli, T., Cuevas, J., & Más, P. 2009. TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. *EMBO Journal* **28**: 3745–3757.

Lin, C. 2002. Blue light receptors and signal transduction. *Plant Cell* **14**: S207–S225.

Locke, J., & Millar, A. 2005. Modelling genetic networks with noisy and varied experimental data: the

circadian clock in *Arabidopsis thaliana*. *Journal of Theoretical Biology* **234**: 383–393.

Locke, J. C. W., Kozma-Bognár, L., Gould, P. D., Fehér, B., Kevei, E., Nagy, F., Turner, M. S., Hall, A., & Millar, A. J. 2006. Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana. Molecular Systems Biology* **2**: 59.

Love, J., Dodd, A. N., & Webb, A. A. R. 2004. Circadian and diurnal calcium oscillations encode photoperiodic information in *Arabidopsis. Plant Cell* **16**: 956–966.

Lowrey, P. L., & Takahashi, J. S. 2004. Mammalian Circadian Biology: Elucidating genome-wide levels of temporal organization. *Annual Review of Genomics and Human Genetics* **5**: 407–441.

Lu, S. X., Knowles, S. M., Webb, C. J., Celaya, R. B., Cha, C., Siu, J. P., & Tobin, E. M. 2011. The Jumonji C Domain-Containing Protein JMJ30 Regulates Period Length in the *Arabidopsis* Circadian Clock. *Plant Physiology* **155**: 906–915.

Makino, S., Matsushika, A., Kojima, M., Yamashino, T., & Mizuno, T. 2002. The APRR1/TOC1 quintet implicated in circadian rhythms of *Arabidopsis thaliana*: I. Characterization with APRR1-overexpressing plants. *Plant & Cell Physiology* **43**: 58–69.

Malan, C., Gregling, M. M., & Gressel, J. 1990. Correlation between CuZn superoxide dismutase and glutathione reductase and environmental and xenobiotic stress tolerance in maize inbreds. *Plant Science* **69**: 157–166.

Marcheva, B., Ramsey, K. M., Buhr, E. D., Kobayashi, Y., Su, H., Ko, C. H., Ivanova, G., Omura, C., Mo, S., Vitaterna, M. H., Lopez, J. P., Philipson, L. H., Bradfield, C. A., Crosby, S. D., JeBailey, L., et al. 2010. Disruption of the clock components *CLOCK* and *BMAL1* leads to hypoinsulinaemia and diabetes. *Nature* **466**: 627–631.

Martin-Tryon, E. L., & Harmer, S. L. 2008. *XAP5 CIRCADIAN TIMEKEEPER* coordinates light signals for proper timing of photomorphogenesis and the circadian clock in *Arabidopsis. Plant Cell* **20**: 1244–1259.

Martin-Tryon, E. L., Kreps, J. A., & Harmer, S. L. 2007. *GIGANTEA* acts in blue light signaling and has biochemically separable roles in circadian clock and flowering time regulation. *Plant Physiology* **143**: 473–486.

Mathur, J. 2004. Cell shape development in plants. *Trends in Plant Science* **9**: 583–590.

Más, P. 2008. Circadian clock function in *Arabidopsis thaliana*: time beyond transcription. *Trends in Cell Biology* **18**: 273–281.

Más, P. 2005. Circadian clock signaling in *Arabidopsis thaliana:* from gene expression to physiology and development. *International Journal of Developmental Biology* **49**: 491–500.

Más, P., & Yanovsky, M. J. 2009. Time for circadian rhythms: plants get synchronized. *Current Opinion in Plant Biology* **12**: 574–579.

McClung, C. R. 2006. Plant circadian rhythms. *Plant Cell* **18**: 792–803.

McWatters, H. G., & Devlin, P. F. 2011. Timing in plants-a rhythmic arrangement. *FEBS letters* **585**: 1474–1484.

McWatters, H., Bastow, R., Hall, A. & Millar, A. 2000. The *ELF3 zeitnehmer* regulates light signalling to the circadian clock. *Nature* **408**: 716–720.

Messerli, G., Partovi Nia, V., Trevisan, M., Kolbe, A., Schauer, N., Geigenberger, P., Chen, J., Davison, A. C., Fernie, A. R., & Zeeman, S. C. 2007. Rapid classification of phenotypic mutants of *Arabidopsis* via metabolite fingerprinting. *Plant Physiology* **143**: 1484–1492.

Mhamdi, A., Hager, J., Chaouch, S., Queval, G., Han, Y., Taconnat, L., Saindrenan, P., Gouia, H., Issakidis-Bourguet, E., & Renou, J. P. 2010. *Arabidopsis GLUTATHIONE REDUCTASE 1* plays a crucial role in leaf responses to intracellular hydrogen peroxide and in ensuring appropriate gene expression through both salicylic acid and jasmonic acid signaling pathways. *Plant Physiology* **153**: 1144–1160.

Miao, Y., Laun, T., Zimmermann, P., & Zentgraf, U. 2004. Targets of the *WRKY53* transcription factor and its role during leaf senescence in *Arabidopsis. Plant Molecular Biology* **55**: 853–867.

Michael, T. P., & McClung, C. R. 2002. Phase-specific circadian clock regulatory elements in *Arabidopsis*. *Plant Physiology* **130**: 627–638.

Michael, T. P., Breton, G., Hazen, S. P., Priest, H., Mockler, T. C., Kay, S. A., & Chory, J. 2008a. A morning-specific phytohormone gene expression program underlying rhythmic plant growth. *PLoS Biology* **6**: 1887–1898.

Michael, T. P., Mockler, T. C., Breton, G., McEntee, C., Byer, A., Trout, J. D., Hazen, S. P., Shen, R., Priest, H. D., & Sullivan, C. M. 2008b. Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genetics* **4**: e14.

Michael, T. P., SalomÈ, P. A., Yu, H. J., Spencer, T. R., Sharp, E. L., McPeek, M. A., Alonso, J. M., Ecker, J. R., & McClung, C. R. 2003. Enhanced fitness conferred by naturally occurring variation in the circadian clock. *Science* **302**: 1049–1053.

Mikkelsen, M., & Thomashow, M. 2009. A role for circadian evening elements in cold-regulated gene

expression in *Arabidopsis*. *Plant Journal* **60**: 328–339.

Millar, A.J. & Kay, S.A. 1996. Integration of circadian and phototransduction pathways in the network controlling *CAB* gene transcription in *Arabidopsis*. *Proceedings of the National Academy of Sciences U S A* **93**: 15491–15496.

Millar, A. J., Carré, I. A., Strayer, C. A., Chua, N. H., & Kay, S. A. 1995. Circadian clock mutants in *Arabidopsis* identified by luciferase imaging. *Science* **267**: 1161–1163.

Millar, A. J., Short, S. R., Chua, N. H., & Kay, S. A. 1992. A novel circadian phenotype based on firefly luciferase expression in transgenic plants. *Plant Cell* **4**: 1075–1087.

Miller, E. W., Dickinson, B. C., & Chang, C. J. 2010. Aquaporin-3 mediates hydrogen peroxide uptake to regulate downstream intracellular signaling. *Proceedings of the National Academy of Sciences U.S.A* **107**: 15681–15686.

Miller, G., Schlauch, K., Tam, R., Cortes, D., Torres, M., Shulaev, V., Dangl, J., & Mittler, R. 2009. The Plant NADPH Oxidase RBOHD Mediates Rapid Systemic Signaling in Response to Diverse Stimuli. *Science Signaling* **2**: ra45.

Mittler, R. 2002. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* **7**: 405–410.

Mittler, R., Kim, Y., Song, L., Coutu, J., Coutu, A., Ciftci-Yilmaz, S., Lee, H., Stevenson, B., & Zhu, J.-K. 2006. Gain- and loss-of-function mutations in Zat10 enhance the tolerance of plants to abiotic stress. *FEBS letters* **580**: 6537–6542.

Mittler, R., Vanderauwera, S., Gollery, M., & Van Breusegem, F. 2004. Reactive oxygen gene network of plants. *Trends in Plant Science* **9**: 490–498.

Mittler, R., Vanderauwera, S., Suzuki, N., Miller, G., Tognetti, V. B., Vandepoele, K., Gollery, M., Shulaev, V., & Van Breusegem, F. 2011. ROS signaling: The new wave? *Trends in Plant Science* **16**: 300–309.

Mizoguchi, T., Wheatley, K., Hanzawa, Y., Wright, L., Mizoguchi, M., Song, H. R., Carré, I. A., & Coupland, G. 2002. *LHY* and *CCA1* are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. *Developmental Cell* **2**: 629–641.

Møller, I. M. 2001. Plant mitochondria and oxidative stress: electron transport, NADPH turnover, and metabolism of reactive oxygen species. *Annual Reviews of Plant Physiology and Plant Molecular Biology* **52**: 561–591.

Monshausen, G. B., Bibikova, T. N., Messerli, M. A., Shi, C., & Gilroy, S. 2007. Oscillations in extracellular pH and reactive oxygen species modulate tip growth of *Arabidopsis* root hairs. *Proceedings of the*

National Academy of Sciences U.S.A **104**: 20996–21001.

Mu, R.-L., Cao, Y.-R., Liu, Y.-F., Lei, G., Zou, H.-F., Liao, Y., Wang, H.-W., Zhang, W.-K., Ma, B., Du, J.-Z., Yuan, M., Zhang, J.-S., & Chen, S.-Y. 2009. An R2R3-type transcription factor gene *AtMYB59* regulates root growth and cell cycle progression in *Arabidopsis*. *Cell Research* **19**: 1291–1304.

Mullineaux, P., & Karpinski, S. 2002. Signal transduction in response to excess light: getting out of the chloroplast. *Current Opinion in Plant Biology* **5**: 43–48.

Møller, I. M., Jensen, P. E., & Hansson, A. 2007. Oxidative modifications to cellular components in plants. *Annual Review of Plant Biology* **58**: 459–481.

Nagy, F., & Schäfer, E. 2002. Phytochromes control photomorphogenesis by differentially regulated, interacting signaling pathways in higher plants. *Annual Review of Plant Biology* **53**: 329–355.

Nakahata, Y., Kaluzova, M., Grimaldi, B., Sahar, S., Hirayama, J., Chen, D., Guarente, L. P., & Sassone-Corsi, P. 2008. The NAD^+-dependent deacetylase SIRT1 modulates CLOCK-mediated chromatin remodeling and circadian control. *Cell* **134**: 329–340.

Nakamichi, N., Kiba, T., Henriques, R., Mizuno, T., Chua, N., & Sakakibara, H. 2010. PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *Plant Cell* **22**: 594–605.

Nakamichi, N., Kusano, M., Fukushima, A., Kita, M., Ito, S., Yamashino, T., Saito, K., Sakakibara, H., & Mizuno, T. 2009. Transcript profiling of an Arabidopsis *PSEUDO RESPONSE REGULATOR* arrhythmic triple mutant reveals a role for the circadian clock in cold stress response. *Plant & Cell Physiology* **50**: 447–462.

Nelson, D. C., Lasswell, J., Rogg, L. E., Cohen, M. A., & Bartel, B. 2000. *FKF1*, a clock-controlled gene that regulates the transition to flowering in *Arabidopsis*. *Cell* **101**: 331–340.

Ni, Z., Kim, E. D., Ha, M., Lackey, E., Liu, J., Zhang, Y., Sun, Q., & Chen, Z. J. 2008. Altered circadian rhythms regulate growth vigour in hybrids and allopolyploids. *Nature* **457**: 327–331.

Nishimura, M., & Dangl, J. L. 2010. *Arabidopsis* and the plant immune system. *Plant Journal* **61**: 1053–1066.

Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., & Shigeoka, S. 2006. Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant Journal* **48**: 535–547.

Nozue, K., Covington, M. F., Duek, P. D., Lorrain, S., Fankhauser, C., Harmer, S. L., & Maloof, J. N. 2007.

Rhythmic growth explained by coincidence between internal and external cues. *Nature* **448**: 358–361.

Nusinow, D. A., Helfer, A., Hamilton, E. E., King, J. J., Imaizumi, T., Schultz, T. F., Farré, E. M., & Kay, S. A. 2011. The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature* **475**: 398–402.

O'Connor, T. R., Dyreson, C., & Wyrick, J. J. 2005. Athena: a resource for rapid visualization and systematic analysis of Arabidopsis promoter sequences. *Bioinformatics* **21**: 4411–4413.

O'Neill, J. S., & Reddy, A. B. 2011. Circadian clocks in human red blood cells. *Nature* **469**: 498–503.

O'Neill, J. S., Van Ooijen, G., Dixon, L. E., Troein, C., Corellou, F., Bouget, F. Y., Reddy, A. B., & Millar, A. J. 2011. Circadian rhythms persist without transcription in a eukaryote. *Nature* **469**: 554–558.

Onai, K., & Ishiura, M. 2005. PHYTOCLOCK 1 encoding a novel GARP protein essential for the *Arabidopsis* circadian clock. *Genes to Cells* **10**: 963–972.

Orozco-Cardenas, M. L., Narvaez-Vasquez, J., & Ryan, C. A. 2001. Hydrogen peroxide acts as a second messenger for the induction of defense genes in tomato plants in response to wounding, systemin and methyl jasmonate. *Plant Cell* **13**: 17-191.

Ort, D., & Baker, N. R. 2002. A photoprotective role for O_2 as an alternative electron sink in photosynthesis? *Current Opinion in Plant Biology* **5**: 193–198.

Overmyer, K., Brosché, M., & Kangasjärvi, J. 2003. Reactive oxygen species and hormonal control of cell death. *Trends in Plant Science* **8**: 335–342.

Panda, S., & Hogenesch, J. B. 2004. It's all in the timing: Many clocks, many outputs. *Journal of Biological Rhythms* **19**: 374–387.

Panda, S., Antoch, M. P., Miller, B. H., Su, A. I., Schook, A. B., Straume, M., Schultz, P. G., Kay, S. A., Takahashi, J. S., & Hogenesch, J. B. 2002. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* **109**: 307–320.

Para, A., Farré, E., Imaizumi, T., Pruneda-Paz, J., Harmon, F., & Kay, S. 2007. PRR3 is a vascular regulator of TOC1 stability in the *Arabidopsis* circadian clock. *Plant Cell* **19**: 3462–3473.

Park, D. H., Somers, D. E., Kim, Y. S., Choy, Y. H., Lim, H. K., Soh, M. S., Kim, H. J., Kay, S. A., & Nam, H. G. 1999. Control of circadian rhythms and photoperiodic flowering by the *Arabidopsis GIGANTEA* gene. *Science* **285**: 1579–1582.

Park, J., Gu, Y., Lee, Yuree, Yang, Z., & Lee, Youngsook. 2004. Phosphatidic acid induces leaf cell death in Arabidopsis by activating the Rho-related small G protein GTPase-mediated pathway of reactive oxygen species generation. *Plant Physiology* **134**: 129–136.

Pignocchi, C., Kiddle, G., Hernández, I., Foster, S. J., Asensi, A., Taybi, T., Barnes, J., & Foyer, C. H. 2006. Ascorbate oxidase-dependent changes in the redox state of the apoplast modulate gene transcript accumulation leading to modified hormone signaling and orchestration of defense processes in tobacco. *Plant Physiology* **141**: 423–435.

Pittendrigh, C. S. 1993. Temporal organization: reflections of a Darwinian clock-watcher. *Annual Review of Physiology* **55**: 16–54.

Plautz, J. D., Straume, M., Stanewsky, R., Jamison, C. F., Brandes, C., Dowse, H. B., Hall, J. C., & Kay, S. A. 1997. Quantitative analysis of *Drosophila period* gene transcription in living animals. *Journal of Biological Rhythms* **12**: 204–217.

Pnueli, L., Liang, H., Rozenberg, M., & Mittler, R. 2003. Growth suppression, altered stomatal responses, and augmented induction of heat shock proteins in cytosolic ascorbate peroxidase (*Apx1*)-deficient *Arabidopsis* plants. *Plant Journal* **34**: 187–203.

Prasad, T. K., Anderson, M. D., Martin, B. A., & Stewart, C. R. 1994. Evidence for chilling-induced oxidative stress in maize seedlings and a regulatory role for hydrogen peroxide. *Plant Cell* **6**: 65–74.

Proost, S., Van Bel, M., Sterck, L., Billiau, K., Van Parys, T., Van de Peer, Y., & Vandepoele, K. 2009. PLAZA, a comparative genomics resource to study gene and genome evolution in plants. *Plant Cell* **21**: 3718–3731.

Pruneda-Paz, J. L., & Kay, S. A. 2010. An expanding universe of circadian networks in higher plants. *Trends in Plant Science* **15**: 259–265.

Pruneda-Paz, J., Breton, G., Para, A., & Kay, S. 2009. A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* **323**: 1481–1485.

Quail, P. H. 2002. Phytochrome photosensory signalling networks. *Nature Reviews Molecular Cell Biology* **3**: 85–93.

Rand, D., & Shulgin, B. 2004. Design principles underlying circadian clocks. *Journal of the Royal Society Interface* **1**: 119–130.

Rawat, R., Schwartz, J., Jones, M. A., Sairanen, I., Cheng, Y., Andersson, C. R., Zhao, Y., Ljung, K., & Harmer, S. L. 2009. REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proceedings of the National Academy of Sciences U.S.A* **106**: 16883–16888.

Rawat, R., Takahashi, N., Hsu, P. Y., Jones, M. A., Schwartz, J., Salemi, M. R., Phinney, B. S., & Harmer, S. L. 2011. REVEILLE8 and PSEUDO-REPONSE REGULATOR5 form a negative feedback loop within the *Arabidopsis* circadian clock. *PLoS Genetics* **7**: e1001350.

Rentel, M. C., Lecourieux, D., Quaked, F., Usher, S. L., Petersen, L., Okamoto, H., Knight, H., Peck, S. C., Grierson, C. S., Hirt, H., & Knight, M. R. 2004. OXI1 kinase is necessary for oxidative burst-mediated signalling in *Arabidopsis*. *Nature* **427**: 858–861.

Rhee, S. G., Bae, Y. S., Lee, S-R., & Kwon, J. 2000. Hydrogen peroxide: a key messenger that modulates protein phosphorylation through cysteine oxidation. *Science's Signal Transduction Knowledge Environment* **2000**: 1.

Ristilä, M., Strid, H., Eriksson, L. A., Strid, Å., & Sävenstrand, H. 2011. The role of the pyridoxine (vitamin B_6) biosynthesis enzyme PDX1 in ultraviolet-B radiation responses in plants. *Plant Physiology and Biochemistry* **49**: 284–292.

Rizhsky, L., Hallak-Herr, E., Van Breusegem, F., Rachmilevitch, S., Barr, J. E., Rodermel, S., Inzé, D., & Mittler, R. 2002. Double antisense plants lacking ascorbate peroxidase and catalase are less sensitive to oxidative stress than single antisense plants lacking ascorbate peroxidase or catalase. *Plant Journal* **32**: 329–342.

Rizhsky, L., Liang, H., & Mittler, R. 2003. The water–water cycle is essential for chloroplast protection in the absence of stress. *Journal of Biological Chemistry* **278**: 38921–38925.

Rohde, A., Morreel, K., Ralph, J., Goeminne, G., Hostyn, V., De Rycke, R., Kushnir, S., Van Doorsselaere, J., Joseleau, J.-P., Vuylsteke, M., Van Driessche, G., Van Beeumen, J., Messens, E., & Boerjan, W. 2004. Molecular phenotyping of the *pal1* and *pal2* mutants of *Arabidopsis thaliana* reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. *Plant Cell* **16**: 2749–2771.

Rosenwasser, S., Rot, I., Sollner, E., Meyer, A. J., Smith, Y., Leviatan, N., Fluhr, R., & Friedman, H. 2011. Organelles contribute differentially to reactive oxygen species-related events during extended darkness. *Plant Physiology* **156**: 185–201.

Rouhier, N. 2010. Plant glutaredoxins: Pivotal players in redox biology and iron-sulphur centre assembly. *New Phytologist* **186**: 365–372.

Rubinovich, L., & Weiss, D. 2010. The *Arabidopsis* cysteine-rich protein GASA4 promotes GA responses and exhibits redox activity in bacteria and in planta. *Plant Journal* **64**: 1018–1027.

Rutter, J., Reick, M., & McKnight, S. L. 2002. Metabolism and the control of circadian rhythms. *Annual*

Review of Biochemistry **71**: 307–331.

Sadacca, L. A., Lamia, K. A., deLemos, A. S., Blum, B., & Weitz, C. J. 2011. An intrinsic circadian clock of the pancreas is required for normal insulin release and glucose homeostasis in mice. *Diabetologia* **54**: 120–124.

Sagi, M., & Fluhr, R. 2006. Production of reactive oxygen species by plant NADPH Oxidases. *Plant Physiology* **141**: 336–340.

Salomé, P.A., Weigel, D. & McClung, C.R. 2010. The role of the *Arabidopsis* morning loop components CCA1, LHY, PRR7, and PRR9 in temperature compensation. *Plant Cell* **22**: 3650–3661.

Santelia, D., Henrichs, S., Vincenzetti, V., Sauer, M., Bigler, L., Klein, M., Bailly, A., Lee, Youngsook, Friml, J., Geisler, M., & Martinoia, E. 2008. Flavonoids redirect PIN-mediated polar auxin fluxes during root gravitropic responses. *Journal of Biological Chemistry* **283**: 31218–31226.

Sargentm, L., Briggs, W. R., & Woodward, D.O. 1966. Circadian nature of a rhythm expressed by an invertaseless strain of *Neutrospora crassa*. *Plant Physiology* **41**: 1343-1349.

Satter, R. L., Morse, M. J., Youngsook, L., Crain, R. C., Cote, G. G., & Moran, N. 1988. Light and clock controlled leaflet movements in *Samanea saman*: a physiological, biological and biochemical analysis. *Botanica Acta* **101**: 205-213.

Sawa, M., & Kay, S. A. 2011. *GIGANTEA* directly activates *Flowering Locus T* in *Arabidopsis thaliana. Proceedings of the National Academy of Sciences U.S.A* **108**: 11698–11703.

Sawa, M., Nusinow, D. A., Kay, S. A., & Imaizumi, T. 2007. FKF1 and GIGANTEA complex formation is required for day-length measurement in *Arabidopsis*. *Science* **318**: 261–265.

Scarpeci, T. E., Zanor, M. I., Carrillo, N., Mueller-Roeber, B., & Valle, E. M. 2008. Generation of superoxide anion in chloroplasts of *Arabidopsis thaliana* during active photosynthesis: a focus on rapidly induced genes. *Plant Molecular Biology* **66**: 361–378.

Schaffer, R., Ramsay, N., Samach, A., Corden, S., Putterill, J., CarrÈ, I.A., *et al.* 1998. The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* **93**: 1219–1229.

Schippers, J.H.M., Nunes-Nesi, A., Apetrei, R., Hille, J., Fernie, A.R. & Dijkwel, P.P. 2008. The *Arabidopsis onset of leaf death5* mutation of quinolinate synthase affects nicotinamide adenine dinucleotide biosynthesis and causes early ageing. *Plant Cell* **20**: 2909–2925.

Schweiger, H.-G. & Schweiger, M. 1977. Circadian rhythms in unicellular organisms: an endeavour to

explain the molecular mechanism. *International Review of Cytology* **51**: 315-342.

Sehgal, A., Price, J. & Young, M.W. 1992. Ontogeny of a biological clock in Drosophila melanogaster. *Proceedings of the National Academy of Sciences U.S.A.* **89**: 1423–1427.

Shirley, B. 1996. Flavonoid biosynthesis: 'new' functions for an “old” pathway. *Trends in Plant Science* **1**: 377–382.

Smith, A. M., & Stitt, M. 2007. Coordination of carbon supply and plant growth. *Plant, Cell & Environment* **30**: 1126–1149.

Smith, S.M., Fulton, D.C., Chia, T., Thorneycroft, D., Chapple, A., Dunstan, H., Hylton, C., Zeeman, S. C., & Smith, A. M. 2004. Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. *Plant Physiology* **136**: 2687–2699.

Somers, D. E., Schultz, T. F., Milnamow, M., & Kay, S. A. 2000. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell* **101**: 319–329.

Stephan, F. K. 2002. The “other” circadian system: food as a Zeitgeber. *Journal of Biological Rhythms* **17**: 284–292.

Stracke, R., Werber, M., & Weisshaar, B. 2001. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* **4**: 447–456.

Strayer, C., Oyama, T., Schultz, T. F., Raman, R., Somers, D. E., M s, P., Panda, S., Kreps, J. A., & Kay, S. A. 2000. Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* **289**: 768–771.

Sulzman, F.M., Ellman, D., Fuller, C.A., Moore-Ede, M.C. & Wassmer, G. 1984. Neurospora circadian rhythms in space: a reexamination of the endogenous-exogenous question. *Science* **225**: 232–234.

Sweeney, B. M. 1971. Circadian rhythms in unicellular organisms. *Proceedings of the International Symposium on Circadian Rhythmicity*, Wageningen, 136-156.

Sweeney, B. M. & Haxo, F. T. 1961. Persistence of a photo-synthetic rhythm in enucleated *Acetabularia*. *Science* **134**: 1361-1363.

Takeda, S., Gapper, C., Kaya, H., Bell, E., Kuchitsu, K., & Dolan, L. 2008. Local positive feedback regulation determines cell shape in root hair cells. *Science* **319**: 1241–1244.

Terauchi, K., Kitayama, Y., Nishiwaki, T., Miwa, K., Murayama, Y., Oyama, T., & Kondo, T. 2007. ATPase activity of KaiC determines the basic timing for circadian clock of cyanobacteria. *Proceedings of the*

National Academy of Sciences U.S.A **104**: 16377–16381.

Thordal Christensen, H., Zhang, Z., Wei, Y., & Collinge, D. B. 1997. Subcellular localization of H_2O_2 in plants. H_2O_2 accumulation in papillae and hypersensitive response during the barley powdery mildew interaction. *Plant Journal* **11**: 1187–1194.

Tognetti, V. B., Van Aken, O., Morreel, K., Vandenbroucke, K., van de Cotte, B., De Clercq, I., Chiwocha, S., Fenske, R., Prinsen, E., Boerjan, W., Genty, B., Stubbs, K. A., Inzé, D., & Van Breusegem, F. 2010. Perturbation of indole-3-butyric acid homeostasis by the UDP-glucosyltransferase *UGT74E2* modulates *Arabidopsis* architecture and water stress tolerance. *Plant Cell* **22**: 2660–2679.

Tomita, J., Nakajima, M., Kondo, T., & Iwasaki, H. 2005. No transcription-translation feedback in circadian rhythm of KaiC phosphorylation. *Science* **307**: 251–254.

Torres, M. A., & Dangl, J. L. 2005. Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. *Current Opinion in Plant Biology* **8**: 397–403.

Torres, M. A., Jones, J. D. G., & Dangl, J. L. 2006. Reactive oxygen species signaling in response to pathogens. *Plant Physiology* **141**: 373–378.

Tóth, R., Kevei, E., Hall, A., Millar, A.J., Nagy, F. & Kozma-Bognár, L. 2001. Circadian clock-regulated expression of phytochrome and cryptochrome genes in *Arabidopsis*. *Plant Physiology* **127**: 1607–1616.

Tsugane, K., Kobayashi, K., Niwa, Y., Ohba, Y., Wada, K., & Kobayashi, H. 1999. A recessive *Arabidopsis* mutant that grows photoautotrophically under salt stress shows enhanced active oxygen detoxification. *Plant Cell* **11**: 1195–1206.

Tsukagoshi, H., Busch, W., & Benfey, P. N. 2010. Transcriptional Regulation of ROS Controls Transition from Proliferation to Differentiation in the Root. *Cell* **143**: 606–616.

Turek, F. W., Joshu, C., Kohsaka, A., Lin, E., Ivanova, G., McDearmon, E., Laposky, A., Losee-Olson, S., Easton, A., Jensen, D. R., Eckel, R. H., Takahashi, J. S., & Bass, J. 2005. Obesity and metabolic syndrome in circadian Clock mutant mice. *Science* **308**: 1043–1045.

Urban, P., Mignotte, C., Kazmaier, M., Delorme, F., & Pompon, D. 1997. Cloning, yeast expression, and characterization of the coupling of two distantly related *Arabidopsis thaliana* NADPH-cytochrome P450 reductases with P450 CYP73A5. *Journal of Biological Chemistry* **272**: 19176–19186.

van der Horst, G.T., Muijtjens, M., Kobayashi, K., Takano, R., Kanno, S., Takao, M., *et al.* 1999. Mammalian Cry1 and Cry2 are essential for maintenance of circadian rhythms. *Nature* **398**: 627–630.

Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L., & Van de Peer, Y. 2009. Unraveling transcriptional control in *Arabidopsis* using cis-regulatory elements and coexpression networks. *Plant Physiology* **150**: 535–546.

Vanderauwera, S., Suzuki, N., Miller, G., van de Cotte, B., Morsa, S., Ravanat, J.-L., Hegie, A., Triantaphylidès, C., Shulaev, V., Van Montagu, M. C. E., Van Breusegem, F., & Mittler, R. 2011. Extranuclear protection of chromosomal DNA from oxidative stress. *Proceedings of the National Academy of Sciences U.S.A* **108**: 1711–1716.

Velez-Ramirez, A. I., van Ieperen, W., Vreugdenhil, D., & Millenaar, F. F. 2011. Plants under continuous light. *Trends in Plant Science* **16**: 310–318.

Vitaterna, M.H., King, D.P., Chang, A.M., Kornhauser, J.M., Lowrey, P.L., McDonald, J.D., *et al.* 1994. Mutagenesis and mapping of a mouse gene, Clock, essential for circadian behavior. *Science* **264**: 719–725.

Walley, J. W., Coughlan, S., Hudson, M. E., Covington, M. F., Kaspi, R., Banu, G., Harmer, S. L., & Dehesh, K. 2007. Mechanical stress induces biotic and abiotic stress responses via a novel cis-element. *PLoS Genetics* **3**: 1800–1812.

Wang, W., Barnaby, J. Y., Tada, Y., Li, H., Tˆr, M., Caldelari, D., Lee, D., Fu, X. D., & Dong, X. 2011. Timing of plant immune responses by a central circadian regulator. *Nature* **470**: 110–114.

Wang, Z. Y., & Tobin, E. M. 1998. Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. *Cell* **93**: 1207–1217.

Wang, Z., Kenigsbuch, D., Sun, L., Harel, E., Ong, M. S., & Tobin, E. M. 1997. A Myb-related transcription factor is involved in the phytochrome regulation of an Arabidopsis Lhcb gene. *Plant Cell* **9**: 491–507.

Welsh, D. K., Imaizumi, T., & Kay, S. A. 2005. Real-time reporting of circadian-regulated gene expression by luciferase imaging in plants and mammalian cells. *Methods in Enzymology* **393**: 269–288.

Wilkins, M. B. 1959. An endogenous rhythm in the rate of carbon dioxide output of *Bryophyllumn*. I. Some preliminary experiments. *Journal of Experimental Botany* **10**: 377-390.

Wilkins, M. B. 1989. On the mechanism of phase control by light in the rhythm of carbon dioxide output in leaves of *Bryophyllum fedtschenkoi*. *Journal of Experimental Botany* **40**: 1315-1321.

Wijnen, H., & Young, M. W. 2006. Interplay of circadian clocks and metabolic rhythms. *Annual Review of Genetics* **40**: 409–448.

Woelfle, M. A., Ouyang, Y., Phanvijhitsiri, K., & Johnson, C. H. 2004. The adaptive value of circadian clocks an experimental assessment in Cyanobacteria. *Current Biology* **14**: 1481–1486.

Wu, J.-F., Wang, Y., & Wu, S.-H. 2008. Two new clock proteins, LWD1 and LWD2, regulate Arabidopsis photoperiodic flowering. *Plant Physiology* **148**: 948–959.

Yanhui, C., Xiaoyuan, Y., Kun, H., Meihua, L., Jigang, L., Zhaofeng, G., Zhiqiang, L., Yunfei, Z., Xiaoxiao, W., Xiaoming, Q., Yunping, S., Li, Z., Xiaohui, D., Jingchu, L., Xing-Wang, D., et al. 2006. The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Molecular Biology* **60**: 107–124.

Yanovsky, M. J., & Kay, S. A. 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. *Nature* **419**: 308–312.

Yanovsky, M. J., & Kay, S. A. 2001. Signaling networks in the plant circadian system. *Current Opinion in Plant Biology* **4**: 429–435.

Yin, L., Wu, N., Curtin, J. C., Qatanani, M., Szwergold, N. R., Reid, R. A., Waitt, G. M., Parks, D. J., Pearce, K. H., Wisely, G. B., & Lazar, M. A. 2007. Rev-erbα, a heme sensor that coordinates metabolic and circadian pathways. *Science* **318**: 1786–1789.

Zeilinger, M., Farré, E., Taylor, S., Kay, S., & Doyle, F. 2006. A novel computational model of the circadian clock in *Arabidopsis* that incorporates *PRR7* and *PRR9*. *Molecular Systems Biology* **2**: 58.

Zhang, X., Chen, Y., Wang, Z.-Y., Chen, Z., Gu, H., & Qu, L.-J. 2007. Constitutive expression of *CIR1* (*RVE2*) affects several circadian-regulated processes and seed germination in *Arabidopsis*. *Plant Journal* **51**: 512–525.

6. Publication

Overview

This section of the thesis describes an additional bioinformatics-based work undertaken during the course of study. The manuscript, entitled '**Positional Information Resolves Structural Variations and Uncovers an Evolutionarily Divergent Genetic Locus in Accessions of *Arabidopsis thaliana***', was published in the third volume of *Genome Biology and Evolution* in 2011. The primary motivation of this research was to use Next-Generation Sequencing (NGS) tools to unravel genetic blueprints of natural *Arabidopsis* populations. By having such blueprints, one can elucidate past evolutionary events through investigating gene distributions among functional families and the divergence among the proteins encoded by these genes. However, assembling genomes from next-generation datasets is particularly challenging due to the inherent read-length limitations of sequence data generated by NGS platforms. Indeed, this has resulted in the mass release of many low quality draft genomes in public repositories, which may not allow for accurate comparative genomics. This work not only aimed to find solutions to genome assembly challenges but also to understand the effects of novel structural variations, i.e., large nucleotide changes such as duplications, deletions and insertions, in natural plant populations. By comparing a selected genetic locus between two *Arabidopsis* natural accessions, Columbia (Col-0) and Landsberg *erecta* (L*er*-0), structural variations were found to occur in genic and non-genic regions. Ultimately, the key point of this manuscript is that 'sequence read positional information' (reads that belong to a known location) could potentially remedy assembly complications arising from duplication events and aid confident mapping of reads in complex genomic regions.

GENOME BIOLOGY AND EVOLUTION
SMBE

GBE

Positional Information Resolves Structural Variations and Uncovers an Evolutionarily Divergent Genetic Locus in Accessions of *Arabidopsis thaliana*

Alvina G. Lai[1], Matthew Denton-Giles[1], Bernd Mueller-Roeber[2,3], Jos H. M. Schippers[2,3], and Paul P. Dijkwel*[,1]

[1]Institute of Molecular BioSciences, Massey University, Private Bag 11-222, Palmerston North 4442, New Zealand

[2]Department of Molecular Biology, Institute of Biochemistry and Biology, University of Potsdam, 14476 Potsdam-Golm, Germany

[3]Max Planck Institute of Molecular Plant Physiology, 14476 Potsdam-Golm, Germany

*Corresponding author: E-mail: p.dijkwel@massey.ac.nz.

Accepted: 17 April 2011

Abstract

Genome sequencing of closely related individuals has yielded valuable insights that link genome evolution to phenotypic variations. However, advancement in sequencing technology has also led to an escalation in the number of poor quality–drafted genomes assembled based on reference genomes that can have highly divergent or haplotypic regions. The self-fertilizing nature of *Arabidopsis thaliana* poses an advantage to sequencing projects because its genome is mostly homozygous. To determine the accuracy of an *Arabidopsis* drafted genome in less conserved regions, we performed a resequencing experiment on a ~371-kb genomic interval in the Landsberg *erecta* (L*er*-0) accession. We identified novel structural variations (SVs) between L*er*-0 and the reference accession Col-0 using a long-range polymerase chain reaction approach to generate an Illumina data set that has positional information, that is, a data set with reads that map to a known location. Positional information is important for accurate genome assembly and the resolution of SVs particularly in highly duplicated or repetitive regions. Sixty-one regions with misassembly signatures were identified from the L*er*-0 draft, suggesting the presence of novel SVs that are not represented in the draft sequence. Sixty of those were resolved by iterative mapping using our data set. Fifteen large indels (>100 bp) identified from this study were found to be located either within protein-coding regions or upstream regulatory regions, suggesting the formation of novel alleles or altered regulation of existing genes in L*er*-0. We propose future genome-sequencing experiments to follow a clone-based approach that incorporates positional information to ultimately reveal haplotype-specific differences between accessions.

Key words: haplotype, allelic variants, drafted genomes, genome partitioning, comparative genomics.

Introduction

The number of genome projects of various scales has increased substantially over the years due to a reduction in sequencing costs as technology advances (Chain et al. 2009). Many laboratories benefit from this impressive technological advancement in terms of rapid generation of high-depth sequence data. However, next-generation sequencing (NGS) platforms are compromised in their ability to generate long reads. Read length reduction is compensated by an increase in coverage where 20- to 30-fold redundancy has been reported as the acceptable criterion by most genome projects (Bentley et al. 2008; Ossowski et al. 2008). Due to the nature of short-read data sets, drafted genomes are assembled based on preexisting published sequences and the quality of the resulting data has yet been sufficiently diagnosed. This has led to the mass release of drafted genomes (Chain et al. 2009), many of whose qualities are only assessed by identifying the number of assembly gaps. Other valuable diagnostic criteria such as the number of errors and misassemblies are potentially missing and can only be revealed with fine-scale analysis.

Computational algorithms have been developed specifically to tackle short-read data sets (Butler et al. 2008; Ossowski et al. 2008; Zerbino and Birney 2008; Simpson

et al. 2009). The identification of single nucleotide polymorphisms (SNPs; Shen et al. 2010) and small insertion–deletion polymorphisms (indels; Krawitz et al. 2010) using a combination of multiple assembly algorithms that are each designed and optimized for different purposes had seemed to be the end goal of genome projects as other forms of deviations relative to the reference genome remain challenging to detect. Resolving SVs, that is, changes that are not single nucleotide variants, such as duplications, inversions, large indels, and copy number variations (CNV) (Feuk et al. 2006; Frazer et al. 2009), have been proven problematic for short-read assemblers (Snyder et al. 2010). Prior to the arrival of NGS technology, comparative genomic hybridization using oligonucleotide arrays have been extensively used as analysis tools for the discovery of submicroscopic SVs (Sebat et al. 2004; Gresham et al. 2008). Recently, several methods have been developed to detect SVs from NGS data sets (Korbel et al. 2007; Chen et al. 2009; Snyder et al. 2010). The accuracy of these techniques remains to be sufficiently tested particularly on highly complex eukaryotic genomes. An example that can potentially result in assembly error is when a tandem duplication spanning across an inversion allele may be interpreted as a de novo complex duplication if only one inversion haplotype is represented in the reference genome (Zhang et al. 2009). The lack of strategies to transverse across rearrangements and co-occurrences of SVs between chromosomal haplotypes can cause assembly gaps as sequence reads from paralogous regions are mistaken as allelic overlaps when they map to a single location (Bailey et al. 2001; Sharp et al. 2006). This problem further complicates accurate variant calling and may hamper large indel detection in such regions. Improper placement of scaffolds may also introduce nonexistence of heretical evolutionary breakages (Lewin et al. 2009).

Arabidopsis thaliana, a flowering plant from the Brassicaceae family, is one of the best studied plant species due to its tractability and the number of research tools available. The self-compatible nature of *Arabidopsis* has allowed each accession or lineage to evolve independently yielding diverse populations that display a multitude of phenotypic variations (Koornneef et al. 2004). Several groups have embarked on the 1001 *A. thaliana* genome project (Weigel and Mott 2009) dedicated to generate genome sequences from numerous accessions of this species. Comparative genomics have frequently been used as a tool to study evolution by natural selection (Feuillet and Keller 2002; Nishiyama et al. 2003; Bowman et al. 2007; Koonin 2009). By comparing two or more genomes, one can infer how natural selection acts in different lineages in driving sequence evolution in genes and nongenic regions and how these changes relate to phenotypic evolution and adaptation (Ellegren 2008). Investigating patterns of divergence around known functional elements could yield insights on the effect that different forces, for example, purifying selection and genetic hitchhiking (Cai et al. 2009), have on genetic polymorphisms (Altshuler et al. 2010).

It has been reported that approximately one quarter of the *A. thaliana* reference genome involves regions that are highly divergent with the presence of rare alleles in at least one accession (Zeller et al. 2008). Genomic SVs underlie phenotypic differences between *A. thaliana* accessions (Fransz et al. 2000; Meyers et al. 2005; Alonso-Blanco et al. 2009). SVs are predominantly multigenic or even multiloci and may not be represented in the reference accession. The role of SVs in chromosomal speciation has been shown in several models (White 1978), an example being the suppressed-recombination model where a genetic barrier is formed between populations. Substitutions linked to these rearranged chromosomes cannot be exchanged, thereby promoting genomic incompatibilities and hence speciation (Rieseberg et al. 1999; Perry et al. 2008; Bikard et al. 2009; Marques-Bonet et al. 2009; Alcázar et al. 2010). Complex SVs also promote genome instability by long-distance nonallelic homologous recombination leading to further CNV (Johnson et al. 2006). Orthologous regions enriched with ancestral segmental duplications may serve as hot spots for constant genomic turnover, and recurrent CNV genesis happens as a result of evolutionarily shared duplications occurring across and within species (Perry et al. 2008).

The stream of drafted genomes released has far outnumbered the small group of high-quality genomes (Chain et al. 2009). Downstream comparative genomics heavily depends on the fidelity of these drafts. A poor quality draft is therefore prone to misinterpretations (Choi et al. 2008; Meader et al. 2010). Here, we performed a fine-scale assessment of the Landsberg *erecta* (L*er*-0) drafted genome at a selected polymorphic locus. We identified and resolved novel SVs in a contiguous L*er*-0 locus using high-coverage Illumina reads that were generated from an experimental method that incorporates positional information. This work not only highlights the importance of rectifying errors on drafted genomes before they are used in downstream applications but also provides an unprecedented view on genomic divergence in an inbred species. We propose future genome projects to proceed in a manner that incorporates positional information in order to improve genome assembly and to reveal large deviations from reference genomes.

Materials and Methods

Genomic DNA Extraction

Arabidopsis thaliana seed stocks for the L*er*-0 accession were obtained from the Nottingham Arabidopsis Stock Centre (ID: NW20). High-quality genomic DNA suited for long-range (LR)–polymerase chain reaction (PCR) amplification was extracted from 21-day-old frozen leaf material according to the modified method of van der Biezen (van

der Biezen et al. 1996). Four grams of leaf tissue was ground in liquid nitrogen and vortexed in 25 ml chilled extraction buffer (0.35 M sorbitol, 0.1 M Tris–HCl, 5 mM ethylenediaminetetraacetic acid [EDTA], pH 7.5, 20 mM $Na_2S_2O_5$). The crude extract was centrifuged at 14,000 revolutions per minute (rpm) for 1 h at 4 °C, and the supernatant was discarded. A 1.25 ml of extraction buffer, 1.75 ml nucleus lysis buffer (0.2 M Tris–HCl, 50 mM EDTA, 2 M NaCl, 2% hexadecyl-trimethyl-ammonium bromide pH 7.5), and 0.6 ml of 5% sarkosyl were used to dissolve the pellet. The mixture was subsequently incubated for 1 h at 65 °C. Chloroform/isoamylalcohol (24:1 v/v) extraction was performed by adding 7.5 ml of the solvent mixture to the tube, followed by centrifugation at 14,000 rpm for 15 min. Clear supernatant was transferred to a clean tube, and DNA was precipitated with an equal volume of chilled isopropanol and incubated on ice for 20 min before centrifugation at 14,000 rpm for 15 min. The isopropanol was decanted, and the pellet was washed with 70% ethanol and air dried for 20 min. The pellet was dissolved in 500 μl Tris–ethylenediaminetetraacetic acid (TE) buffer containing 10 μl of 10 mg/ml RNaseA. Genomic DNA was stored at 4 °C to prevent multiple freeze-thaw sessions that might hamper LR-PCR amplifications.

LR-PCR Amplification and Illumina Sequencing

Primers for LR-PCR were designed using Primer3 (Rozen and Skaletsky 2000) to amplify overlapping genomic fragments of 647–13,702 bp, spanning an ~371-kb contiguous locus in L*er*-0 (supplementary table 2A, Supplementary Material online). LR-PCR amplifications (milliQ water: 75.6 μl; 10× buffer: 10 μl; deoxyribonucleotide triphosphate [2.5 mM]: 8 μl; forward primer [10 μM]: 2 μl; reverse primer [10 μM]: 2 μl; high-fidelity Takara ExTaq enzyme [5 units/μl]: 0.4 μl; DNA template [90 ng/μl]: 2 μl for 100 μl reaction) were performed using an autosegment extension program (3 min 94 °C/30 s 94 °C, 30 s 62 °C, 5–10 min 68 °C, 30× cycles/5 min 68 °C), increasing the extension time for 15 s each cycle after 14 cycles in the PalmCycler. PCR products were separated on 0.8% (for fragments larger than 10 kb) or 1.0% (for fragments smaller than 10 kb) 1× Tris-acetate-EDTA gel for amplicon size confirmation, followed by purification using the QIAquick PCR Purification Kit. Concentration of each purified PCR product was quantified. L*er*-0 amplicons were pooled in equal molarity to yield DNA in the concentration of 5 μg/50 μl TE. Sequencing was performed on the GAII to generate a 75-bp single-read data set.

Pipeline Analysis and Read Trimming

Illumina Pipeline version 1.6 was used for pipeline analysis. Off-Line Basecaller programs, Firecrest and Bustard, were used for image analysis and base calling, respectively. Approximately 87.9% of clusters passed filtering. The GERALD module in CASAVA 1.6 was used to combine tile-based .qseq files into a single .txt file. File conversion from .qseq to .fastq was done using SSAKE (Warren et al. 2007) qseq2-fastq.pl script. Reads were trimmed according to a Phred score of 20 using the TQSfastq.py script. SSAKE was further utilized to generate de novo contigs under the following parameters—*m*: 15 (minimum number of overlapping bases with the seed during overhang consensus build up) and *x*: 15 (minimum overlap between contigs to merge adjacent contigs in a scaffold).

Detection of Misassemblies and Variant Identification

Trimmed reads were assembled to either Col-0 reference or L*er*-0 draft sequence using Geneious assembler (Drummond et al. 2010) by allowing 4–6 mismatches and 5–50 bp gaps to account for indels. Misassemblies were identified by detecting aberrant assembly signatures in Geneious. Two hundred bases at the left and right flanks of the ambiguous regions were extracted and used as references for targeted iterative read mapping described in the Results section. A SHORE consensus analysis (Ossowski et al. 2008) was performed to obtain GC content and errors in read positions. SNPs were identified using the Find Variations/SNPs option in Geneious by setting the minimum coverage parameter to 100 and minimum variant frequency parameter to 0.8. Locus alignments of Col-0, L*er*-0 draft, and L*er*-0 revised sequences were generated using the progressiveMauve aligner (Darling et al. 2010).

Validation of SVs by Sanger Sequencing

Several resolved indels were randomly chosen for validation by Sanger sequencing. Genomic DNA was PCR amplified using primers designed by Primer3. Both L*er*-0 and Col-0 alleles were amplified, and size differences were visualized on an agarose gel. The L*er*-0 allele was subjected to dideoxy sequencing by the ABI-Sanger instrument followed by alignment of the sequence trace to the iteratively resolved indel for validation purposes.

Data Deposition

L*er*-0_chromosome_3_locus.fasta (GenBank: HQ698308).

Results

LR-PCR Amplification of a Polymorphic L*er*-0 Genomic Interval

An ~371-kb genomic interval on chromosome 3 (Col-0 position: 16653794–17025087) that spans six Col-0 bacterial artificial chromosomes (BACs), that is, F18N11, F9K21, T6D9, F16L2, F12M12, and F18L15, was selected for the study of the prevalence of SVs between a reference (Col-0) and a nonreference (L*er*-0) *Arabidopsis* accession. LR-PCR was used to amplify overlapping genomic fragments

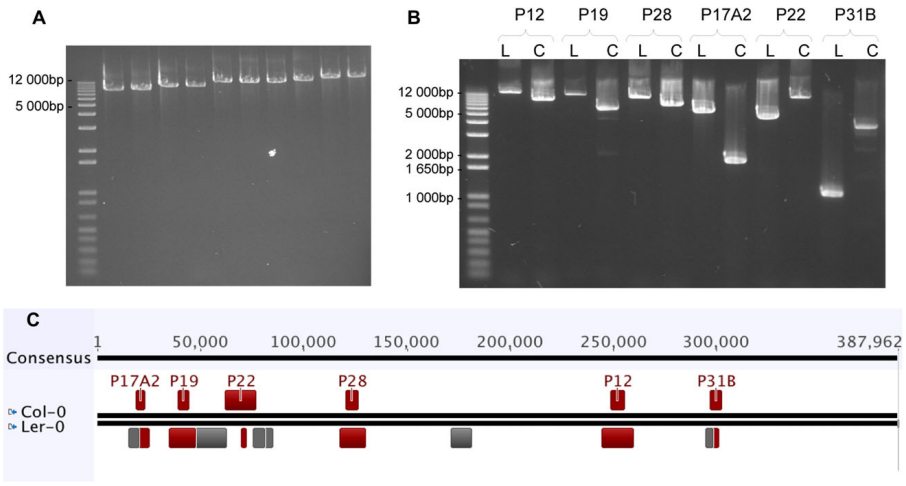


Fig. 1.—LR-PCR amplification and large indel polymorphisms between L*er*-0 and Col-0. L*er*-0 locus that corresponds to 16653794-17025087 positions on Col-0 chromosome 3 is amplified in 49 overlapping fragments using LR-PCR. (*A*) Several examples of LR-PCR amplicons are shown on the gel. (*B*) Gel image depicts large indel polymorphisms between L*er*-0 and Col-0. Amplicon identifier: L, L*er*-0 allele; C, Col-0 allele; P*X*, primer identifier. (*C*) Illustration of locus-specific genetic architecture between L*er*-0 and Col-0. Putative large indels are represented as red blocks and six unamplified gaps as gray blocks.

within the locus with amplicon sizes ranging from 647 to 13,702 bp (fig. 1*A*). LR-PCR was performed in two steps. In the first step, 40 primer pairs were used for amplification, and we were able to obtain 29 out of 40 amplicons. In the second step, an additional 26 primer pairs were designed to divide regions that were not obtained in the first round into 2 or 3 smaller fragments (supplementary table 2A, Supplementary Material online). The second round of amplification is crucial to rule out chances of obtaining no amplicons due to misannealing of the first primer pairs to polymorphic L*er*-0 sites because Col-0 is used as the reference for primer design. From the second round, 20 additional amplicons were obtained. The entire locus is spanned by 49 amplicons, including six gaps that were not covered by PCR (fig. 1*C*). In an attempt to bridge the gaps, additional primers that spanned those gaps were designed. However, we were still unable to obtain any amplicon for the six gaps, suggesting the presence of large insertions in these regions that are beyond amplifiable range, that is, larger than 25 kb (data not shown). The locus of study is partitioned into 49 amplicons that represent genomic fragments obtained from known locations and thus having positional information. Comparison between Col-0 and L*er*-0 amplicon lengths revealed that at least six amplicon pairs harbor large indel polymorphisms (fig. 1*B*). We hypothesized that in addition to these six large indels, a considerable number of indels of significant size remained undetected due to limited gel resolution. Overall, the PCR results suggest that large SVs exist within the selected genomic region between the two *Arabidopsis* accessions.

High-Coverage Sequencing of Amplicons to Detect Interaccession SVs

Illumina Sequencing and Read Mapping to the Col-0 Reference. To precisely capture the sequence context of these SVs, we proceeded to sequence the ~371-kb contiguous locus in L*er*-0 using the Illumina Genome Analyzer II platform. L*er*-0 amplicons were pooled in equal molarity (supplementary fig. 1*A*, Supplementary Material online) and sequenced to generate a 75-bp single-read data set (supplementary fig. 1*B*, *C*, and *D*, Supplementary Material online) with positional information. The filtered and quality-trimmed reads were assembled to the Col-0 reference locus by allowing up to four mismatches and gaps of up to 50 bp to permit small indel detection. Using the Geneious software (Drummond et al. 2010), misassembly signatures, indicative of SVs, were identified. Deletions in L*er*-0 were seen as gaps in the assembly and insertions as arrays of consecutive mismatches (supplementary fig. 2, Supplementary Material online). To locate the region of the six large indels as observed from differences in Col-0/L*er*-0 amplicon

lengths (fig. 1*B*), all primer sequences were aligned to the Col-0 reference. By tracking the flanking primers for the six large indel amplicons, misassembly signatures found at those regions confirmed the occurrences of indels in L*er*-0. In addition, 123 non-SNP misassembly signatures were found (excluding the six unamplified gaps) that corroborated our initial speculation on the presence of additional indels that fall below the range of gel-based detection.

Read Mapping to the Ler-0 Draft. The Wellcome Trust Centre for Human Genetics (WTCHG) has generated an L*er*-0 draft genome from 36- to 51-bp paired-end Illumina libraries of approximately 40-fold coverage. As part of our analysis, we subsequently used the L*er*-0 draft as the reference for read mapping based on the assumption that the draft sequence would be a better reference than Col-0. In parallel, our analysis will also serve as an indicator of L*er*-0 draft sequence quality. Reads were assembled to the L*er*-0 draft by allowing up to six mismatches and 5 bp gaps. The number of non-SNP misassembly signatures was reduced from 123 misassemblies down to 61 misassemblies when the draft sequence was used. However, the large SVs detected from PCR amplicon sizing were not represented in the L*er*-0 draft sequence (table 1). The L*er*-0 draft was generated by a combination of de novo assembly and reference-based mapping. Hence, a large pool of de novo contigs could not be incorporated in the draft due to lack of sequence context from the Col-0 reference and the lack of positional information for these contigs. Therefore, it is expected that SV sequence information remained in the pool of unmapped contigs.

Improving Local Assembly to Reveal SVs Using Targeted Iterative Read Mapping. The PCR-based approach provides us with the information that reads obtained originate from the target locus and not from other genomic regions. We assumed that the pool of unmapped reads (~7%) accounted for SVs. Geneious assembler was used to perform a targeted iterative read-mapping step to map these reads to their designated regions. Each misassembled region was flagged, and their left and right flanking sequences were extracted for iterative mapping. Iterative read mapping consists of the following five steps (fig. 2): 1) Extract 200-bp sequences that flank the misassembled region (these flanks serve as reference sequences for subsequent iterative mapping). 2) Map all reads to both flanks independently. 3) After each round of iteration, reads that assembled to the border of the flank will have sequences extended beyond this flank. The extended sequence is then incorporated to the border of the initial flank to produce a longer flank that is the combination of the initial flank and the assembled read sequence (approximately 45–50 bp for each iteration). Reads are then remapped to the new reference flank. 4) Repeat steps 2 and 3 until the left and right iteratively "extended" flanks overlap and can be aligned. 5) Incorporate the new local consensus sequence into the reference sequence followed by realigning all original reads to the modified reference.

Table 1
L*er*-0 Amplicon Size Estimates Correlate with the Actual Lengths in the L*er*-0 Revised Sequence

Primer ID	Gel-Estimated L*er*-0 Amplicon Length (bp)	Col-0 Length (bp)	L*er*-0 Draft Length (bp)	L*er*-0 Revised Length (bp)
P12	13,500	9,816	9,816	14,526
P19	12,000	6,811	6,989	13,389
P28	11,000	8,683	8,755	12,285
P17A2	7,000	2,084	2,084	6,782
P22	5,000	13,702	13,741	5,691
P31B	1,300	4,282	4,340	1,248

Note.—L*er*-0 amplicon lengths were estimated on an agarose gel, and Col-0 lengths were obtained from TAIR. The corresponding lengths of these amplicons were determined by mapping flanking primer sequences to the L*er*-0 draft and L*er*-0 revised sequence. *PX*, primer identifier.

Manual iterative steps allowed us to pinpoint problematic regions that could not be resolved by automated assembly programs. In a particular iterative step when there was more than one possible read option for subsequent contig extension (fig. 3*A*), we could not proceed onto the next iteration. Instead, iterative read mapping was performed from the opposite flank until it could be aligned to the previous flank. Regions were flagged as unresolved when more than one read option was obtained from both left and right flank extensions, as selecting any one of these possible read options would ultimately result in an incorrect final consensus sequence. This step is crucial to prevent the generation of incorrect chimeric contigs that occur when attempting to assemble duplicated or conserved regions. By referring to each amplicon size, the newly assembled sequence can be cross-checked with the estimated PCR product length.

Out of the 61 misassembled regions, 57 were resolved (supplementary table 1, Supplementary Material online) by local iterative mapping, whereas the remaining four regions could not be confidently determined. For the first three regions, more than one option in iterative extensions from both flanks was present (Fig. 3*A*). Nevertheless, initial iterative results suggested the presence of duplications in these regions. We subsequently attempted to resolve these regions by making use of de novo contigs generated by SSAKE (Warren et al. 2007) using only unmapped reads. In the first two regions, a single de novo contig mapped to each of the corresponding iterative flanks. The contigs were incorporated into the flanks, and iterative mapping was performed to validate the contig sequence. In the third region, more than one contig mapped to the flanks, and the correct one could therefore not be confidently identified without further analysis. Thus, 2 out of the 3 regions were resolved by de novo contig mapping combined with an iterative validation step. In the fourth region, we encountered stretches of long CT-AG inverted repeat sequences from

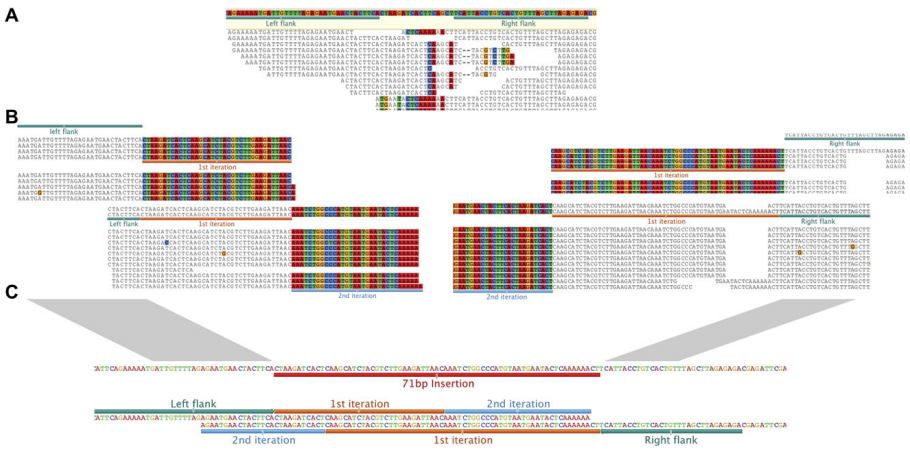


Fig. 2.—Draft sequence correction by iterative read mapping. (*A*) An insertion site is identified by detecting misassembly patterns as described in supplementary figure 2 (Supplementary Material online). Left and right sequences that flank the incorrect region on the draft are used as references for local iterative read mapping. (*B*) In this particular case, two rounds of iterative mapping from both flanks are sufficient to span the insertion. (*C*) Alignment between iteratively extended left and right flanks.

both left and right flank directions (fig. 3*B*). This region is estimated to be 2 kb in length by cross-checking to its corresponding amplicon size (P19 in table 1). An ~1-kb de novo contig flanked by CT and AG sequences was identified and was confirmed to be present within the region by restriction digestion on the PCR amplicon from this region. Because the CT-GA repeats extended beyond the read length (reads that consist entirely of these dinucleotide repeat sequences were identified), the actual length of the repeats could not be deduced. Nevertheless, the results suggest that the total length of the combined CT and AG repeats is close to 1 kb. Repeat expansion has been found

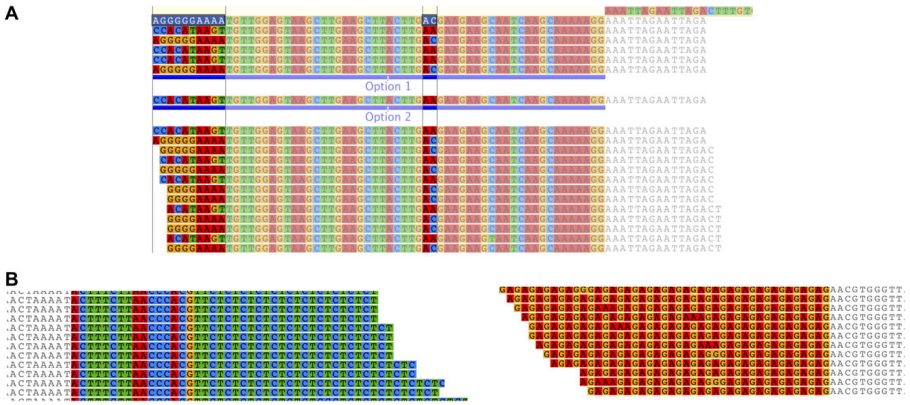


Fig. 3.—Limitations of iterative read mapping. (*A*) Figure illustrates more than one possible read option obtained during iterative mapping. Iterative extension is then performed from the opposite flank to prevent the generation of chimeric contigs. (*B*) Figure depicts a stretch of long inverted dinucleotide repeat in *Ler*-0 that is absent from the Col-0 genome. Further iterative steps are not possible in this region as repeat length is longer than the read length. This region is estimated to be 2 kb in length based on PCR amplicon size.

Table 2
Comparative Analysis of Col-0, Ler-0 Draft (WTCHG), and Ler-0 Revised Sequence Using Locus-Specific and Whole-Genome Ler-0 Reads

	No. of Aligned Locus-Specific Ler-0 Reads (Mean Coverage)	No. of Aligned WTCHG Whole-Genome Ler-0 Reads (Mean Coverage)
Col-0	2,002,286 (375.4)	161,312 (16.1)
Ler-0 draft (WTCHG)	3,096,868 (595.5)	210,349 (21.9)
Ler-0 revised	3,432,240 (643.6)	220,178 (22.5)

NOTE.—Locus-specific reads and whole-genome reads are aligned to the Ler-0 draft and revised sequences.

to cause environment-dependent genetic defects in *Arabidopsis* (Sureshkumar et al. 2009). Interestingly, TAIR Blast (http://arabidopsis.org/Blast/index.jsp) revealed that this stretch of long inverted repeats (CT and GA) is not found anywhere in the Col-0 genome, hence not represented in the Ler-0 draft sequence either. In total, 60 (98.4%) out of 61 misassembled sites were resolved to generate a revised Ler-0 sequence of 375,893 bp in length. The largest insertion and deletion resolved by iterative read mapping were 4,819 and 5,139 bp, respectively. By accounting for the size of the six unamplified gaps, the Ler-0 locus was estimated to be considerably larger than its Col-0 counterpart.

To evaluate the accuracy of the Ler-0 revised sequence, locus-specific reads were mapped to all three sequences (Col-0, Ler-0 draft, and Ler-0 revised) using the most stringent parameters (no gaps and no ambiguities were allowed). Because only reads that have no errors were included, the mean coverage decreased from ~930-fold (when one error is allowed) to ~643-fold (only perfect reads allowed). The same process was repeated using Ler-0 whole-genome reads from WTCHG. In comparison to the Ler-0 draft, the Ler-0 revised sequence is a better reference (table 2). From the stringent alignment of locus-specific reads to the Ler-0 revised sequence, seven gaps were identified (six gaps corresponding to unamplified regions and one gap to the aforementioned unresolved region). Similarly, seven gaps were present when Ler-0 whole-genome reads were aligned to the revised sequence. However, size differences were observed in the gaps when either locus-specific reads or whole-genome reads were used. This is due to the fact that PCR primers were designed to amplify regions that are spanned by the forward and reverse primers. However, variations in Ler-0 do not always start and end at the primer-binding sites. The absence of misassembly signatures overall demonstrates that the Ler-0 revised sequence is superior to the draft sequence. Moreover, Sanger sequencing on 16 random corrections subsequently confirmed that all were

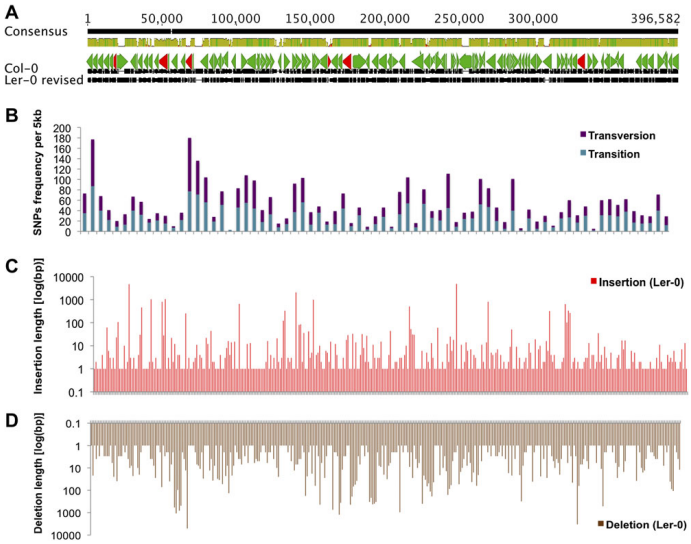


FIG. 4.—Schematic diagram of polymorphisms on the selected Ler-0 locus. Variations between Ler-0 and Col-0 are indicated on the diagram. (*A*) Figure illustrates the pairwise alignment between Ler-0 and Col-0. TAIR10 annotated genes (green arrows) and transposable element genes (red arrows) are indicated, respectively. Detailed representations of the variations between Ler-0 and Col-0, (*B*) SNPs, (*C*) insertions, and (*D*) deletions, are indicated in the diagram.

Table 3
Large Indels that Overlap Genes and Regulatory Regions

Figure ID	Col-0 Gene	Gene Description[a]
Figure 5*A*	(TAIR:At3G45500)	RING/U-box protein with C6HC-type zinc finger
Figure 5*B*	(TAIR:At3G45490)	RING/U-box superfamily protein
Figure 5*C*	(TAIR:At3G45840)	Protein binding/zinc ion binding
Figure 5*D*	(TAIR:At3G45955)	tRNA-Val
Figure 5*E*	(TAIR:At3G46110)	Unknown protein
Figure 6*A*	(TAIR:At3G45990)	Cofilin/tropomyosin-type actin-binding protein
Supplementary figure 4*A* (Supplementary Material online)	(TAIR:At3G46060)	Small GTP-binding protein
Supplementary figure 4*B* (Supplementary Material online)	(TAIR:At3G45910)	Unknown protein
Supplementary figure 4*C* (Supplementary Material online)	(TAIR:At3G45540)	RING/U-box protein with C6HC-type zinc finger
	(TAIR:At3G45550)	Non-LTR retrotransposon family (LINE)
	(TAIR:At3G45555)	Zinc finger (C3HC4-type RING finger) family protein
Supplementary figure 4*D* (Supplementary Material online)	(TAIR:At3G45750)	Nucleotidyltransferase family protein
	(TAIR:At3G45755)	Transposable element gene
	(TAIR:At3G45760)	Nucleotidyltransferase family protein
Supplementary figure 4*E* (Supplementary Material online)	(TAIR:At3G45673)	Unknown protein

NOTE.—LTR, long terminal repeat.
[a]Information obtained from the TAIR10 genome annotation

accurate (supplementary table 2B, Supplementary Material online).

To investigate the feasibility of our method for low-coverage whole-genome data sets, targeted iterative read assembly was performed on a random unmapped contig obtained from WTCHG's L*er*-0 N50 de novo contigs. Using the L*er*-0 whole-genome reads from WTCHG that has a modest coverage of 40-fold, a selected 478-bp contig was iteratively extended to a 2,091-bp sequence. This sequence was validated by Sanger sequencing (Lai AG, Dijkwel PP, unpublished data) and does not align to any region of the L*er*-0 draft, suggesting that it is present within a haplotype-specific insertion in L*er*-0. In an attempt to fill in the six unamplified gaps in the locus of interest, iterative read mapping was done using WTCHG L*er*-0 whole-genome reads as well as the de novo contigs. However, we were mostly unsuccessful for several reasons. The relatively low-coverage data set along with the lack of read positional information did not allow accurate iterative mapping particularly when the region is duplicated or is highly repetitive. Furthermore, because of the lack of positional information, the correct de novo contig that aligns to the border of the gap could not be selected when there is more than one possible match.

The previously predicted SVs were resolved by iterative read mapping using a high-coverage data set aided by PCR-based positional information. In total, 31 large indels (>100 bp), 52 smaller (<100 bp) indel-like misassemblies, and 722 novel SNPs that were not present in the L*er*-0 draft sequence were identified. On average, one SNP per 97 bp (10 SNPs/kb) and one indel per 507 bp (2 indels/kb) were detected between L*er*-0 and Col-0. Novel variations identified from this study were not represented in the L*er*-0 draft presumably because they occurred in duplicated or highly conserved regions where these regions can hamper accurate variant calling. Alignment between the L*er*-0 and Col-0 loci yielded a pairwise identity of 84.2%. In addition, we provide a snapshot of variations between L*er*-0 and Col-0 at this small genomic interval (fig. 4, supplementary fig. 3 and table 1, Supplementary Material online).

Biological and Evolutionary Significance of SVs

We next determined whether the SVs could have effects on genes. According to TAIR10 annotation, the 371-kb locus on chromosome 3 comprises 102 genes, 3 transfer RNA genes, and 6 transposable element genes (fig. 4*A*). Fifteen novel large indels were found to be present within genes and regulatory regions (table 3). These large indels are grouped into three categories: 1) SVs that alter predicted open reading frames, 2) SVs located in regulatory regions, and 3) SVs affecting clusters of genes with similar functions (fig. 5; supplementary fig. 4, Supplementary Material online).

In the first category, SVs were found to either disrupt genes or, as observed in several cases, predicted to produce new transcripts. Figure 5*A* depicts a copia-like retrotransposon insertion within the second intron of (TAIR:At3G45500), whereas in figure 5*B*, a transposon insertion before the first exon of (TAIR:At3G45490) is illustrated. In another example, an 812-bp deletion was identified in a 3.4-kb Col-0 cofilin/tropomyosin-type actin-binding gene (TAIR:At3G45990) (fig. 6*A*). Interestingly, TAIR10 Gbrowse (http://gbrowse.arabidopsis.org/cgi-bin/gbrowse/arabidopsis/) revealed that no expressed sequence tag was found for this gene. A gene prediction program (Stanke and Morgenstern 2005) predicted a 1.1-kb gene from the revised L*er*-0 sequence that was subsequently validated by PCR amplification and Sanger sequencing. TAIR BLASTP results of the putative L*er*-0 allele suggest that it is an *ACTIN-DEPOLYMERIZING FACTOR 4*–like gene (fig. 6*B*). In addition, insertions within

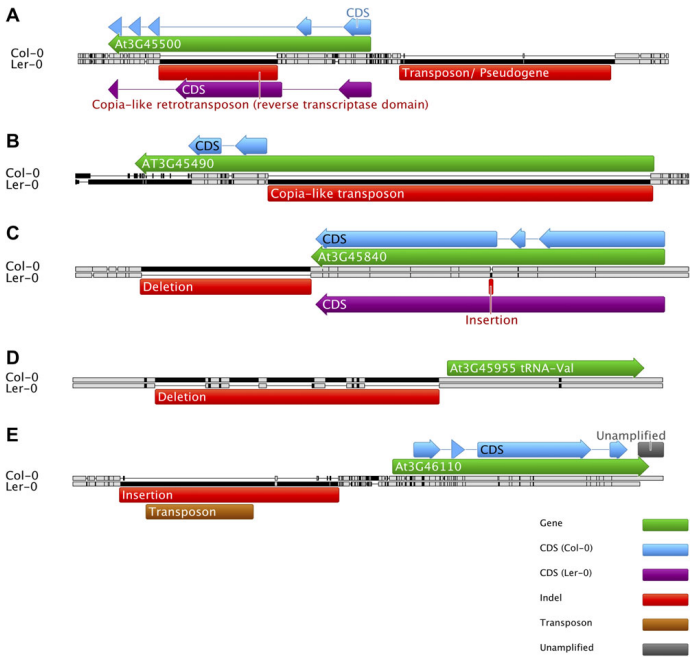


Fig. 5.—SVs that overlap genes. (*A* and *B*) depict copia-like retrotransposon sequences inserted in the corresponding Ler-0 allele. (*C*) A 10-bp insertion within a gene resulted in an inferred intronless transcript variant. (*D*) illustrates a deletion and (*E*) a transposon-like insertion in regulatory regions. Augustus program (Stanke and Morgenstern 2005) is used to predict coding sequences (CDS) of the Ler-0 alleles.

genes can also lead to the formation of inferred new transcripts, for example, an intronless variant (fig. 5*C*). A further noteworthy observation is the insertion in the third intron of At3G46060 (supplementary fig. 4*A*, Supplementary Material online) encoding a GTP-binding protein involved in ethylene signaling (Zimmerli et al. 2008).

SVs occurring in regulatory regions can influence gene expression through numerous positional effects (Feuk et al. 2006). Deletion of regulatory elements (fig. 5*D*) or insertion within such elements (fig. 5*E*) might affect expression of the immediate downstream gene and also the successive gene if both genes share the same *cis*-regulatory

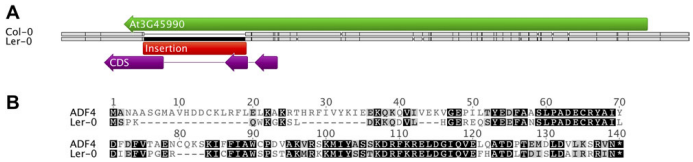


Fig. 6.—Large deletion within a putative Col-0 gene suggests the formation of a novel allelic variant in Ler-0. (*A*) An 812-bp Col-0 deletion is found to be located within a cofilin/tropomyosin-type actin-binding gene (TAIR:At3G45990). Augustus program predicted a 1.1-kb gene model from the Ler-0 allele, which differs from the 3.4-kb Col-0 gene. (*B*) BLASTP revealed that the protein encoded by the Ler-0 allele has 45% pairwise amino acid identity to known ACTIN-DEPOLYMERIZING FACTOR 4 (ADF4) protein encoded by (TAIR:At5G59890). Identical amino acid motifs are highlighted in black and similar motifs in gray.

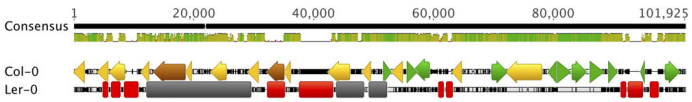


Fig. 7.—Zinc-binding protein gene cluster is enriched in SVs. Diagram illustrates the presence of large indels (red blocks) in a region where a cluster of zinc-binding protein genes (yellow arrows) is located. Gray blocks, brown arrows, and green arrows indicate unamplified regions, transposable element genes, and other protein-coding genes, respectively.

elements (Cordaux and Batzer 2009). In the third category, a high number of SVs were found in a region enriched with genes that encode zinc-binding proteins (fig. 7). Col-0 has two transposable element genes within this region, and we hypothesize that additional transposons are present in the unamplified gaps (data not shown).

Our findings confirm that transposable elements do not merely cause genetic perturbations; they participate in gene regulatory networks in ways that SNPs could not achieve (Heard et al. 2010). In this work, fundamental challenges in SV detection were tackled using an LR-PCR–based resequencing approach that yielded valuable read positional information. Genome assemblies could be improved to show SVs if the experiment is planned in a way that incorporates positional information to the reads. This work emphasizes the importance of detecting SVs as they can have significant implications on downstream biological inferences, particularly on the identification and the study of evolutionarily shared allelic variants.

Discussion

Many agree that the real excitement in whole-genome–sequencing experiments only starts when another genome of a closely related individual is sequenced (Ossowski et al. 2008; Hoberman et al. 2009; McKernan et al. 2009). The human 1000 genomes project (Collins et al. 2003) and the *Arabidopsis* 1001 genomes project (Weigel and Mott 2009) are two examples of joint international collaborations to create a catalogue of intraspecific genetic variations. Together with the rapid advancement in NGS technology and the reduction in sequencing costs, there has been a massive proliferation in the number of drafted genomes produced. However, the inability of current assembly programs to address problematic areas has resulted in the generation of many poor quality drafts (Chain et al. 2009). Capturing large genomic SVs has been particularly challenging (Chen et al. 2009; Kidd et al. 2010).

In an attempt to identify problematic regions and find methods for improving draft genomes, we performed a resequencing experiment at a selected genomic interval of the *Arabidopsis* L*er*-0 accession. Fine-scale sequence analysis at this target locus suggests that *A. thaliana* L*er*-0 and Col-0 genomes are highly variable. From our analysis, it was observed that the L*er*-0 draft sequence accurately incorporates Col-0/L*er*-0 polymorphisms if they are short in length and/or located in regions that are not conserved, duplicated, or repetitive. On the contrary, large SVs that lie in conserved, duplicated, or repetitive regions such as variations in gene families and transposon-like indels were not incorporated in the L*er*-0 draft. Nevertheless, those SVs may affect gene integrity and expression. Over 700 indels (supplementary table 1, Supplementary Material online) between L*er*-0 and Col-0 and 15 large indels (figs. 5, 6, and 7; supplementary fig. 4, Supplementary Material online) present in genes and regulatory regions were identified. Seven of these indels involve transposon-like sequences. Although once thought to be "junk" DNA, an increasing number of studies have shed new light on the functional role of these jumping genes (Lippman et al. 2004; Wheelan et al. 2005). Transposons represent a dynamic portion of genomes, where some can mediate rearrangements of adjacent DNA (Bennetzen 2005), present new regulatory effects on nearby genes (Michaels et al. 2003; Blewitt et al. 2005; Weil and Martienssen 2008; Lisch 2009), and contribute to gene expression divergence between closely related species (Hollister et al. 2011). The presence of transposons could also affect recombination in adjacent genes by heterochromatic effects (He and Dooner 2009).

Our results also suggest the occurrence of a potential synteny break (Al-Shahrour et al. 2010) between L*er*-0 and Col-0 within a zinc-binding protein gene cluster (fig. 7). Ten large indels that include three transposon-like insertions, one transposon deletion in L*er*-0, and three unamplified gaps further imply that this neighborhood has been dynamically reorganized in L*er*-0. Indeed, functional clusters in mammals are significantly enriched by SINE elements as they contribute to the rearrangement process (Zhao et al. 2004). The prevalence of SVs in genic regions can potentially lead to the formation of natural allelic variants or alter gene expression and function altogether. Thus, it is imperative for drafted genomes to incorporate SVs in both coding and noncoding regions so that accurate biological and evolutionary inferences can be drawn from comparative genomics studies on closely related individuals.

Positional Information Allows Correct Assembly of SVs

Whole-genome sequencing is now a routine practice, thanks to the advancements in sequencing technology.

Assembling large and complex genomes is unfortunately a less straightforward task. For example, it is particularly challenging to deduce large insertions in nonreference accessions, variations within conserved or duplicated regions, and variations in microsatellite repeat lengths. Moreover, if the reference accession has a reduced genome (Schmuths et al. 2004), it can significantly impair insertion-based SV detection in nonreference accessions (supplementary fig. 5, Supplementary Material online). Using a combination of wet lab and dry lab approaches, we demonstrated the feasibility in resolving regions that have marked deviations from the reference genome. Amplicon size information was employed to identify the location of large SVs. Once the approximate location was identified, it can be narrowed down to the point where the variation starts by looking for misassembly signatures. Local iterative read mapping was then performed to resolve the variation in question, and the length of the newly deduced sequence was then compared with its respective amplicon size. Algorithms for iterative gap closure have been described elsewhere (Tsai et al. 2010). However, these algorithms detect gaps in assemblies and are not suitable for insertions that do not manifest as assembly gaps (supplementary fig. 2*B*, Supplementary Material online).

Conserved or duplicated regions can affect variant detection, for example, large deletions in conserved regions, transposon-like indels, and polymorphisms within gene families. Santuari and colleagues have recently demonstrated the combined use of tiling array hybridizations with NGS to detect large deletions by identifying regions that have weak hybridization signals along with the absence of short reads (Santuari and Hardtke 2010; Santuari et al. 2010). Here we show that fine-scale manual inspection can resolve regions that are conserved, duplicated, or repetitive. Information contained in a single read is significantly limited by its length and can result in ambiguous placement of reads to homologous regions (Young et al. 2010). Aberrant alignments of homologous reads may inflate the number of false-positive detections (Pool et al. 2010). In particular, we observed ambiguous placements of transposons in the L*er*-0 draft. Although deletions are easier to detect, we have nevertheless identified large deletions absent from the L*er*-0 draft. Deletions that lie in conserved regions will be missed (false negatives) as reads from homologous regions can map to the reference sequence although it is not present in the study accession. Because our work was targeted to a specific locus, regions that are duplicated elsewhere in the genome will not interfere with the iterative mapping step, unless a particular region is duplicated within the locus itself. Therefore, we emphasize the importance of having positional information that assists sorting of reads to their respective locations and allows the resolution of duplications independently without interference from other homologous sequence reads.

Previously, an indel prediction has been performed using the 2-fold redundant L*er*-0 shotgun contigs generated by Cereon Genomics (Ziolkowski et al. 2009). Thirteen out of the 19 predicted indels that fall within the locus of interest were found to be false positives, the largest being a 7.8-kb insertion. The high rate of false-positive predictions can be attributed to the assignment of incorrect chimeric Cereon contigs (Lai AG, Dijkwel PP, unpublished data) that have partial sequence homology to a particular region. The incorrect placement of contigs is therefore exacerbated by the absence of positional information.

Another challenge in whole-genome assembly is the accurate deductions of microsatellite repeat lengths from short-read data sets (supplementary table 1E, Supplementary Material online). In theory, paired-end mapping should mitigate this problem if the gap spanned by the paired reads is larger than the repeat itself. Most paired-end libraries, however, lack sufficient coverage to enable reliable sequence predictions (Schatz et al. 2010). Therefore, positional information is useful for the sorting of repetitive sequences to their respective genomic locations. Furthermore, accurate deduction of repeat length is crucial in order to reveal rare allelic variants (Sureshkumar et al. 2009). In short, significant progress can be made on genome assembly if the experimental design prior to sequencing is modified such that positional information is incorporated into data sets.

Genes Associated with SVs May Evolve New Functions

The organization of SVs has two implications on genome evolution. First, structural changes can be observed in regions that have high rates of evolutionary turnover and second it allows genes that are duplicated or transposed to new chromosomal regions to be free from selective constraints and evolve independently, giving rise to genes with altered functions or altered regulation (Samonte and Eichler 2002). The most common type of SVs that affect genes are segmental duplications where a likely outcome would be the accumulation of partial gene structures or pseudogenes (Lynch and Conery 2000; Zhang 2003). These paralogous genomic copies have been often treated as "dead on arrival." Recent studies on whole-genome tiling analysis, however, revealed that pseudogenes can be expressed (Akama et al. 2009). Expressed pseudogenes also play a role in the regulation of the messenger RNA stability of its homologous coding gene (Hirotsune et al. 2003). Gene density greatly correlates with segmental duplication density and in comparison to unique genes; genes in segmental duplicated regions are more likely to display inter- and intraspecific CNV (Tuzun et al. 2005) along with signatures of positive selection (Johnson et al. 2001; Birtle et al. 2005). Although genes affected by SVs are most likely associated with subtle

phenotypic alterations due to selective constraints, they can nevertheless have an influence on the phenotype by altering gene dosage (Sharp et al. 2006). Genes involved in environmental interaction and host defense have been found to be enriched with SVs (Emes et al. 2003; Tuzun et al. 2005). Examining structurally dynamic regions of the genome may provide clues on lineage-specific adaptation patterns (Emes et al. 2003; Sharp et al. 2006) that are under diversifying positive selection pressure.

Harnessing Positional Information to Boost Comparative Genomics

Our work suggests that positional information is important for obtaining reliable ordering of scaffolds on chromosomes and improving genome assembly to unveil dynamic genome architectures. Likewise, the development of high-resolution physical maps (Lewin et al. 2009) are indispensible to the ordering of contigs in whole-genome alignments and also for the discovery of evolutionary break point regions based on comparative physical maps (Larkin et al. 2009). A comparison between two forms of genome assembly, that is, hierarchical sequencing of large insert clones and whole-genome shotgun sequence assembly (WGSA) of reads, revealed that the WGSA method yields a 20-Mb shorter sequence than the clone-based assembly (Marques-Bonet et al. 2009). Length discrepancy is caused by the failure of many whole-genome shotgun reads to map to a locus containing a highly duplicated and rapidly evolving gene family (Johnson et al. 2006). This problem will be further aggravated when significantly shorter NGS reads are used (Marques-Bonet et al. 2009).

A fail proof method that accurately detects SVs is still potentially missing. We envisage genome-sequencing experiments to proceed in a clone-based manner that allows the incorporation of positional information to the generated reads. This technique is comparable to the "first-map, then sequence" strategy that uses a BAC-based scaffolding method (Kuhl et al. 2010), which has been successfully implemented in various sequencing projects (Fujiyama et al. 2002; Larkin et al. 2009; Lewin et al. 2009). Construction of large DNA insert libraries will be useful for genome-sequencing projects. This form of genome partitioning will undoubtedly require more work than generating reduced representation libraries from restriction digestions (Young et al. 2010). Although reduced representation libraries can simplify assembly and potentially yield larger contigs, it lacks the positional information required to tease out duplicated regions.

With the current capacity, an entire genome can be sequenced on a single flow cell by making pools of large insert clones and subsequently multiplexing these pools. These reads will have positional information and can then be assigned to their corresponding genomic intervals where de novo contigs can subsequently be generated from these region-specific reads. Using the combinatorial pooling and multiplexing strategy, tens of thousands of different samples can be analyzed with only several hundred appended barcodes (Erlich et al. 2009). Different levels of multiplexing can also be performed to achieve the desired resolution based on resource availability (Wood et al. 2010). With such positional information available, it is possible to elucidate more complex forms of polymorphisms that include segmental duplications, transversions, and transposition events. The size of each clone can be used to validate the accuracy of the assembled contigs. Indeed, several groups have started to follow the clone-based sequencing approach at a low-resolution scale in order to capture a more representative depiction of large intraspecific variations (Kidd et al. 2008; Hurwitz et al. 2010).

NGS platforms have been widely used in targeted resequencing experiments on selected genomic intervals (Martinez Barrio et al. 2009; Turner et al. 2010). Resequencing experiments are often required for the study of intraspecific polymorphisms in regions suspected to host vast amounts of variations. Discovering beneficial or heterotic genetic traits in crop species is primarily performed using a quantitative trait loci (QTL) mapping strategy. Because reference genomes per se may not contain the locus of interest, our approach can successfully identify such SV-related QTL. A majority of drafted genomes fail to provide sufficient granularity for comparative genomics in this sense. Hence, most studies on haplotypic variants still rely on clone-based Sanger shotgun sequencing (Alcázar et al. 2009; Heuer et al. 2009). An undistorted view of data quality is important for end users; hence, each genome should be independently assembled to reveal haplotypic differences. Furthermore, the accuracy of downstream gene annotations relies on the fidelity of the initial assembly. The resulting annotations will not only mislead end users but also defy the initial justification of comparative genomics. Elucidation of complex and dynamic regions of the genome should be the end goal of NGS projects apart from cataloguing small variations such as SNPs. The full benefit of comparative genomics can only be realized when high-quality genome sequences are available.

Supplementary Material

Supplementary figures 1–5 and tables 1 and 2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Professor Richard Mott of the WTCHG for granting access to the prereleased L*er*-0 draft and sequence, whose work was supported by the Biotechnology and Biological Sciences Research Council (BB/F022697/1). We thank Massey Genome Service staff, Lorraine Berry for performing the Illumina sample preparation and sequencing run, Maurice

Collins for running the Illumina pipeline analysis, and Dr Murray Cox for critically reviewing the manuscript. This work was supported by the Massey University Research Fund to P.P.D., Institute of Molecular Biosciences PhD studentship to A.G.L., and the Federal Ministry of Education and Research for Research Units for Systems Biology (BMBF FORSYS) Systems Biology Research Initiative Funding to B.M.-R. (FKZ 0313924).

Literature Cited

Akama T, et al. 2009. Whole-genome tiling array analysis of *Mycobacterium leprae* RNA reveals high expression of pseudogenes and noncoding regions. J Bacteriol. 191:3321–3327.
Al-Shahrour F, et al. 2010. Selection upon genome architecture: conservation of functional neighborhoods with changing genes. PLoS Comput Biol. 6:e1000953. doi:1000910.1001371/journal.pcbi.1000953
Alcázar R, Garcia AV, Parker JE, Reymond M. 2009. Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. Proc Natl Acad Sci U S A. 106:334–339.
Alcázar R, et al. 2010. Natural variation at Strubbelig Receptor Kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. Nat Genet. 42:1135–1139.
Alonso-Blanco C, et al. 2009. What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell. 21:1877–1896.
Altshuler DL, et al. 2010. A map of human genome variation from population-scale sequencing. Nature. 467:1061–1073.
Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE. 2001. Segmental duplications: organization and impact within the current human genome project assembly. Genome Res. 11:1005–1017.
Bennetzen JL. 2005. Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev. 15:621–627.
Bentley DR, et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 456:53–59.
Bikard D, et al. 2009. Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. Science. 323:623–626.
Birtle Z, Goodstadt L, Ponting C. 2005. Duplication and positive selection among hominin-specific *PRAME* genes. BMC Genomics. 6:120.
Blewitt ME, et al. 2005. An *N*-ethyl-*N*-nitrosourea screen for genes involved in variegation in the mouse. Proc Natl Acad Sci U S A. 102:7629–7634.
Bowman JL, Floyd SK, Sakakibara K. 2007. Green genes—comparative genomics of the green branch of life. Cell. 129:229–234.
Butler J, et al. 2008. ALLPATHS: de novo assembly of whole-genome shotgun microreads. Genome Res. 18:810–820.
Cai JJ, Macpherson JM, Sella G, Petrov DA. 2009. Pervasive hitchhiking at coding and regulatory sites in humans. PLoS Genet. 5:e1000336. doi:1000310.1001371/journal.pgen.1000336
Chain PSG, et al. 2009. Genome project standards in a new era of sequencing. Science. 326:236–237.
Chen K, et al. 2009. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods. 6:677–681.
Choi JH, et al. 2008. A machine-learning approach to combined evidence validation of genome assemblies. Bioinformatics. 24:744–750.
Collins FS, Morgan M, Patrinos A. 2003. The human genome project: lessons from large-scale biology. Science. 300:286–290.
Cordaux R, Batzer MA. 2009. The impact of retrotransposons on human genome evolution. Nat Rev Genet. 10:691–703.
Darling A, Mau B, Perna N. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 5:e11147. doi:11110.11371/journal.pone.0011147
Drummond A, et al. 2010. Geneious v5.1. [cited 2011 Feb]. Available from: http://www.geneious.com/.
Ellegren H. 2008. Comparative genomics and the study of evolution by natural selection. Mol Ecol. 17:4586–4596.
Emes RD, Goodstadt L, Winter EE, Ponting CP. 2003. Comparison of the genomes of human and mouse lays the foundation of genome zoology. Hum Mol Genet. 12:701–709.
Erlich Y, et al. 2009. DNA Sudoku-harnessing high-throughput sequencing for multiplexed specimen analysis. Genome Res. 19:1243–1253.
Feuillet C, Keller B. 2002. Comparative genomics in the grass family: molecular characterization of grass genome structure and evolution. Ann Bot. 89:3–10.
Feuk L, Carson AR, Scherer SW. 2006. Structural variation in the human genome. Nat Rev Genet. 7:85–97.
Fransz PF, et al. 2000. Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. Cell. 100:367–376.
Frazer KA, Murray SS, Schork NJ, Topol EJ. 2009. Human genetic variation and its contribution to complex traits. Nat Rev Genet. 10:241–251.
Fujiyama A, et al. 2002. Construction and analysis of a human-chimpanzee comparative clone map. Science. 295:131–134.
Gresham D, Dunham MJ, Botstein D. 2008. Comparing whole genomes using DNA microarrays. Nat Rev Genet. 9:291–302.
He L, Dooner HK. 2009. Haplotype structure strongly affects recombination in a maize genetic interval polymorphic for Helitron and retrotransposon insertions. Proc Natl Acad Sci U S A. 106:8410–8416.
Heard E, et al. 2010. Ten years of genetics and genomics: what have we achieved and where are we heading? Nat Rev Genet. 11:723–733.
Heuer S, et al. 2009. Comparative sequence analyses of the major quantitative trait locus *phosphorus uptake 1 (Pup1)* reveal a complex genetic structure. Plant Biotechnol J. 7:456–471.
Hirotsune S, et al. 2003. An expressed pseudogene regulates the messenger-RNA stability of its homologous coding gene. Nature. 423:91–96.
Hoberman R, et al. 2009. A probabilistic approach for SNP discovery in high-throughput human resequencing data. Genome Res. 19:1542–1552.
Hollister JD, et al. 2011. Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proc Natl Acad Sci U S A. 108:2322–2327.
Hurwitz BL, et al. 2010. Rice structural variation: a comparative analysis of structural variation between rice and three of its closest relatives in the genus Oryza. Plant J. 63:990–1003.
Johnson ME, et al. 2001. Positive selection of a gene family during the emergence of humans and African apes. Nature. 413:514–519.
Johnson ME, et al. 2006. Recurrent duplication-driven transposition of DNA during hominoid evolution. Proc Natl Acad Sci U S A. 103:17626–17631.
Kidd JM, et al. 2008. Mapping and sequencing of structural variation from eight human genomes. Nature. 453:56–64.
Kidd JM, et al. 2010. Characterization of missing human genome sequences and copy-number polymorphic insertions. Nat Methods. 7:365–371.
Koonin EV. 2009. Darwinian evolution in the light of genomics. Nucleic Acids Res. 37:1011–1034.
Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occurring genetic variation in *Arabidopsis thaliana*. Annu Rev Plant Biol. 55:141–172.
Korbel JO, et al. 2007. Paired-end mapping reveals extensive structural variation in the human genome. Science. 318:420–426.
Krawitz P, et al. 2010. Microindel detection in short-read sequence data. Bioinformatics 26:722–729.

Kuhl H, et al. 2010. The European sea bass *Dicentrarchus labrax* genome puzzle: comparative BAC-mapping and low coverage shotgun sequencing. BMC Genomics. 11:68.
Larkin DM, et al. 2009. Breakpoint regions and homologous synteny blocks in chromosomes have different evolutionary histories. Genome Res. 19:770–777.
Lewin HA, Larkin DM, Pontius J, O'Brien SJ. 2009. Every genome sequence needs a good map. Genome Res. 19:1925–1928.
Lippman Z, et al. 2004. Role of transposable elements in heterochromatin and epigenetic control. Nature. 430:471–476.
Lisch D. 2009. Epigenetic regulation of transposable elements in plants. Annu Rev Plant Biol. 60:43–66.
Lynch M, Conery JS. 2000. The evolutionary fate and consequences of duplicate genes. Science. 290:1151.
Marques-Bonet T, Ryder OA, Eichler EE. 2009. Sequencing primate genomes: what have we learned? Annu Rev Genomics Hum Genet. 10:355–386.
Martinez Barrio A, et al. 2009. Targeted resequencing and analysis of the diamond-blackfan anemia disease locus RPS19. PLoS One. 4:e6172. doi:6110.1371/journal.pone.0006172
McKernan KJ, et al. 2009. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. Genome Res. 19:1527–1541.
Meader S, Hillier LW, Locke D, Ponting CP, Lunter G. 2010. Genome assembly quality: assessment and improvement using the neutral indel model. Genome Res. 20:675–684.
Meyers BC, Kaushik S, Nandety RS. 2005. Evolving disease resistance genes. Curr Opin Plant Biol. 8:129–134.
Michaels SD, He Y, Scortecci KC, Amasino RM. 2003. Attenuation of FLOWERING LOCUS C activity as a mechanism for the evolution of summer-annual flowering behavior in *Arabidopsis*. Proc Natl Acad Sci U S A. 100:10102–10107.
Nishiyama T, et al. 2003. Comparative genomics of *Physcomitrella patens* gametophytic transcriptome and *Arabidopsis thaliana*: implication for land plant evolution. Proc Natl Acad Sci U S A. 100:8007–8012.
Ossowski S, et al. 2008. Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Res. 18:2024–2033.
Perry GH, et al. 2008. Copy number variation and evolution in humans and chimpanzees. Genome Res. 18:1698–1710.
Pool JE, Hellmann I, Jensen JD, Nielsen R. 2010. Population genetic inference from genomic sequence variation. Genome Res. 20:291–300.
Rieseberg LH, Whitton J, Gardner K. 1999. Hybrid zones and the genetic architecture of a barrier to gene flow between two sunflower species. Genetics. 152:713–727.
Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 132:365–386.
Samonte RV, Eichler EE. 2002. Segmental duplications and the evolution of the primate genome. Nat Rev Genet. 3:65–72.
Santuari L, Hardtke CS. 2010. The case for resequencing studies of *Arabidopsis thaliana* accessions: mining the dark matter of natural genetic variation. F1000 Biol Rep. 2:85. doi: 10.3410/B2-85.
Santuari L, et al. 2010. Substantial deletion overlap among divergent *Arabidopsis* genomes revealed by intersection of short reads and tiling arrays. Genome Biol. 11:R4.
Schatz MC, Delcher AL, Salzberg SL. 2010. Assembly of large genomes using second-generation sequencing. Genome Res. 20:1165–1173.
Schmuths H, Meister A, Horres R, Bachmann K. 2004. Genome size variation among accessions of *Arabidopsis thaliana*. Ann Bot. 93:317–321.
Sebat J, et al. 2004. Large-scale copy number polymorphism in the human genome. Science. 305:525–528.
Sharp AJ, Cheng Z, Eichler EE. 2006. Structural variation of the human genome. Annu Rev Genomics Hum Genet. 7:407–442.
Shen YF, et al. 2010. A SNP discovery method to assess variant allele probability from next-generation resequencing data. Genome Res. 20:273–280.
Simpson JT, et al. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19:1117–1123.
Snyder M, Du J, Gerstein M. 2010. Personal genome sequencing: current approaches and challenges. Genes Dev. 24:423–431.
Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 33:W465–W467.
Sureshkumar S, et al. 2009. A genetic defect caused by a triplet repeat expansion in *Arabidopsis thaliana*. Science. 323:1060–1063.
Tsai IJ, Otto TD, Berriman M. 2010. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol. 11:R41. doi:10.1186/gb-2010-1111-1184-r1141
Turner TL, Bourne EC, Von Wettberg EJ, Hu TT, Nuzhdin SV. 2010. Population resequencing reveals local adaptation of *Arabidopsis lyrata* to serpentine soils. Nat Genet. 42:260–263.
Tuzun E, et al. 2005. Fine-scale structural variation of the human genome. Nat Genet. 37:727–732.
van der Biezen EA, Brandwagt BF, van Leeuwen W, Nijkamp HJJ, Hille J. 1996. Identification and isolation of the *FEEBLY* gene from tomato by transposon tagging. Mol Gen Genet. 251:267–280.
Warren RL, Sutton GG, Jones SJM, Holt RA. 2007. Assembling millions of short DNA sequences using SSAKE. Bioinformatics 23:500–501.
Weigel D, Mott R. 2009. The 1001 Genomes Project for *Arabidopsis thaliana*. Genome Biol. 10:107. doi:110.1186/gb-2009-1110-1185-1107
Weil C, Martienssen R. 2008. Epigenetic interactions between transposons and genes: lessons from plants. Curr Opin Genet Dev. 18:188–192.
Wheelan SJ, Aizawa Y, Han JS, Boeke JD. 2005. Gene-breaking: a new paradigm for human retrotransposon-mediated gene evolution. Genome Res. 15:1073–1078.
White MJD. 1978. Chain processes in chromosomal speciation. Syst Biol. 27:285–298.
Wood HM, et al. 2010. Using next-generation sequencing for high resolution multiplex analysis of copy number variation from nanogram quantities of DNA from formalin-fixed paraffin-embedded specimens. Nucleic Acids Res. 38:e151. doi:110.1093/nar/gkq1510.
Young AL, et al. 2010. A new strategy for genome assembly using short sequence reads and reduced representation libraries. Genome Res. 20:249–256.
Zeller G, et al. 2008. Detecting polymorphic regions in *Arabidopsis thaliana* with resequencing microarrays. Genome Res. 18:918–929.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18:821–829.
Zhang F, Gu W, Hurles ME, Lupski JR. 2009. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet. 10:451–481.
Zhang J. 2003. Evolution by gene duplication: an update. Trends Ecol Evol. 18:292–298.
Zhao S, et al. 2004. Human, mouse, and rat genome large-scale rearrangements: stability versus speciation. Genome Res. 14:1851–1860.
Zimmerli L, et al. 2008. The xenobiotic B-aminobutyric acid enhances *Arabidopsis* thermotolerance. Plant J. 53:144–156.
Ziolkowski PA, Koczyk G, Galganski L, Sadowski J. 2009. Genome sequence comparison of Col and L*er* lines reveals the dynamic nature of *Arabidopsis* chromosomes. Nucleic Acids Res. 37:3189–3201.

Associate editor: Michael Purugganan

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: ALVINA GRACE LAI

Name/Title of Principal Supervisor: DR. PAUL DIJKWEL

Name of Published Research Output and full reference:

Lai, A. G., Denton-Giles, M., Mueller-Roeber, B., Schippers, J. H. M., & Dijkwel, P. P. 2011. Positional information resolves structural variations and uncovers an evolutionarily divergent genetic locus in accessions of Arabidopsis thaliana. Genome biology and evolution 3: 627–640.

In which Chapter is the Published Work: 7. Publications

Please indicate either:

- The percentage of the Published Work that was contributed by the candidate: 80%

  and / or

- Describe the contribution that the candidate has made to the Published Work:

  AGL participated in project design, carried out sample preparations, LR-PCR, sequence analysis, data interpretation, validation experiments and drafted the manuscript.

Alvina Grace Lai
Digitally signed by Alvina Grace Lai
DN: cn=Alvina Grace Lai, o=Massey University, ou, email=a.g.lai@massey.ac.nz, c=NZ
Date: 2012.02.09 01:05:08 Z

Candidate's Signature

9 February 2012

Date

Paul Dijkwel
Digitally signed by Paul Dijkwel
DN: cn=Paul Dijkwel, o=Massey University, ou=IMBS, email=p.dijkwel@massey.ac.nz, c=NZ
Date: 2012.02.27 11:27:16 +13'00'

Principal Supervisor's signature

29/2/2012

Date