An early-morning gene network controlled by phytochromes and cryptochromes regulates photomorphogenesis pathways in Arabidopsis

Martin Balcerowicz1,\*, Mahiar Mahjoub1,2\*, Duy Nguyen1,3, Hui Lan1,4, Dorothee Stoeckle1,5, Susana Conde6,7, Katja E. Jaeger8, Philip A. Wigge8,9, Daphne Ezer7,10

1 Sainsbury Laboratory, University of Cambridge, Cambridge, United Kingdom  
2 current address: Royal Prince Alfred Hospital, Central Clinical School, University of Sydney, NSW, Australia   
3 current address: Plant Systems Physiology, Institute for Water and Wetland Research, Radboud University, Nijmegen, The Netherlands.   
4 current address: Zhejiang Normal University, Wucheng District, Jinhua, Zhejiang, China  
5 current address: Department of Plant and Microbial Biology, University of Zürich, Switzerland  
6 Department of Statistics, University of Warwick, United Kingdom  
7 Alan Turing Institute, London, United Kingdom  
8 Leibniz-Institute of Vegetable and Ornamental Crops (IGZ), 14979 Großbeeren, Germany  
9 Institute of Biochemistry and Biology, University of Potsdam, Germany  
10 Department of Biology, University of York  
\* equal contribution

#### Word counts: Abstract: 195; Introduction: 696; Results: 2829; Discussion: 1322; Materials and methods: 1372 Total words: 6414 Number of figures: 5 Number of supplementary figures: 8 Number of supplementary tables: 6

**Running title:** Multiple pathways control the “dawn burst” in gene expression

**Keywords:** BBX family, cryptochrome, dawn, HY5, network, morning, phytochrome, transcriptome

#### 

#### Abstract:

Light perception at dawn plays a key role in coordinating multiple molecular processes and in entraining the plant circadian clock. An Arabidopsis mutant lacking the main photoreceptors however still shows clock entrainment, indicating that the integration of light into the morning transcriptome is not well understood. We performed a high-resolution RNA-seq time series experiment, sampling every two-minutes from dawn. In parallel experiments, we perturbed temperature, the circadian clock, photoreceptor signalling and chloroplast-derived light signalling.  We used this data to infer a gene network that describes the gene expression dynamics after light stimulus in the morning, and then we validated key edges. By sampling time points at high density, we are able to identify three light- and temperature-sensitive bursts of transcription factor activity, one of which lasts for only about eight minutes.  Phytochrome and cryptochrome mutants cause a delay in the transcriptional bursts at dawn, and completely remove a burst of expression in key photomorphogenesis genes (HY5 and BBX family).  Our complete network is available online (<http://www-users.york.ac.uk/~de656/dawnBurst/dawnBurst.html>). Taken together, our results show that phytochrome and cryptochrome signaling is required for fine-tuning the dawn transcriptional response to light, but separate pathways can robustly activate much of the programme in their absence.

# Introduction

Light at dawn is a major cue enabling plants to entrain circadian gene expression programmes, coordinate developmental processes such as elongation growth and anticipate likely daytime stresses. Transcripts that are induced at dawn include genes involved in responses to drought (Grundy et al., 2015), and heat (Dickinson et al., 2018), as well as in anthocyanin biosynthesis (Seaton et al., 2018). In addition to abiotic stresses, Arabidopsis is also less susceptible to biotic stresses such as certain fungal pathogens in the morning, which is linked to jasmonic acid signalling (Ingle et al., 2015). Phytohormone genes in general represent another set of dawn-expressed genes and are essential for controlling the diurnal rhythm of growth, and many of these genes feature a G-box motif in their promoter (Michael et al., 2008). Consistent with this, a large set of genes with G-box promoter motifs that are expressed within one hour of dawn including many that are involved in phytohormone signalling and in response to metals (Ezer et al., 2017b).

Light at dawn plays a key role for entraining the circadian clock (Millar, 2004; P. A. Salomé and McClung, 2005; Webb et al., 2019), a process that is further contributed to by changes in temperature (Gil and Park, 2019) and even humidity (Mwimba et al., 2018). Despite the physiological importance of these responses, how light at dawn entrains the clock is not fully understood. Red- and blue-light sensors – including phytochrome A (phyA), phyB and cryptochrome 1 (cry1) – are involved in entrainment (Devlin and Kay, 2000; Hall et al., 2002; Salomé et al., 2002; Somers et al., 1998), while blue-light receptors of the zeitlupe (ztl) family also have clock-associated functions, but appear to be more relevant at later times during the day and at dusk (Christie et al., 2015; Suetsugu and Wada, 2013). In addition to clock entrainment, phytochromes and cryptochromes are also core regulators of photoperiodism and photomorphogenesis (Franklin et al., 2005; Jackson, 2009; Kami et al., 2010). phyA in particular has been reported as a key sensor of dawn and photoperiod and is essential for the induction of the core circadian clock component *PSEUDO RESPONSE REGULATOR 9* (*PRR9*) (Seaton et al., 2018). PRR9 and PRR7 together with the MYB transcription factors CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) and LATE ELONGATED HYPOCOTYL (LHY) constitute the morning loop of the circadian clock (Creux and Harmer, 2019).

Many higher order mutants of photoreceptor or clock genes disrupt the gating of hypocotyl growth and photoperiodic regulation of flowering, but have limited effects on entrainment: a *phyA phyB cry1 cry2* mutant still displays a functional clock (Yanovsky et al., 2000) while clock mutants such as *cca1 lhy* and *prr7 prr9* retain rhythmicity in many clock outputs under light:dark photocycles (Farré et al., 2005; Mizoguchi et al., 2002), suggesting that other pathways contribute to the entrainment and dawn-induced gene expression. Indeed, photosynthesis and sugar production can also entrain the circadian clock, producing a second ‘metabolic dawn’ (Haydon et al., 2013) and a chloroplast-derived light signal is known to gate the expression of thermotolerance genes that peak after dawn via HEAT SHOCK FACTOR A1 (HSFA1) transcription factors (Dickinson et al., 2018).

To understand the burst of gene expression in response to light in the morning, we performed a highly resolved RNA-seq time series. We find five distinct transcriptional waves within the first two hours of the morning and characterise how each wave of expression responds to temperature elevation and light signals during the night and subjective day. By analysing a photoreceptor mutant (*phyA phyB cry1 cry2*), a circadian clock mutant (*prr5 prr7 prr9*), and a mutant impaired in chloroplast light/temperature signalling (*hsfa1QK*), we reveal that the dawn burst is not globally controlled by a single pathway, but that partially distinct subsets of dawn genes are perturbed in each of these mutant backgrounds. Furthermore, we infer a gene regulatory network and validate edges using DNA binding data. We find that ELONGATED HYPOCOTYL 5 (HY5) and B-BOX PROTEIN 31 (BBX31) are among the TFs that are predicted to regulate multiple expression waves, and that HY5 and multiple BBX genes constitute a dawn subnetwork that is controlled by phytochromes and cryptochromes to finetune photomorphogenic and photoperiodic responses in the early morning.

# Results

## There is a burst of gene expression of DNA binding proteins at dawn

Since many groups of genes have been shown to have a dawn peak, we investigated how extensive this behaviour is, and the extent to which genes encoding transcriptional regulators also exhibit a peak in expression (**Table S1**).  In previous 24 h RNA-seq time courses (Ezer et al., 2017a; Jung et al., 2016), we observe that 39% of all genes encoding DNA binding proteins peak immediately before (ZT0) or an hour after dawn (ZT1) (**Figure S1A**). Genes related to abiotic stresses (Grundy et al., 2015; Dickinson et al., 2018) and the phytohormones abscisic acid (ABA), auxin and brassinosteroid (BR) (Michael et al. (2008) were previously found to display increased expression in the morning. Consistent with these studies, we found that there is a significant enrichment for ABA and ethylene-linked DNA binding proteins that have maximal expression in this time period compared to other DNA binding proteins (58% and 55%, p<0.005 in both cases using a Fisher Exact test with Bonferroni correction), (**Figure 1A i, ii**).  There was also an enrichment for DNA-binding proteins that are associated with GO terms related to light (53%, p<0.02) and stress (55%, p<0.002) (**Figure 1A iii, iv)**.  However, we found no significant enrichment for auxin or BR in the morning, which differs from observations by Michael et al. (2008), and neither were circadian clock-associated DNA binding proteins enriched, which is expected since only a subset of circadian genes are active in the morning (**Figure S1A**).

Intriguingly, while phytochromes are important for light entrainment (Somers et al., 1998), light entrainment of most transcripts is not lost in the *phyABCDE* background. DNA binding factors with a peak in gene expression in the hour after dawn still largely show a peak in expression in *phyABCDE*, although these are somewhat perturbed and several individual genes show a large fold-change of expression **(Figure 1B)**. Since genes with a peak in expression immediately before dawn show increased expression in *phyABCDE* before and after dawn we conclude that phytochromes may play a more important role in inhibiting nighttime genes than activating morning genes. We observe the same trends in a mutant of the evening clock gene *EARLY FLOWERING 3* (*ELF3*), which is consistent with both pathways mainly controlling nighttime gene expression (Ezer et al., 2017a) (**Figure S1B**).

## There are multiple waves of gene expression after light exposure at dawn

While many regulatory genes have peak expression at the onset of light, we do not know whether these genes are all expressed in one cluster, or whether there are separate individually resolved waves of morning gene expression. To address this, we performed a high-resolution time series experiment at the onset of light, with 25 time points sampled at 22°C (**Table S2, Figure S2**).

To visualise these data, we clustered the gene expression values from the 22°C time series, focusing on genes with GO terms associated with DNA binding (see **Methods** and **Table S3)**. The dawn burst consists of multiple coordinated transcript waves of genes associated with DNA binding. While there are genes whose expression levels decrease (clusters 1-2) or increase (clusters 3-4) over the time course, we observe three distinct bursts of expression. For the purposes of describing these clusters in a concise way, we will refer to them as the clusters that peak at 16-24 minutes (clusters 7-8), 18-45 minutes (clusters 9-10), and 45-105 minutes (clusters 5-6) (**Figure 2A)**.  The full gene expression profiles can be viewed as box plots in **Figure S3**.

These clusters contain many genes that are relevant to light signalling, temperature response, and immunity. Clusters 7-8 include many genes that are related to temperature, such as *HEAT SHOCK FACTOR A1A* (*HSFA1A*) and immunity genes such as the WRKY transcription factor *WRKY33* and *BASIC LEUCINE ZIPPER 10* (*bZIP10*).  This is consistent with other research that suggests that plants may be less sensitive to biotic stress at dawn (Ingle et al., 2015; Wang et al., 2011).  Interestingly, clusters 5-6 include *HEAT SHOCK FACTOR B2B* (*HSFB2B*) that suppresses the heat shock response and inhibits *HSFA1A* (Ikeda et al., 2011). The clusters of genes that peak at 45-105 minutes include genes that appear to be downstream of red (*RVE7, HY5, JMJ22, PNT1*) and blue (*MYC2, MYC4, CIB2, CRY3*) light signalling.

## The dawn gene expression waves are sensitive to temperature and light

Given that light and temperature strongly influence the circadian clock and that many dawn burst genes are associated with both light and temperature responses, we investigated the effect of temperature and light perturbations on dawn expression waves. We repeated the high resolution RNA-seq experiment at 27°C (**Table S2, Figure S2**), and found that many genes continue to peak at the same time, although we detect an earlier rise in expression for clusters 7-8. In addition, we observe that the earlier expressed genes (clusters 1-2 and 7-8) have elevated expression at 27°C compared to 22°C while peak expression was lower for the later expressed genes (clusters 3-6) at 27°C (**Figure 2A**). Please recall that the clusters were determined using the 22°C time series only.  These results are shown in a complementary way using boxplots in **Figure S4**.  
Previously, Rugnone et al. (2013) found sets of genes that were induced or repressed by light, and moreover they identified genes whose sensitivity to light was dependent on whether the light treatment occurred at night or during the subjective day (after an extended night).  Comparing their gene lists with ours (**Figure 2B**, **Table S4)**, we find significant enrichment for light-repressed genes among early expressed genes in clusters 1-2 (p-value < 1e-14, based on Fisher exact test with Bonferroni correction).  In contrast, those genes that were expressed late in the time series or that peaked at 45-105 minutes or 18-45 minutes tended to be light-induced. We identified a number of key light signalling genes in clusters 5-6 of our high resolution time course and we find that over a third of genes within this group are light-induced.  Intriguingly, the clusters of genes that includes *HSFA1A* (clusters 7-8) are not enriched for light-induced genes, except for a slight enrichment for genes whose expression is induced by light at night (p-value <0.037).

These results show that there are multiple bursts of gene expression at dawn that are differentially regulated by temperature and light, and moreover many of the dawn-expressed genes are regulated by light in a time-of-day dependent manner.

## Late dawn gene expression waves are delayed in both *phyA phyB cry1 cry2* (light signalling) and *prr5 prr7 prr9* (circadian) mutants

DNA binding proteins appear to have multiple coordinated bursts of expression at dawn and these bursts are either light- or temperature-dependent or both.  To characterise the gene regulatory mechanisms that underlie these distinct bursts, we performed RNA-seq on a series of mutant strains that perturbed red- and blue-light sensitivity (*phyA phyB cry1 cry2*, Ler background), the circadian clock morning loop (*prr5 prr7 prr9*, Col-0 background) and chloroplast-derived light/temperature signals (*hsfA1QK***,** mostlyCol-0 background), all of which have been previously characterised (Nakamichi et al., 2005; Yanovsky et al., 2000; Yoshida et al., 2011). These were sampled at the most informative time points according to NITPicker (Ezer and Keir, 2019). We observe similar patterns of expression in this subset of time points as we were able to see in the original time series in Col-0 (**Figure 2C).** There was strong temporal autocorrelation within each time series **(Figure S2A),** suggesting that gene expression varied smoothly over time, and consistent gene expression patterns across Col-0 biological replicates **(Figure S2B)**, despite different sequencing protocols being used. To be able to detect common patterns of regulation among these genes, we re-clustered each of the five main gene groups separately (**Figure S5A, Table S5).** Of 760 transcription factor genes retained in this analysis, we find that roughly 50% show similar trends in expression changes in *phyA phyB cry1 cry2* and *prr5 prr7 prr9* compared to the respective wild types, i.e. they either show up- or downregulation in both mutants (e.g. the large sub-clusters 1-2-III and 5-6-IX). On the other hand, approximately 200 genes show distinct expression changes in one of the two mutants, i.e. they are misregulated in only one background or show opposite trends in the two mutants (e.g. sub-clusters 3-4-IX and 5-6-X). Compared to the photoreceptor and clock mutants, fewer sub-clusters show clear misexpression in *hsfa1QK*, which is consistent with HSFA1 regulating many non-TF targets, in particular heat shock proteins (Dickinson et al., 2018). The mutant-specific differences in gene expression are described in depth in the legend of **Figure S5A**.

Thus, it appears that the circadian clock, classical light signalling via phytochromes/cryptochromes and, to a lesser degree, chloroplast-derived light signals via HSFA1 contribute to the coordination of TF gene expression in the early morning. We also note that very few genes among the early expression bursts (clusters 7-10) show reduced expression in any mutant background. This could be attributed to the fact that multiple signalling routes need to be disrupted for clear reduction to occur; however, it is likely that additional light input pathways exist that are not covered in our analysis and contribute particularly to the control of the early dawn burst.

## An early morning gene regulatory network

Next, we wished to identify possible regulatory links by inferring a gene regulatory network.  Since transcription and translation are distinct events, it follows that there will be a time delay between the appearance of a transcript and its cognate protein corresponding to the time required to translate the mRNA.  We therefore applied a time delay network inference method called dynGenie3 (Huynh-Thu and Geurts, 2018).  Because all large-scale network inference algorithms have large false-positive rates, we wished to confirm as many edges as possible using available DNA binding data via DAP-seq from O’Malley et al. (2016).

Among TFs in our network with available binding data, we found that 34.3% of our high confidence edges (top 1% highest scoring edges) were consistent with DAP-seq.  In contrast, we only observed a 12.6% overlap with DAP-seq data among edges in our network that received a score of ‘0’ using dynGenie3 among TF with available DAP-seq data.  This suggests that dynGenie3 is successfully enriching for biologically plausible edges.

From now on, we only consider the subset of the dynGenie3 network that is consistent with DAP-seq data (**Figure 3A**): the network consistently forms a U-shape under a force-directed layout, with early expressed genes primarily on the left side of the U and late expressed genes primarily on the right side.  These genes are enriched for photoperiodism-related GO terms, including GO terms associated with light sensing, heat response and circadian rhythms **(Figure 3B).**

A number of transcription factors that are known to play a role in light signalling and photomorphogenesis sit at the base of the U, linking early and late expressed genes;  these include  *HY5, BBX31*, *CYCLING DOF FACTOR 5 (CDF5)*and *REVEILLE 1 (RVE1)* **(Figure 3A)**.  These transcription factors are all involved in light-controlled hypocotyl elongation (Chattopadhyay et al., 1998; Heng et al., 2019; Martín et al., 2018; Rawat et al., 2009). While *RVE1* and *CDF5* expression is under strong circadian control and peaks at dawn (Henriques et al., 2017; Rawat et al., 2009), *HY5* and *BBX31* are light-induced genes (Heng et al., 2019; Oyama et al., 1997) and their transcript levels display a strong peak post dawn in our time course data **(Figure S5B)**. Based on the timing of their expression and their position within our network, RVE1 and CDF5 may promote expression of growth-related genes before dawn, while HY5 and BBX31 appear to coordinate the down-regulation of early dawn genes as well as the induction of subsequent transcriptional waves.  Moreover, betweenness (HY5 and BBX31 have the 12th and 13th highest values) and page rank (HY5 and BBX31 have the 6th and 3rd highest values) indicate the important position of HY5 and BBX31 within the network structure (**Table S2**).

HY5 is one of the most thoroughly characterised components of the light signalling pathway and, aside from its role in promoting photomorphogenesis, has been found to affect the circadian clock, temperature responses, chlorophyll and anthocyanin biosynthesis as well as nutrient uptake (Gangappa and Botto, 2016). In agreement with its multifaceted roles, we find genes related to light signalling (*PHYTOCHROME INTERACTING FACTOR 4*/*PIF4*, *EARLY LIGHT INDUCIBLE PROTEIN 1/ELIP1*, *BBX16*), multiple heat shock genes, the circadian clock genes *TIMING OF CAB EXPRESSION 1 (TOC1)* and *ALTERED SEED GERMINATION 4* (*ASG4*)/*RVE3* as well as several UDP glucosyltransferase genes involved in flavonoid metabolism among its predicted targets **(Figure 3C)**. In contrast to HY5, BBX31 represents a negative regulator of photomorphogenesis (Heng et al., 2019). Its targets do not show strong enrichment for specific biological processes but include the related BBX transcription factors *BBX18* and *BBX30* as well as *HY5* itself **(Figure 3D)**. It has previously been shown that HY5 represses *BBX31* transcription (Heng et al., 2019), but based on our data a reciprocal regulation of *HY5* by BBX31 may occur as well.

## Phytochromes and cryptochromes control morning photomorphogenesis pathways via BBX family proteins and HY5

A regulatory network of BBX transcription factors, HY5 and the related bZIP transcription factor HY5 HOMOLOG (HYH) has been postulated to control a large portion of the light-regulated transcriptome (Xu, 2019). We observe that multiple members of this network feature strongly during the dawn burst: not only *HY5* and *BBX31*, but also *HYH*, *BBX16*, *BBX24/SALT TOLERANCE (STO)*, *BBX25/SALT TOLERANCE HOMOLOG (STH)*, *BBX30* and *BBX32* transcript levels peak within an hour after dawn,while these peaks are completely abolished in the *phyA phyB cry1 cry2* mutant background **(Figure 4)** but only partially reduced in *phyA phyB* and *cry1 cry2* double mutants (**Figure S6**). For many of these genes, expression still peaked at dawn in *prr5 prr7 prr9*, but was generally higher and persisted over a longer time, while their expression remained largely unaffected in *hsfa1QK* **(Figure S7)**. This implies that phytochromes and cryptochromes act in parallel to induce *HY5*, *HYH* and the *BBX* genes (hereafter referred to as BBX Set A). Another set of *BBX* genes (BBX Set B), comprised of *BBX20* and *BBX21*, does not exhibit a dawn burst in expression, but their transcript levels are strongly upregulated in the *phyA phyB cry1 cry2* mutant **(Figure 4**). Interestingly, apart from BBX16 all transcription factors included in Set A have been shown to negatively regulate photomorphogenesis (Gangappa et al., 2013, 2013; Heng et al., 2019; Holtan et al., 2011; Indorf et al., 2007; Job et al., 2018), while HY5, HYH as well as BBX20 and BBX21 represent photomorphogenesis-promoting transcription factors (Bursch et al., 2020; Chattopadhyay et al., 1998; Fan et al., 2012; Holm et al., 2002; Job et al., 2018).

We sought to further examine the BBX/HY5 transcriptional network and its light regulation with a set of qPCR experiments and monitored gene expression (a) during an extended night, (b) in *phyA phyB cry1 cry2*, *prr5 prr7 prr9* and *hsfa1QK* mutants (to confirm the RNA-seq results), and (c) in *hy5 hyh* and *bbx30 bbx31* mutants as well as in a *BBX31* overexpressor (*BBX31-Ox*). We investigated the expression patterns of *BBX* and *HY5*/*HYH* genes shown in Figure 4 as well as six putative downstream targets of BBX31/HY5 predicted in our network – *BBX27* and *PIF4,* which show a drop in transcript levels after dawn, as well as *BBX18, POPEYE* (*PYE*), *ASG4*/*RVE3* and *COLD CIRCADIAN RHYTHM AND RNA-BINDING 1* (*CCR1*).

Notably, dawn peaks in the transcripts of *BBX* set A, *HY5* and *HYH* were completely absent during an extended night as well as in the *phyA phyB cry1 cry2* mutant (**Figure 5A, B; S8A, B**), while expression levels of *BBX* set B remained high under these conditions (**Figure 5C; S8C**). Putative targets that display induced expression after dawn mimicked the pattern observed for BBX set A, while putative targets with decreasing expression after dawn (BBX27, PIF4) displayed a similar pattern to BBX set B (**Figure 5D, S8D**). These observations confirm that it is a light stimulus, and not the circadian clock, that tightly controls this part of the dawn burst, and it can largely be accounted for by light sensing via phytochromes and cryptochromes.

Expression patterns in *bbx30 bbx31* were largely indistinguishable from those in wild type (**Figure 5; S8**), likely because of the high degree of redundancy among BBX proteins and/or residual BBX30 function as we detected high levels of truncated *BBX30* transcript in the mutant. However, many dawn-induced genes, including the putative direct targets *BBX18*, *ASG4* and *CCR1*, displayed substantially lower transcript levels in both the *hy5 hyh* mutant and the *BBX31-Ox* line (**Figure 5A, B, D; S8A, B, D**), indicating that HY5 and HYH act as activators of the dawn burst, while BBX31 represents a negative regulator. Only one exception to this pattern was found: *PYE* expression was unchanged in *BBX31-Ox*, but reduced in both *hy5 hyh* and *bbx30 bbx31*, suggesting positive effects of both upstream regulators.

Taken together, our results suggest that BBX and HY5/HYH factors constitute a highly orchestrated transcriptional network that controls photomorphogenic and photoperiodic responses at dawn downstream of phytochrome and cryptochrome photoreceptors.

# Discussion

## A high-resolution time series is required to capture complex transcriptional dynamics at dawn

While several previous reports indicated a burst of gene expression in the early morning, here we were able to fully characterise the transcriptional dynamics in this time period using an RNA-seq time course with very high sampling frequency.  We were able to identify multiple coordinated waves of gene expression. These included clusters of early-expressed genes which were enriched for light-repressed genes (clusters 1-2) as well as clusters enriched for light-induced genes (clusters 5-6), which peaked during the 2 hour window after dawn and also displayed decreased peak expression levels at high ambient temperature of 27°C. We also identified gene sets that were not enriched for light-regulated genes (as defined by Rugnone et al., 2013), but showed clearly elevated expression at high temperature and included many genes involved in heat stress response and thermotolerance (clusters 7-8).

In general, it is a useful strategy to conduct a few high resolution time course experiments, and then use this data to select informative time point experiments for future experiments. Following this approach we find that a large proportion of TFs change their expression in the early morning, and the expression of many of these genes is perturbed in similar ways in both the circadian clock and photoreceptor mutants. Specifically, we observed (i) elevated levels of temperature-responsive TF genes and (ii) a time delay in the expression in most clusters. These findings agree with the notion that phytochromes and cryptochromes provide light input for circadian entrainment  (Devlin and Kay, 2000; Hall et al., 2002; Salomé et al., 2002; Somers et al., 1998) and are thus tightly linked to the circadian clock morning loop, constituted by CCA1, LHY and the PRRs. Phytochromes also act as *bona fide* temperature sensors (Jung et al., 2016; Legris et al., 2016), while cryptochromes can dampen temperature responses (Ma et al., 2016) and both PRR9 and PRR7 are essential for temperature entrainment of the clock (Salomé and McClung, 2005).

Whereas many late dawn burst TF genes show abolished or delayed induction in *phyA phyB cry1 cry2* and *prr5 prr7 prr9*, early dawn burst genes are either unaffected or display higher expression in these genetic backgrounds. Phytochromes, cryptochromes and PRRs have all been found to act as transcriptional (co-)repressors (Chen et al., 2014; Jung et al., 2016; Liu et al., 2016; Pedmale et al., 2016), which agrees with their direct targets being upregulated in the respective mutants. They may hence regulate late dawn bursts indirectly, by repressing early expressed negative regulators or through post-transcriptional mechanisms. Phytochromes and cryptochromes in particular control many TFs post-translationally through inhibition of the CONSTITUTIVELY PHOTOMORPHOGENIC 1 (COP1)/SUPPRESSOR OF PHYA-101 (SPA) complex, a key light signalling repressor that promotes degradation of many TFs involved in photomorphogenesis (Hoecker, 2017; Podolec and Ulm, 2018). This leaves us with the question of how early dawn burst genes are induced upon light exposure. Other blue light receptors are still active in *phyA phyB cry1 cry2*: while ztl proteins have core roles in circadian processes, phototropins are not directly linked to entrainment, but trigger stomatal opening in response to light (Christie et al., 2015; Suetsugu and Wada, 2013), which may in turn be perceived as a change in water potential within the plant. Moreover, not all light-triggered retrograde signals from chloroplasts are disrupted in the *hsfa1QK* mutant. While HSFA1 TFs act in response to light-induced changes in the chloroplast’s redox state (Dickinson et al., 2018), other signals may bypass this route. The increase in primary photosynthates after dawn for instance can be perceived by energy sensors such as Snf1-RELATED KINASE 1 (SnRK1) and TARGET OF RAPAMYCIN (TOR), which are also known to affect clock entrainment (Shin et al., 2017; Zhang et al., 2019).

It is important to note that we have characterised gene expression changes that result from the kind of sudden onset of light that occur in growth chambers, rather than those found in natural conditions, which have for instance been investigated by Annunziata et al. (2018).  Nevertheless, there are a number of benefits of this approach. Firstly, this allows us to more directly compare our results to other studies of biotic and abiotic responses in the early morning that were also performed in artificial growth chambers with sudden onset of light (Cortijo et al., 2019; Dickinson et al., 2018; Ezer et al., 2017b; Grundy et al., 2015; Ingle et al., 2015; Michael et al., 2008; Seaton et al., 2018). Secondly, this kind of study allows us to directly address how the transcriptome responds to a *light stimulus* in the morning. Specifically, since we know the exact time that plants were first able to detect light, we can measure the exact time delay between this exposure and a change in gene expression.   Finally, farming in artificial lights is becoming common, and it is important to understand how crops respond to lighting conditions that are reminiscent of growth chambers (Dutta Gupta and Agarwal, 2017; Ibaraki, 2016; Olvera-Gonzalez et al., 2013).

## Phytochromes and cryptochromes coordinate light responses in the early morning via HY5 and BBX transcription factors

Our time course data allowed us to infer a high-confidence transcriptional network that controls the dawn burst in gene expression. HY5 and BBX family proteins appear at the core of this network downstream of phytochrome and cryptochrome photoreceptors and are likely to coordinate the early and late waves of gene expression in the early morning. Notably, while loss of phytochrome function on its own strongly affects the pre-dawn peak in DNA binding genes, phytochromes and cryptochromes can compensate for each other’s function in controlling the post-dawn HY5/BBX subnetwork, at least under white light. They induce both positive (*HY5*, *HYH*) and negative (*BBX* Set A) regulators of light responses, while they simultaneously repress other photomorphogenesis-promoting factors (*BBX* Set B). *HY5* and *HYH* are not only induced by light at the transcriptional level, but the respective proteins are also stabilised through light-induced inactivation of the COP1/SPA repressor complex (Holm et al., 2002; Osterlund et al., 2000; Saijo et al., 2003). Boosting HY5 and HYH protein levels via both mechanisms will trigger light responses in the early morning, but this process may require fine-tuning through other factors including BBX proteins. BBX24, BBX25 and BBX32 interfere with HY5 transcriptional activity through direct interaction (Gangappa et al., 2013; Holtan et al., 2011; Job et al., 2018), while BBX30 and BBX31 act downstream of HY5 to promote elongation growth (Heng et al., 2019). BBX20 and BBX21 on the other hand promote HY5 function by increasing *HY5* transcript level and post-translationally enhancing HY5 activity (Bursch et al., 2020; Job et al., 2018; Wei et al., 2016; Xu et al., 2018), but these effects seem to be largely suppressed at dawn. Another layer of complexity is added by the fact that the COP1/SPA complex also promotes degradation of several BBX proteins (Indorf et al. 2007; Yan et al. 2011; Xu et al. 2016; Gangappa et al. 2013).

HY5 and BBX proteins interdependently tune photomorphogenic responses such as the reduction in elongation growth after dawn and may provide light input to the circadian clock via *TOC1* and *ASG4*, but may also exert additional functions independently of each other. BBX proteins have not yet been implicated in other HY5-mediated responses such as biosynthesis of secondary metabolites. On the other hand, BBX30, BBX31 and BBX32 act as negative regulators of the floral transition independently of HY5 through the repression of *FLOWERING LOCUS T* (*FT*) (Graeff et al., 2016; Tripathi et al., 2017). Strong induction of these genes in the early morning may be involved in gating *FT’s* responsiveness to inductive signals in the photoperiodic control of flowering.

In summary, our results reveal a large transcriptional regulatory network that controls the dawn burst of gene expression downstream of photoreceptor, circadian clock and chloroplast-derived signals. Phytochrome and cryptochrome photoreceptors coordinate a subset of this network to adjust photomorphogenesis, photoperiodism and clock entrainment in accordance with the plant’s light environment.

# Methods

## Plant material and growth conditions

## All Arabidopsis thaliana mutants used in this study have been described previously. *phyA-201 phyB-5* (Reed et al., 1998), *phyA-201 phyB-5 cry1-201 cry2*/*fha-1* (Yanovsky et al., 2000) and *phyA-201 phyB-1 phyC-1 phyD-1 phyE-1* (*phyABCDE*) (Hu et al., 2013) are in the Ler background, *bbx30-1 bbx31-1* (Yadav et al., 2019), *BBX31-Ox* (Graeff et al., 2016), *elf3-1* (Zagotta et al., 1996), *cry1-104 cry2-1* (Mao et al., 2005), *hy5-2 hyh* (Zhang et al., 2017) and *prr5-11 prr7-11 prr9-10* (Nakamichi et al., 2005) are in the Col-0 background, and *hsfA1a hsfA1b hsfA1d hsfA1e* (*hsfa1QK*) (Liu et al., 2011) is in a mixed Col-0/Ws background.

## For all experiments, seeds were surface-sterilised and sown on ½ Murashige and Skoog (MS) agar plates at pH5.7 without sucrose; the agar was topped with nylon mesh that allowed quick removal of seedlings for sample collection. They were stratified for 3 days at 4°C in the dark and then allowed to germinate for 24 h at 22°C under cool-white fluorescent light at 75-85 µmol m-2 s-1. Plates were then shifted into chambers at constant 22°C or 27°C and cool-white fluorescent light at 90-100 µmol m-2 s-1 with short photoperiods (8 h light, 16 h dark). The plates from which we collected were picked randomly, not based on position. We also mixed-up plate positions in the chamber on several days to reduce positional effects.

## RNA isolation and quantitative PCR (qPCR)

For expression analysis, 15-20 7-d-old seedlings were collected per sample by snap-freezing in liquid nitrogen, disrupted using a TissueLyser II (Qiagen) and RNA was extracted using the MagMaxTM-96 Total RNA Isolation Kit (Thermo Fisher Scientific). The RNA was reverse transcribed into cDNA using the Transcriptor First Strand cDNA Synthesis Kit (Roche). Transcript levels were determined by qPCR on a LightCycler® 480 II instrument (Roche) using LightCycler® 480 SYBR Green I Master Mix (Roche). Data were analysed by the ΔΔCT method and two to three biological replicates were used. Primers used for qPCR are listed in **Table S6**.

## RNA-seq: High resolution time series

For the high-resolution time couse, 15-20 7-d-old whole seedlings were collected per sample. Samples were collected 10 min before dawn as well as in 2 min intervals for the first half hour after dawn followed by 15 min intervals for the following 90 min. RNA was extracted as described above. RNA quality and integrity was assessed on an Agilent 2200 TapeStation. Library preparation was performed with 1 µg of high integrity total RNA using the NEBNext® UltraTM Directional RNA Library Prep Kit for Illumina® (New England Biolabs). The libraries were sequenced on a NextSeq® 500 (Illumina) using paired-end sequencing of 75 bp with a NextSeq® 500/550 High Output v2 Kit for 150 cycles.

## Quant-seq: Low resolution time series

For the low-resolution time course, 15-20 7-d-old whole seedlings were collected per sample. Samples were collected 10 min before dawn as well as 10, 24, 30, 45, 60, 105 and 120 min after dawn. RNA was extracted as described above. RNA quality and integrity was assessed on an Agilent 2200 TapeStation. Library preparation was performed with 500 ng of high integrity total RNA using the QuantSeq 3’ mRNA-Seq Library Prep Kit FWD for Illumina® (Lexogen). The libraries were sequenced on a NextSeq® 500 (Illumina) using single-end sequencing of 75 bp with a NextSeq® 500/550 High Output v2 Kit for 75 cycles.

## Bioinformatics analysis of RNA-seq data

All data is available in SRA: **PRJNA687210.** The Illumina RNA-seq data was analysed using the same bioinformatics pipeline used in Ezer et al. (2017b), using Trimmomatic-0.32 to trim reads (Bolger et al., 2014), Tophat for mapping to the TAIR10 annotated genome (Trapnell et al., 2009), HTseq-count to find the raw counts after duplicates were removed (Anders et al., 2015), and Cufflinks to calculate Fragments Per Kilobase Million (FPKM), which was then converted into Transcripts Per Million (TPM) (Trapnell et al., 2013).  The time points for the Lexogen Quant-seq experiment were selected using NITPicker (Ezer and Keir, 2019).

The Lexogen Quant-seq data was analysed using the Integrated Data Analysis Pipeline on Bluebee® platform, which maps the reads to the genome using the STAR Aligner (Dobin et al., 2013) and counts reads using HTseq-count (Anders et al., 2015).  Quant-seq expression values are expressed as Reads per Million (RPM), because Quant-seq does not require normalisation by gene length.

## Clustering

Clustering of the 24-hour time series (Figure 1) were drawn using hierarchical clustering (default parameters of hclust in R).

Clustering of the high resolution time gene expression time series data (in Figure 2) was performed using the CLUST algorithm (Abu-Jamous and Kelly, 2018), using recommended settings for RNA-seq TPM data as per the reference manual (i.e. log2, Z- and quantile normalisation of TPM values).

The clustering in Figure 2A was performed only on the 22°C high resolution time series data.  The other gene expression time series in Figure 2A and 2B were drawn using the same gene order.  The z-scores were calculated across all samples within each row of Figure 2A and Figure 2B.

## Gene list curation for network inference

A number of criteria were used to generate a gene list for performing network inference.

In the beginning, all lowly or non-expressing genes were not included in the network inference analysis. The criteria for removal each sample were: [1] rowSum(RPMj,t) < 7 x nj,t x 1.05 (RPMj,t is RPM value and nj,t the total number of time-points for gene j at timepoint t) and [2] genes that had <5 time-points where the TPMj < 7.  Genes with at least one time-point containing TPM = 0 were removed (requirement of dynGENIE3 package).

The final gene list used for network inference was obtained from the following sources: (1) GO categories, (2) consensus cluster and (3) DE analysis.  A gene only needed to meet one of these three criteria in order to be included in the analysis. Note that prior to the analysis in each of the three approaches, only genes satisfying the mentioned expression thresholds were included.

Firstly, all genes that had GO categories that were of biological interest to us were initially included.  Specifically, this referred to genes that had Biological Process GO terms that included the words ’stress’, ‘light’, ‘auxin’, ‘abscisic’, ‘ethylene’, ‘circadian’, as well as all genes that had the Molecular Function ‘DNA binding’.

Secondly, we selected for genes that had similar expression patterns as other genes in the data set.  The reason we chose this criteria was that we did not want to include lots of genes that had extremely noisy patterns of expression.  Clustering of the time-course gene expression was performed for each sample Col-0, Ler, *prr5 prr7 prr9*, *phyA phyB cry1 cry2*, *hsfa1QK*at 22°C using the CLUST algorithm using the recommended settings (mentioned previously).  Recall that the CLUST algorithm filters genes that do not cluster well with any of the clusters.  Any gene that appeared in any of the clusters detected by the CLUST algorithm was included in the analysis.

Thirdly, differential expression analysis of gene was performed based on the time-course expression using the *odp* method from the package edge (R Bioconductor) (Storey et al. 2005). Significant genes were chosen based on having an adjusted p-value (q-value) < 0.05. The following WT and mutant pairs at 22°C were compared: Col-0 vs *prr5 prr7 prr9*, Col-0 vs *hsfa1QK*and Ler vs *phyA phyB cry1 cry2*.

This analysis produced a gene list that was too expansive for dynGenie3.  The filtered gene list was ranked based on decreasing CV (sd/mean) with the top 1500 chosen for network inference. All the 1500 selected genes were confirmed to satisfy the expression cutoffs outlined earlier.

## Network inference and analysis

Network inference on the curated gene list was performed using dynGENIE3 R package (Huynh-Thu and Geurts 2018). For each gene, the time-course gene expression was supplied from the following samples: Col-0/Ler/*prr5 prr7 prr9/phyA phyB cry1 cry2/hsfa1QK*at 22°C and Col-0/Ler/*prr5 prr7 prr9/phyA phyB cry1 cry2*at 27°C. For downstream analysis of network, only edges that had scores within the top 1% and that were consistent with DAP-seq data (O’Malley et al., 2016) were included.  GO term enrichment within the network was determined using the PAFway package in CRAN.  Figures were drawn using the network visualisation tools developed for Ara-BOX-cis (Ezer et al., 2017b), and is available here: <http://www-users.york.ac.uk/~de656/dawnBurst/dawnBurst.html>

# Acknowledgements

# We thank Stephan Wenkel for sharing of *bbx30 bbx31* and *BBX31-Ox* seeds and discussion of unpublished results. This project was funded by the Alan Turing Institute Research Fellowship under EPSRC Research grant (TU/A/000017) to DE; EPSRC/BBSRC Innovation Fellowship (EP/S001360/1) to DE and SC; and an EMBO fellowship (ALTF 1418-2015) to MB.

# Author Contributions

MB, DE, PAW planned and designed the research. MB, DN, DS, KEJ designed and performed the experiments. MM, HL, SC, DE designed and performed the data analysis. MB, MM, PAW, DE wrote the manuscript.

# Competing interests

# The authors declared that they have no conflict of interest.

# References

Abu-Jamous B, Kelly S. 2018. Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. *Genome Biology* **19**:172. doi:10.1186/s13059-018-1536-8

Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169. doi:10.1093/bioinformatics/btu638

Annunziata MG, Apelt F, Carillo P, Krause U, Feil R, Koehl K, Lunn JE, Stitt M. 2018. Response of Arabidopsis primary metabolism and circadian clock to low night temperature in a natural light environment. *J Exp Bot* **69**:4881–4895. doi:10.1093/jxb/ery276

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114–2120. doi:10.1093/bioinformatics/btu170

Bursch K, Toledo-Ortiz G, Pireyre M, Lohr M, Braatz C, Johansson H. 2020. Identification of BBX proteins as rate-limiting cofactors of HY5. *Nature Plants* 1–8. doi:10.1038/s41477-020-0725-0

Chattopadhyay S, Ang L-H, Puente P, Deng X-W, Wei N. 1998. Arabidopsis bZIP Protein HY5 Directly Interacts with Light-Responsive Promoters in Mediating Light Control of Gene Expression. *Plant Cell* **10**:673–683. doi:10.1105/tpc.10.5.673

Chen F, Li B, Li G, Charron J-B, Dai M, Shi X, Deng XW. 2014. Arabidopsis Phytochrome A Directly Targets Numerous Promoters for Individualized Modulation of Genes in a Wide Range of Pathways. *The Plant Cell* **26**:1949–1966. doi:10.1105/tpc.114.123950

Christie JM, Blackwood L, Petersen J, Sullivan S. 2015. Plant Flavoprotein Photoreceptors. *Plant Cell Physiol* **56**:401–413. doi:10.1093/pcp/pcu196

Cortijo S, Aydin Z, Ahnert S, Locke JC. 2019. Widespread inter-individual gene expression variability in Arabidopsis thaliana. *Molecular Systems Biology* **15**:e8591. doi:10.15252/msb.20188591

Creux N, Harmer S. 2019. Circadian Rhythms in Plants. *Cold Spring Harb Perspect Biol* **11**:a034611. doi:10.1101/cshperspect.a034611

Devlin PF, Kay SA. 2000. Cryptochromes Are Required for Phytochrome Signaling to the Circadian Clock but Not for Rhythmicity. *The Plant Cell* **12**:2499–2509. doi:10.1105/tpc.12.12.2499

Dickinson PJ, Kumar M, Martinho C, Yoo SJ, Lan H, Artavanis G, Charoensawan V, Schöttler MA, Bock R, Jaeger KE, Wigge PA. 2018. Chloroplast Signaling Gates Thermotolerance in Arabidopsis. *Cell Reports* **22**:1657–1665. doi:10.1016/j.celrep.2018.01.054

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21. doi:10.1093/bioinformatics/bts635

Dutta Gupta S, Agarwal A. 2017. Artificial Lighting System for Plant Growth and Development: Chronological Advancement, Working Principles, and Comparative Assessment In: Dutta Gupta S, editor. Light Emitting Diodes for Agriculture: Smart Lighting. Singapore: Springer. pp. 1–25. doi:10.1007/978-981-10-5807-3\_1

Ezer D, Jung J-H, Lan H, Biswas S, Gregoire L, Box MS, Charoensawan V, Cortijo S, Lai X, Stöckle D, Zubieta C, Jaeger KE, Wigge PA. 2017a. The evening complex coordinates environmental and endogenous signals in *Arabidopsis*. *Nature Plants* **3**:17087. doi:10.1038/nplants.2017.87

Ezer D, Keir J. 2019. NITPicker: selecting time points for follow-up experiments. *BMC Bioinformatics* **20**:166. doi:10.1186/s12859-019-2717-5

Ezer D, Shepherd SJK, Brestovitsky A, Dickinson P, Cortijo S, Charoensawan V, Box MS, Biswas S, Jaeger KE, Wigge PA. 2017b. The G-Box Transcriptional Regulatory Code in Arabidopsis. *Plant Physiology* **175**:628–640. doi:10.1104/pp.17.01086

Fan X-Y, Sun Yu, Cao D-M, Bai M-Y, Luo X-M, Yang H-J, Wei C-Q, Zhu S-W, Sun Ying, Chong K, Wang Z-Y. 2012. BZS1, a B-box Protein, Promotes Photomorphogenesis Downstream of Both Brassinosteroid and Light Signaling Pathways. *Molecular Plant* **5**:591–600. doi:10.1093/mp/sss041

Farré EM, Harmer SL, Harmon FG, Yanovsky MJ, Kay SA. 2005. Overlapping and Distinct Roles of PRR7 and PRR9 in the Arabidopsis Circadian Clock. *Current Biology* **15**:47–54. doi:10.1016/j.cub.2004.12.067

Franklin KA, Larner VS, Whitelam GC. 2005. The signal transducing photoreceptors of plants. *Int J Dev Biol* **49**:653–664. doi:10.1387/ijdb.051989kf

Gangappa SN, Botto JF. 2016. The Multifaceted Roles of HY5 in Plant Growth and Development. *Molecular Plant* **9**:1353–1365. doi:10.1016/j.molp.2016.07.002

Gangappa SN, Crocco CD, Johansson H, Datta S, Hettiarachchi C, Holm M, Botto JF. 2013. The Arabidopsis B-BOX Protein BBX25 Interacts with HY5, Negatively Regulating BBX22 Expression to Suppress Seedling Photomorphogenesis[C][W]. *Plant Cell* **25**:1243–1257. doi:10.1105/tpc.113.109751

Gil K-E, Park C-M. 2019. Thermal adaptation and plasticity of the plant circadian clock. *New Phytologist* **221**:1215–1229. doi:10.1111/nph.15518

Graeff M, Straub D, Eguen T, Dolde U, Rodrigues V, Brandt R, Wenkel S. 2016. MicroProtein-Mediated Recruitment of CONSTANS into a TOPLESS Trimeric Complex Represses Flowering in Arabidopsis. *PLOS Genetics* **12**:e1005959. doi:10.1371/journal.pgen.1005959

Grundy J, Stoker C, Carré IA. 2015. Circadian regulation of abiotic stress tolerance in plants. *Front Plant Sci* **6**. doi:10.3389/fpls.2015.00648

Hall A, Kozma‐Bognár L, Bastow RM, Nagy F, Millar AJ. 2002. Distinct regulation of CAB and PHYB gene expression by similar circadian clocks. *The Plant Journal* **32**:529–537. doi:10.1046/j.1365-313X.2002.01441.x

Haydon MJ, Mielczarek O, Robertson FC, Hubbard KE, Webb AAR. 2013. Photosynthetic entrainment of the Arabidopsis thaliana circadian clock. *Nature* **502**:689–692. doi:10.1038/nature12603

Heng Y, Lin F, Jiang Y, Ding M, Yan T, Lan H, Zhou H, Zhao X, Xu D, Deng XW. 2019. B-Box Containing Proteins BBX30 and BBX31, Acting Downstream of HY5, Negatively Regulate Photomorphogenesis in Arabidopsis. *Plant Physiology* **180**:497–508. doi:10.1104/pp.18.01244

Henriques R, Wang H, Liu J, Boix M, Huang L-F, Chua N-H. 2017. The antiphasic regulatory module comprising CDF5 and its antisense RNA FLORE links the circadian clock to photoperiodic flowering. *New Phytologist* **216**:854–867. doi:10.1111/nph.14703

Hoecker U. 2017. The activities of the E3 ubiquitin ligase COP1/SPA, a key repressor in light signaling. *Current Opinion in Plant Biology*, 37 Physiology and metabolism 2017 **37**:63–69. doi:10.1016/j.pbi.2017.03.015

Holm M, Ma L-G, Qu L-J, Deng X-W. 2002. Two interacting bZIP proteins are direct targets of COP1-mediated control of light-dependent gene expression in Arabidopsis. *Genes Dev* **16**:1247–1259. doi:10.1101/gad.969702

Holtan HE, Bandong S, Marion CM, Adam L, Tiwari S, Shen Y, Maloof JN, Maszle DR, Ohto M, Preuss S, Meister R, Petracek M, Repetti PP, Reuber TL, Ratcliffe OJ, Khanna R. 2011. BBX32, an Arabidopsis B-Box Protein, Functions in Light Signaling by Suppressing HY5-Regulated Gene Expression and Interacting with STH2/BBX21[W][OA]. *Plant Physiol* **156**:2109–2123. doi:10.1104/pp.111.177139

Hu W, Franklin KA, Sharrock RA, Jones MA, Harmer SL, Lagarias JC. 2013. Unanticipated regulatory roles for Arabidopsis phytochromes revealed by null mutant analysis. *Proc Natl Acad Sci U S A* **110**:1542–1547. doi:10.1073/pnas.1221738110

Huynh-Thu VA, Geurts P. 2018. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. *Scientific Reports* **8**:1–12. doi:10.1038/s41598-018-21715-0

Ibaraki Y. 2016. Lighting Efficiency in Plant Production Under Artificial Lighting and Plant Growth Modeling for Evaluating the Lighting Efficiency In: Kozai T, Fujiwara K, Runkle ES, editors. LED Lighting for Urban Agriculture. Singapore: Springer. pp. 151–161. doi:10.1007/978-981-10-1848-0\_11

Ikeda M, Mitsuda N, Ohme-Takagi M. 2011. Arabidopsis HsfB1 and HsfB2b Act as Repressors of the Expression of Heat-Inducible Hsfs But Positively Regulate the Acquired Thermotolerance1[C][W][OA]. *Plant Physiol* **157**:1243–1254. doi:10.1104/pp.111.179036

Indorf M, Cordero J, Neuhaus G, Rodríguez‐Franco M. 2007. Salt tolerance (STO), a stress-related protein, has a major role in light signalling. *The Plant Journal* **51**:563–574. doi:10.1111/j.1365-313X.2007.03162.x

Ingle RA, Stoker C, Stone W, Adams N, Smith R, Grant M, Carré I, Roden LC, Denby KJ. 2015. Jasmonate signalling drives time-of-day differences in susceptibility of Arabidopsis to the fungal pathogen Botrytis cinerea. *The Plant Journal* **84**:937–948. doi:10.1111/tpj.13050

Jackson SD. 2009. Plant responses to photoperiod. *New Phytologist* **181**:517–531. doi:10.1111/j.1469-8137.2008.02681.x

Job N, Yadukrishnan P, Bursch K, Datta S, Johansson H. 2018. Two B-Box Proteins Regulate Photomorphogenesis by Oppositely Modulating HY5 through their Diverse C-Terminal Domains. *Plant Physiology* **176**:2963–2976. doi:10.1104/pp.17.00856

Jung J-H, Domijan M, Klose C, Biswas S, Ezer D, Gao M, Khattak AK, Box MS, Charoensawan V, Cortijo S, Kumar M, Grant A, Locke JCW, Schäfer E, Jaeger KE, Wigge PA. 2016. Phytochromes function as thermosensors in Arabidopsis. *Science* **354**:886–889. doi:10.1126/science.aaf6005

Kami C, Lorrain S, Hornitschek P, Fankhauser C. 2010. Chapter Two - Light-Regulated Plant Growth and Development In: Timmermans MCP, editor. Current Topics in Developmental Biology, Plant Development. Academic Press. pp. 29–66. doi:10.1016/S0070-2153(10)91002-8

Legris M, Klose C, Burgie ES, Rojas CCR, Neme M, Hiltbrunner A, Wigge PA, Schäfer E, Vierstra RD, Casal JJ. 2016. Phytochrome B integrates light and temperature signals in Arabidopsis. *Science* **354**:897–900. doi:10.1126/science.aaf5656

Liu H-C, Liao H-T, Charng Y-Y. 2011. The role of class A1 heat shock factors (HSFA1s) in response to heat and other stresses in Arabidopsis. *Plant, Cell & Environment* **34**:738–751. doi:10.1111/j.1365-3040.2011.02278.x

Liu TL, Newton L, Liu M-J, Shiu S-H, Farré EM. 2016. A G-Box-Like Motif Is Necessary for Transcriptional Regulation by Circadian Pseudo-Response Regulators in Arabidopsis. *Plant Physiol* **170**:528–539. doi:10.1104/pp.15.01562

Ma D, Li X, Guo Y, Chu J, Fang S, Yan C, Noel JP, Liu H. 2016. Cryptochrome 1 interacts with PIF4 to regulate high temperature-mediated hypocotyl elongation in response to blue light. *PNAS* **113**:224–229. doi:10.1073/pnas.1511437113

Mao J, Zhang Y-C, Sang Y, Li Q-H, Yang H-Q. 2005. A role for Arabidopsis cryptochromes and COP1 in the regulation of stomatal opening. *PNAS* **102**:12270–12275. doi:10.1073/pnas.0501011102

Martín G, Rovira A, Veciana N, Soy J, Toledo-Ortiz G, Gommers CMM, Boix M, Henriques R, Minguet EG, Alabadí D, Halliday KJ, Leivar P, Monte E. 2018. Circadian Waves of Transcriptional Repression Shape PIF-Regulated Photoperiod-Responsive Growth in Arabidopsis. *Current Biology* **28**:311-318.e5. doi:10.1016/j.cub.2017.12.021

Michael TP, Breton G, Hazen SP, Priest H, Mockler TC, Kay SA, Chory J. 2008. A Morning-Specific Phytohormone Gene Expression Program underlying Rhythmic Plant Growth. *PLOS Biology* **6**:e225. doi:10.1371/journal.pbio.0060225

Millar AJ. 2004. Input signals to the plant circadian clock. *J Exp Bot* **55**:277–283. doi:10.1093/jxb/erh034

Mizoguchi T, Wheatley K, Hanzawa Y, Wright L, Mizoguchi M, Song H-R, Carré IA, Coupland G. 2002. LHY and CCA1 Are Partially Redundant Genes Required to Maintain Circadian Rhythms in Arabidopsis. *Developmental Cell* **2**:629–641. doi:10.1016/S1534-5807(02)00170-3

Mwimba M, Karapetyan S, Liu L, Marqués J, McGinnis EM, Buchler NE, Dong X. 2018. Daily humidity oscillation regulates the circadian clock to influence plant physiology. *Nature Communications* **9**:1–10. doi:10.1038/s41467-018-06692-2

Nakamichi N, Kita M, Ito S, Yamashino T, Mizuno T. 2005. PSEUDO-RESPONSE REGULATORS, PRR9, PRR7 and PRR5, Together Play Essential Roles Close to the Circadian Clock of Arabidopsis thaliana. *Plant Cell Physiol* **46**:686–698. doi:10.1093/pcp/pci086

Olvera-Gonzalez E, Alaniz-Lumbreras D, Ivanov-Tsonchev R, Villa-Hernández J, Olvera-Olvera C, González-Ramírez E, Araiza-Esquivel M, Torres-Argüelles V, Castaño V. 2013. Intelligent lighting system for plant growth and development. *Computers and Electronics in Agriculture* **92**:48–53. doi:10.1016/j.compag.2012.11.012

O’Malley RC, Huang SC, Song L, Lewsey MG, Bartlett A, Nery JR, Galli M, Gallavotti A, Ecker JR. 2016. Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292. doi:10.1016/j.cell.2016.04.038

Osterlund MT, Hardtke CS, Wei N, Deng XW. 2000. Targeted destabilization of HY5 during light-regulated development of Arabidopsis. *Nature* **405**:462–466. doi:10.1038/35013076

Oyama T, Shimura Y, Okada K. 1997. The Arabidopsis HY5 gene encodes a bZIP protein that regulates stimulus-induced development of root and hypocotyl. *Genes Dev* **11**:2983–2995. doi:10.1101/gad.11.22.2983

Pedmale UV, Huang SC, Zander M, Cole BJ, Hetzel J, Ljung K, Reis PAB, Sridevi P, Nito K, Nery JR, Ecker JR, Chory J. 2016. Cryptochromes Interact Directly with PIFs to Control Plant Growth in Limiting Blue Light. *Cell* **164**:233–245. doi:10.1016/j.cell.2015.12.018

Podolec R, Ulm R. 2018. Photoreceptor-mediated regulation of the COP1/SPA E3 ubiquitin ligase. *Current Opinion in Plant Biology*, Cell signalling and gene regulation **45**:18–25. doi:10.1016/j.pbi.2018.04.018

Rawat R, Schwartz J, Jones MA, Sairanen I, Cheng Y, Andersson CR, Zhao Y, Ljung K, Harmer SL. 2009. REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc Natl Acad Sci U S A* **106**:16883–16888. doi:10.1073/pnas.0813035106

Reed JW, Elumalai RP, Chory J. 1998. Suppressors of an Arabidopsis thaliana phyB Mutation Identify Genes That Control Light Signaling and Hypocotyl Elongation. *Genetics* **148**:1295–1310.

Rugnone ML, Soverna AF, Sanchez SE, Schlaen RG, Hernando CE, Seymour DK, Mancini E, Chernomoretz A, Weigel D, Más P, Yanovsky MJ. 2013. LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. *PNAS* **110**:12120–12125. doi:10.1073/pnas.1302170110

Saijo Y, Sullivan JA, Wang H, Yang J, Shen Y, Rubio V, Ma L, Hoecker U, Deng XW. 2003. The COP1–SPA1 interaction defines a critical step in phytochrome A-mediated regulation of HY5 activity. *Genes Dev* **17**:2642–2647. doi:10.1101/gad.1122903

Salomé P. A., McClung CR. 2005. What makes the Arabidopsis clock tick on time? A review on entrainment. *Plant, Cell & Environment* **28**:21–38. doi:10.1111/j.1365-3040.2004.01261.x

Salomé Patrice A., McClung CR. 2005. PSEUDO-RESPONSE REGULATOR 7 and 9 Are Partially Redundant Genes Essential for the Temperature Responsiveness of the Arabidopsis Circadian Clock. *Plant Cell* **17**:791–803. doi:10.1105/tpc.104.029504

Salomé PA, Michael TP, Kearns EV, Fett-Neto AG, Sharrock RA, McClung CR. 2002. The out of phase 1 Mutant Defines a Role for PHYB in Circadian Phase Control in Arabidopsis. *Plant Physiology* **129**:1674–1685. doi:10.1104/pp.003418

Seaton DD, Toledo-Ortiz G, Ganpudi A, Kubota A, Imaizumi T, Halliday KJ. 2018. Dawn and photoperiod sensing by phytochrome A. *Proc Natl Acad Sci U S A* **115**:10523–10528. doi:10.1073/pnas.1803398115

Shin J, Sánchez-Villarreal A, Davis AM, Du S-X, Berendzen KW, Koncz C, Ding Z, Li C, Davis SJ. 2017. The metabolic sensor AKIN10 modulates the Arabidopsis circadian clock in a light-dependent manner. *Plant Cell Environ* **40**:997–1008. doi:10.1111/pce.12903

Somers DE, Devlin PF, Kay SA. 1998. Phytochromes and Cryptochromes in the Entrainment of the Arabidopsis Circadian Clock. *Science* **282**:1488–1490. doi:10.1126/science.282.5393.1488

Suetsugu N, Wada M. 2013. Evolution of Three LOV Blue Light Receptor Families in Green Plants and Photosynthetic Stramenopiles: Phototropin, ZTL/FKF1/LKP2 and Aureochrome. *Plant Cell Physiol* **54**:8–23. doi:10.1093/pcp/pcs165

Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. 2013. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* **31**. doi:10.1038/nbt.2450

Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**:1105–1111. doi:10.1093/bioinformatics/btp120

Tripathi P, Carvallo M, Hamilton EE, Preuss S, Kay SA. 2017. Arabidopsis B-BOX32 interacts with CONSTANS-LIKE3 to regulate flowering. *Proc Natl Acad Sci U S A* **114**:172–177. doi:10.1073/pnas.1616459114

Wang W, Barnaby JY, Tada Y, Li H, Tör 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. doi:10.1038/nature09766

Webb AAR, Seki M, Satake A, Caldana C. 2019. Continuous dynamic adjustment of the plant circadian oscillator. *Nature Communications* **10**:1–9. doi:10.1038/s41467-019-08398-5

Wei C-Q, Chien C-W, Ai L-F, Zhao J, Zhang Z, Li KH, Burlingame AL, Sun Y, Wang Z-Y. 2016. The Arabidopsis B-BOX Protein BZS1/BBX20 Interacts with HY5 and Mediates Strigolactone Regulation of Photomorphogenesis. *J Genet Genomics* **43**:555–563. doi:10.1016/j.jgg.2016.05.007

Xu D. 2019. COP1 and BBXs-HY5-mediated light signal transduction in plants. *New Phytol*. doi:10.1111/nph.16296

Xu D, Jiang Y, Li J, Holm M, Deng XW. 2018. The B-Box Domain Protein BBX21 Promotes Photomorphogenesis. *Plant Physiology* **176**:2365–2375. doi:10.1104/pp.17.01305

Yadav A, Bakshi S, Yadukrishnan P, Lingwan M, Dolde U, Wenkel S, Masakapalli SK, Datta S. 2019. The B-Box-Containing MicroProtein miP1a/BBX31 Regulates Photomorphogenesis and UV-B Protection1. *Plant Physiol* **179**:1876–1892. doi:10.1104/pp.18.01258

Yanovsky MJ, Mazzella MA, Casal JJ. 2000. A quadruple photoreceptor mutant still keeps track of time. *Current Biology* **10**:1013–1015. doi:10.1016/S0960-9822(00)00651-5

Yoshida T, Ohama N, Nakajima J, Kidokoro S, Mizoi J, Nakashima K, Maruyama K, Kim J-M, Seki M, Todaka D, Osakabe Y, Sakuma Y, Schöffl F, Shinozaki K, Yamaguchi-Shinozaki K. 2011. Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. *Mol Genet Genomics* **286**:321–332. doi:10.1007/s00438-011-0647-7

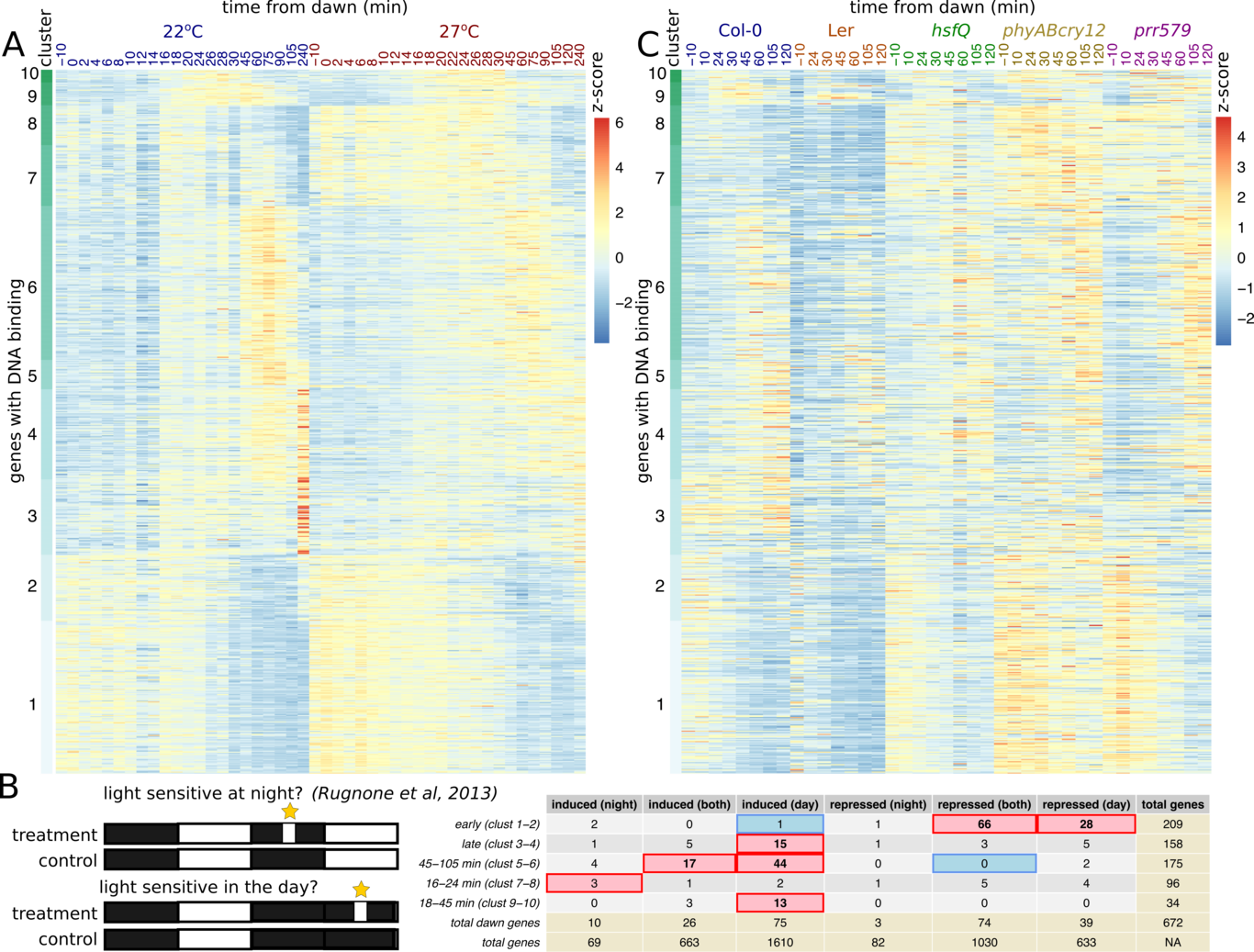
Zagotta MT, Hicks KA, Jacobs CI, Young JC, Hangarter RP, Meeks‐Wagner DR. 1996. The Arabidopsis ELF3 gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. *The Plant Journal* **10**:691–702. doi:10.1046/j.1365-313X.1996.10040691.x

Zhang N, Meng Y, Li X, Zhou Y, Ma L, Fu L, Schwarzländer M, Liu H, Xiong Y. 2019. Metabolite-mediated TOR signaling regulates the circadian clock in Arabidopsis. *Proc Natl Acad Sci USA* **116**:25395–25397. doi:10.1073/pnas.1913095116

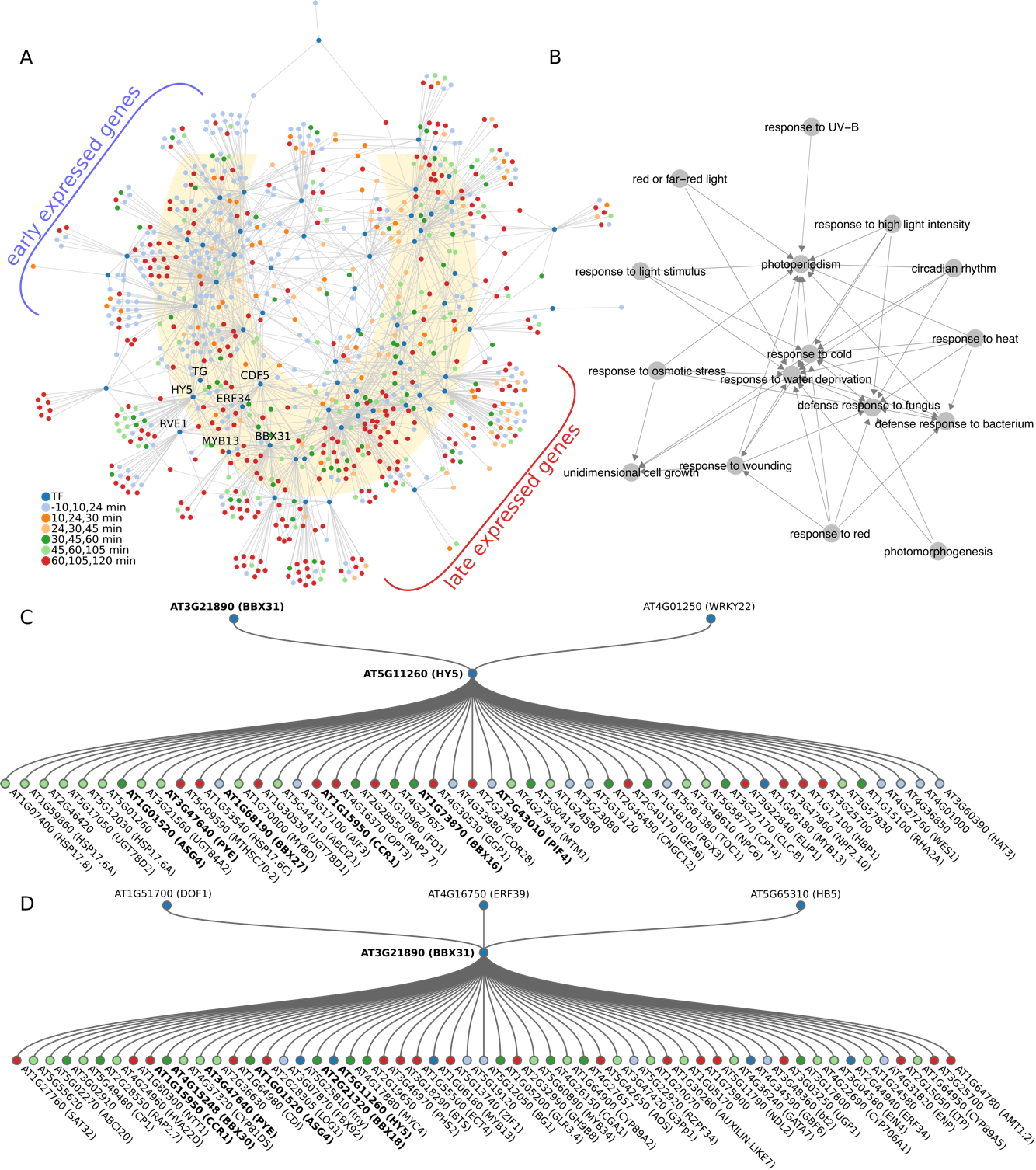
Zhang Y, Li C, Zhang J, Wang J, Yang J, Lv Y, Yang N, Liu J, Wang X, Palfalvi G, Wang G, Zheng L. 2017. Dissection of HY5/HYH expression in Arabidopsis reveals a root-autonomous HY5-mediated photomorphogenic pathway. *PLOS ONE* **12**:e0180449. doi:10.1371/journal.pone.0180449

****

**Figure 1: DNA binding proteins exhibit a burst of expression at dawn.** (A) The gene expression pattern of DNA binding proteins in ethylene (i), ABA (ii) , stress (iii), and light (iv) were clustered based on their expression in 8 time points under short day conditions in Col-0 at 22°C. The time point right before dawn (ZT0) and an hour after dawn (ZT1) are highlighted for clarity.  Similar figures for ‘auxin’, ’brassinosteroid’, ‘circadian clock’ and ‘all DNA binding’ are found in Fig S1.  (B) DNA-binding proteins whose genes have maximal expression at ZT0 (right before dawn) or ZT1 (an hour after dawn) in Col-0 at 22°C were identified, based on the analysis in (A).  We compare the expression levels of these genes in phyABCDE backgrounds compared to Ler (the background strain for phyABCDE).  Similar figures comparing elf3-1 and Col-0 are found in Fig S2.

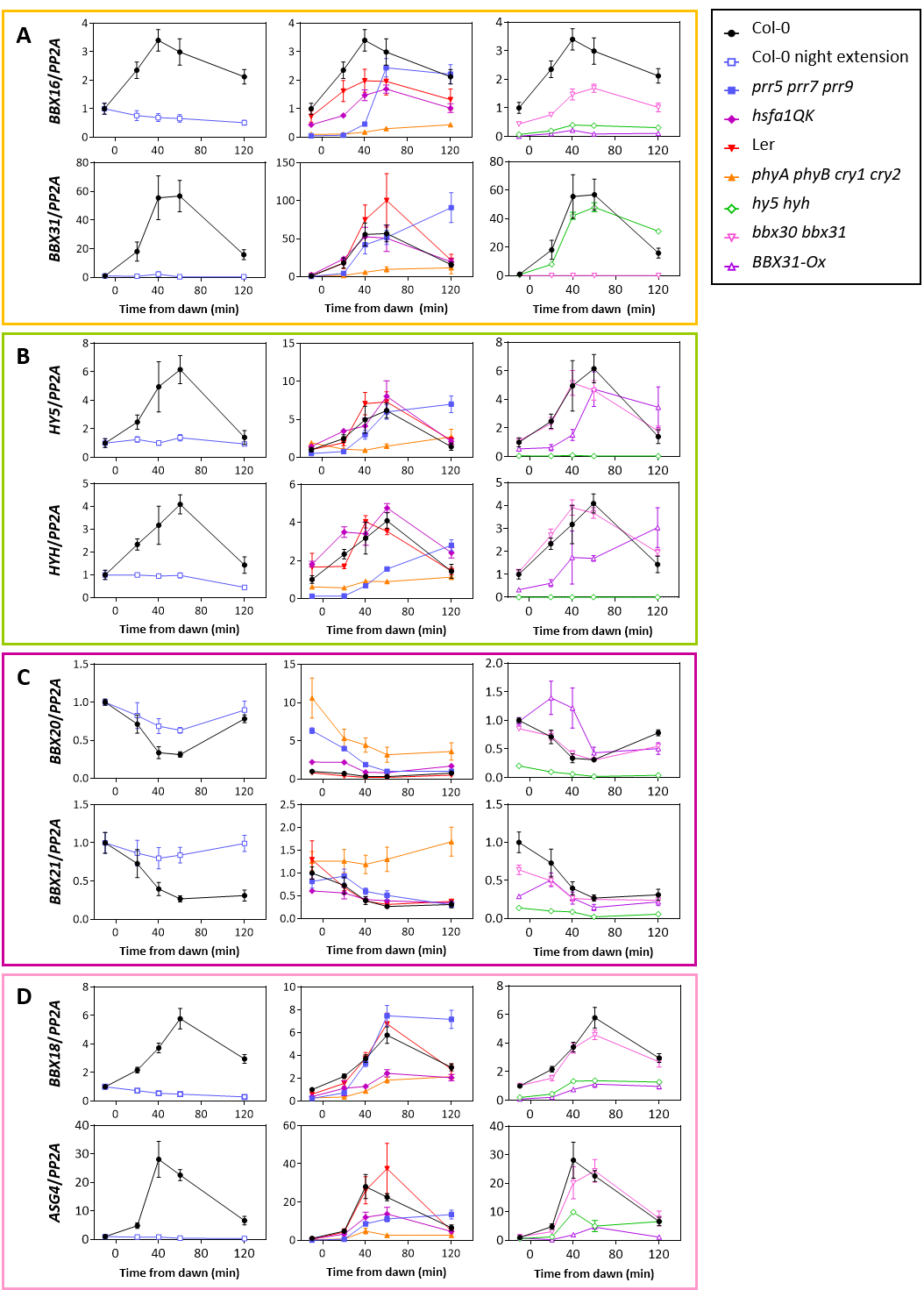


**Figure 2: Multiple waves of gene expression at dawn**.  Genes encoding DNA binding proteins were clustered into 10 groups using CLUST based on their gene expression pattern at 22°C in Col-0 grown in short day conditions.  (A) The gene expression pattern of these clusters is shown for Col-0 at both 22°C and 27°C.  (B) Previous experiments by Rugnone et al, 2013 identified genes that are sensitive to light in a time-of-day dependent way.  Many of these genes that were identified as being light induced or repressed in the day and night were also found to be expressed within the clusters in (A).  We use a table instead of a Venn diagram for clarity.  A coloured block indicates that there is a p-value < 0.05 based on a Fisher exact test with Bonferroni correction (red if it is more overlap than expected and blue if there is less than expected).  Bold indicates that the p-value <0.001.  (C) The gene expression pattern for those clusters in (A) for additional strains at 22°C.  Note that this experiment used fewer time points and transcripts were quantified using end-sequencing (QuantSeq 3’ sequencing, Lexogen) instead of whole transcript sequencing (Illumina RNA-seq).  Due to space, the names of the strains were abbreviated as follows: hsfA1QK was labelled as hsfQ, phyA phyB cry1 cry2 was labelled as phyABcry12, and prr5 prr7 prr9 was labelled as prr579.  Note that z-scores were calculated from the gene expression values in TPM.

 ***Figure 3****:* ***Analysis of higher confidence morning gene network****.  (A) This depicts the complete network of gene regulatory relationships that are predicted by dynGenie3 and consistent with DAP-seq data.  All genes that regulate at least one downstream target are coloured dark blue, while other genes are colour-coded by the time period in which they have maximal expression (as expressed as the average gene expression in 3 adjacent time points).  The complete network can be viewed here:* [*http://www-users.york.ac.uk/~de656/dawnBurst/dawnBurst.html*](http://www-users.york.ac.uk/~de656/dawnBurst/dawnBurst.html)*. (B) This graph depicts enrichment for pairs of GO terms within the network, as determined by the PAFway package.  The specific upstream and downstream targets predicted for HY5 (C) and BBX31 (D) are shown.   Genes that are discussed in the text are highlighted in bold for clarity.  The colour key is the same as in (A).*



**Figure 4:** **Gene expression profiles for genes related to key BBX/HY5 subnetwork in wild-type and phyA phyB cry1 cry2 seedlings**. Gene expression is expressed as transcripts per million (TPM) as determined by the RNA-seq experiments. The colour of the box that surrounds the line plot corresponds to the corresponding element in our model of network, illustrated in the top right.



**Figure 5:** **Gene expression profiles monitored by qPCR for genes related to the key BBX/HY5 subnetwork**. Expression was analysed in 7-d-old seedlings grown in short day conditions at 22°C. Genes analysed cover representatives of BBX set A (BBX16 and BBX31) (A), HY5/HYH (B), BBX set B (C) and the putative HY5/BBX target genes BBX18 and ASG4 (D). Their expression was analysed during an extended night at the start of day 8 (night extension, left column), in phyA phyB cry1 cry2, prr5 prr7 prr9 and hsfa1QK mutants (middle column), in hy5 hyh and bbx30 bbx31 mutants and a BBX31 overexpressor (BBX31-Ox) line (right column). Data represent the mean ± standard error of the mean (SEM) of three biological replicates. Note the Col-0 pattern is shown in all panels for comparison. Data for additional genes of BBX set A and four additional targets can be found in Figure S8.