RESEARCH ARTICLE

American Society of Plant Biologists · SOCIETY FOR EXPERIMENTAL BIOLOGY · WILEY

Effects of light regimes on circadian gene co-expression networks in *Arabidopsis thaliana*

Quentin Rivière[1,2] | Virginie Raskin[1] | Romário de Melo[1] |
Stéphanie Boutet[3] | Massimiliano Corso[3] | Matthieu Defrance[4] |
Alex A. R. Webb[5] | Nathalie Verbruggen[1] | Armand D. Anoman[1]

[1]Laboratory of Plant Physiology and Molecular Genetics, Université Libre de Bruxelles, Brussels, Belgium

[2]Biology Centre, Czech Academy of Sciences, Institute of Plant Molecular Biology, České Budějovice, Czech Republic

[3]Université Paris-Saclay, INRAE, AgroParisTech, Institut Jean-Pierre Bourgin (IJPB), Versailles, France

[4]Interuniversity Institute of Bioinformatics in Brussels, Université Libre de Bruxelles, Brussels, Belgium

[5]Department of Plant Sciences, University of Cambridge, Cambridge, UK

Correspondence
Armand D. Anoman and Nathalie Verbruggen, Laboratory of Plant Physiology and Molecular Genetics, Université Libre de Bruxelles, 1050 Brussels, Belgium.
Email: armandanom@gmail.com and nathalie.verbruggen@ulb.be

Funding information
Fonds National de la Recherche Scientifique FNRS-FRS, Grant/Award Number: PDR T 0085.16; Université Paris-Saclay (University of Paris-Saclay), Grant/Award Number: ANR-17-EUR-0007

Abstract

Light/dark (LD) cycles are responsible for oscillations in gene expression, which modulate several aspects of plant physiology. Those oscillations can persist under constant conditions due to regulation by the circadian oscillator. The response of the transcriptome to light regimes is dynamic and allows plants to adapt rapidly to changing environmental conditions. We compared the transcriptome of Arabidopsis under LD and constant light (LL) for 3 days and identified different gene co-expression networks in the two light regimes. Our studies yielded unforeseen insights into circadian regulation. Intuitively, we anticipated that gene clusters regulated by the circadian oscillator would display oscillations under LD cycles. However, we found transcripts encoding components of the flavonoid metabolism pathway that were rhythmic in LL but not in LD. We also discovered that the expressions of many stress-related genes were significantly increased during the dark period in LD relative to the subjective night in LL, whereas the expression of these genes in the light period was similar. The nocturnal pattern of these stress-related gene expressions suggested a form of "skotoprotection." The transcriptomics data were made available in a web application named *Cyclath*, which we believe will be a useful tool to contribute to a better understanding of the impact of light regimes on plants.

KEYWORDS
Arabidopsis thaliana, circadian rhythms, co-expression networks, flavonoids, light regime, skotoprotection

1 | INTRODUCTION

Light provided by the sun makes life possible on earth as it powers the production of organic matter through photosynthesis, a process that first evolved in bacteria and has been evolutionarily conserved in plants (Hohmann-Marriott & Blankenship, 2011; Xiong & Bauer, 2002). The mechanisms by which light energy is harvested and how plants mitigate light harmful effects have been studied intensively (Bassi & Dall'Osto, 2021; Rochaix, 2014; Shen, 2015). Reactive oxygen species (ROS), which are produced upon normal plant metabolism, can be generated in excess when too much light is absorbed, and this can damage the photosynthetic apparatus (Long et al., 1994). To prevent ROS accumulation and harmful effects, plants employ various photoprotection mechanisms, including the movement of chloroplasts

and leaves, dissipation of excess energy into heat, enzymatic defenses such as superoxide dismutase and ascorbate peroxidases, and non-enzymatic components such as carotenoids, ascorbate, and glutathione, and nonphotochemical quenching reactions involving light harvesting complexes (LHCs) or xanthophylls (Bassi & Dall'Osto, 2021; Demmig-Adams & Adams III, 1992).

Plants possess the capability to synthesize specialized metabolites for scavenging ROS, such as flavonoids (Falcone Ferreyra et al., 2021). These molecules play multifaceted roles, extending beyond merely safeguarding against oxidative stress to encompass attracting pollinators, facilitating hormone transport, aiding in seed dispersal, and fortifying defenses against pathogens (Corso & de la Torre, 2020; Pérez-García et al., 2015; Tanaka et al., 2008).

Light signals, which regulate photomorphogenesis, are decoded by several photoreceptors, such as phytochromes, cryptochromes, phototropins, and UVR8 receptors (Franklin & Quail, 2010; Roeber et al., 2021). Light regulation results in transcriptional changes, which can be easily observed by the diurnal pattern of expression of many genes. In addition to direct light regulation across the diel cycle, there is also circadian regulation of gene expression as an output of the circadian oscillator, which is most easily observed in constant light (LL). In Arabidopsis, up to 37% of the transcriptome is thought to be circadian regulated (Covington et al., 2008; Hsu & Harmer, 2012; Romanowski et al., 2020). The coordinated diel and circadian regulation of gene transcript abundance results in rhythmic biological processes including photosynthesis, hypocotyl growth, mineral nutrition, immune responses, and light signaling (Haydon et al., 2015; Millar, 2016; Romanowski et al., 2020).

The circadian oscillator orchestrates transcriptional changes by integrating various transcriptional regulators, each exhibiting a 24-h expression cycle owing to their involvement in interconnected transcriptional/translational feedback loops. *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) and *LATE ELONGATED HYPOCOTYL* (*LHY*) peak near dawn; *PSEUDORESPONSE REGULATOR* (*PRR*) *5*, *7*, and *9* are expressed throughout the day; *TIMING OF CAB EXPRESSION 1* (*TOC1*) is expressed at dusk; *GIGANTEA* (*GI*), *EARLY FLOWERING 3* (*ELF3*), *ELF4*, and *LUX ARRHYTHMO* (*LUX*) peak at night (Hsu & Harmer, 2014b; Webb et al., 2019). Other genes like *ZEITLUPE* (*ZTL*), *LIGHT-REGULATED WD1* (*LWD1*), *TEOSINTE BRANCHED 1* (*TCP20*), and *CCA1 HIKING EXPEDITION 1* (*CHE*) are also associated with circadian regulation (Hsu & Harmer, 2014b; Nohales & Kay, 2020).

The study of the interactions between the core circadian clock components and light input pathways was facilitated by large-scale transcriptomics, mathematical modeling, and gene co-expression networks (Dalchau et al., 2010; De Caluwé et al., 2017; Hsu et al., 2013; Hsu & Harmer, 2014a, 2014b; W. Huang et al., 2012). Gene co-expression networks are especially powerful to identify genes with similar biological functions (Contreras-López et al., 2018; Cortijo et al., 2020). However, their effectiveness relies on the accurate computation of correlations between transcript abundance, which requires large datasets. For the analysis of diel and/or circadian rhythms, the quality of gene co-expression networks mainly depends on the technology for gene expression profiling as well as on the duration of the

experiment and the sampling frequency. The recent works of Romanowski et al. (2020) and Bonnot and Nagel (2021) provided new insights into the circadian rhythmicity of genes and proteins. However, their studies did not include diel transcriptomes.

In this work, we performed RNA-sequencing (RNA-seq) to produce time-series data in Arabidopsis seedlings under diel and LL regimes. We analyzed transcript rhythms, differential expression, and co-expression networks, and we were able to identify rhythmic networks rather than solely rhythmic transcripts. The striking expression pattern of some stress-related genes led us to introduce the term "skotoprotection" to refer to those genes that could help plants face the adverse consequences of nighttime. We have made the transcriptomics data available in a web application we developed called *Cyclath*, dedicated to the study of diel and circadian rhythms.

2 | RESULTS

2.1 | Rhythmic changes in transcript abundance in light/dark (LD) cycles and LL

The study was designed to identify rhythmic Arabidopsis transcripts in 12/12-h LD cycles and under LL, as well as differentially expressed genes (DEGs) between these two conditions (Figure 1). In both conditions, 8-day plantlets were harvested after 30 min, 1 h, 2 h, 4 h, and then every 4 h until 72 h.

From a total of 21,487 expressed genes, 9053 (42.1%) and 2402 (11.2%) were oscillating in LD and LL, respectively, as determined by JTK_cycle in R (Hughes et al., 2010; Wu et al., 2016) (Bonferroni-corrected *p*-value < .01) (Figure 2, Table 1, and Dataset S1). A total of 7569 transcripts oscillated in LD but not in LL, while 918 transcripts oscillated in LL but not in LD (Figure 2). Being rhythmic in LD is not a direct indicator of circadian control because only around 16% of the rhythmic genes in LD were also rhythmic in LL (1484 out of 9053). Conversely, around 62% of rhythmic genes in LL were also rhythmic in LD (1484 out of 2402). Comparison with previously published circadian experiments (Bonnot & Nagel, 2021; Covington et al., 2008; Hsu & Harmer, 2014a; Romanowski et al., 2020) revealed that our study yielded a generally lower number of rhythmic transcripts under LL conditions compared to those reported in the referenced studies (Figure S1 and Table S1).

Of the oscillating transcripts in LD, 3893 were most abundant during the day (day-phased, LD-DP), while 5160 were most abundant at night (night-phased, LD-NP) (Figure 2 and Dataset S2). Day-phased transcripts in LD were enriched in chlorophyll metabolic processes, photosynthesis, translation, and RNA metabolic processes, whereas night-phased transcripts were enriched in response to stress and protein ubiquitination (Dataset S2).

Of the oscillating transcripts in LL, 1075 were more abundant during the subjective day (day-phased, LL-DP) and 1327 were more abundant during the subjective night (night-phased, LL-NP) (Figure 2 and Dataset S2). The biological processes of photosynthesis and circadian rhythm were overrepresented in both subjective day- and
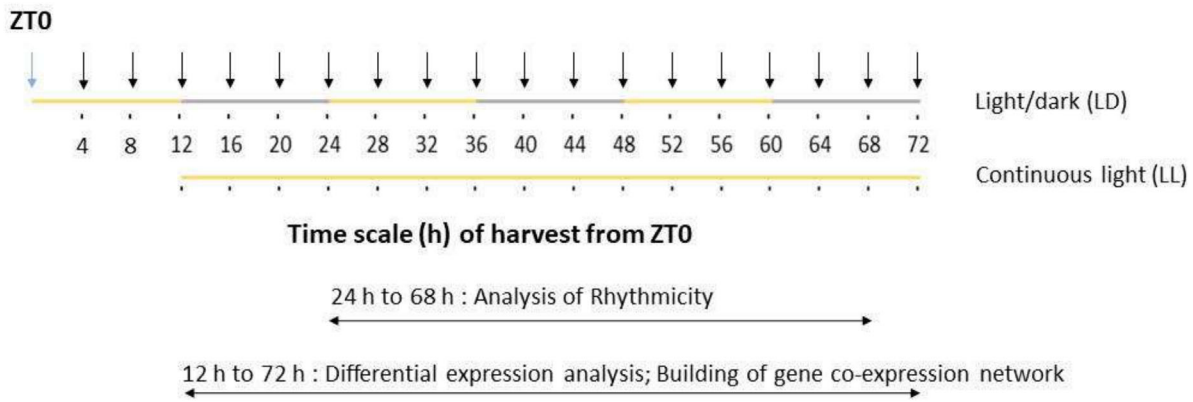
FIGURE 1 Experimental design. After 8 days of entrainment, Arabidopsis plants were either kept under 12/12-h light/dark (LD) cycles or transferred to constant light (LL). Samples were harvested every 4 h until 72 h. Two biological replicates were performed at each time point. Blue arrow, zeitgeber time 0 (ZT0); black arrows, sampling time point.

FIGURE 2 Upset plot of rhythmic genes in light/dark (LD) cycles and in constant light (LL). The number of genes in each intersection is indicated above the columns. Orange and black dots represent day-phased and night-phased genes, respectively. Rhythmicity (Bonferroni-corrected *p*-value < .01) was assessed from 24 to 68 h, both in LD and in LL. LD-DP and LD-NP, day-phased and night-phased genes in LD, respectively; LL-DP and LL-NP, subjective day- and subjective night-phased genes in LL, respectively.
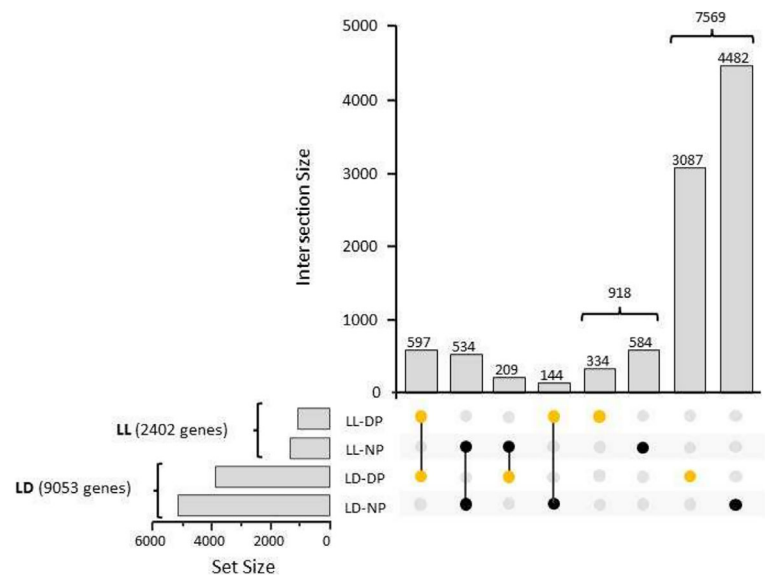
TABLE 1 Number of rhythmic transcripts under light/dark (LD) cycles and constant light (LL) and distribution according to their phase. Transcripts were considered rhythmic if the adjusted *p*-value was below .01 with JTK_cycle. Two replicates were performed per time point and per condition.

	LD	LL
Phase during the day or subjective day	3893	1075
Phase at night or subjective night	5160	1327
Total	9053	2402

subjective night-phased genes in LL. Additionally, flavonoid metabolic processes and terpenoid metabolic processes were also overrepresented in subjective night-phased transcripts (Dataset S2).

The time of the peak phase sometimes differed between LL and LD. No overrepresented biological process was observed in rhythmic genes shifting from LD-NP to LL-DP. However, the response to light stimulus, especially the response to blue light, was significantly overrepresented in genes shifting from LD-DP to LL-NP (Dataset S2).

2.2 | Alterations in transcript abundance in LL relative to LD cycles

Beyond variations in the temporal patterns of transcript abundance between LD and LL conditions, our findings align with previous research indicating that both conditions impact the absolute abundance of transcripts.

At each time point, we identified the DEGs in LL relative to LD using DESeq2. In total, 5921 DEGs were detected (Dataset S1).

The number of DEGs was consistently higher during the first subjective night compared to the next subjective day, leading to an oscillatory pattern in DEG numbers over time (Figure 3a).

Clustering the DEGs allowed us to categorize them into two distinct clusters based on their abundance patterns during the subjective night in LL compared to the night in LD (Figure 3b). Cluster 1 comprised transcripts with higher abundance during the subjective night in LL, while Cluster 2 contained transcripts with lower abundance during the same period. Visualization of the relative expression of

transcripts in both clusters using their medoids (Figure 3c,d) revealed that transcript expression during the light period in LD mirrored that of LL. This observation was further suppo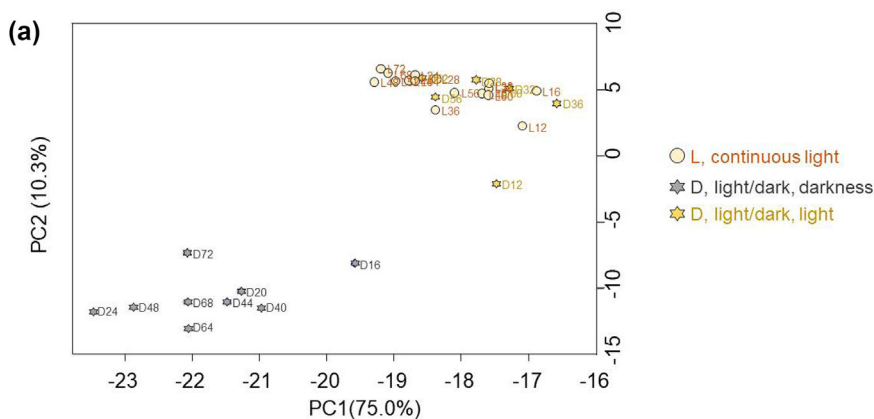rted by a principal component analysis (PCA), grouping samples harvested during the light period in LD with those in LL (Figure 4a).

The heatmap of the DEGs at each time point (Figure 3b) helped us identify genes with the highest and lowest differential expression in LL relative to LD. In the upper part of the heatmap (Cluster 1 in Figure 3b), *FLOWERING LOCUS T* (*FT*; AT1G65480) and its homolog

TWIN SISTER OF FT (*TSF*; AT4G20370) or *EARLY LIGHT-INDUCIBLE PROTEIN* (*ELIP1*; AT3G22840) were among the genes showing the highest differential expression over time (Figure 4b and Dataset S1). Additionally, several genes related to terpenes or flavonoids specialized metabolites, such as *TERPENE SYNTHASE 04* (*TPS04*; AT1G61120) and *PRODUCTION OF ANTHOCYANIN PIGMENT 2* (*PAP2/MYB90*; AT1G66390), were identified among the top more expressed DEGs. These genes were differentially expressed as early as 16 h after the start of the experiment (Figures 4b and 5a). Finally,
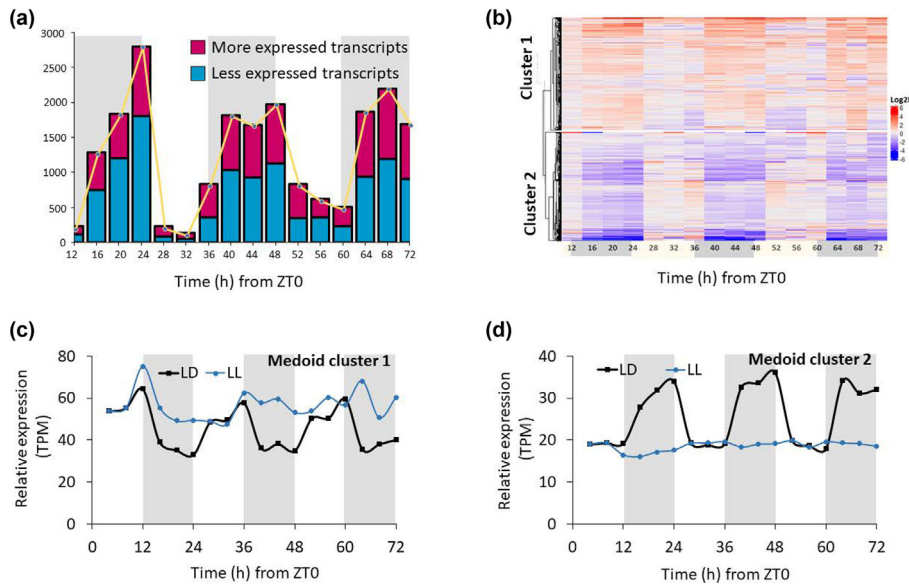

FIGURE 3 Differential expression analysis. (a) Number of differentially expressed genes (DEGs) in constant light (LL) relative to light/dark (LD) cycles at each time point from 12 to 72 h. The yellow line indicates the trend of the total number of DEGs. (b) Differential expression over time (log$_2$ fold change, log$_2$ FC) of DEGs (|log$_2$ FC > 1|, p < .01). DEGs were separated into two clusters, and medoids for Clusters 1 and 2 are represented in (c) and (d), respectively. Shaded areas in (a), (c), and (d) and shaded time points in (b) correspond to night or subjective night. ZT, zeitgeber time.

FIGURE 4 Principal component analysis (PCA) and heatmap of some differentially expressed genes (DEGs). (a) PCA of samples in light/dark (LD) and constant light (LL) according to PC1 and PC2, which contributed the most to separation between samples. (b) DEGs in LL relative to LD, among the more expressed or the less expressed according to the heatmap in Figure 3b. Asterisks indicate a significant difference in LL relative to LD (p-value < .01 and absolute log$_2$ FC > 1). n = 2 per time point and per condition. Full data are available in Dataset S1. Gray time points and white time points indicate subjective day and subjective night, respectively.
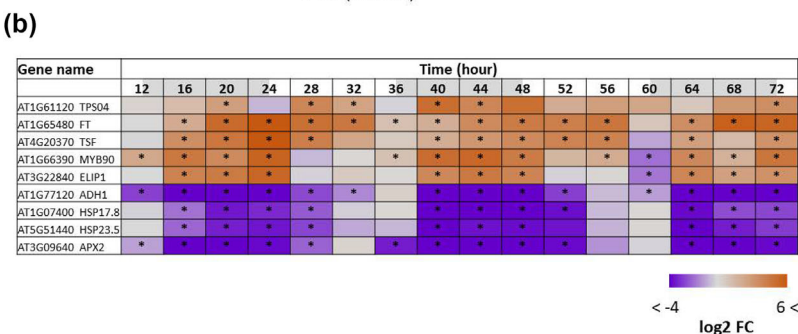
FIGURE 5 Relative expression in transcripts per million (TPM) of some of the more expressed (a) or less expressed (b) genes in constant light (LL) relative to light/dark (LD) cycles. White area, day (in LD) or subjective day (in LL); gray area, night (in LD) or subjective night (in LL). The full dataset is available in Dataset S1.
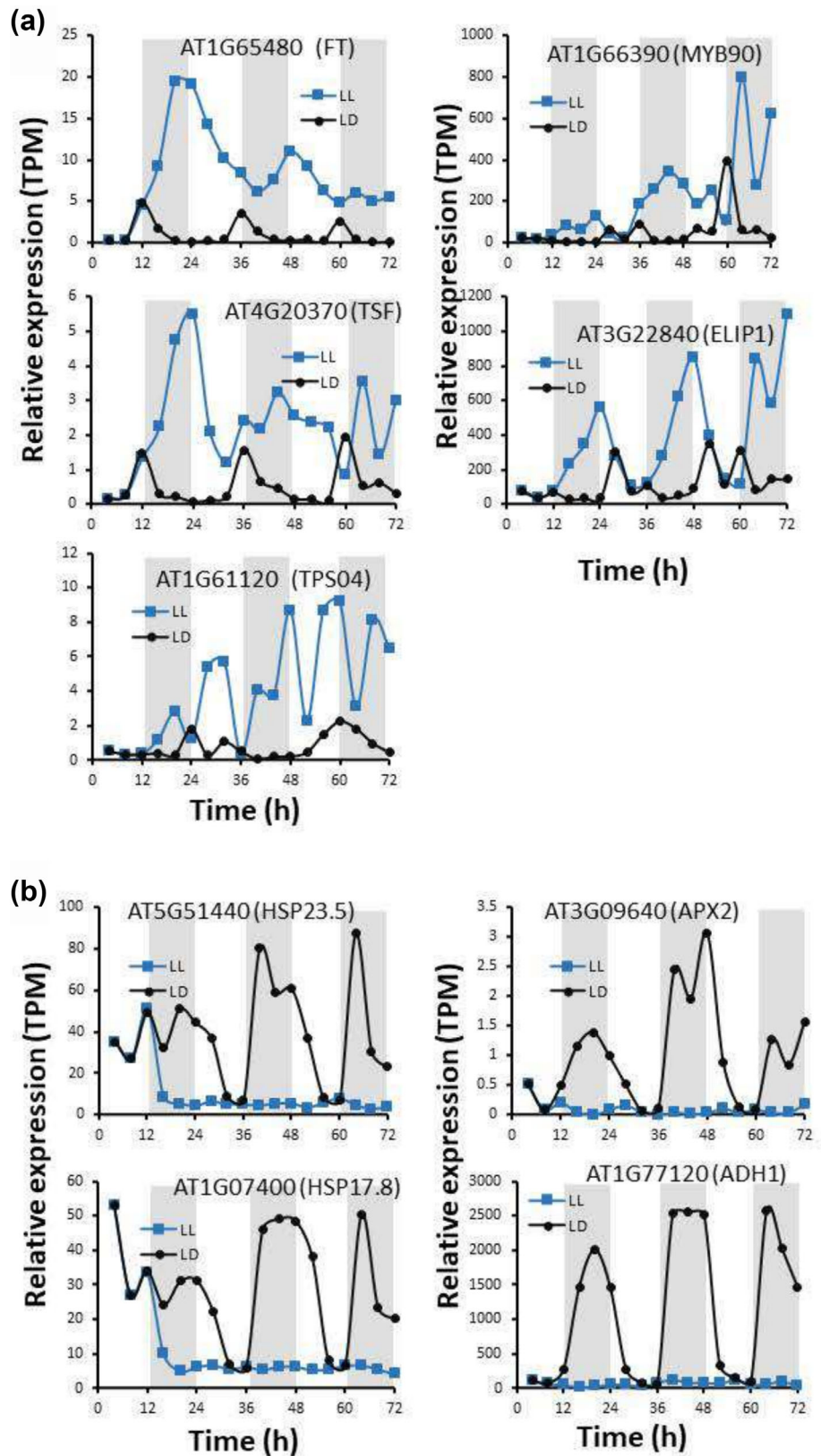
genes involved in starch metabolism, such as those encoding the iso-forms of *ADP-GLUCOSE PYROPHOSPHORYLASE* (*APL2, APL3,* and *APL4*), were globally more abundant in LL relative to LD (Dataset S1).

In the bottom part of Cluster 2 (Figure 3b), the least expressed genes in LL relative to LD were mainly related to stress, such as

ASCORBATE PEROXIDASE 2 (*APX2*; AT3G09640), *ALCOHOL DEHY-DROGENASE 1* (*ADH1*; AT1G77120), and several *HEAT SHOCK PRO-TEINS* (*HSPs*) (Figure 4b and Dataset S1). Their expression was low and relatively constant under LL, while peaking at night, showing diel oscillations in LD (Figure 5b).

To gain deeper insights into the differential expression between LL and LD, we conducted a comprehensive analysis to identify the overrepresented biological processes among DEGs at each time point (Fisher test, Bonferroni-corrected p-value < .05) (Figure S2A,B).

DEGs exhibiting higher abundance in LL relative to LD were associated with overrepresented categories linked to plant response to stress, particularly during the subjective day, and flavonoid biosynthetic processes during the subjective night.

Conversely, DEGs displaying lower abundance in LL relative to LD were predominantly observed during the subjective night and were mostly linked to plant responses to stress, similarly to the more abundant genes. These stress-related DEGs were plotted on a heatmap, and they could be separated into five clusters (Figure 6).

2.3 | Gene co-expression networks as tools to investigate diel and circadian control of biological processes

To identify co-expressed genes, we built two distinct networks, one for LD (11,079 co-expressed transcripts) and another for LL (8962 co-
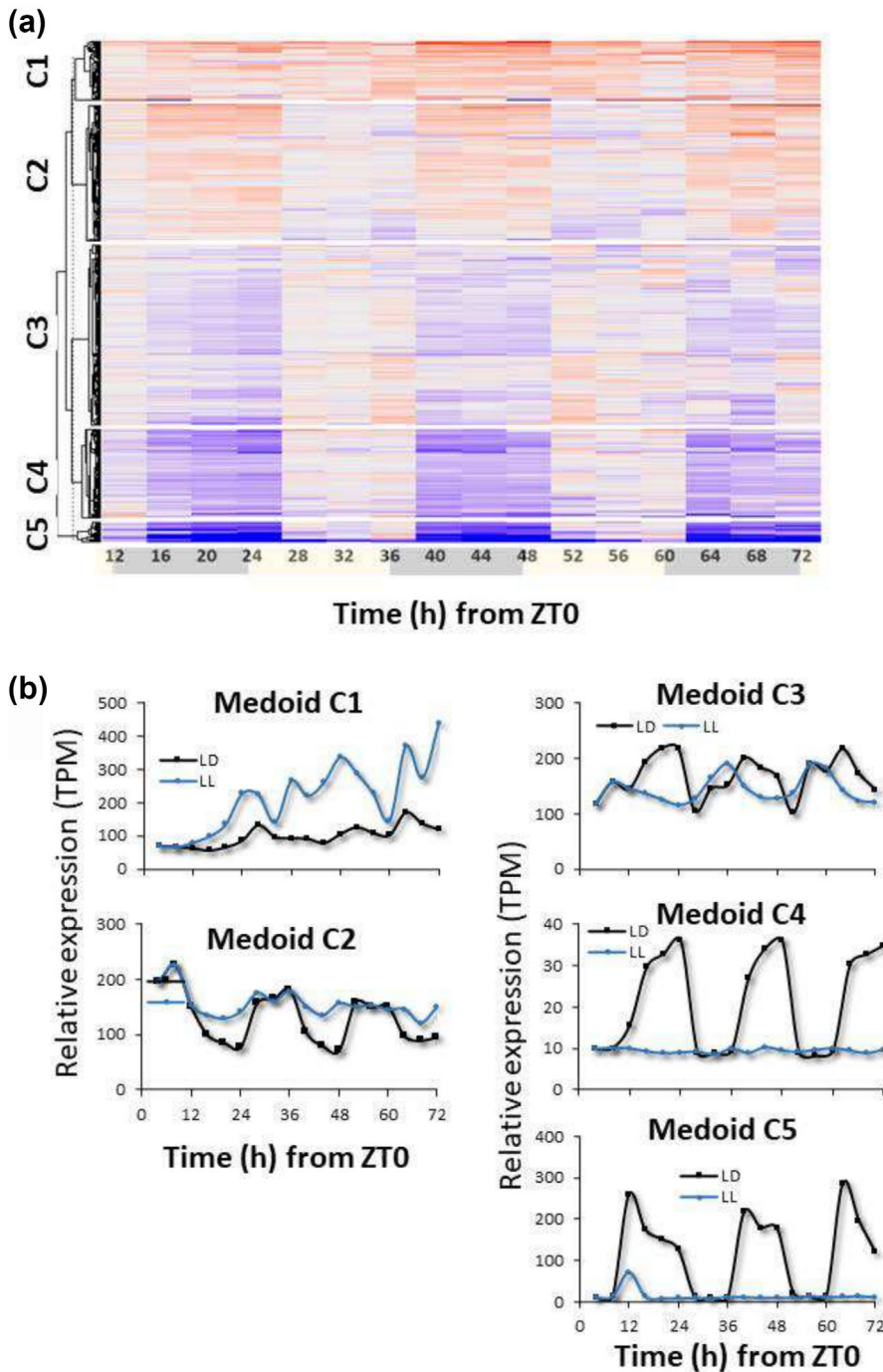
FIGURE 6 Subset of differentially expressed genes (DEGs) in constant light (LL) relative to light/dark (LD) and involved in response to stress. (a) Heatmap representing the $\log_2$ fold change ($\log_2$ FC) of stress-related DEGs. The letter "C" refers to "cluster." Shaded time points correspond to the subjective night. (b) Medoids of the clusters in (a). Their relative expression is indicated in transcripts per million. Clusters 1 and 2 corresponded to more expressed genes, and Clusters 3–5 corresponded to less expressed genes in LL relative to LD.

expressed transcripts) (correlation > .9, adjusted *p*-value < .01; Figures S3 and S4).

Statistics associated with each gene, including degree (number of connections of a node) and betweenness (number of times that a path passes through the node), were calculated, and genes were ranked based on these parameters. Highly ranked genes (nodes in the network) had a high degree, a high betweenness, or both (Dataset S1). Then, genes were clustered according to their expression patterns, with the largest clusters named after the biological processes that were most overrepresented. We could distinguish between LD and LL clusters such as "photosynthesis and circadian rhythm," "response to stress," and "RNA metabolic processes, transcription and translation" (Datasets S3 and S4 and Figures S3 and S4).
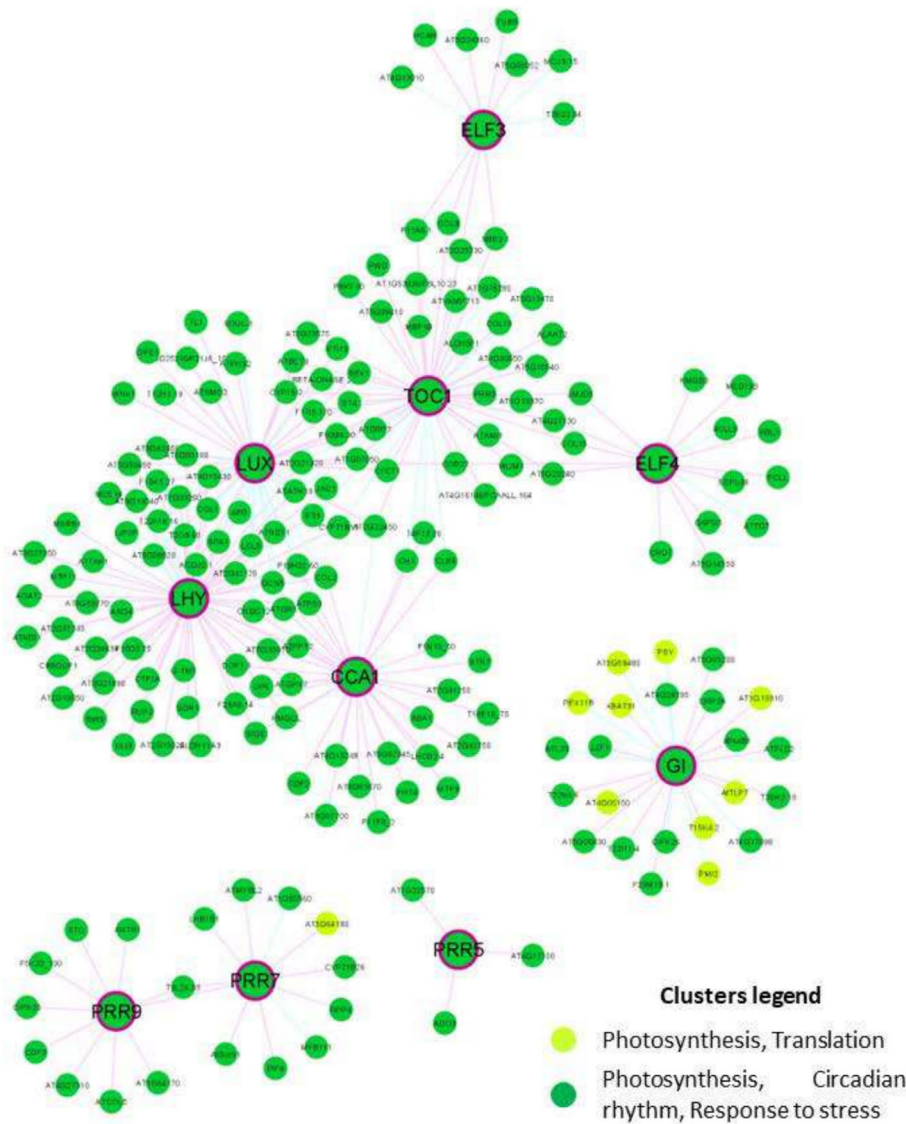
In the LD network, genes of the core oscillator were present in the "photosynthesis and circadian rhythm" cluster, except *PRR7* and *GI*, which were present in the "response to stress" cluster (Figure 7). Additionally, the first neighbors of *CCA1* and *LHY* were involved in the response to blue light; the first neighbors of *PRR9* were mostly involved in photosynthesis and response to light stimulus; and the first neighbors of *ELF4* were involved in the starch metabolic process and response to temperature stress (Dataset S3). *GI* emerged as a hub as it was co-expressed with nearly 400 genes. The subcluster formed by *GI* and its first neighbors was associated with biological processes such as response to stimulus, response to stress, and response to oxygen levels (Dataset S3).

The LL condition provoked several changes in the gene co-expression networks. In the LL network, all core genes of the oscillator belonged to the cluster "photosynthesis, circadian rhythm, response to stress" (Figure 8 and Dataset S1). Additionally, the first neighbors of *LHY* were involved in photomorphogenesis, and the first neighbors of *LUX* were involved in the regulation of photomorphogenesis (Dataset S4). However, by far the most striking change was that *GI* no longer served as a central hub (connected to only 24 genes), and no functional category was significantly overrepresented among its
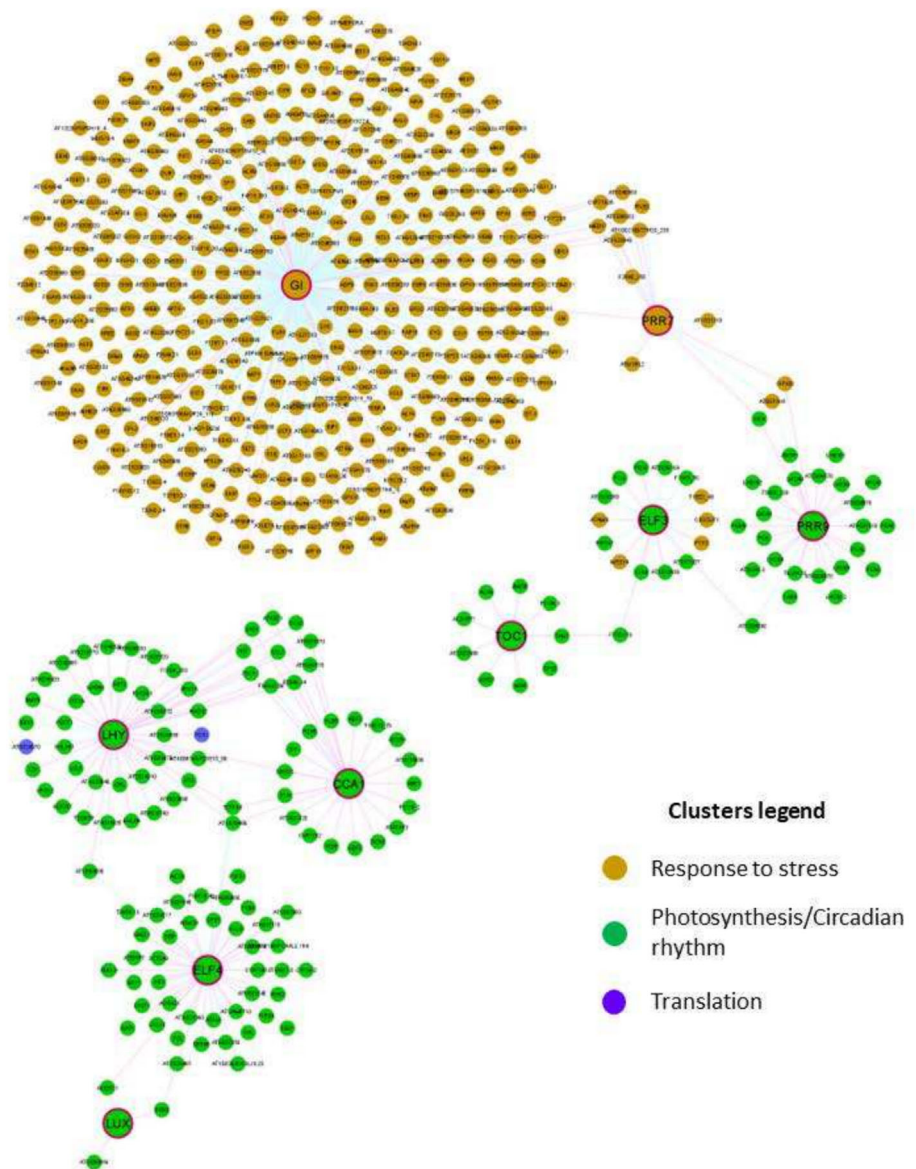
FIGURE 7 Co-expression network of the core oscillator genes in the light/dark (LD) network. The genes of the core oscillator that are represented have a bigger size and a red border. Blue edges and red edges represent negative and positive correlations, respectively. The threshold for significant correlation was set to .9, *p*-value < .01.

Clusters legend

● Response to stress

● Photosynthesis/Circadian rhythm

● Translation

FIGURE 8 Co-expression network of the core oscillator genes in the constant light (LL) network. The genes of the core oscillator that are represented have a bigger size with a red border. Blue edges and red edges represent negative and positive correlations, respectively. The threshold for significant correlation was set to .9, *p*-value < .01.

first neighbors. LL also promoted new interactions between genes in the core oscillator. For instance, *TOC1* and *LUX* were now in direct or indirect connections with *CCA1* and *LHY* (Figures 7 and 8).

LL also revealed a strong connection between *PRR7* and *MYB DOMAIN PROTEIN 111* (*MYB111*; AT5G49330), a gene involved in the regulation of flavonol biosynthesis. While *MYB111* was not oscillating in LD, it showed rhythmicity in LL (Dataset S1). Additionally, 24 genes involved in the flavonoid metabolic process behaved similarly to *MYB111*, belonging to the cluster of rhythmic genes in LL. Indeed, the number of DEGs in the flavonoid metabolic process within the "photosynthesis/circadian rhythm" cluster increased from 12 in the LD to 27 in the LL network (Figure S5). The JTK_cycle analysis confirmed this observation, as rhythmic DEGs related to flavonoid metabolism increased from 13 in LD to 30 in LL (Dataset S1).

Finally, *AT2G25530*, *AT4G30470*, or *AT5G62210*, which are not yet annotated, displayed strong connections to flavonoid-related genes, suggesting a potential role in the flavonoid metabolic process.

2.4 | Predictions from a third network

We built a third network, the LD–LL network, by combining gene expression data from LD and LL conditions (Figures 9 and S6). In this network, 8362 genes were co-expressed (correlation > .9, adjusted *p*-value < .01). Genes that had the three highest degrees and that belonged to the most highly ranked were *MA3 DOMAIN-CONTAINING TRANSLATION REGULATORY FACTOR* (*MRF1*; AT5G63190), *LIGHT-RESPONSE BTB 1* (*LRB1*; AT2G46260), and *OBERON1* (*OBE1*; AT3G07780) (Dataset S1). The *MRF1* transcript was shown to be induced by light and starvation (Lee et al., 2017). *LRB1* is involved in the regulation of cryptochrome-dependent and phyB-dependent photoresponses (Chen et al., 2021). *OBE1* participates in the maintenance of the root and shoot apical meristem (Saiga et al., 2008) and may contribute to the early flowering phenotype of Arabidopsis plants grown under LL (Eddy & Hahn, 2008). Among genes that had a positive correlation with *OBE1*, we observed an overrepresentation of
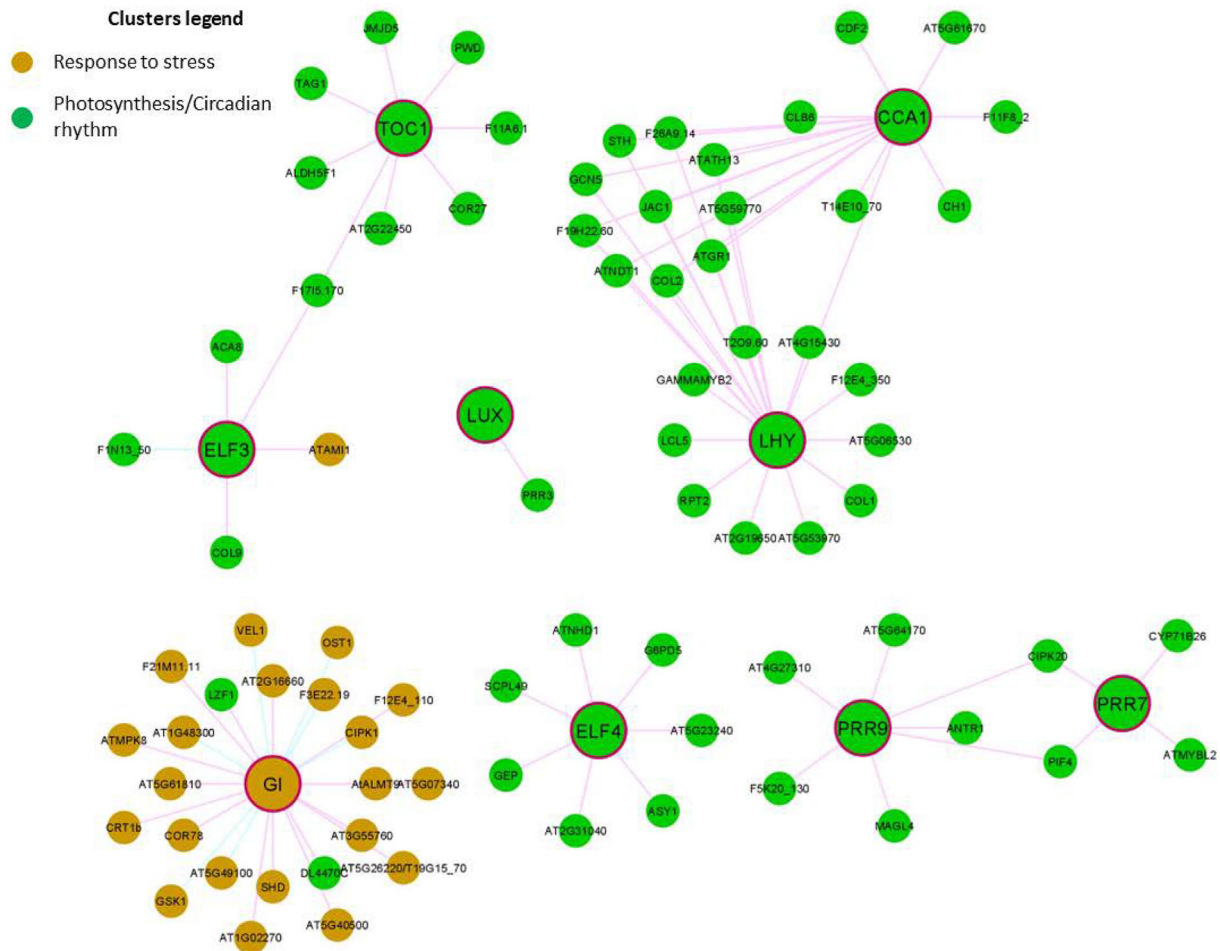
FIGURE 9 Co-expression network of the core oscillator genes in the light/dark and constant light (LD–LL) networks. The genes of the core oscillator that are represented have a bigger size with a red border. Blue edges and red edges represent negative and positive correlations, respectively. The threshold for significant correlation was set to .9, *p*-value < .01.

genes involved in the vegetative-to-reproductive phase transition of the meristem (adjusted *p*-value = .041). Moreover, *OBE1* was co-expressed with *REBELOTE* (*RBL*; AT3G55510), a gene encoding a protein known to physically interact with OBE1, which is involved in floral development (Prunet et al., 2008).

2.5 | Relationship between transcript oscillations and specialized metabolite content

Noting the augmented number of rhythmic genes associated with flavonoid metabolism under LL compared to LD cycles, along with their high connectivity within the rhythmic cluster, we investigated whether the content of flavonoids might be induced upon transfer to LL. Our findings revealed a significant increase in the relative content of several flavonoids just 24 h after plants were transferred from LD to LL (Table 2 and Figure S7). However, despite the observed increase in flavonoid content, we did not observe oscillations corresponding to the fluctuating transcript abundance under LL. Furthermore, we observed a significant upregulation of transcripts associated with

flavonoid metabolism, such as *PAP2* and *FLAVONOL SYNTHASE 2* (*FLS2*; AT5G63580) (Dataset S1). Similarly, flavonoid abundance did not exhibit oscillations in LD, where fewer fluctuations in flavonoid-related transcripts were observed.

2.6 | The Cyclath web application

We have developed the "Cyclath" web application, accessible at https://cyclath.shinyapps.io/cyclath/, providing easy access to the rhythmicity analyses, differential expressions, and gene co-expression networks presented in this study. Figure 10 showcases a glimpse of the application homepage. Cyclath offers users the ability to input the AGI number of a gene of interest, select the desired light regime, and, upon submission, explore analyses for LD, LL, and LD–LL conditions. Users also have the option to download the results upon completion.

Key features in Cyclath include (i) the construction of networks from data obtained in a single experimental design; (ii) the utilization of RNA-seq data; (iii) the visualization of the dynamic nature of co-expression networks under varying light regimes; (iv) the

TABLE 2 Relative content of flavonoids in light/dark (LD) and constant light (LL) and their rhythmicity assessment.

Condition	Flavonoid	Time (h)							Rhythmicity			
		24	28	32	36	40	44	48	p-value	Period	Phase	Amplitude
LD	Kaempferol-3-H-7-H	17.39 ± .16	16.02 ± .00	17.78 ± .21	16.73 ± .42	16.68 ± .61	16.64 ± .46	15.86 ± .29	1.00	24.00	6.00	.25
LD	Kaempferol-3-H-deoxyH	11.61 ± .30	10.71 ± .03	1.90 ± .34	11.26 ± .38	10.87 ± .02	11.21 ± .17	10.88 ± .57	1.00	20.00	16.00	.28
LD	Kaempferol-3-H-H-7-H	9.11 ± .50	8.21 ± .01	8.66 ± .38	8.48 ± .40	8.12 ± .74	8.74 ± .12	8.20 ± .61	1.00	20.00	16.00	.28
LD	Quercetin-3-G-R	1.33 ± .06	1.23 ± .06	1.20 ± .01	1.36 ± .13	1.41 ± .00	1.62 ± .17	1.34 ± .07	.07	24.00	20.00	.13
LD	Kaempferol-diH-R and quercetin-diH-R	.47 ± .02	.46 ± .04	.53 ± .06	.62 ± .09	.52 ± .05	.52 ± .04	.53 ± .02	1.00	20.00	12.00	.06
LD	Isorhamnetin-deoxyH-deoxyH	.45 ± .02	.34 ± .01	.41 ± .02	.40 ± .01	.53 ± .01	.59 ± .04	.45 ± .03	.08	24.00	20.00	.10
LD	Kaempferol-H	.96 ± .03	1.02 ± .01	1.03 ± .00	.91 ± .02	.95 ± .03	1.00 ± .01	1.02 ± .01	.14	20.00	6.00	.06
LD	Quercetin-3-G-R-R	.49 ± .03	.52 ± .04	.57 ± .05	.63 ± .04	.68 ± .05	.52 ± .11	.54 ± .06	.21	24.00	14.00	.08
LL	Kaempferol-3-H-7-H	23.47 ± 1.42	24.46 ± .38	23.57 ± .26	23.11 ± .56	23.32 ± .09	23.15 ± .13	24.52 ± .09	1.00	20.00	6.00	.49
LL	Kaempferol-3-H-deoxyH	14.67 ± .00	14.95 ± .56	14.66 ± .32	14.65 ± .24	14.35 ± .07	13.98 ± .21	15.29 ± .07	.55	24.00	8.00	.22
LL	Kaempferol-3-H-H-7-H	10.70 ± .50	10.88 ± .41	11.16 ± .18	11.78 ± .51	10.89 ± .11	10.34 ± .22	11.51 ± .07	.13	20.00	12.00	.29
LL	Quercetin-3-G-R	1.38 ± .07	1.41 ± .14	1.41 ± .01	1.77 ± .21	1.33 ± .02	1.39 ± .05	1.37 ± .23	1.00	20.00	10.00	.09
LL	Kaempferol-diH-R and quercetin-diH-R	.58 ± .06	.49 ± .03	.61 ± .02	.53 ± .01	.56 ± .00	.59 ± .01	.55 ± .01	1.00	24.00	22.00	.03
LL	Isorhamnetin-deoxyH-deoxyH	.44 ± .01	.45 ± .00	.43 ± .01	.54 ± .09	.40 ± .06	.46 ± .12	.49 ± .15	1.00	24.00	18.00	.01
LL	Kaempferol-H	1.36 ± .00	1.27 ± .04	1.33 ± .03	1.39 ± .08	1.27 ± .01	1.23 ± .07	1.35 ± .07	1.00	20.00	10.00	.03
LL	Quercetin-3-G-R-R	.74 ± .01	.84 ± .04	.73 ± .01	.84 ± .05	.83 ± .10	.86 ± .11	.75 ± .03	1.00	24.00	18.00	.05

Note: At each time point, the relative content ± standard error (SE) is shown, $n = 2$. A paired Student's t test was performed over the time course to evaluate significantly different metabolites in LL relative to LD, which are highlighted (p-value < .05).

Abbreviations: G, glucose; H, hexose; R, rhamnose.

RIVIÈRE ET AL.

American Society of Plant Biologists · Cultivating a better future through plant biology research · S·E·B SOCIETY FOR EXPERIMENTAL BIOLOGY · WILEY | 11 of 18

FIGURE 10 Homepage of the Cyclath web application. Upon entering an Arabidopsis AGI number, the user can choose to display the data for "light/dark cycles," "constant light," or "light/dark and constant light." In this last option, the app will display whether the gene is differentially expressed in constant light relative to light/dark. The outputs are rhythmicity, differential expression, and co-expression networks.

implementation of a stringent threshold for significant interactions; (v) the centralization of rhythmic analyses, differential expression, and co-expression networks within a unified platform; and (vi) the independent construction of LD and LL networks from separate datasets.

Several other databases exist for Arabidopsis co-expression networks, such as Genemania, Genevestigator, and Arabidopsis Co-expression Tool (ACT). The ACT database and Genevestigator integrated multiple public transcriptomic data to build the co-expression networks. Genemania networks were built from transcriptomics and proteomics data from multiple publicly available datasets.

We compared our networks with the ones from these databases, taking as examples the networks of *CCA1* and *LHY* (Dataset S5). *LHY* and *LNK2* were the only genes consistently co-expressed with *CCA1* in all six network databases. Among genes co-expressed with *LHY*, *B-BOX DOMAIN PROTEIN 19* (*BBX19*), *CONSTANS-LIKE* (*COL*) *1* (*COL1*), *COL2*, and *REVEILLE 8* (*RVE8*) were present in five out of six networks.

3 | DISCUSSION

3.1 | The light regime affects transcriptome rhythmicity

In this study, we investigated the effect of light regimes on changes in transcript rhythmicity, differential transcript expression, and co-expression of transcripts with the view of identifying new players for light-related biological processes and providing a database. For this purpose, Arabidopsis plants exposed to LL were compared to plants in LD both transcriptionally and metabolically.

In LL, we found a lower number of rhythmic transcripts in comparison to other studies in Arabidopsis based on Illumina sequencing (Bonnot & Nagel, 2021; Romanowski et al., 2020). Variation between the amount and specific circadian-regulated transcripts is common between laboratories and may result from different experimental conditions, including light intensity, quality, and media composition, all known to affect circadian dynamics and gene expression (Webb et al., 2019). Another possible explanation for the lower number of rhythmic transcripts in our study compared to others might be the lower magnesium (Mg) concentration in our media (200 μM) compared to most other studies (1500 μM), as a reduction in Mg content was associated with a reduced circadian rhythmicity (de Melo et al., 2021; Feeney et al., 2016). However, 200 μM of Mg is sufficient for normal Arabidopsis growth and sustains the normal circadian period for at least 8 days (de Melo et al., 2021).

As expected and in line with previous research (e.g., Müller et al., 2020), a decrease in the abundance of rhythmic transcripts was observed under LL compared to LD. This disparity may be attributed to the fact that rhythmicity in LL primarily mirrors the outputs of the circadian oscillator, whereas in LD, it is influenced by both circadian rhythms and direct light regulation. While 1484 rhythmic genes in LD remained rhythmic in LL, unexpectedly, 918 genes that were not rhythmic in LD became rhythmic in LL. These genes were associated with responses to light stimuli and flavonoid metabolism. In strong support of this finding, genes that were rhythmic only in LL (Mockler

et al., 2007) in the DIURNAL database showed an overrepresentation in the same biological processes (Dataset S2). However, DIURNAL, which studies gene expression under various light regimes, does not analyze gene differential expression between these regimes.

We found that in LL, there was a significant induction of genes related to starch content and flavonoid metabolism. The induction of starch metabolism-related genes in LL relative to LD is consistent with the increased starch content in LL relative to conditions of fluctuating light observed by Gollan et al. (2023). The induction of flavonoid gene expression is likely associated with light stress caused by LL exposure (J. Huang et al., 2019; Kleine et al., 2007), and once this expression has been activated, the connection to circadian regulation becomes evident (Hildreth et al., 2022; Pérez-García et al., 2015).

3.2 | Flavonoid-related transcripts that are differentially expressed are more closely connected to the circadian clock under LL conditions

Major classes of flavonoids include, but are not limited to, chalcones, flavonols, and anthocyanins (Corso et al., 2020; Falcone Ferreyra et al., 2021). The first genes involved in the anthocyanin biosynthesis pathway include *TRANSPARENT TESTA* (*TT*) *4* (*TT4/CHS*), *TT5* (*CHI*), and *FLAVANONE 3-HYDROXYLASE* (F3H), which are later followed by *TT3* (*DFR*), *TT18* (*LDOX*), and *UDP-GLUCOSYL TRANSFERASE 79B1* (*UF3GT/UFGT79B1*) (Tanaka et al., 2008). Several transcription factors (TFs) are responsible for the spatiotemporal regulation of flavonoid production (Morita et al., 2006), with some like *REVEILLE8* (*RVE8*) belonging to the core oscillator of the circadian clock (Pérez-García et al., 2015).

Pérez-García et al. (2015) notably showed the interaction between *REVEILLE 8* (*RVE8*; AT3G09600) and several *NIGHT LIGHT-INDUCIBLE AND CLOCK-REGULATED GENES* (*LNKs*) to regulate anthocyanin content. The presence of oscillations in the expression of flavonoid genes in LL but not in LD might rely on *SnRK1* (AT3G01090). Broucke et al. (2023) demonstrated the role of *SnRK1* in tightly regulating anthocyanin content to avoid wasting energy during the daytime. However, this tight regulation is expected to loosen when light is on during the subjective night, provoking an increased expression of anthocyanin-related genes during this period (Pérez-García et al., 2015). Under LL, circadian regulation became prominent, with up to 22 flavonoid-related genes that were not rhythmic in LD becoming rhythmic in LL (Dataset S1). In strong support, in the LL network, genes involved in anthocyanin biosynthesis were observed to subcluster together in the circadian rhythm cluster (Figure S5). Along with these transcriptomic changes, we observed a significant increase in flavonoid content under LL relative to LD (Table 2 and Figure S7). Consistent with these results, Maier et al. (2013) found a stabilizing effect of light on *PRODUCTION OF ANTHOCYANIN PIGMENT* (*PAP*) *1* (*PAP1* or *MYB75*; AT1G56650) and *PAP2* (*MYB90*; AT1G66390). *PAP1* was additionally present in the subcluster of anthocyanin-related genes in the LL network.

3.3 | Stress-related genes are induced under darkness to confer "skotoprotection"

Excessive light stress has been associated with an induction of genes involved in hormone biosynthesis, terpenoids, and flavonoid metabolism (J. Huang et al., 2019). We made a similar observation for genes exhibiting higher abundance in LL relative to LD (Cluster 1 of Figure 3b and Dataset S1). However, among genes showing less abundance in LL relative to LD (Cluster 2 of Figure 3b and Dataset S1), we observed an overrepresentation of genes involved in response to stress and related categories, such as response to oxidative stress and response to external stimulus. These overrepresented categories were mainly observed for DEGs during the subjective night, suggesting that it is the unexpected light during the subjective night that results in the stress.

In addition to our observation of more DEGs during the subjective night as compared to the subjective day (Figure 3a,b), we found that around 600 stress-related genes were repressed in LL relative to LD (Figures 5b and 6a). For most of these genes, the peak of relative expression occurred during the night in LD (medoids of Clusters 3–5 of Figure 6b). These genes peaking in the dark in LD might be associated with the stress caused by darkness, which does not occur in LL. Our analysis pipeline has allowed the reporting of this observation for the first time, to our knowledge. We propose the term "skotoprotection" to refer to those stress-related genes that are specifically induced under darkness. These genes were characterized by three criteria: (i) They were rhythmic in LD and non-rhythmic in LL; (ii) they were not differentially repressed during the subjective day (in LL) relative to the day (in LD); and (iii) they were upregulated during the night (in LD) relative to the subjective night (in LL). Examples of such genes include *APX2*, *ADH1*, *HSP17.8*, and *HSP23.5* (Figure 5b).

3.4 | Gene regulation dynamics change under different light regimes

We explored the dynamic response of the transcriptome in response to different light regimes by building co-expression networks (Contreras-López et al., 2018). Our analysis revealed shifts in the connections between nodes (genes) in the LD and LL networks, as demonstrated in Figures S3 and S4. Notably, core circadian oscillator components exhibited altered connections across LD and LL networks (see Figures 7–9). Co-expression networks of *CCA1* and *LHY* circadian oscillators highlighted an overrepresentation of rhythmic processes in all three networks (LD, LL, and LD–LL) (Dataset S5), aligning with their known role in regulating rhythmic processes. The shared first neighbors between *CCA1* and *LHY* underscored their partial functional redundancy (Mizoguchi et al., 2002). Furthermore, another component of the core oscillator, *GI* (Martin-Tryon et al., 2007; Nohales & Kay, 2020), which is involved in response to stress (Mishra & Panigrahi, 2015), emerged as a hub linked to stress-response genes specifically in the LD network, hinting at a dependency of *GI* expression and co-expression networks on light regimes.

We applied a stringent threshold to determine the first neighbors (Pearson correlation coefficient > .9 and adjusted *p*-value < .01), so we cannot discard the hypothesis that a lower threshold would have led to a different outcome, which is a limitation of most network reconstruction approaches. Comparisons with other Arabidopsis co-expression databases such as Genemania, Genevestigator, and ACT (Dataset S5) provided insights into the consistency of our networks. For instance, *LHY* and *LNK2* consistently co-expressed with *CCA1* across all databases, while genes like *BBX19*, *COL1*, *COL2*, and *RVE8* exhibited high consistency in co-expression with *LHY* in five out of six networks.

Our analysis not only uncovered mechanisms underlying plant responses to LL, such as the early flowering phenotype in Arabidopsis (Eddy & Hahn, 2008), but also identified potential new players in light response and gene function. Genes such as *FT* and *TSF*, known to promote flowering (Kardailsky et al., 1999; Yamaguchi et al., 2005), were among the most induced genes in LL compared to LD conditions. Additionally, *OBE1*, implicated in meristem maintenance, exhibited positive co-expression with flowering-related genes in the LD–LL network, suggesting its involvement in flowering regulation.

Furthermore, observations regarding genes like *ELIP1*, co-expressed with eight genes related to flavonoid metabolism, and top DEGs like *TPS04*, involved in plant defense against abiotic stress through terpene metabolism, underscore the utility of our approach in predicting gene functions and having novel biological insights.

We provide a web application called *Cyclath*, which hosts all the data presented in this study. This database contains the relative expression of Arabidopsis genes under two different light regimes, namely, diel cycles and LL. This database also shows differential expression between both conditions, as well as an analysis of their rhythmicity. The results obtained through Cyclath can thus complement other databases, such as DIURNAL.

Cyclath can additionally be used in combination with other available transcriptomic resources, such as *Genemania*, *ACT*, and *Genevestigator*. However, it is important to point out that those databases were built combining data from several different studies, while the high amount of data generated in our study allowed us to build Cyclath networks from a unique experimental design. An advantage of our approach is the attainment of a Pearson correlation coefficient exceeding 90%. Predicted co-expressed genes derived from Cyclath networks can thus help build new research hypotheses and validate co-expression predictions. These hypotheses may also contribute to the functional validation of genes associated with skotoprotection. Ultimately, we envision Cyclath as a valuable tool for devising strategies to improve crop resilience to stressors.

4 | EXPERIMENTAL PROCEDURE

4.1 | Plant material and growth conditions

Arabidopsis thaliana (Arabidopsis) seed stock, ecotype Columbia-0 (Col-0), was supplied by the Nottingham Arabidopsis Stock Center (NASC) (Scholl et al., 2000).

Seeds were sterilized as previously described (De Caluwé et al., 2017) and sown individually on modified Murashige and Skoog (MS) media, solidified with .5% w/v Mg-free high gel strength agar (Sigma-Aldrich, Germany), containing 200 μM of Mg and 1% w/v sucrose (de Melo et al., 2021). After stratification for 2 days in the dark at 4°C, seedlings were entrained for 8 days at 20°C with a 12-h day/12-h night photoperiod with white light (~100 μmol photons m^{-2} sec^{-1}) in a growth cabinet (Panasonic MLR-352-PE, the Netherlands). Following entrainment, plantlets were transferred to liquid-modified MS media either in a day/night photoperiod (LD) or in LL. Pools of 20–50 plantlets were harvested after 30 min, 1 h, 2 h, 4 h, and then every 4 h until 72 h, with two replicates at each time point. The experimental design is shown in Figure 1. A total of 74 independent samples were harvested and sequenced by RNA-seq. Zeitgeber time 0 (ZT0) indicates the onset of light on the day the experiment started. Accordingly, all time points indicate the time after ZT0.

4.2 | RNA extraction

Total RNA was isolated from 100 mg of frozen ground tissues of whole plantlets. RNA was extracted with the Maxwell 16 LEV Plant RNA Kit (Promega, Benelux BV) using the Maxwell 16 AS2000 Instrument (Promega) according to the manufacturer's instructions. The quality and purity of the samples were verified with a NanoDrop 2000 UV–Vis Spectrophotometer (Thermo Scientific, UK).

4.3 | Library preparation and sequencing

Sequencing was conducted at VIB Nucleomics Core (VIB, Belgium, www.nucleomics.vib.be). Library preparation was performed with the TruSeq RNA Sample Preparation Kit (Illumina, USA), and the library was sequenced using the NextSeq High Output Kit in single-end mode (1 × 75 cycles) (Illumina).

4.4 | Read preprocessing

First, the fastq files were preprocessed to remove the adapter sequences by hard-clipping, using cutadapt 1.15 (Martin, 2011). Then, reads shorter than 35 bp after adapter trimming were removed. Using FastX 0.0.14 (HannonLab, 2010) and ShortRead 1.36.0 (Morgan et al., 2009), polyA-reads, ambiguous reads, low-quality reads, and artifact reads were also removed. Finally, reads that were aligned to phix_illumina (control for sequencing) were identified and removed with bowtie 2.3.3.1 (Langmead & Salzberg, 2012).

4.5 | Mapping of RNA-seq data

The preprocessed reads were aligned to the reference genome of *A. thaliana* (TAIR10) using STAR 2.5.2b (Dobin et al., 2013).

Subsequently, mapped reads that were multi-mappings or with low quality (more than 90% of the bases equal A, reads containing N, and more than 50% of the bases < Q25) were removed, and the reads were sorted according to the chromosomes with samtools 1.5 (Li et al., 2009).

4.6 | Summarization of expression levels

The number of reads in the alignment that overlapped with the gene features was determined using featureCounts 1.5.3 (Liao et al., 2014). Reads that were ambiguous or that could not be attributed to any gene were not counted. The raw counts per gene were extracted in an Excel sheet, and genes for which all samples had less than 1 count per million were removed.

Finally, the EDASeq package (Risso et al., 2011) in R (R-4.0.3) was used for within-sample normalization, followed by a between-sample normalization. The normalized counts were used to calculate the relative expression of each transcript in transcripts per million (TPM). Genes with total relative expression over the time course (from 4 to 72 h) below 1 TPM were excluded, resulting in a final list of 21,487 expressed transcripts used for all the following analyses.

4.7 | Gene differential expression

Differential expression was assessed with DESeq2 in R (Love et al., 2014), comparing gene expression in LL (treatment) with the expression in LD (control). At each time point, a gene was considered differentially expressed if $\log_2$ fold change ($\log_2$ FC) > 1 or $\log_2$ FC < −1, and adjusted p-value < .01 (Benjamini–Hochberg correction; Benjamini & Hochberg, 1995). Gene ontology enrichment analysis was performed with the Protein ANalysis THrough Evolutionary Relationships (PANTHER) classification system (http://pantherdb.org), using the Fisher exact test and the Bonferroni correction for multiple testing. Heatmaps were generated with Complexheatmap 2.6.2 (Gu et al., 2016), and genes were clustered with Cluster 2.1.4 in R (Struyf et al., 1997) using function *partitioning around medoids* (PAM). In our case, a medoid is a gene that represents the trend of other genes in the same cluster (Reynolds et al., 1992). Because samples harvested from 30 min to 8 h were identical between LD and LL, the differential expression analysis started at 12 h.

4.8 | Rhythmicity analysis

We assessed gene rhythmicity with the JTK_cycle algorithm included in the Metacycle package in R (Hughes et al., 2010; Wu et al., 2016). Minimum and maximum periods were set to 20 and 28 h, respectively. Relative expression in TPM of the two replicates per time point from 24 to 68 h was used as input, and genes were considered rhythmic if their adjusted p-value (Benjamini–Hochberg correction) was strictly lower than .01.

Rhythmic transcripts with phases between 0 and 12 h (0 ≤ phase < 12) were considered day phased, while those between 12 and 24 h (12 ≤ phase < 24) were considered night phased. The phase was recalculated according to Romanowski et al. (2020), where "recalculated phase = JTK_cycle phase × 24/JTK_cycle period."

We performed a comparison of our results with other published circadian studies (Covington et al., 2008; Hsu & Harmer, 2012; Romanowski et al., 2020), as well as with data present in the DIURNAL database (http://diurnal.mocklerlab.org; Mockler et al., 2007), which gathers diurnal and circadian microarray data for Arabidopsis and other model species. Among the seven datasets included in that database, we selected the light/dark dataset (LDHC) and the continuous light dataset (LLHC) for comparison as they had the most similar design to our study. Gene rhythmicity in LDHC and LLHC datasets was recalculated using the JTK_cycle algorithm with a significance threshold set to .05 (adjusted p-value), which is the value commonly used with microarray data.

To visualize intersections between rhythmic genes from LD and LL datasets, we used the UpsetR package in R (Conway et al., 2017; Lex et al., 2014).

4.9 | Gene co-expression networks

Gene co-expression networks were constructed in R, according to Contreras-López et al. (2018). Briefly, raw counts from 12 to 72 h in LD cycles, in LL, or a combination of both (LD time series followed by LL time series) were used to generate three distinct co-expression networks. Read counts were median normalized using R package EBSeq 1.30.0 (Leng et al., 2013), and the whole list of 21,487 expressed transcripts was analyzed. The value and significance of correlations between transcripts were then calculated using R package psych 2.2.9, selecting Pearson correlation as the method to analyze the normalized data (Revelle, 2017). To build the network, we selected only the genes for which an absolute correlation threshold strictly greater than .9 and an adjusted p-value strictly lower than .01. Before visualizing the networks, degree and betweenness centrality were calculated with the R package igraph 1.2.6 (Csardi & Nepusz, 2006) and used to compute a ranking statistic. Finally, Cytoscape software 3.8.21 was used to visualize the networks, and the cytoscape plug-in clusterMaker 2.0 (Morris et al., 2011) was employed to group genes and infer possible biological functions for those with unknown annotation.

4.10 | Determination of metabolite content

For the determination of metabolite content, Arabidopsis plantlets were grown on modified MS solid media in a 12-h photoperiod (LD) at 100 μmol m^{-2} sec^{-1}. After 17 days, they were either kept in LD cycles or transferred to LL. Whole leaves were harvested at 24, 28, 32, 36, 40, 44, and 48 h after transfer and flash frozen in liquid N$_2$. After grinding, they were separated into two aliquots of 50 mg of leaf

fresh weight per time point and condition. Polar and semipolar metabolites were then extracted. A total of 1.5 ml of methanol/water/acetone/TFA, V/V 40/32/28/.05%, and 200 ng of apigenin (used as an internal standard) were added to each sample. The mixtures were placed in an ultrasonic bath for 5 min at 25 kHz at 4°C and shaken for 30 min at 4°C using a ThermoMixer™ C (Eppendorf). These samples were then centrifuged at 20,000 g for 20 min. Extraction was performed twice with the same sample, and the two resulting supernatants were pooled and dried in a rotary evaporator and freeze dryer. The dry pellet was dissolved in 150 μl of water ULC/MS grade (Biosolve) with 10% of ACN ULC/MS grade (Biosolve) and filtered through a Whatman filter. Chromatography (ultra-performance liquid chromatography and tandem mass spectrometry [UPLC–MS/MS]) analysis of polar/semipolar specialized metabolites, UPLC–MS/MS data processing, and metabolite annotation were carried out according to Boutet et al. (2022), normalizing metabolite content per apigenin and per fresh weight.

We removed the metabolites with low quality (QC > 40%) and further normalized each metabolite content by the total ion count per sample as described elsewhere (da Silva et al., 2021; Deininger et al., 2011). We finally multiplied the result for each metabolite by a thousand. This normalization method is similar to the *TPM* normalization of RNA-seq. A paired Student's t test was performed to evaluate significant differences between metabolite content (p-value < .05).

4.11 | Comparison with other network databases

Co-expression networks for Arabidopsis were obtained from other databases, namely, *Genemania* (https://genemania.org/; Warde-farley et al., 2010), *Genevestigator* (GENEVESTIGATOR 9.6.1; Hruz et al., 2008), and *ACT* (https://www.michalopoulos.net/act/; Zogopoulos et al., 2021).

On *Genemania*, we selected genes that were co-expressed transcriptionally with our queries.

On *Genevestigator*, we selected the top 25 genes that were positively and negatively correlated with our queries, giving a final list of around 50 genes.

On *ACT*, we selected the default output.

AUTHOR CONTRIBUTIONS

ADA and NV conceived the project; QR and VR made the Cyclath app; ADA, RdM, and SB performed the experiments; ADA, SB, MC, MD, and NV analyzed the data; ADA, AARW, and NV wrote the paper, with contributions from all authors.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

PEER REVIEW

The peer review history for this article is available in the supporting information for this article.

DATA AVAILABILITY STATEMENT

The RNA-seq transcriptomic data and metadata have been deposited at NCBI and are available under the bioproject names PRJNA996904 and PRJNA1004191, respectively, at the following addresses: https://dataview.ncbi.nlm.nih.gov/object/PRJNA996904?reviewer=251hkm9c8m8i7qb4ld3eljlo14 and https://dataview.ncbi.nlm.nih.gov/object/PRJNA1004191?reviewer=h9lbp911etfr81evo908ld3ocm. The LC–MS/MS untargeted metabolomic data and metadata have been deposited at the Data Gouv (data INRAE) repository portal with the following identifier: https://doi.org/10.57745/2YJCCU.

ORCID

Quentin Rivière https://orcid.org/0000-0003-3522-6256
Virginie Raskin https://orcid.org/0009-0006-7094-085X
Romário de Melo https://orcid.org/0000-0002-3887-2301
Massimiliano Corso https://orcid.org/0000-0002-3243-1660
Alex A. R. Webb https://orcid.org/0000-0003-0261-4375
Nathalie Verbruggen https://orcid.org/0000-0003-2296-5404
Armand D. Anoman https://orcid.org/0000-0003-0043-2180

REFERENCES

Bassi, R., & Dall'Osto, L. (2021). Dissipation of light energy absorbed in excess: The molecular mechanisms. *Annual Review of Plant Biology*, *72*, 47–76. https://doi.org/10.1146/annurev-arplant-071720-015522

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bonnot, T., & Nagel, D. H. (2021). Time of the day prioritizes the pool of translating mRNAs in response to heat stress. *Plant Cell*, *33*, 2164–2182. https://doi.org/10.1093/plcell/koab113

Boutet, S., Barreda, L., Perreau, F., Totozafy, J., Gakière, B., Delannoy, É., Martin-Magniette, M-L., Monti, A., North, H., Zanetti F., & Corso, M. (2022). Untargeted metabolomic analyses reveal the diversity and plasticity of the specialized metabolome in seeds of different Camelina sativa genotypes. *The Plant Journal*, *110*, 147–165. https://doi.org/10.1111/tpj.15662

Broucke, E., Dang, T. T. V., Li, Y., Hulsmans, S., Van Leene, J., De Jaeger, G., Hwang, I., Van den Ende, W., & Rolland, F. (2023). SnRK1 inhibits anthocyanin biosynthesis through both transcriptional

regulation and direct phosphorylation and dissociation of the MYB/-bHLH/TTG1 MBW complex. *The Plant Journal*, *115*, 1193–1213. https://doi.org/10.1111/tpj.16312

Chen, Y., Hu, X., Liu, S., Su, T., Huang, H., Ren, H., Gao, Z., Wang, X., Lin, D., Wohlschlegel, J. A., Wang, Q., & Lin, C. (2021). Regulation of Arabidopsis photoreceptor CRY2 by two distinct E3 ubiquitin ligases. *Nature Communications*, *12*, 2155. https://doi.org/10.1038/s41467-021-22410-x

Contreras-López, O., Moyano, T. C., Soto, D. C., & Gutiérrez, R. A. (2018). Step-by-step construction of gene co-expression networks from high-throughput Arabidopsis RNA sequencing data. In *Methods in molecular biology* (Vol. 1761) (pp. 275–301). Humana Press Inc. https://doi.org/10.1007/978-1-4939-7747-5_21

Conway, J. R., Lex, A., & Gehlenborg, N. (2017). UpSetR: An R package for the visualization of intersecting sets and their properties. *Bioinformatics*, *33*, 2938–2940. https://doi.org/10.1093/bioinformatics/btx364

Corso, M., & de la Torre, V. S. G. (2020). Biomolecular approaches to understanding metal tolerance and hyperaccumulation in plants. *Metallomics*, *12*, 840–859. https://doi.org/10.1039/d0mt00043d

Corso, M., Perreau, F., Mouille, G., & Lepiniec, L. (2020). Specialized phenolic compounds in seeds: Structures, functions, and regulations. *Plant Science*, *296*, 110471. https://doi.org/10.1016/j.plantsci.2020.110471

Cortijo, S., Bhattarai, M., Locke, J. C. W., & Ahnert, S. E. (2020). Co-expression networks from gene expression variability between genetically identical seedlings can reveal novel regulatory relationships. *Frontiers in Plant Science*, *11*, 599464. https://doi.org/10.3389/fpls.2020.599464

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. https://doi.org/10.1186/gb-2008-9-8-r130

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems*.

da Silva, V. C. H., Martins, M. C. M., Calderan-Rodrigues, M. J., Artins, A., Monte Bello, C. C., Gupta, S., Sobreira, T. J. P., Riaño-Pachón, D. M., Mafra, V., & Caldana, C. (2021). Shedding light on the dynamic role of the "target of rapamycin" kinase in the fast-growing C4 species *Setaria viridis*, a suitable model for biomass crops. *Frontiers in Plant Science*, *12*, 492. https://doi.org/10.3389/fpls.2021.637508

Dalchau, N., Hubbard, K. E., Robertson, F. C., Hotta, C. T., Briggs, H. M., Stan, G.-B., Goncalves, J. M., & Webb, A. A. R. (2010). Correct biological timing in Arabidopsis requires multiple light-signaling pathways. *Proceedings of the National Academy of Sciences*, *107*, 13171–13176. https://doi.org/10.1073/pnas.1001429107

De Caluwé, J., de Melo, J. R. F., Tosenberger, A., Hermans, C., Verbruggen, N., Leloup, J. C., & Gonze, D. (2017). Modeling the photoperiodic entrainment of the plant circadian clock. *Journal of Theoretical Biology*, *420*, 220–231. https://doi.org/10.1016/j.jtbi.2017.03.005

de Melo, J. R. F., Gutsch, A., De Caluwé, T., Leloup, J. C., Gonze, D., Hermans, C., Webb, A. A. R., & Verbruggen, N. (2021). Magnesium maintains the length of the circadian period in Arabidopsis. *Plant Physiology*, *185*, 519–532. https://doi.org/10.1093/plphys/kiaa042

Deininger, S., Cornett, D. S., Paape, R., Becker, M., Pineau, C., Rauser, S., Walch, A., & Wolski, E. (2011). Normalization in MALDI-TOF imaging datasets of proteins: Practical considerations. *Analytical and Bioanalytical Chemistry*, *401*, 167–181. https://doi.org/10.1007/s00216-011-4929-z

Demmig-Adams, B., & Adams, W. W. III (1992). Photoprotection and other responses of plants to high light stress. *Annual Review of Plant Physiology and Plant Molecular Biology*, *43*, 599–626. https://doi.org/10.1146/annurev.pp.43.060192.003123

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635

Eddy, R., & Hahn, D. T. (2008). 101 ways to try to grow Arabidopsis: Does growth under 24-hour light hasten production? Can plants be damaged by it? *Purdue Methods for Arabidopsis Growth*. Paper 16. http://docs.lib.purdue.edu/pmag/16

Falcone Ferreyra, M. L., Serra, P., & Casati, P. (2021). Recent advances on the roles of flavonoids as plant protective molecules after UV and high light exposure. *Physiologia Plantarum*, *173*, 736–749. https://doi.org/10.1111/ppl.13543

Feeney, K. A., Hansen, L. L., Putker, M., Olivares-Yañez, C., Day, J., Eades, L. J., Larrondo, L. F., Hoyle, N. P., O'Neill, J. S., & Van Ooijen, G. (2016). Daily magnesium fluxes regulate cellular timekeeping and energy balance. *Nature*, *532*, 375–379. https://doi.org/10.1038/nature17407

Franklin, K. A., & Quail, P. H. (2010). Phytochrome functions in Arabidopsis development. *Journal of Experimental Botany*, *61*, 11–24. https://doi.org/10.1093/jxb/erp304

Gollan, P. J., Grebe, S., Roling, L., Grimm, B., Spetea, C., & Aro, E.-M. (2023). Photosynthetic and transcriptome responses to fluctuating light in Arabidopsis thylakoid ion transport triple mutant. *Plant Direct*, *7*, e534. https://doi.org/10.1002/pld3.534

Gu, Z., Eils, R., & Schlesner, M. (2016). Genome analysis. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*, 2847–2849. https://doi.org/10.1093/bioinformatics/btw313

HannonLab. (2010). *Fastx-toolkit*. Fastx-Toolkit.

Haydon, M. J., Román, Á., & Arshad, W. (2015). Nutrient homeostasis within the plant circadian network. *Frontiers in Plant Science*, *6*, 299. https://doi.org/10.3389/fpls.2015.00299

Hildreth, S. B., Littleton, E. S., Clark, L. C., Puller, G. C., Kojima, S., & Winkel, B. S. J. (2022). Mutations that alter Arabidopsis flavonoid metabolism affect the circadian clock. *The Plant Journal*, *110*, 932–945. https://doi.org/10.1111/tpj.15718

Hohmann-Marriott, M. F., & Blankenship, R. E. (2011). Evolution of photosynthesis. *Annual Review of Plant Biology*, *62*, 515–548. https://doi.org/10.1146/annurev-arplant-042110-103811

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., & Zimmermann, P. (2008). Genevestigator V3: A reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics*, *2008*, 420747. https://doi.org/10.1155/2008/420747

Hsu, P. Y., Devisetty, U. K., & Harmer, S. L. (2013). Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *eLife*, *2*, e00473. https://doi.org/10.7554/eLife.00473

Hsu, P. Y., & Harmer, S. L. (2012). Circadian phase has profound effects on differential expression analysis. *PLoS ONE*, *7*, e49853. https://doi.org/10.1371/journal.pone.0049853

Hsu, P. Y., & Harmer, S. L. (2014a). Global profiling of the circadian transcriptome using microarrays. *Methods in Molecular Biology*, *1158*, 45–56. https://doi.org/10.1007/978-1-4939-0700-7_3

Hsu, P. Y., & Harmer, S. L. (2014b). Wheels within wheels: The plant circadian system. *Trends in Plant Science*, *19*, 240–249. https://doi.org/10.1016/j.tplants.2013.11.007

Huang, J., Zhao, X., & Chory, J. (2019). The Arabidopsis transcriptome responds specifically and dynamically to high light stress. *Cell Reports*, *29*, 4186–4199. https://doi.org/10.1016/j.celrep.2019.11.051

Huang, W., Pérez-García, P., Pokhilko, A., Millar, A. J., Antoshechkin, I., Riechmann, J. L., & Mas, P. (2012). Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. *Science*, *335*, 75–79. https://doi.org/10.1126/science.1219075

Hughes, M. E., Hogenesch, J. B., & Kornacker, K. (2010). JTK_CYCLE: An efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets. *Journal of Biological Rhythms*, *25*, 372–380. https://doi.org/10.1177/0748730410379711

Kardailsky, I., Shukla, V. K., Ahn, J. H., Dagenais, N., Christensen, S. K., Nguyen, J. T., Chory, J., Harrison, M. J., & Weigel, D. (1999). Activation tagging of the floral inducer FT. *Science*, *286*, 1962–1965. https://doi.org/10.1126/science.286.5446.1962

Kleine, T., Kindgren, P., Benedict, C., Hendrickson, L., & Strand, A. (2007). Genome-wide gene expression analysis reveals a critical role for CRYPTOCHROME1 in the response of Arabidopsis to high irradiance. *Plant Physiology*, *144*, 1391–1406. https://doi.org/10.1104/pp.107.098293

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Lee, D., Park, S. J., Ahn, C. S., & Pai, H. (2017). MRF family genes are involved in translation control, especially under energy-deficient conditions, and their expression and functions are modulated by the TOR signaling pathway. *The Plant Cell*, *29*, 2895–2920. https://doi.org/10.1105/tpc.17.00563

Leng, N., Dawson, J. A., Thomson, J. A., Ruotti, V., Rissman, A. I., Smits, B. M. G., Haag, J. D., Gould, M. N., Stewart, R. M., & Kendziorski, C. (2013). Gene expression EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*, *29*, 1035–1043. https://doi.org/10.1093/bioinformatics/btt087

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R., & Pfister, H. (2014). UpSet: Visualization of intersecting sets. *IEEE Transactions on Visualization and Computer Graphics*, *20*, 1983–1992. https://doi.org/10.1109/TVCG.2014.2346248

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Data, G. P., & Sam, T. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656

Long, S. P., Humphries, S., & Falkowski, P. G. (1994). Photoinhibition of photosynthesis in nature. *Annual Review of Plant Physiology and Plant Molecular Biology*, *45*, 633–662. https://doi.org/10.1146/annurev.pp.45.060194.003221

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*, 550. https://doi.org/10.1186/s13059-014-0550-8

Maier, A., Schrader, A., Kokkelink, L., Falke, C., Welter, B., Iniesto, E., Rubio, V., Uhrig, J. F., Hülskamp, M., & Hoecker, U. (2013). Light and the E3 ubiquitin ligase COP1/SPA control the protein stability of the MYB transcription factors PAP1 and PAP2 involved in anthocyanin accumulation in Arabidopsis. *The Plant Journal*, *74*, 638–651. https://doi.org/10.1111/tpj.12153

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. Journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

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. https://doi.org/10.1104/pp.106.088757

Millar, A. J. (2016). The intracellular dynamics of circadian clocks reach for the light of ecology and evolution. *Annual Review of Plant Biology*, *67*, 595–618. https://doi.org/10.1146/annurev-arplant-043014-115619

Mishra, P., & Panigrahi, K. C. (2015). GIGANTEA—An emerging story. *Frontiers in Plant Science*, *6*, 8. https://doi.org/10.3389/fpls.2015.00008

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. https://doi.org/10.1016/S1534-5807(02)00170-3

Mockler, T., Michael, T., Priest, H., Shen, R., Sullivan, C., Givan, S., McEntee, C., Kay, S., & Chory, J. (2007). The diurnal project: Diurnal and circadian expression profiling, model-based pattern matching, and promoter analysis. *Cold Spring Harbor Symposia on Quantitative Biology*, *72*, 353–363. https://doi.org/10.1101/sqb.2007.72.006

Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pagès, H., & Gentleman, R. (2009). ShortRead: A bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics*, *25*, 2607–2608. https://doi.org/10.1093/bioinformatics/btp450

Morita, Y., Saitoh, M., Hoshino, A., Nitasaka, E., & Iida, S. (2006). Isolation of cDNAs for R2R3-MYB, bHLH and WDR transcriptional regulators and identification of c and ca mutations conferring white flowers in the Japanese morning glory. *Plant and Cell Physiology*, *47*, 457–470. https://doi.org/10.1093/pcp/pcj012

Morris, J. H., Apeltsin, L., Newman, A. M., Baumbach, J., Wittkop, T., Su, G., Bader, G. D., & Ferrin, T. E. (2011). clusterMaker: A multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics*, *12*, 436. https://doi.org/10.1186/1471-2105-12-436

Müller, L. M., Mombaerts, L., Pankin, A., Davis, S. J., Webb, A. A. R., Goncalves, J., & von Korff, M. (2020). Differential effects of day-/night cues and the circadian clock on the barley transcriptome. *Plant Physiology*, *183*, 765–779. https://doi.org/10.1104/pp.19.01411

Nohales, M. A., & Kay, S. A. (2020). Molecular mechanisms at the core of the plant circadian oscillator. *Nature Structural & Molecular Biology*, *23*, 1061–1069.

Pérez-García, P., Ma, Y., Yanovsky, M. J., & Mas, P. (2015). Time-dependent sequestration of RVE8 by LNK proteins shapes the diurnal oscillation of anthocyanin biosynthesis. *PNAS*, *112*, 5249–5253. https://doi.org/10.1073/pnas.1420792112

Prunet, N., Morel, P., Thierry, A. M., Eshed, Y., Bowman, J. L., Negrutiu, I., & Trehin, C. (2008). Rebelote, Squint, and Ultrapetala1 function redundantly in the temporal regulation of floral meristem termination in *Arabidopsis thaliana*. *Plant Cell*, *20*, 901–919. https://doi.org/10.1105/tpc.107.053306

Revelle, W. (2017). psych: Procedures for psychological. *Psychometric, and Personality Research. Software*.

Reynolds, A. P., Richards, G., de la Iglesia, B., & Rayward-Smith, V. J. (1992). Clustering rules: A comparison of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and Algorithms*, *5*, 475–504.

Risso, D., Schwartz, K., Sherlock, G., & Dudoit, S. (2011). GC-content normalization for RNA-Seq data. *BMC Bioinformatics*, *12*, 17. https://doi.org/10.1186/1471-2105-12-480

Rochaix, J. D. (2014). Regulation and dynamics of the light-harvesting system. *Annual Review of Plant Biology*, *65*, 287–309. https://doi.org/10.1146/annurev-arplant-050213-040226

Roeber, V. M., Bajaj, I., Rohde, M., Schmülling, T., & Cortleven, A. (2021). Light acts as a stressor and influences abiotic and biotic stress responses in plants. *Plant Cell and Environment*, *44*, 645–664. https://doi.org/10.1111/pce.13948

Romanowski, A., Schlaen, R. G., Perez-Santangelo, S., Mancini, E., & Yanovsky, M. J. (2020). Global transcriptome analysis reveals circadian control of splicing events in *Arabidopsis thaliana*. *Plant Journal*, *103*, 889–902. https://doi.org/10.1111/tpj.14776

Saiga, S., Furumizu, C., Yokoyama, R., Kurata, T., Sato, S., Kato, T., Tabata, S., Suzuki, M., & Komeda, Y. (2008). The Arabidopsis OBERON1 and OBERON2 genes encode plant homeodomain finger proteins and are required for apical meristem maintenance. *Development*, *135*, 1751–1759. https://doi.org/10.1242/dev.014993

Scholl, R. L., May, S. T., & Ware, D. H. (2000). Seed and molecular resources for Arabidopsis. *Plant Physiology*, *124*, 1477–1480. https://doi.org/10.1104/pp.124.4.1477

Shen, J.-R. (2015). The structure of photosystem II and the mechanism of water oxidation in photosynthesis. *Annual Review of Plant Biology*, *66*, 23–48. https://doi.org/10.1146/annurev-arplant-050312-120129

Struyf, A., Hubert, M., & Rousseeuw, P. J. (1997). Clustering in an object-oriented environment. *Journal of Statistical Software*, *1*, 1–30.

Tanaka, Y., Sasaki, N., & Ohmiya, A. (2008). Biosynthesis of plant pigments: Anthocyanins, betalains and carotenoids. *The Plant Journal*, *54*, 733–749. https://doi.org/10.1111/j.1365-313X.2008.03447.x

Warde-farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D., & Morris, Q. (2010). The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research*, *38*, 214–220. https://doi.org/10.1093/nar/gkq537

Webb, A. A. R., Seki, M., Satake, A., & Caldana, C. (2019). Continuous dynamic adjustment of the plant circadian oscillator. *Nature Communications*, *10*, 550. https://doi.org/10.1038/s41467-019-08398-5

Wu, G., Anafi, R. C., Hughes, M. E., Kornacker, K., & Hogenesch, J. B. (2016). MetaCycle: An integrated R package to evaluate periodicity in large scale data. *Bioinformatics*, *32*, 3351–3353. https://doi.org/10.1093/bioinformatics/btw405

Xiong, J., & Bauer, C. E. (2002). Complex evolution of photosynthesis. *Annual Review of Plant Biology*, *53*, 503–521. https://doi.org/10.1146/annurev.arplant.53.100301.135212

Yamaguchi, A., Kobayashi, Y., Goto, K., Abe, M., & Araki, T. (2005). TWIN SISTER OF FT (TSF) acts as a floral pathway integrator redundantly with FT. *Plant & Cell Physiology*, *46*, 1175–1189. https://doi.org/10.1093/pcp/pci151

Zogopoulos, V. L., Saxami, G., Malatras, A., Angelopoulou, A., Jen, C.-H., Duddy, W. J., Daras, G., Hatzopoulos, P., Westhead, D. R., & Michalopoulos, I. (2021). Arabidopsis Coexpression Tool: A tool for gene coexpression analysis in *Arabidopsis thaliana*. *ISCIENCE*, *24*, 102848. https://doi.org/10.1016/j.isci.2021.102848

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.