

SCIENTIFIC REPORTS



OPEN

Genome-wide mapping of DNase I hypersensitive sites reveals chromatin accessibility changes in *Arabidopsis* euchromatin and heterochromatin regions under extended darkness

Yue Liu¹, Wenli Zhang², Kang Zhang¹, Qi You¹, Hengyu Yan¹, Yuannian Jiao³, Jiming Jiang⁴, Wenying Xu¹ & Zhen Su¹

Light, as the energy source in photosynthesis, is essential for plant growth and development. Extended darkness causes dramatic gene expression changes. In this study, we applied DNase-seq (DNase I hypersensitive site sequencing) to study changes of chromatin accessibility in euchromatic and heterochromatic regions under extended darkness in *Arabidopsis*. We generated 27 Gb DNase-seq and 67.6 Gb RNA-seq data to investigate chromatin accessibility changes and global gene expression under extended darkness and control condition in *Arabidopsis*. We found that ~40% DHSs (DNase I hypersensitive sites) were diminished under darkness. In non-TE regions, the majority of DHS-changed genes were DHS-diminished under darkness. A total of 519 down-regulated genes were associated with diminished DHSs under darkness, mainly involved in photosynthesis process and retrograde signaling, and were regulated by chloroplast maintenance master regulators such as GLK1. In TE regions, approximately half of the DHS-changed TEs were DHS-increased under darkness and were primarily associated with the LTR/Gypsy retrotransposons in the heterochromatin flanking the centromeres. In contrast, DHS-diminished TEs under darkness were enriched in Copia, LINE, and MuDR dispersed across chromosomes. Together, our results indicated that extended darkness resulted in more increased chromatin compaction in euchromatin and decompaction in heterochromatin, thus further leading to gene expression changes in *Arabidopsis*.

Light is one of the essential environmental inputs for plant growth and development. Several important genes (e.g., *PHYA*, *PHYB*, *CRY2*) have been shown to be responsible for plant adaptation to light conditions¹, and to lead to adaptive changes in cells and even the whole organism^{2,3}. Thus, the mechanism by which light affects plants has been extensively investigated. Abnormal light conditions are generally thought to induce significant phenotypic changes. For example, low red/far-red ratios of shade light induce dramatic phenotypic changes such as hypocotyl growth, petiole elongation, leaf area reduction, hyponastic leaf movement, fewer branches, and leaf senescence³⁻⁶. Shade light also causes early flowering, thus indicating that extended darkness might accelerate the transition from vegetative growth to reproductive growth^{3,7}. In addition, extended periods of darkness trigger leaf senescence in *Arabidopsis*⁸⁻¹⁰.

¹College of Biological Sciences, China Agricultural University, State key Laboratory of Plant Physiology and Biochemistry, Beijing, China. ²Nanjing Agricultural University, State Key Laboratory for Crop Genetics and Germplasm Enhancement, JCIC-MCP, Nanjing, China. ³State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, The Chinese Academy of Sciences, Beijing, China. ⁴University of Wisconsin-Madison, Department of Horticulture, Madison, WI, USA. Yue Liu and Wenli Zhang contributed equally to this work. Correspondence and requests for materials should be addressed to Z.S. (email: zhensu@cau.edu.cn)

Received: 7 February 2017

Accepted: 16 May 2017

Published online: 22 June 2017

Complex transcriptional networks have been shown to mediate plant development in response to light, including seedling photomorphogenesis, seed germination, shade avoidance, and photoperiod response¹¹. Light modulates photoreceptor activity, which in turn triggers transcription factor (TF) regulation. For example, the phytochrome-interacting TFs, PIF4 and PIF5, negatively regulate phytochrome B (phyB)-mediated red light responses, and are required for dark-induced leaf senescence^{12, 13}. The expression of the major senescence-promoting NAC transcription factor ORESARA 1 (ORE1), which is a downstream factor of *PIF4/5*, is activated in the dark^{12, 13}. Under darkness, PIFs repress GOLDEN2-LIKE1 (*GLK1*), which is an important chloroplast maintenance master regulator involved in the PIFs-dependent regulatory network and influences photosynthesis-associated genes and plastid retrograde signalling^{14–16}.

The regulation of various TFs may be affected by the changes in chromatin structure under different light conditions. Recently, light signalling has been reported to control nuclear architecture reorganization during establishment¹⁷. Changes in chromatin structure affect the binding of TFs to regulatory elements, thus altering the expression of the associated genes¹⁸. Consequently, DNase I hypersensitive sites (DHSs) represent chromatin regions that are accessible for TF binding and, thereby, can be used to predict the presence of TFs^{19, 20}. DHS mapping analysis has been broadly performed for the global identification of cis-regulatory elements (CREs) and TF binding sites^{21–24}. Currently, DHSs are known to be closely associated with gene expression in different species, such as *Arabidopsis*²⁴, rice²⁵, human^{21, 23}, and other species. DNase-seq (DNase I treatment coupled with high-throughput DNA sequencing) is a powerful technique for identifying cis-regulatory elements across the genome and is highly effective for comparing TF binding profiles in different developmental stages and environmental conditions. For example, Sullivan *et al.* have conducted DNase-seq against of 7-day-old seedlings and have uncovered light-cued regulatory DNA dynamics and TF network remodeling during photomorphogenesis²⁶.

Moreover, extended darkness has been reported to cause dramatic changes in gene expression in *Arabidopsis*. A large number of differentially expressed genes (DEGs) under extended darkness and control condition in *Arabidopsis* have been identified using ATH1 GeneChip or large-scale quantitative real-time PCR (qRT-PCR)^{8, 27, 28}. Lin and Wu have examined global gene expression in response to 1-day and 5-day dark treatment by ATH1 GeneChip⁸. Parlitz *et al.* revealed the differential expression of senescence-associated genes and TFs after 2-day and 4-day dark treatment by qRT-PCR²⁷. We have also performed ATH1 GeneChip-based transcriptional profiling data for extended darkness treatment²⁸. Integration analysis of gene expression datasets generated from different laboratories has demonstrated that the DEGs have ~70% or even more overlap across different days of extended darkness (Supplementary Fig. 1), thus indicating that DEGs at different time-points after exposure to extended darkness are highly consistent.

To study the possible regulatory elements of DEGs under extended darkness in *Arabidopsis*, we conducted DNase-seq and RNA-seq to study chromatin accessibility and the association between DHSs and gene expression in *Arabidopsis* plants under extended darkness and control conditions. We also used combinatorial analyses of DNase-seq and sRNA-seq to study the relationship between DHSs and siRNA changes in transposable element (TE) regions under these two conditions.

Results

Genome-wide mapping of DNase I hypersensitive sites under extended darkness treatment and control condition.

Little is known about chromatin dynamics in response to extended darkness, despite the significant number of differentially expressed genes under extended darkness and normal conditions. Changes in chromatin accessibility affect TF binding to regulatory elements, thereby affecting the expression of the associated genes¹⁸. To investigate chromatin dynamics under extended darkness and control condition, we performed DNase-seq using 15-day-old *Arabidopsis* whole plants exposed to 3-day darkness treatment with normal light condition as a control. We identified a total of 10,380 DHSs and 5,948 DHSs under control and darkness conditions, respectively (Supplementary Table 1). Forty percent (40%) of the DHSs were diminished after dark treatment, thus indicating that the chromatin might have been condensed. We examined the genome distribution of the DHSs under each condition. We divided the *Arabidopsis* genome into six subclasses: promoter, 5'UTR, 3'UTR, coding exon, intron, and intergenic regions. The distributions of the DHS locations relative to *Arabidopsis* genes differed between the two conditions (Fig. 1a). Generally, a higher percentage of DHSs were present in promoter regions than in other genomic regions, in agreement with previous findings showing that promoters are associated with open chromatin in mammalian and plant genomes^{21, 24, 25}. However, the percentage of DHSs located in promoter regions under darkness decreased by 18.9% compared with that under control condition (28.5% vs. 47.4%). In addition, exon regions occupied 29.1% under darkness, compared with 14.9% under control condition.

DNA sequences in chromatin include TEs and non-TEs. TEs can change their positions within a genome. The regulation of TEs may differ from that of protein-coding genes. To investigate the effect of extended darkness on chromatin changes in both TE and non-TE genes, we divided DHSs into non-TE-related DHSs and TE-related DHSs. Meta-gene profiles of the DHSs generated along the generic region of non-TE genes showed that the profile of the DHS distribution upstream of transcription start sites (TSSs) was similar under both conditions (Fig. 1b), but the signal intensities of DHSs were higher under control condition than under darkness, especially for DHSs located in promoter regions near TSSs. Thus, extended darkness might induce chromatin condensation and decrease chromatin accessibility near promoter regions. However, the distribution of TE-related DHSs displayed very different patterns (Fig. 1c). The signals for DNase I sensitivity in TE regions were higher under extended darkness than under control condition. Under dark treatment, DHSs preferentially localized to TE bodies, in contrast to the distribution of non-TE genes.

To identify non-TE genes associated with DHSs under darkness and control conditions, we used CEAS software²⁹ to statistically analyse DHSs located within the 3 kb upstream or 1 kb downstream of TSSs, which were defined as DHS-associated genes. We identified 15,244 non-TE genes under control condition and 8,831 non-TE

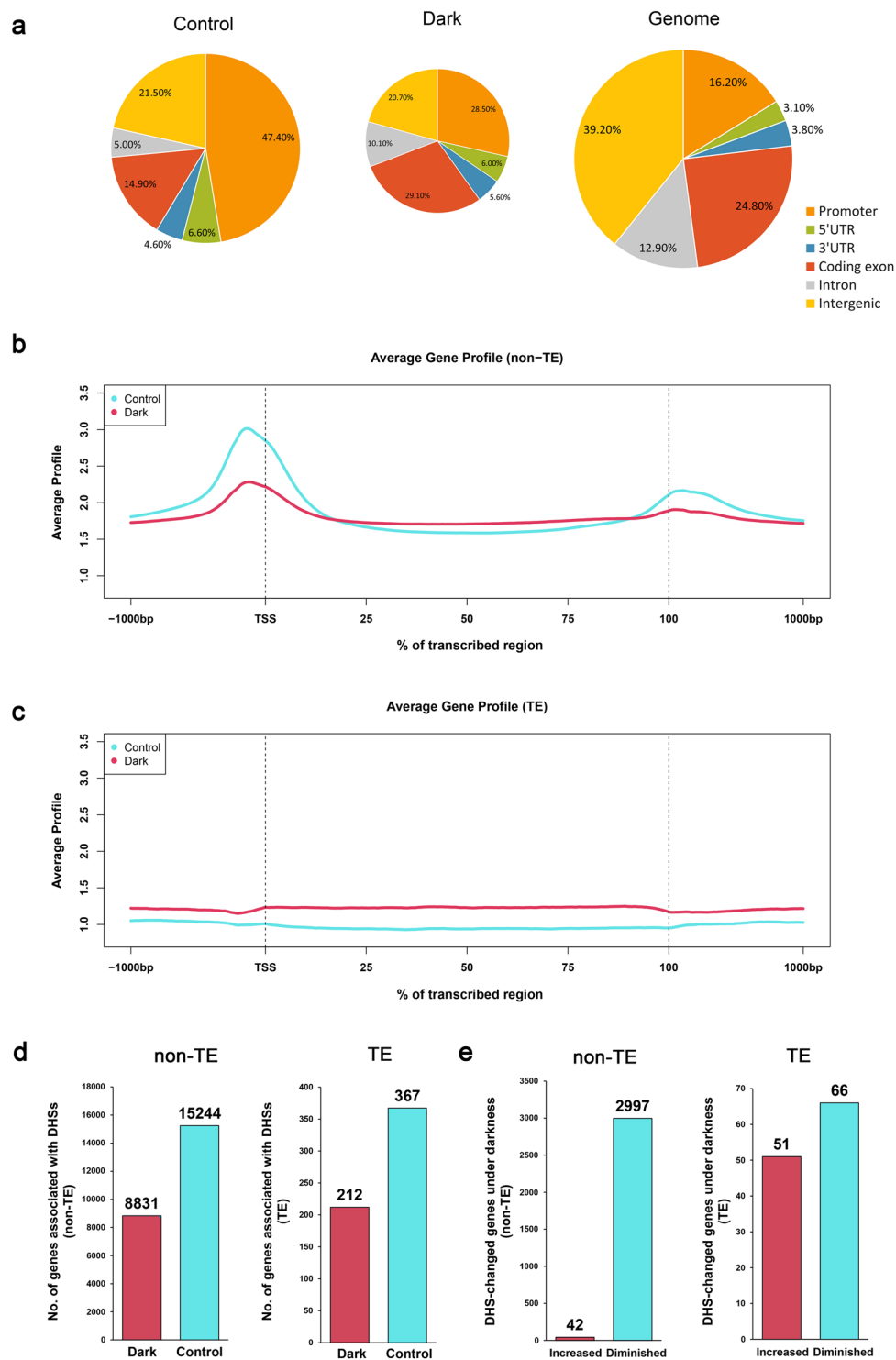


Figure 1. Distribution of DHSs under extended darkness and control condition. **(a)** Percentage of DHSs within different regions of the *Arabidopsis* genome under extended darkness and control conditions. The genome was divided into six classes: promoter, 5'UTR, 3'UTR, coding exon, intron and intergenic regions. The “Genome” panel represents the proportion of these six classes in *Arabidopsis* genome. **(b,c)** The profile of the DHS distribution in all *Arabidopsis* non-TE **(b)** and TE **(c)** genes including 1 kb upstream and downstream sequencing density of DHSs. The X-axis represents the conversion of the gene body into a percentage to standardize genes of different lengths. The 1-kb upstream and downstream regions of the gene are included. The Y-axis represents average values of DNase hypersensitivities. **(d)** Number of non-TE and TE genes associated with DHSs under darkness and light. **(e)** Number of DHS-changed non-TE and TE genes under darkness and light.

genes under darkness (Fig. 1d). The number of DHS-associated genes significantly decreased under darkness. Collectively, the DHS characterization revealed that changes in chromatin accessibility are associated with gene regions under extended darkness. To further investigate differences in chromatin accessibility between the two conditions, we identified differential DHS regions (by using MACS) located in the regions 3 kb upstream and 1 kb downstream of TSSs as DHS-changed genes. We identified 42 DHS-increased and 2,997 DHS-diminished non-TE genes under extended darkness (Fig. 1e and Supplementary Data 1), thus suggesting that darkness might decrease chromatin accessibility within non-TE regions. Meanwhile, we identified 367 and 212 TE-related genes under normal condition and extended darkness, respectively (Fig. 1d). We further identified 51 DHS-increased and 66 DHS-diminished TE genes under extended darkness (Fig. 1e and Supplementary Data 2). We also used public DNase-seq data (GSM847326-28, 2-week-old seedling plants) as a control to verify the DHS-changed genes under extended darkness (Supplementary Data 3), and approximately 76.6% of all the DHS-changed genes (including TE and non-TE genes) were confirmed.

To understand the functions of these DHS-changed genes, gene ontology (GO) enrichment analysis of 2,997 DHS-diminished genes was conducted by using agriGO³⁰ and REVIGO³¹ (Supplementary Fig. 2 and Supplementary Table 2). The GO term associated with the regulation of transcription was significantly enriched, including *MYB*, *WRKY*, *NAC* family genes. DHS-diminished genes were strongly associated with the GO terms of photosynthesis and light, response to stimulus and stress, and hormone-mediated signaling pathway. GO terms of anatomical structure development, such as leaf morphogenesis, epidermal cell differentiation, cell cycle and cytoskeleton organization, were also enriched. Interestingly, GO terms such as gene silencing and ncRNA metabolic process were significantly enriched in these DHS-diminished genes under extended darkness.

Differential chromatin accessibility and transcriptional activity of non-TE genes under extended darkness and control condition.

To investigate the relationship between DHSs and gene expression, we conducted RNA-seq using the same materials as those used for the DNase-seq experiments with three independent biological replicates (Supplementary Table 3). Differentially expressed genes under extended darkness and control condition were identified using Cufflinks software³² (Supplementary Data 4). In total, we identified 4,259 genes with significantly altered expression (fragments per kilobase of transcript per million mapped reads [FPKM] value > 2-fold change, p-value < 0.05), including 2,111 up-regulated genes and 2,148 down-regulated genes, after dark treatment (Supplementary Fig. 3a). To verify the RNA-seq results, we selected some genes for real-time RT-PCR validation. The additional biological samples were collected under the same conditions used for the RNA-seq experiments. The real-time RT-PCR results for the majority of the selected genes confirmed the RNA-seq results (Supplemental data 5). Moreover, we also compared our RNA-seq data with the published ATH1 GeneChip results^{8,28} under extended darkness (Supplemental data 4). Among the 4,259 changed genes, more than 70% genes were also detected with similar trends in previously published microarray results.

To understand the involvement of these DEGs in biological processes and functions, we conducted GO enrichment analysis by using agriGO³⁰ and REVIGO³¹ (Supplementary Fig. 3b,c and Supplementary Tables 4,5). As expected, under extended darkness, the GO terms response to absence of light and shade avoidance were significantly enriched; in contrast, under the control condition, photosynthesis and light response related GO terms were highly enriched. In addition, GO terms associated with anatomical structure development and hormone-mediated signal pathways were also enriched under the control condition. The most significant GO terms were associated with transcriptional regulation under both extended darkness and control condition, a result consistent with previous findings in which TFs have been found to be differentially expressed in response to dark treatment^{8,27}.

We integrated DNase-seq with RNA-seq datasets under darkness and control condition together. Genes with higher expression displayed higher levels of DNase I sensitivity under both darkness and normal conditions (Fig. 2a,b). Next we compared the groups of DHS-changed genes and differentially expressed non-TE genes, and discovered that most (98.6%) of DHS-changed genes were DHS-diminished genes under extended darkness treatment (Fig. 2c). There were 519 DHS-diminished and down-regulated genes (Fig. 2d,e and Supplementary Data 6) under extended darkness, thereby indicating a positive relationship between DNase I hypersensitivity and gene expression for these 519 genes. In contrast, there were 180 DHS-diminished genes (Supplementary Data 7) under extended darkness treatment showing negative relationship between DNase I hypersensitivity and gene expression, with lower DNase I hypersensitivity values (Fig. 2f) and higher FPKM values (Fig. 2g). This antagonistic relationship suggested that other regulatory mechanisms such as non-coding RNAs, may be involved in the differential gene expression under extended darkness and control condition.

To gain insight into the potential functions of these 519 DHS-diminished genes under extended darkness, we performed GO enrichment analysis (Fig. 2h and Supplementary Table 6) and found that photosynthesis (FDR P-value = 1.5E-26) was the most enriched GO term. Some processes closely related to photosynthesis, such as photosynthetic electron transport chain (FDR P-value = 3.00E-06), were also significantly enriched. GO terms associated with response to light, such as red light and far-red light (FDR P-value = 5.90E-07), blue light (FDR P-value = 4.00E-06), and UV (FDR P-value = 0.03) were enriched. In addition to light-related GO terms, development-related GO terms such as floral organ development (FDR P-value = 0.00077), as well as hormone-related terms such as salicylic acid biosynthetic process (FDR P-value = 9.00E-06), were also enriched. In addition to GO enrichment analysis, we performed gene sets enrichment analysis (GSEA) of gene family, PlantCyc and KEGG by using PlantGSEA³³. Significant gene sets were all associated with photosynthesis and some closely related gene sets such as chlorophyll and starch (Table 1). We compared the enriched GO terms for these 519 DHS-diminished and down-regulated genes with those of all DHS-diminished genes and all down-regulated genes (Supplementary Data 8). All three groups were enriched in photosynthesis- and light-related GO terms. There were relatively more development-related GO terms enriched in the group of all DHS-diminished genes and the group of all down-regulated genes. Other GO terms, such as gene silencing,

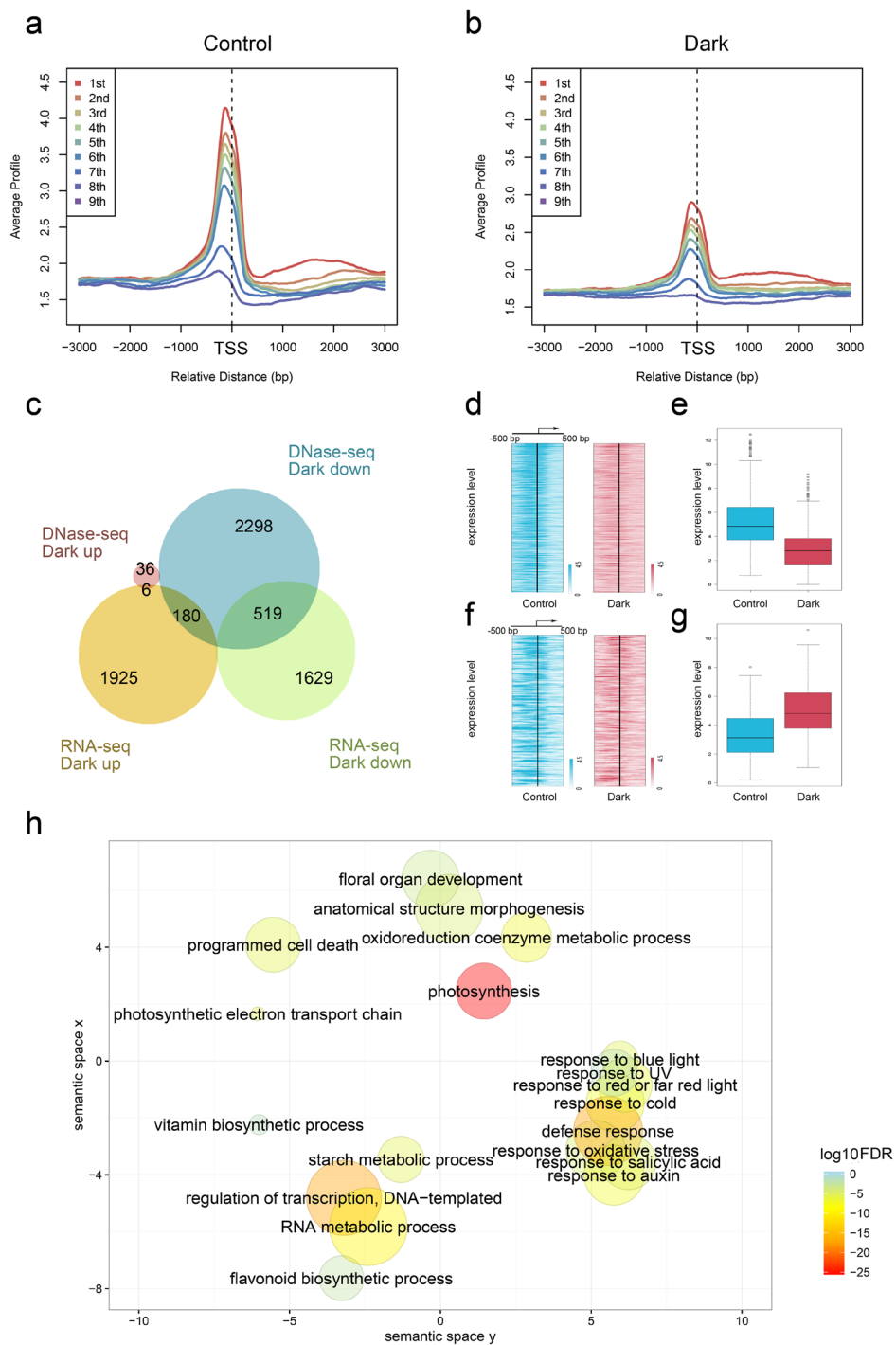


Figure 2. Integrated analysis of DHS-changed and differentially expressed genes. **(a,b)** The distribution of DHS along 3 kb upstream and downstream of TSSs of *Arabidopsis* genes with different expression level under control **(a)** and extended darkness **(b)**, respectively. All expressed genes were divided into 9 bins, which were expressed genes (FPKM > 0) from high expression (1st) to low expression (9th) based on the expression levels. **(c)** Venn diagrams between DHS-changed genes and differentially expressed genes under extended darkness. **(d)** DHS intensities around TSSs of the group of 519 genes from **(c)** under extended darkness and control conditions, respectively. The genes were sorted according to the expression level determined by FPKM under control condition. For each gene, the DHS intensities are displayed along -500 bp to 500 bp regions around the TSS. **(e)** The gene expression values are shown for the group of 519 genes from **(c)**. **(f)** DHS intensities around TSSs of the group of 180 genes from **(c)** under extended darkness and control conditions, respectively. The genes were sorted according to the expression level determined by FPKM under control condition. For each gene, the DHS intensities are displayed along -500 bp to 500 bp regions around the TSS. **(g)** The gene expression values are shown for the group of 180 genes from **(c)**. **(h)** GO enrichment analysis of 519 genes in **(c)** by agriGO and REVIGO. The scatter plot shows the cluster representatives in a two-dimensional space derived by applying multidimensional scaling to a matrix of the significant GO terms with semantic similarities. Bubble color and size indicates the \log_{10} (FDR p-value) (legend in bottom right-hand corner).

Gene Set Name	No. Genes	No. Genes in the Overlap	FDR
Gene Family			
Chloroplast and Mitochondria gene families, Chlorophyll a/b-binding protein family	17	6	1.57E-03
PlantCyc			
photosynthesis light reactions	38	11	5.41E-07
starch degradation	25	6	5.33E-03
KEGG			
Photosynthesis - antenna proteins	19	7	3.01E-05
Photosynthesis	73	11	3.01E-05
Starch and sucrose metabolism	88	11	1.21E-04
One carbon pool by folate	15	4	0.0196
Carbon fixation in photosynthetic organisms	75	7	0.034

Table 1. Enriched gene sets among DHS-diminished and down-regulated genes under extended darkness.

ncRNA processing, cytoskeleton organization and cell cycle process, tended to be enriched in the group of the all DHS-diminished genes. We also analysed GO enrichment for those 180 genes with a negative association between DHSs and expression (Supplementary Table 7). The regulation of transcription (FDR P-value = 2.3E-12) was mostly enriched, but the GO terms related to light were not enriched in these 180 genes.

Identification of regulatory elements for DHS-diminished and down-regulated photosynthesis-associated genes under extended darkness. As described above, the 519 DHS-diminished and down-regulated genes included many photosynthesis-associated genes. We found that these genes were involved in different biological processes related to photosynthesis, such as photosynthetic electron transport chain, light harvesting, light reaction, chloroplast organization, and chlorophyll metabolic process (Supplementary Data 9). We also found some genes involved in retrograde signaling (RS), such as *GUN4*, *GUN5*, and *CHI1*.

During motif analysis of the differential DHSs around photosynthesis-associated genes within the 519 DHS-diminished and down-regulated genes, we also found that in the group associated with photosynthesis, the putative GLK-binding site “CCAATC”¹⁶ was significantly enriched (Z-score = 2.22; P-value = 0.01) (Supplementary Table 8). According to previous studies, *GLK1*, as a chloroplast maintenance master regulator, is involved in a complex regulatory network during darkness¹⁵. To further elucidate the relationships among DHSs, gene expression, and motif positions, we used the UCSC genome browser³⁴ to analyse *GUN5*, *CHI1*, *LHCB2.4*, *LHCB3* and *LHCB4.2* involved in photosynthesis (Fig. 3). The GLK binding motif and surrounding regions were located in the promoter regions of these down-regulated photosynthesis-associated genes under darkness, which were also the regions with lower DNase I hypersensitivity under darkness. Our RNA-seq analysis showed that *GLK1* was down-regulated under extended darkness (Fig. 4a). We analysed the *GLK1*-induced genes reported previously¹⁶ together with our identified DHSs and gene expression levels. These *GLK1*-induced genes displayed lower expression and lower DNase I sensitivities under darkness (Fig. 4b,c). Our results indicated that extended darkness might block *GLK1* binding sites and further repress the expression of photosynthesis-related genes.

GLK1 has been reported to be a crucial factor regulating photosynthesis-associated nuclear genes (PhANGs) expression in retrograde signalling^{35,36}. Dark-grown seedlings undergoing photomorphogenesis have recently been reported to resemble seedlings grown on lincomycin (a drug affecting retrograde signaling by inhibiting chloroplast biogenesis) and norflurazon (an herbicide that inhibits carotenoid biosynthesis, thereby influencing retrograde signaling)³⁷. Analyzing of the changes in DHSs and expression of genes involved in retrograde signaling³⁵, revealed that the related genes tended to be down-regulated and DHS-diminished under extended darkness, especially *GUN4*, *GUN5* and *FCI1*, which are involved in chloroplast biogenesis, as well as *SIG2* and *SIG6*, which are important components of chloroplast transcription (Supplementary Table 9). The *gun4*, *gun5*, and *sig6* mutants showed lower chlorophyll accumulation than did wild type^{38,39}. To study the influence of DHS on retrograde signaling under extended darkness, we found that down-regulated genes associated with retrograde signaling³⁷ were significantly enriched among these 519 genes under extended darkness (Fig. 4d). Additionally, among these 519 genes, those located in chloroplasts were significantly enriched (Fig. 4e). Furthermore, we found that more than half of the down-regulated genes related to chloroplast retrograde signaling³⁷ overlapped with those that were down-regulated under extended darkness (Fig. 4f). Together with the changes in genes involved in retrograde signaling, we speculated that darkness influenced retrograde signaling by down-regulating *GUN4*, *GUN5*, *FCI1*, *SIG2* and *SIG6* in chloroplasts, thereby repressing the transcriptional regulation of PhANGs by *GLK1* (Fig. 4g).

Identification of DHSs in different TE families and related siRNAs under extended darkness and control condition. A total of 367 and 212 TE-related genes were identified under normal condition and extended darkness, respectively, including 51 DHS-increased and 66 DHS-diminished TE genes under darkness (Fig. 1e and Supplementary Data 2). The TE families differed in terms of DHSs under extended darkness and control condition (Table 2). The DHS-diminished TEs under extended darkness belonged primarily to the LTR/Copia, LINE/L1 and DNA/MuDR families, which were located dispersedly on chromosomes⁴⁰. By contrast, most

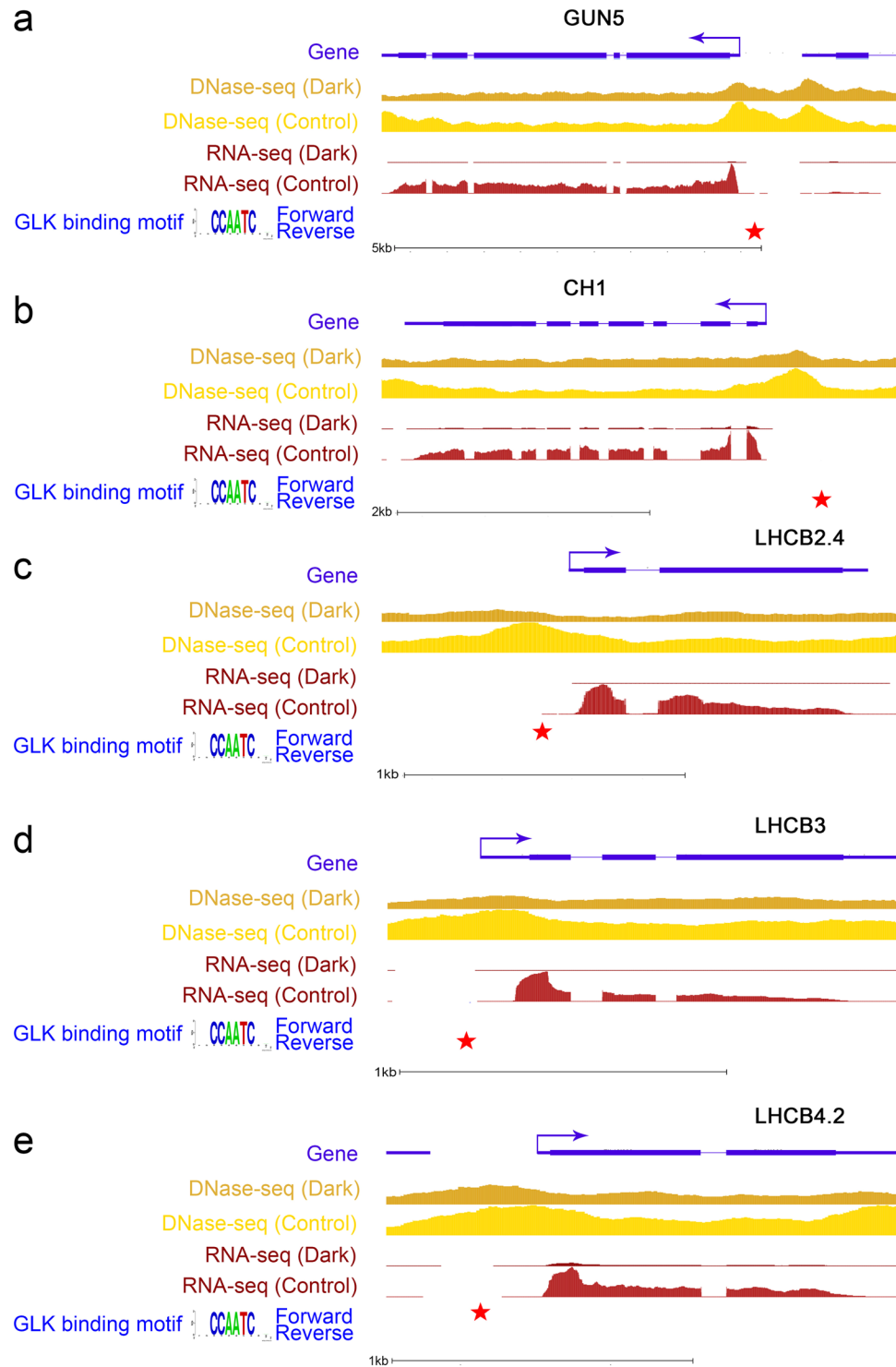


Figure 3. Selected genes associated with DHSs, expression and motifs in the UCSC genome browser. The gene models are shown in purple, with the direction marked by an arrow. The brown and yellow colors represent DHS signals under extended darkness and control condition, respectively. The red five-pointed stars represent the location of motifs. The dark red and light red colors represent expression signals under extended darkness and control condition, respectively. **(a)** *GENOMES UNCOUPLED 5 (GUN5)* encodes a magnesium chelatase involved in plastid-to-nucleus signal transduction. **(b)** *CHLORINA 1 (CH1)* encodes chlorophyllide a oxygenase which converts chlorophyllide a to chlorophyllide b by catalysing two successive hydroxylations at the 7-methyl group of chlorophyllide a. **(c)** *LIGHT-HARVESTING CHLOROPHYLL B-BINDING 2 (LHCB2.4)* belongs to the Lhc supergene family encodes the light-harvesting chlorophyll a/b-binding (LHC) proteins that constitute the antenna system of the photosynthetic apparatus. **(d)** *LIGHT-HARVESTING CHLOROPHYLL B-BINDING PROTEIN 3 (LHCB3)* is a component of the main light-harvesting chlorophyll a/b-protein complex of Photosystem II (LHC II). **(e)** *LIGHT HARVESTING COMPLEX PHOTOSYSTEM II (LHCB4.2)* is a protein involved in the light-harvesting complex of photosystem II. The mRNA has cell-to-cell mobility.

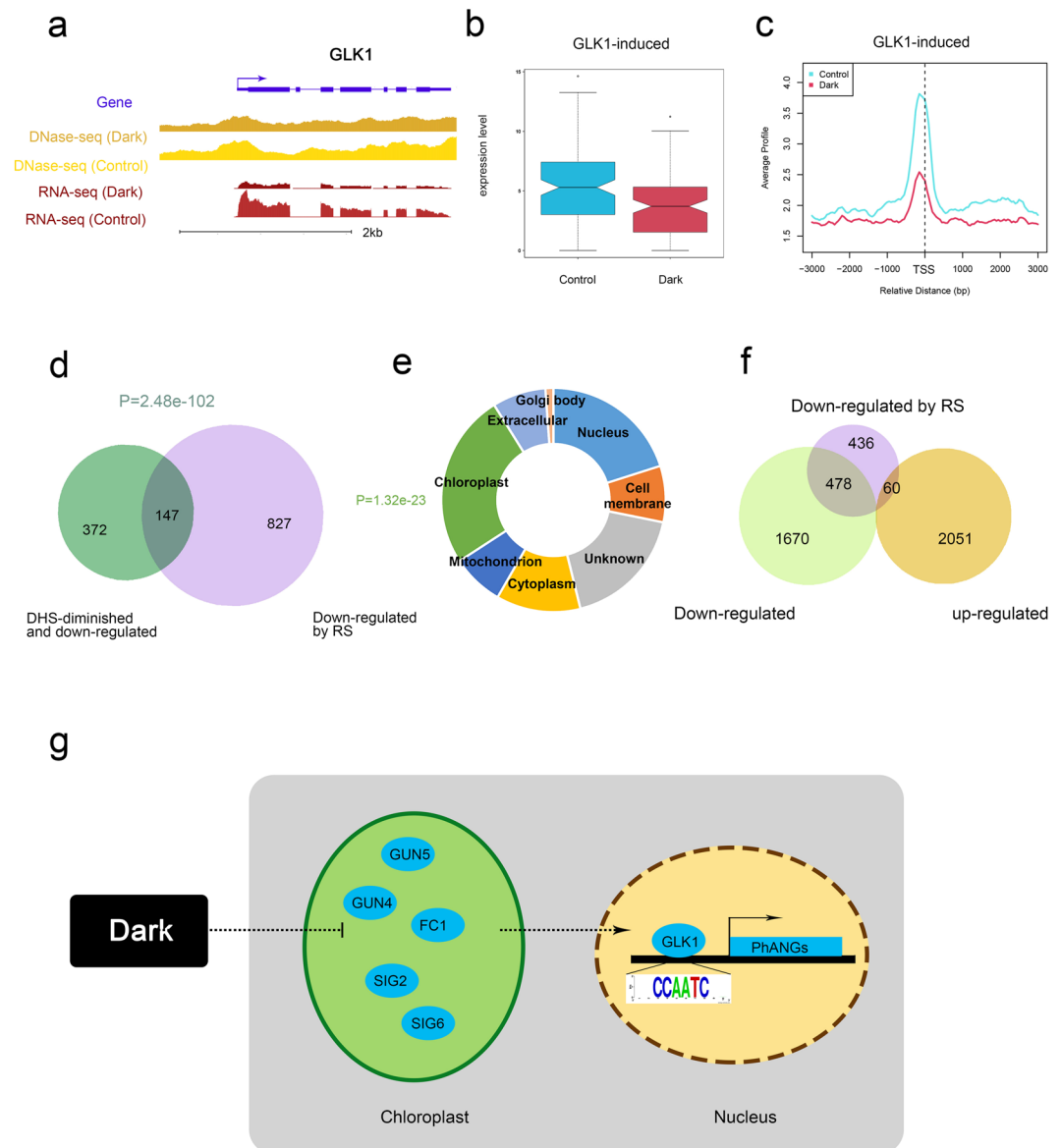


Figure 4. Changes in genes involved in retrograde signaling under extended darkness. **(a)** DNase-seq and RNA-seq tracks indicate a gene model for GLK1 in the UCSC genome browser. The gene models are shown in purple, with the direction marked by an arrow. The brown and yellow colors represent DHS signals under extended darkness and control condition, respectively. The dark red and light red colors represent that expression signals under extended darkness and control condition, respectively. **(b)** Expression changes in GLK1-induced genes under extended darkness. **(c)** DHS changes in GLK1-induced genes under extended darkness. **(d)** Venn diagram showing down-regulated genes by RS and 519 down-regulated and DHS-diminished genes under extended darkness. The P-value for the overlap in these 519 genes (hypergeometric distribution) is shown at the top. **(e)** Subcellular location of 519 DHS-diminished and down-regulated genes. The P-value for chloroplast in these 519 genes (hypergeometric distribution) is shown to the left. **(f)** Venn diagram showing down-regulated genes by RS and up- or down-regulated under extended darkness. **(g)** A model for darkness-influenced retrograde signaling. Darkness can repress some genes in chloroplasts in terms of expression or diminished DHS, which regulate GLK1 binding to conserved motifs in the promoters of PhANGs to promote gene expression by retrograde signaling.

of the DHS-increased TEs were LTR/Gypsy retrotransposons located in the heterochromatin flanking the centromeres⁴⁰ (Supplementary Fig. 4). This finding indicated that chromatin tended to loosen in heterochromatin under extended darkness, in agreement with results from previous studies showing that light signalling affects chromatin organization¹⁷.

Transposons have been reported to become activated in response to environmental stress through the repression of gene silencing, such as via RNA-directed DNA methylation (RdDM)⁴¹. As described above, the GO terms of gene silencing and ncRNA metabolism process were significantly enriched in 2997 DHS-diminished

type	DHS-increased under extended darkness		DHS-diminished under extended darkness		Genome	
	count	percentage	count	percentage	count	percentage
LTR/Gypsy	37*	72.55%*	5	7.58%	1277	32.65%
LTR/Copia	4	7.84%	18*	27.27%*	514	13.14%
LINE/L1	2	3.92%	19*	28.79%*	398	10.18%
DNA/MuDR	1	1.96%	17*	25.76%*	725	18.54%
DNA/HAT	0	0.00%	1	1.52%	84	2.15%
DNA/En-Spm	3	5.88%	2	3.03%	356	9.10%
DNA/Harbinger	0	0.00%	0	0.00%	43	1.10%
RC/Helitron	0	0.00%	1	1.52%	166	4.24%
null/SADHU	0	0.00%	1	1.52%	17	0.43%
Unassigned	4	7.84%	2	3.03%	326	8.34%

Table 2. Differential TE families in DHSs under extended darkness and control condition. *Indicates a significant difference. Significance was calculated by hypergeometric distribution with a cut-off of 0.05.

genes. Subsequently, we analysed changes in expression and DHS among genes involved in the RdDM pathway under extended darkness. We found that the genes involved in RdDM tended to be down-regulated and/or DHS-diminished under extended darkness (Supplementary Table 10), including *AGO4*, *DMS3*, and *NRPD1A*. Thus, we speculated that the RdDM pathway might be influenced by extended darkness treatment.

TEs are closely related to small interfering RNAs (siRNAs), in that they are transcribed into non-protein-coding transcripts that are immediately converted into double-stranded RNA^{41,42}. These short double-stranded RNAs are cleaved into siRNAs involved in siRNA-mediated gene silencing. Therefore, we studied the changes in siRNA under extended darkness and control condition via siRNA-seq using the same materials as those were used for the RNA-seq and DNase-seq experiments (Supplementary Table 11). The siRNAs associated with TEs were mainly 21- to 24-nucleotide siRNAs, and the length distribution of the sequenced small RNAs was investigated^{41,42}. Our results showed that 21–24 nt siRNA had relatively high abundance, especially 21 nt and 24 nt siRNAs (Supplementary Fig. 5).

We identified candidate siRNA reads by mapping them to repeat regions for subsequent analysis and then overlapping the siRNA reads with the location of the TEs. These siRNAs were found to be located in 974 and 1031 TE regions under control and extended darkness conditions, respectively. We then identified siRNA-changed TEs based by calculating reads per million (RPM) for one TE (fold change ≥ 1.6 and $P < 0.05$), respectively. We identified 183 siRNA-increased and 131 siRNA-decreased TEs under darkness (Supplementary Data 10). Furthermore, the families of siRNA-changed TEs were very similar to dynamically changed TEs in DHSs (Supplementary Table 12). For example, the siRNA-increased TEs under extended darkness mainly belonged primarily to the Gypsy families. As shown in Fig. 5, some siRNA-changed TEs were DHS-changed under extended darkness and control condition. The siRNA-increased Gypsy-like TEs such as *AT4G06664*, *AT4G06712* and *AT5G31087* were also increased in DHSs in their promoter or gene regions, whereas siRNA-decreased Copia-like TEs such as *AT2G03080*, *AT4G20180* and *AT1G70010* were diminished in DHSs upstream of TSSs. Thus, we proposed that differential DHSs in TE regions are strongly associated with siRNA changes under extended darkness and control condition.

Discussion

Influence of extended darkness on chromatin accessibility. Extended darkness causes dramatic changes in transcriptional activity, which is controlled by light-responsive *cis*-elements. DNase I hypersensitivity is an efficient method for charting chromatin accessibility and DHSs in chromatin, and can further be used to map functional elements. To study transcriptional regulation, we conducted DNase-seq analysis under extended darkness and control condition in mature *Arabidopsis* plants. Through genome-wide mapping of chromatin accessibility, we revealed the chromatin dynamics from control condition to extended darkness. We identified a total of 10,380 DHSs corresponding to control condition, and 5,948 DHSs under extended darkness. Approximately 40% of the DHSs were diminished after extended darkness, thus indicating that chromatin might become compacted after extended dark treatment. Genomic distribution analysis of the DHSs under both conditions showed that in the promoter regions, chromatin under control condition appeared to be more accessible than that under darkness, whereas DHSs tended to be open in heterochromatin regions under darkness. Extended darkness thereby resulted in more closed chromatin in euchromatic regions and decompaction in heterochromatic regions. We showed that the distribution and proportion of non-TE and TE-related DHSs were quite different. Under extended darkness, DHSs were preferentially located in TE bodies, in contrast to the distribution of the non-TE genes. Bourbousse and colleagues have showed that prolonged exposure to light triggers nuclear growth and the condensation of some heterochromatic regions during *Arabidopsis* postembryonic development, whereas both nuclear growth and low heterochromatin condensation are repressed in the dark¹⁷. Our results were consistent with light signals controlling nuclear architecture reorganization during *Arabidopsis* seedling establishment, demonstrating that different light conditions influence chromatin accessibility during seedling establishment and in mature plants.

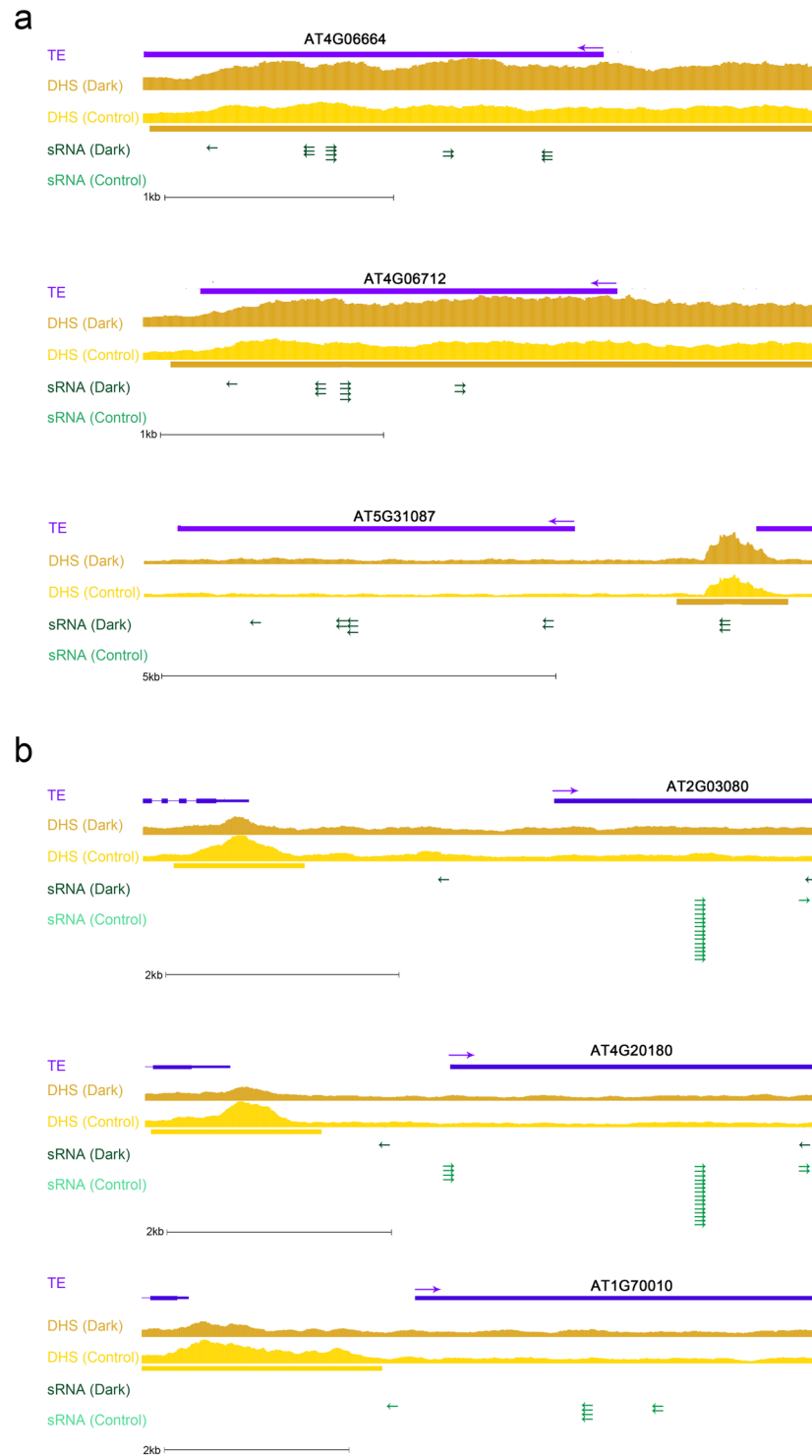


Figure 5. Selected TEs associated with DHSs and mapped siRNA reads in the UCSC genome browser. The gene models are shown in purple, with the direction marked by an arrow. The brown and yellow colors represent that DHS signals under extended darkness and control condition, respectively. The DHS-increased region is marked by a brown bar, and the DHS-diminished region is marked by a yellow bar. The dark and light green colors represent siRNA reads. The direction of mapping is marked by an arrow. **(a)** siRNA-increased Gypsy-like TEs. **(b)** siRNA-decreased Copia-like TEs.

Association of DHSs with expression regulation of non-TE genes under extended darkness. Through high-resolution mapping of DHSs, we found that extended darkness affected chromatin accessibility. The altered gene expression profiles revealed how plants respond to extended darkness and control

condition, thus raising questions concerning whether the chromatin accessibility changes under darkness are associated with differential gene expression. The overlap of differential chromatin accessibility and transcriptional activity of non-TE genes were calculated, and 519 genes with dark-diminished DHSs were identified with down-regulated transcriptional levels under darkness. Photosynthesis and chlorophyll-related genes were significantly enriched among these 519 genes, on the basis of GO term and GSEA, as expected. This finding was consistent with the observation that extended darkness resulted in leaf senescence and lower chlorophyll content compared with control condition. These enriched photosynthesis-related GO terms suggested that under the control condition, the light triggers open chromatin with high DNase I hypersensitivity near genes associated with photosynthesis and chlorophyll and further promotes the expression of these genes. However, under extended darkness, most chromatin regions for non-TE genes tend to be condensed, and the darkness diminishes DHSs near light-responsive genes and represses gene expression. Some other enriched GO terms included leaf morphogenesis and floral organ development, thus potentially explaining the different phenotypes observed between control condition and extended darkness. For example, extended darkness caused stem/petiole elongation and early flowering. Some GO terms, such as “salicylic acid biosynthetic process” and “response to jasmonic acid stimulus” also suggested association between extended darkness and hormone biosynthesis/signal transduction pathways. Therefore, we conclude that extended darkness could trigger chromatin condensation near TSSs of the genes affecting photosynthesis, and leading to changes in plant anatomy and differentially regulated hormone-related genes via the down-regulation of their transcriptional activities, thereafter influencing the *Arabidopsis* growth and development.

Particularly, in the promoter regions of the photosynthesis-related genes, the GLK-binding site “CCAATC” was significantly enriched. GLK (GOLDEN 2-like) is a nuclear-encoded transcription factor and a key regulator in the light-induced transcriptional network. GLK1 is a transcriptional factor in chloroplast retrograde signaling, a process whereby chloroplast components give rise to one or more retrograde signals from the chloroplast to nucleus, to evoke changes in nuclear gene expression that consequently modulate chloroplast function and photosystem^{35,36}. Our results showed that GLK1-induced genes were DHS-diminished and down-regulated, thus leading us to investigate whether chloroplast retrograde signaling was influenced by extended darkness. Through differential expression analysis, we found that many genes located in chloroplasts such as *GUN4*, *GUN5*, *SIG2* and *SIG6* were down-regulated or DHS-diminished under extended darkness. These genes were associated with chloroplast biogenesis and transcription in chloroplasts^{35,36}. DHSs located in the GLK1 binding motif were diminished in the promoter regions of photosynthesis-related genes, which potentially explained how repressing GLK1 might also affect the transcriptional regulation of photosynthesis-related genes under extended darkness. This result was consistent with findings from a previously report showing that light influences on retrograde signaling during photomorphogenesis. In the darkness, *GLK1* expression is also repressed, and the retrograde signaling affects deetiolation³⁷. Our results revealed that extended darkness could influence chloroplast retrograde signaling.

Differential DHSs in TEs and possible association with siRNA changes under extended darkness.

By identifying DHSs in different TE families under extended darkness treatment and control condition, we observed a relatively high proportion of DHS-increased TEs under extended darkness, consisting of LTR/Gypsy retrotransposons located in the heterochromatin flanking the centromeres. However the DHSs-diminished TEs under extended darkness were enriched in the Copia, LINE, and MuDR families dispersed throughout the chromosomes. Our results indicated that extended darkness might affect chromatin accessibility in TE regions. Recently, Bourbousse *et al.* have reported the effects of light on chromatin structure by cytogenetics analyses¹⁷. Their results have shown that the RNA level of *AtGPI*, a Gypsy retrotransposon, is decreased after light exposure. Here, we proposed that extended darkness might activate Gypsy retrotransposons and increase accessibility in heterochromatic regions.

TEs can produce transcripts that are immediately converted into double-stranded RNA and cleaved into siRNA⁴². Through siRNA-seq analysis, we found that changes of TEs in siRNAs were consistent with those of the DHSs. In Gypsy regions, siRNAs were more enriched under extended darkness than under control condition, whereas some siRNAs were more enriched in Copia regions under control condition than under darkness. The consistent changes between siRNA and DHSs indicated that the siRNAs may be derived from TEs, which may be affected by the chromatin structure.

Extended darkness might affect the RdDM pathway.

Recently, transposons have been reported to be activated in response to environmental stress through a combination of loss of RNA-directed DNA methylation (RdDM)⁴¹. Transcriptional gene silencing (TGS) consists of light-stimulated concurrent processing during seedling development¹⁷. Bourbousse *et al.* have found that many genes associated with RdDM, such as *DDM1*, *AGO4* and *AGO6*, have lower RNA levels in darkness than in light¹⁷. In addition, the leaves of *drd1* and *ddm1* mutants lost less chlorophyll content than wild-type leaves after extended darkness treatment⁴³. In the present study, we analysed the changes in the expression and DHS of genes involved in the RdDM pathway under extended darkness. We found that the genes involved in RdDM tended to be down-regulated and/or DHS-diminished under extended darkness treatment (Supplementary Table 10), including *AGO4*, *DMS3*, *NRPD1A*, *CMT3*, *MET1* and *VIMI*. Thus, we speculated that the RdDM pathway might be influenced by extended darkness. To study changes in DNA methylation under extended darkness, we generated bisulfite-seq (BS-seq) data, but we identified no differences in DNA methylation globally at the whole-genome level between extended darkness and control condition (data not shown), in agreement with the results of Bourbousse¹⁷.

In summary, we conducted DNase-seq and mRNA-seq to identify differential DHSs and differentially expressed genes, and we revealed chromatin accessibility changes in euchromatic and heterochromatic regions under extended darkness in *Arabidopsis* (Fig. 6). Our results showed that in non-TE regions, the majority of the

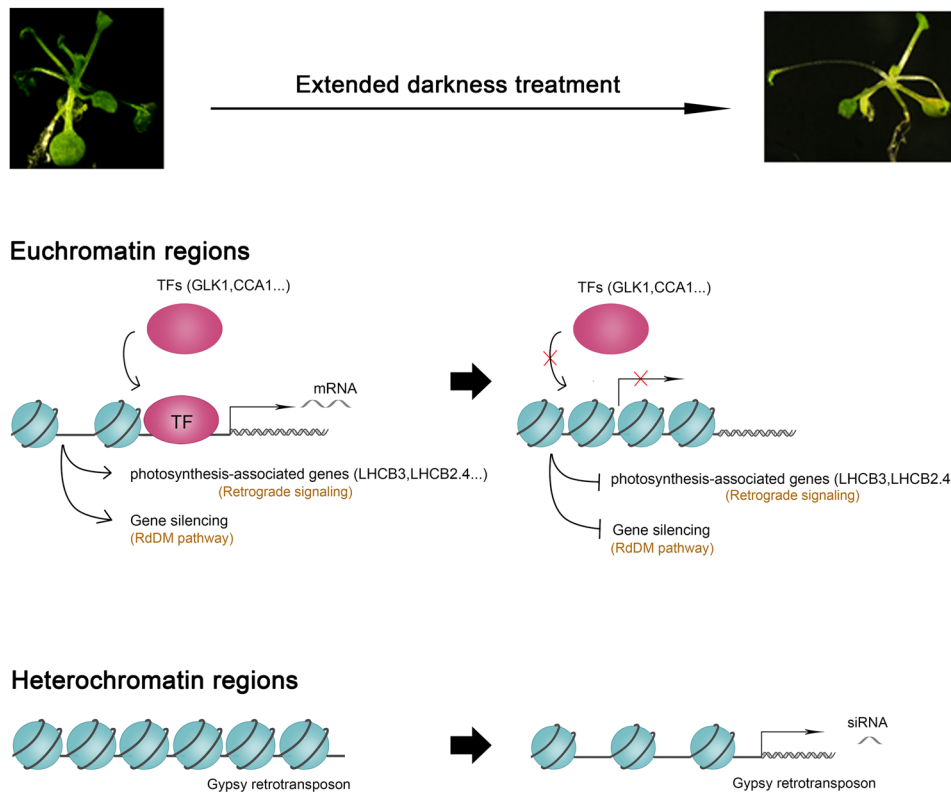


Figure 6. Model for changes in chromatin accessibility during extended darkness. In euchromatin, chromatin becomes closed during extended darkness. This change represses the binding of light-associated TFs, such as GLK1 and CCA1, to DHSs in the promoters of downstream photosynthesis-associated genes. The repression of photosynthesis-associated genes may be influenced by retrograde signaling. In addition, gene silencing, such as the RdDM pathway, is also repressed by closed chromatin. In heterochromatin, chromatin tends to open during extended darkness, and LTR/Gypsy transposons are transcribed into siRNAs.

DHS-changed genes were DHS-diminished genes under extended darkness. The association analysis of DHSs and expression regulation indicated that light might promote TF binding to DHSs containing conserved motifs within chromatin that is loose and accessible. The modified chromatin structure potentially regulated downstream target gene expression, especially genes involved in photosynthesis and chloroplast retrograde signaling. In TE regions, there was a relatively high proportion of DHS-increased TEs under extended darkness treatment, particularly Gypsy retrotransposons. In addition, the differentially changed DHSs in TE regions might be strongly associated with siRNA changes under extended darkness and control condition. In addition, we found that some genes had increased expression levels but were DHS-diminished during extended darkness treatment. We speculate that other factors are likely to be involved in regulating the extended darkness response, such as miRNA targets and histone modifications. Overall, our findings revealed the changes in chromatin accessibility under extended darkness and might provide answers to fundamental questions about how light signals influence transcription at the genomic scale, but the complicated regulatory mechanism requires further study.

Materials and Methods

Plant materials and growth conditions. *Arabidopsis thaliana* (Col-0) seeds were surface-sterilized and sown on half-strength Murashige and Skoog (MS) medium in Petri plates. The seeds were stratified for 3 d at 4 °C and then transferred to a conditioning chamber under 16 h light (22 °C)/8 h dark (19 °C) cycles. *Arabidopsis* plants were grown in Petri plates for 15 days. All the plants subjected to dark treatment in Petri dishes were wrapped in aluminium foil. *Arabidopsis* plants were grown under control condition (see above) as a control. After 3 days, whole plants with and without extended dark treatment were harvested for mRNA-seq, DNase-seq, and sRNA-seq.

RNA-seq and analysis. Total RNA was extracted using TRIZOL reagent (Invitrogen, CA, USA) and purified using an RNeasy Mini Kit (Qiagen). The RNA samples were extracted from whole *Arabidopsis* whole plants under extended darkness treatment and control condition. RNA-seq libraries were developed from three biological replicates from whole plants under extended darkness treatment and control condition, and sequenced with a HiSeq 2500 Sequencing System (Illumina, San Diego, CA, USA) in the Research Technology Support Facility at Michigan State University.

Sequencing reads of RNA-seq were aligned to the *Arabidopsis* genome (TAIR10) by using TopHat software⁴⁴. Calculation of FPKM (fragments per kilobase of transcript per million mapped reads) and the identification of differentially expressed genes were performed using Cuffdiff in Cufflinks software³². Genes showing statistically significant differential expression, on the basis of [FPKM] value > 2-fold change and p-value < 0.05 were identified as differentially expressed genes (DEGs). Gene ontology (GO) enrichment analysis was performed using the agriGO website³⁰ and REVIGO³¹.

DNase-seq and analysis. DNase-seq experiments, including nuclei isolation, DNase I digestion, and DNase-seq library construction, were performed as previously described²⁵. DNase-seq libraries were developed from whole plants under extended darkness and normal conditions, and were sequenced by HiSeq 2500 Sequencing System (Illumina, San Diego, CA, USA) in the Research Technology Support Facility at Michigan State University.

Bowtie2 software⁴⁵ was used to align the sequencing reads of DNase-seq to the *Arabidopsis* reference genome (TAIR10) by using default parameters. The peak in different conditions and differentially changed peaks were called by MACS software⁴⁶. The nomodel parameter was set, the d-value parameter was set as 200, and the m-fold was set as 8–30. The CEAS software²⁹ was used to analyse the distance between TSSs of genes and the nearest called peaks. We then identified differentially changed DHS peaks located in the region of 3 kb upstream and 1 kb downstream of TSSs as DHS-changed genes. GO enrichment analysis was performed using the agriGO website³⁰ and gene set enrichment analysis using plantGSEA³³.

Motif analysis. We extracted sequences from the +/– 500 bp regions surrounding the DHS-changed peak summits in 519 DHS-diminished and down-regulated genes under extended darkness as search sequences, and the sequences of 1000 bp regions at upstream and downstream of the DHS-changed peak as background sequences. We analysed the enrichment of motifs from several groups, including the Plant Cis-acting Regulatory DNA Elements (PLACE) database⁴⁷, AthaMap webserver⁴⁸ and PlantCARE database⁴⁹. Significance was determined using an algorithm based on Z score and P-value filtering with a cut-off of 0.05⁵⁰.

sRNA-seq and analysis. Total RNA was extracted using TRIZOL reagent (Invitrogen) and purified using RNeasy Mini Kits (Qiagen). RNA samples were extracted from *Arabidopsis* whole plants under extended darkness treatment and control condition. Construction of the sRNA libraries and deep-sequencing were carried out by the Beijing Genomics Institute (BGI, Hong Kong, China): isolated total RNA from each sample was separated on 15% denaturing polyacrylamide gels for size selection. Small RNAs (18 to 30 nt) were selected and the developed sRNA libraries were directly sequenced with a HiSeq 2000 Sequencing System (Illumina, San Diego, CA, USA). The sequenced sRNAs were aligned to the genome with SOAP software⁵¹. These sRNAs were annotated by mapping to sequences including miRNAs, repeats, snRNA, rRNAs, and tRNAs in miRBase, Rfam and GeneBbank. Candidate siRNA reads were identified by mapping to repeat regions. siRNA reads located on TEs were overlapped by using BED Tools software⁵². siRNA-changed TEs were identified by calculating reads per million (RPM) of siRNA for one TE with a cut-off of 1.6 and 0.05 in fold change and P-value, respectively.

References

- Chaiwanon, J., Wang, W., Zhu, J. Y., Oh, E. & Wang, Z. Y. Information Integration and Communication in Plant Growth Regulation. *Cell* **164**, 1257–1268, doi:10.1016/j.cell.2016.01.044 (2016).
- Casal, J. J. Shade avoidance. *Arabidopsis Book* **10**, e0157, doi:10.1199/tab.0157 (2012).
- Casal, J. J. Photoreceptor signaling networks in plant responses to shade. *Annu Rev Plant Biol* **64**, 403–427, doi:10.1146/annurev-arplant-050312-120221 (2013).
- Casal, J. J. Phytochrome A enhances the promotion of hypocotyl growth caused by reductions in levels of phytochrome B in its far-red-light-absorbing form in light-grown *Arabidopsis thaliana*. *Plant physiology* **112**, 965–973 (1996).
- Finlayson, S. A., Krishnareddy, S. R., Kebrom, T. H. & Casal, J. J. Phytochrome regulation of branching in *Arabidopsis*. *Plant Physiol* **152**, 1914–1927, doi:10.1104/pp.109.148833 (2010).
- Kozuka, T. *et al.* Involvement of auxin and brassinosteroid in the regulation of petiole elongation under the shade. *Plant Physiol* **153**, 1608–1618, doi:10.1104/pp.110.156802 (2010).
- Cerdan, P. D. & Chory, J. Regulation of flowering time by light quality. *Nature* **423**, 881–885, doi:10.1038/nature01636 (2003).
- Lin, J. F. & Wu, S. H. Molecular events in senescing *Arabidopsis* leaves. *Plant J* **39**, 612–628, doi:10.1111/j.1365-313X.2004.02160.x (2004).
- Weaver, L. M. & Amasino, R. M. Senescence is induced in individually darkened *Arabidopsis* leaves but inhibited in whole darkened plants. *Plant Physiology* **127**, 876–886, doi:10.1104/Pp.010312 (2001).
- Weaver, L. M., Gan, S. S., Quirino, B. & Amasino, R. M. A comparison of the expression patterns of several senescence-associated genes in response to stress and hormone treatment. *Plant Molecular Biology* **37**, 455–469, doi:10.1023/A:1005934428906 (1998).
- Jiao, Y., Lau, O. S. & Deng, X. W. Light-regulated transcriptional networks in higher plants. *Nat Rev Genet* **8**, 217–230, doi:10.1038/nrg2049 (2007).
- Sakuraba, Y. *et al.* Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in *Arabidopsis*. *Nat Commun* **5**, 4636, doi:10.1038/ncomms5636 (2014).
- Zhang, Y., Liu, Z., Chen, Y., He, J. X. & Bi, Y. PHYTOCHROME-INTERACTING FACTOR 5 (PIF5) positively regulates dark-induced senescence and chlorophyll degradation in *Arabidopsis*. *Plant Sci* **237**, 57–68, doi:10.1016/j.plantsci.2015.05.010 (2015).
- Oh, E., Zhu, J. Y. & Wang, Z. Y. Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. *Nat Cell Biol* **14**, 802–809, doi:10.1038/ncb2545 (2012).
- Liebsch, D. & Keech, O. Dark-induced leaf senescence: new insights into a complex light-dependent regulatory pathway. *New Phytologist* **212**, 563–570, doi:10.1111/nph.14217 (2016).
- Waters, M. T. *et al.* GLK transcription factors coordinate expression of the photosynthetic apparatus in *Arabidopsis*. *Plant Cell* **21**, 1109–1128, doi:10.1105/tpc.108.065250 (2009).
- Bourbousse, C. *et al.* Light signaling controls nuclear architecture reorganization during seedling establishment. *Proc Natl Acad Sci USA* **112**, E2836–2844, doi:10.1073/pnas.1503512112 (2015).
- Zhang, W. L., Zhang, T., Wu, Y. F. & Jiang, J. M. Open Chromatin in Plant Genomes. *Cytogenetic and Genome Research* **143**, 18–27, doi:10.1159/000362827 (2014).

19. Wu, C., Wong, Y. C. & Elgin, S. C. The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity. *Cell* **16**, 807–814 (1979).
20. He, H. H. *et al.* Differential DNase I hypersensitivity reveals factor-dependent chromatin dynamics. *Genome Res* **22**, 1015–1025, doi:10.1101/gr.133280.111 (2012).
21. Boyle, A. P. *et al.* High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**, 311–322, doi:10.1016/j.cell.2007.12.014 (2008).
22. Hesselberth, J. R. *et al.* Global mapping of protein-DNA interactions *in vivo* by digital genomic footprinting. *Nat Methods* **6**, 283–289, doi:10.1038/nmeth.1313 (2009).
23. Song, L. *et al.* Open chromatin defined by DNaseI and FAIRE identifies regulatory elements that shape cell-type identity. *Genome Res* **21**, 1757–1767, doi:10.1101/gr.121541.111 (2011).
24. Zhang, W., Zhang, T., Wu, Y. & Jiang, J. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. *Plant Cell* **24**, 2719–2731, doi:10.1105/tpc.112.098061 (2012).
25. Zhang, W. *et al.* High-resolution mapping of open chromatin in the rice genome. *Genome Res* **22**, 151–162, doi:10.1101/gr.131342.111 (2012).
26. Sullivan, A. M. *et al.* Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. *Cell Rep* **8**, 2015–2030, doi:10.1016/j.celrep.2014.08.019 (2014).
27. Parltitz, S., Kunze, R., Mueller-Roeber, B. & Balazadeh, S. Regulation of photosynthesis and transcription factor expression by leaf shading and re-illumination in Arabidopsis thaliana leaves. *J Plant Physiol* **168**, 1311–1319, doi:10.1016/j.jplph.2011.02.001 (2011).
28. Yu, J. *et al.* JAZ7 negatively regulates dark-induced leaf senescence in Arabidopsis. *J Exp Bot*, doi:10.1093/jxb/erv487 (2015).
29. Shin, H., Liu, T., Manrai, A. K. & Liu, X. S. CEAS: cis-regulatory element annotation system. *Bioinformatics* **25**, 2605–2606, doi:10.1093/bioinformatics/btp479 (2009).
30. Du, Z., Zhou, X., Ling, Y., Zhang, Z. & Su, Z. agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* **38**, W64–70, doi:10.1093/nar/gkq310 (2010).
31. Supek, F., Bosnjak, M., Skunca, N. & Smuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800, doi:10.1371/journal.pone.0021800 (2011).
32. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**, 511–515, doi:10.1038/nbt.1621 (2010).
33. Yi, X., Du, Z. & Su, Z. PlantGSEA: a gene set enrichment analysis toolkit for plant community. *Nucleic Acids Res* **41**, W98–103, doi:10.1093/nar/gkt281 (2013).
34. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res* **12**, 996–1006, doi:10.1101/gr.229102. Article published online before print in May 2002 (2002).
35. Chan, K. X., Phua, S. Y., Crisp, P., McQuinn, R. & Pogson, B. J. Learning the Languages of the Chloroplast: Retrograde Signaling and Beyond. *Annu Rev Plant Biol* **67**, 25–53, doi:10.1146/annurev-arplant-043015-111854 (2016).
36. Chi, W., Feng, P., Ma, J. & Zhang, L. Metabolites and chloroplast retrograde signaling. *Curr Opin Plant Biol* **25**, 32–38, doi:10.1016/j.pbi.2015.04.006 (2015).
37. Martin, G. *et al.* Phytochrome and retrograde signalling pathways converge to antagonistically regulate a light-induced transcriptional network. *Nat Commun* **7**, 11431, doi:10.1038/ncomms11431 (2016).
38. Ishizaki, Y. *et al.* A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J* **42**, 133–144, doi:10.1111/j.1365-313X.2005.02362.x (2005).
39. Mochizuki, N., Brusslan, J. A., Larkin, R., Nagatani, A. & Chory, J. Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. *Proc Natl Acad Sci USA* **98**, 2053–2058, doi:10.1073/pnas.98.4.2053 (2001).
40. Zemach, A. *et al.* The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205, doi:10.1016/j.cell.2013.02.033 (2013).
41. Matzke, M. A. & Moshier, R. A. RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* **15**, 394–408, doi:10.1038/nrg3683 (2014).
42. Sigman, M. J. & Slotkin, R. K. The First Rule of Plant Transposable Element Silencing: Location, Location, Location. *Plant Cell* **28**, 304–313, doi:10.1105/tpc.15.00869 (2016).
43. Cho, E. J. *et al.* A Mutation in Plant-Specific SWI2/SNF2-Like Chromatin-Remodeling Proteins, DRD1 and DDM1, Delays Leaf Senescence in Arabidopsis thaliana. *PLoS One* **11**, e0146826, doi:10.1371/journal.pone.0146826 (2016).
44. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111, doi:10.1093/bioinformatics/btp120 (2009).
45. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359, doi:10.1038/nmeth.1923 (2012).
46. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137, doi:10.1186/gb-2008-9-9-r137 (2008).
47. Higo, K., Ugawa, Y., Iwamoto, M. & Korenaga, T. Plant cis-acting regulatory DNA elements (PLACE) database: 1999. *Nucleic Acids Res* **27**, 297–300 (1999).
48. Hehl, R. & Bulow, L. AthaMap web tools for the analysis of transcriptional and posttranscriptional regulation of gene expression in Arabidopsis thaliana. *Methods Mol Biol* **1158**, 139–156, doi:10.1007/978-1-4939-0700-7_9 (2014).
49. Rombauts, S., Dehais, P., Van Montagu, M. & Rouze, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res* **27**, 295–296 (1999).
50. Yu, J. *et al.* SFGD: a comprehensive platform for mining functional information from soybean transcriptome data and its use in identifying acyl-lipid metabolism pathways. *BMC Genomics* **15**, 271, doi:10.1186/1471-2164-15-271 (2014).
51. Li, R., Li, Y., Kristiansen, K. & Wang, J. SOAP: short oligonucleotide alignment program. *Bioinformatics* **24**, 713–714, doi:10.1093/bioinformatics/btn025 (2008).
52. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842, doi:10.1093/bioinformatics/btq033 (2010).

Acknowledgements

We thank Qunlian Zhang for technical support and Zhou Du for critical suggestions. This work was supported by grants from the National Natural Science Foundation of China (grants 31371291 and 90817006).

Author Contributions

Conceived and designed the experiments: W.X., Z.S. Performed the experiments: Y.L., W.Z., K.Z., H.Y., W.X. Analysed the data: Y.L., W.Z., X.Y., Q.Y., W.X., Z.S. Contributed bioinformatics platform and analysis tools: Y.L., X.Y., H.Y., Z.S. Wrote the paper: Y.L., W.Z., Y.J., J.J., W.X., Z.S.

Additional Information

Supplementary information accompanies this paper at doi:10.1038/s41598-017-04524-9

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017