https://doi.org/10.1093/plphys/kiad487 PLANT PHYSIOLOGY 2023: 193: 2287–2291

Plant Physiology®

Letter

Molecular evidence for enhancer–promoter interactions in light responses of soybean seedlings

Mingkun Huang,[1,2,*] **Ling Zhang**,[1] **Wai-Shing Yung**,[2] **Yufang Hu**,[1] **Zhili Wang**,[2] **Man-Wah Li**[2] **and Hon-Ming Lam**[2,*]

1 Plant Functional Genomics and Bioinformatics Centre, Lushan Botanical Garden Jiangxi Province and Chinese Academy of Sciences, 332900 Jiujiang, Jiangxi, P.R. China
2 School of Life Sciences and Centre for Soybean Research of the State Key Laboratory of Agrobiotechnology, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, P.R. China

*Author for correspondence: honming@cuhk.edu.hk (H.-M.L.), huangmk@lsbg.cn (M.H.)

Dear Editor,

Enhancers are noncoding DNA sequences that can activate distal gene expression via transcription factor–mediated interactions with proximal promoters (Weber et al. 2016). Being generally located in poorly mapped regions of the genome, knowledge about these noncoding sequences and their mechanisms of action is limited, especially in crop species. Light is the primary energy source for plants and an essential environmental cue for plant development and survival (Chen et al. 2004). Thus, understanding how plants regulate gene expression in response to light and the potential roles of enhancers, is crucial for improving crop production. Using the assay for transposase-accessible chromatin with sequencing (ATAC-seq) and other molecular techniques, we generated molecular evidence for interactions among specific enhancers, promoters, and transcription factors (TFs) in the soybean (*Glycine max*) genome for modulating gene expression in response to different light regimens.

Ten-day-old soybean seedlings grown under a normal light/dark cycle (light grown) or in complete darkness (dark grown) differed drastically in their hypocotyl lengths and cotyledon chlorophyll contents (Fig. 1A; Supplemental Methods), making them convenient for investigating possible light-responsive enhancers. Since ATAC-seq has been well adopted for identifying distal regulatory elements (e.g. enhancers) in plants (Lu et al. 2019), we performed ATAC-seq and RNA-seq on light- and dark-grown cotyledons (LC and DC, respectively) and hypocotyls (LH and DH, respectively) (Fig. 1B; Supplemental Figs. S1 and S2 and Table S1; Supplemental Methods). ATAC-seq signals were enriched at the transcription start site (TSS) and were positively associated with gene expression levels in both light- and dark-grown cotyledons and hypocotyls, as previously reported (Supplemental Fig. S3A and Table S2; Huang et al. 2021a, 2021b; Huang et al. 2022). We identified 38,441, 36,688, 46,504, and 46,196 enriched accessible chromatin regions (ACRs) in LH, DH, DC, and LC, respectively (Supplemental Fig. S3B and Table S3).

Comparing ACRs between light- and dark-grown tissues, we identified 5,884 (3,738 upregulated in LC and 2,146 upregulated in DC) and 5,388 (3,093 upregulated in DH and 2,295 upregulated in LH) differentially accessible chromatin regions (DACRs) (Fig. 1C; Supplemental Fig. S4A and Tables S4 and S5). "Photosynthesis" was the only Gene Ontology (GO) term associated with DACR-associated genes (DACRGs) enriched in both hypocotyls and cotyledons (Supplemental Fig. S4B and S4C), supporting the importance of DACRs for regulating downstream gene expression related to carbon fixation in the presence of light. To pinpoint TF binding preferences at these DACRs, we conducted ATAC-seq footprint analysis (Bentsen et al. 2020) for genome-wide TF motif scanning (Supplemental Fig. S5 and Table S6; Supplemental Methods). Several TF binding motifs (TFBMs) showed significantly different ATAC-seq footprint scores between light and dark conditions, suggesting these TFs might play important roles in forming the DACRs (Supplemental Figs. S6 and S7 and Tables S7 and S8).

Received July 28, 2023. Accepted August 18, 2023. Advance access publication September 5, 2023

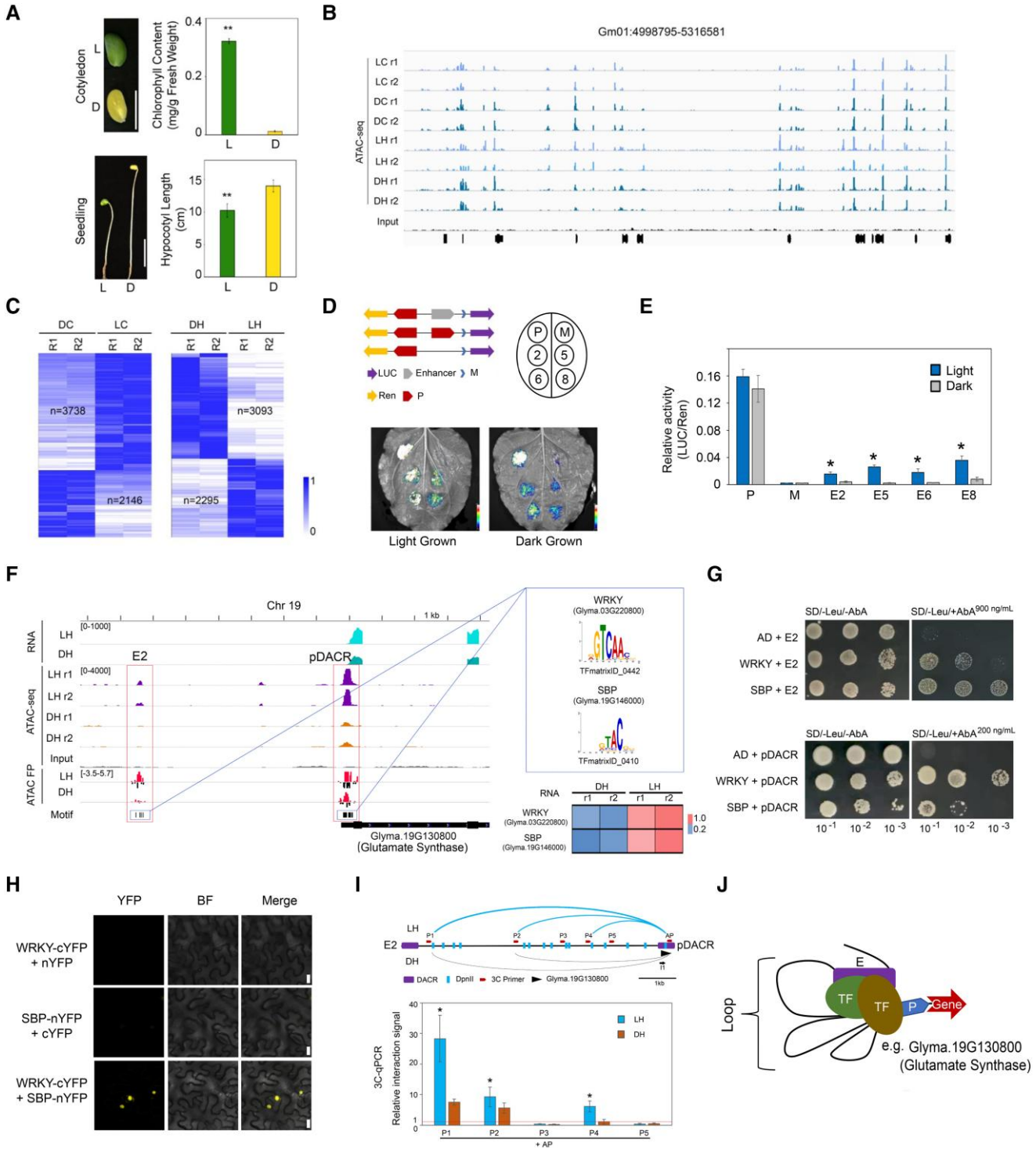

Figure 1. Transcription factor–mediated enhancer–promoter interactions during light response in soybean. **A)** Effects of light on chlorophyll content in cotyledons and hypocotyl lengths of 10-day-old soybean seedlings. Upper panel: light (L)- and dark (D)-grown cotyledons of soybean seedlings and their respective chlorophyll contents. Lower panel: light (L)- and dark (D)-grown soybean seedlings and their respective hypocotyl lengths. Scale bar = 4 cm. Error bar: standard deviation of 15 seedlings. $^{**}P < 0.01$, by Student's *t*-test. **B)** Examples of ATAC-seq signal distributions in chromosome 1 (Gm01). Bottom track: gene models. r1 and r2 indicate 2 biological replicates. LC, light-grown cotyledon; LH, light-grown hypocotyl; DC, dark-grown cotyledon; DH, dark-grown hypocotyl. **C)** Heatmap showing the clustering of DACRs between different light conditions in the cotyledon or the hypocotyl. **D)** Transcription activities of 4 selected distal DACR candidates in transient expression experiments. Upper left panel: schematics of various constructs. LUC, firefly luciferase reporter; Ren, Renilla luciferase reporter; M, empty vector containing the mini 35S only; P, positive control with the *firefly luciferase* reporter driven by the full-length 35S promoter. Upper right panel: diagram of the layout of the *Agrobacterium*

(continued)

Among these DACRs, 35.2% of those found in cotyledons and 40.1% of those found in hypocotyls were localized at the distal intergenic region (>2 kb from the nearest gene) and thus classified as light-responsive enhancer candidates (Supplemental Tables S4 and S5). We identified conserved TFBMs in enhancers associated with homologous genes in soybean and Arabidopsis (*Arabidopsis thaliana*; Sullivan et al. 2014), partially supporting evolutionary conservation of light-responsive enhancers in these species (Supplemental Figs. S8 and Table S9). Long terminal repeat (LTR) insertion was associated with increased distance from the enhancer to the associated gene (Supplemental Fig. S8), consistent with previous observations (Lu et al. 2019). Many light-responsive enhancers in hypocotyls (1,661) and cotyledons (1,735) overlap with enhancer loci in several soybean tissues (Huang et al. 2021a, 2022), indicating that these might also be involved in light-responsive or tissue-specific functions (Supplemental Fig. S9 and Table S10). To further verify whether these distal DACR sequences possess enhancer activities, we selected 10 candidates (labeled E1–E10) with greater enrichment of the ATAC-seq signal under light conditions than in the dark (Supplemental Fig. S10). These candidates were cloned into a STARR-like vector (Zhang et al. 2022) upstream of a *firefly luciferase* (*LUC*) reporter gene and agroinfiltrated into *Nicotiana benthamiana* (*N. benthamiana*) leaves (Supplemental Methods). Compared with the empty vector control (*Mini35S:LUC* without the addition of an enhancer sequence; M), 7 candidates strongly activated *LUC* expression under light conditions (Supplemental Fig. S10). Activities of 4 candidates (E2, E5, E6, and E8) with high-*LUC* activation levels were further tested under dark conditions. As expected, these 4 candidates had significantly higher activities in the light than in the dark, supporting that they were *bona fide* enhancers that play important roles in regulating gene expression in the light (Fig. 1D and E).

Enhancer–promoter interactions mediated by TFs are important for downstream gene regulation. We therefore investigated the E2 enhancer candidate associated with the promoter DACR (pDACR) of *Glutamate Synthase* (*GS*), a key gene in light responses (Forde and Lea 2007). According to the ATAC-seq footprint, both E2 and its associated pDACR harbor binding motifs for WRKY (Glyma.03G220800) and SBP (Glyma.19G146000) TFs (Fig. 1F). Similar to the *GS* gene, expression of genes encoding this *WRKY* and *SBP* was also higher in LH than in DH (Fig. 1F).

Both WRKY and SBP bound directly to E2 and pDACR in yeast 1-hybrid assays (Fig. 1G). WRKY and SBP could also induce E2 activity following transient expression in *N. benthamiana* leaves, supporting the enhancer role of E2 (Supplemental Fig. S11). We observed protein–protein interactions between WRKY and SBP using bimolecular fluorescence complementation (BiFC) (Fig. 1H; Supplemental Methods). Quantitative analysis of chromosome conformation capture (3C-qPCR; Hagege et al. 2007) using an anchor primer (AP) on the pDACR and several primers (P1-P5) between E2 and the pDACR (Fig. 1I; Supplemental Methods) detected the chromatin interactions P1-AP, P2-AP, and P4-AP in LH but only P1-AP and P2-AP interactions in DH (Fig. 1I). Moreover, P1-AP, P2-AP, and P4-AP interaction signals were significantly higher in LH compared with those in DH (Fig. 1I), possibly resulting from greater enhancer activity of E2 mediated by more highly expressed TFs (e.g. *WRKY* and *SBP*) under light-grown conditions. Reanalysis of soybean leaf Hi-C data from a recent study (Wang et al. 2021) revealed that E2 and pDACR are located in the same topologically associating domain (TAD), a genomic region with high frequency of sequence interaction. This partially supports the E2-pDACR interaction (Supplemental Fig. S12). E2 might also interact with other genes in this TAD through DNA loops (Supplemental Fig. S12), indicating multiregulatory roles for E2 in gene expression.

Figure 1. (Continued)
infiltration (agroinfiltration) sites of the different constructed vectors on a *N. benthamiana* leaf. P, positive control; M, empty vector. The numbers 2, 5, 6, and 8 correspond to the selected enhancer candidates. The vertical line represents the midrib of the leaf. Lower panel: bioluminescence images showing LUC activities driven by the enhancer candidates under light or dark conditions after agroinfiltration. **E)** Measurement of the relative LUC/Ren activity under light or dark conditions after agroinfiltration. Asterisk (*) indicates a significant difference between light and dark conditions (Student's *t*-test, $P < 0.05$). Error bars indicate the standard deviations of 5 biological replicates. **F)** IGV screenshot showing the locations of E2 (rectangle outline on the left) and pDACR (rectangle outline on the right) in the region ~8 kb upstream of the *Glutamate Synthase* (*GS*) gene. The right upper panel showed the WRKY (Glyma.03G220800) and SBP (Glyma.19G146000) binding motifs (TFmatrixID_0442 and TFmatrixID_0410, respectively) predicted by the ATAC footprint (ATAC FP). In the right bottom panel, the expression levels (in FPKM) of *WRKY* and *SBP* in LH and DH are shown in the heatmap. r1 and r2 indicate 2 biological replicates. **G)** Yeast 1-hybrid results showing the binding of WRKY and SBP to pDACR and E2. AD, transcription activation domain; SD-Leu, synthetic defined media without leucine. **H)** Bimolecular fluorescence complementation (BiFC) results showing the interactions between WRKY and SBP in vivo. Scale bar = 100 μm. **I)** The chromatin loop between E2 and pDACR associated to *GS* can be detected by chromosome conformation capture (3C)-qPCR in both light- and dark-grown hypocotyls (LH and DH, respectively). Upper panel: the genomic structure between E2 and pDACR. Blue curve indicated the loops (P1-AP, P2-AP, and P4-AP) detected in LH, while the gray curve (P1-AP and P2-AP) indicated the loops detected in DH in 3C-qPCR. Primers on the pDACR served as the anchor primer (AP), while primers P1–P5 were designed to avoid the *Dpn*II restriction site. Lower panel: the relative DNA interaction signal in 3C-qPCR. Total genomic DNA was used as the control template. The relative DNA interaction signal was normalized to the amplification signal of I1 and AP. Asterisks (*) indicate significant differences between light and dark conditions ($P < 0.05$, by Student's *t* test). Error bar, standard deviation of 3 biological replicates. **J)** A simple model describing how the transcription factor (TF)–mediated promoter–enhancer interaction can regulate gene (e.g. *Glutamate Synthase*) expression. E, enhancer; P, promoter.

Collectively, this study provides clear molecular evidence for enhancer–promoter and enhancer/promoter–TF, and TF–TF interactions, in mediating enhancer functions and chromatin loop formation to regulate downstream gene expression (e.g. *GS*) in response to changes in environmental stimuli such as light availability (Fig. 1J).

Acknowledgments

We thank Jee Yan Chu for copyediting this manuscript.

Author contributions

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plphys/pages/General-Instructions) is H.-M.L.

H.-M.L. and M.K.H. coordinated this research and designed the details of the experiments. M.K.H., L.Z., Y.F.H., and Z.L.W. performed the wet-lab experiments and M.K.H. conducted the bioinformatics analyses. M.K.H., M.W.L., W.S.Y., and H.-M.L. cowrote the manuscript.

Supplemental data

The following materials are available in the online version of this article.

Supplemental Figure S1. Pearson correlations between 2 replicates of RNA-seq and assay for transposase-accessible chromatin with sequencing (ATAC-seq) data obtained from soybean seedling hypocotyls and cotyledons grown under different light conditions.

Supplemental Figure S2. Overview of the RNA-seq data and assay for transposase-accessible chromatin with sequencing (ATAC-seq) data from soybean seedling hypocotyls and cotyledons grown under different light conditions.

Supplemental Figure S3. Genomic enrichment of assay for transposase-accessible chromatin with sequencing (ATAC-seq) signals in soybean seedling cotyledons and hypocotyls grown under different light conditions.

Supplemental Figure S4. Differentially accessible chromatin regions (DACRs) in soybean seedling hypocotyls and cotyledons due to different light conditions.

Supplemental Figure S5. Overview of assay for transposase-accessible chromatin (ATAC) footprint analysis using ATAC-seq data in soybean.

Supplemental Figure S6. Identification of transcription factor binding motifs (TFBMs) with differential assay for transposase-accessible chromatin (ATAC) footprint scores between light and dark conditions.

Supplemental Figure S7. Traces of transcription factor (TF) binding events at differentially accessible chromatin region-associated genes (DACRGs) according to the assay for transposase-accessible chromatin (ATAC) footprint at the differential TF binding motifs (TFBMs).

Supplemental Figure S8. Homologous genes regulated by light-responsive distal enhancer-like elements in Arabidopsis and soybean harbor a conserved *cis* DNA binding motif.

Supplemental Figure S9. Light-responsive distal differentially accessible chromatin regions (DACRs) potentially involved in regulating gene expression in different tissues.

Supplemental Figure S10. Transcription activities of 10 selected enhancer candidates in transient expression experiments.

Supplemental Figure S11. WRKY and SBP can activate E2 enhancer activity.

Supplemental Figure S12. Topologically associating domain containing E2 enhancer and its potential regulatory targets.

Supplemental Table S1. Data summary of ATAC-seq and RNA-seq.

Supplemental Table S2. FPKM values and differentially expressed genes calculated using RNA-seq data.

Supplemental Table S3. Accessible chromatin regions (ACRs) identified in light- or dark-grown soybean cotyledons or hypocotyls.

Supplemental Table S4. Differentially accessible chromatin regions (DACRs) identified in light-grown cotyledons (LC) versus dark-grown cotyledons (DC).

Supplemental Table S5. Differentially accessible chromatin regions (DACRs) identified in light-grown hypocotyls (LH) versus dark-grown hypocotyls (DH).

Supplemental Table S6. Assay for transposase-accessible chromatin with sequencing (ATAC-seq) footprint analysis results.

Supplemental Table S7. Transcription factor (TF) binding motifs with differential ATAC-seq footprint scores.

Supplemental Table S8. Transcription factor (TF) binding motif information in the associated differentially accessible chromatin regions (DACRs).

Supplemental Table S9. Motif information for soybean distal regulatory elements associated with homologous genes in Arabidopsis.

Supplemental Table S10. Overlap between light-responsive enhancers and accessible chromatin regions (ACRs) in each tissue.

Supplemental Table S11. Primer information.

Supplemental Methods.

Funding

This research was supported by the Hong Kong Research Grants Council Area of Excellence Scheme (AoE/M-403/16), Lo Kwee-Seong Biomedical Research Fund awarded to H.-M.L., the Lushan Botanical Garden Research Funds (2022ZWZX01), and Jiangxi Provincial Natural Science Foundation (20232BAB205015). Any opinions, findings, conclusions, or recommendations expressed in this publication do not reflect the views of the Government of the Hong Kong Special Administrative Region or the Innovation and Technology Commission.

Conflict of interest statement. The authors declared no conflict of interest.

Data availability

The raw data of ATAC-seq and RNA-seq in this study have been submitted to NCBI under the BioProject accession number PRJNA751745. The analyzed data were also available on the JBrowse in Wildsoydb Datahub (https://datahub.wildsoydb.org/) (Xiao et al. 2022).

References

Bentsen M, Goymann P, Schultheis H, Klee K, Petrova A, Wiegandt R, Fust A, Preussner J, Kuenne C, Braun T, et al. ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. Nat Commun. 2020:**11**(1):4267. https://doi.org/10.1038/s41467-020-18035-1

Chen M, Chory J, Fankhauser C. Light signal transduction in higher plants. Annu Rev Genet. 2004:**38**(1):87–117. https://doi.org/10.1146/annurev.genet.38.072902.092259

Forde BG, Lea PJ. Glutamate in plants: metabolism, regulation, and signalling. J Exp Bot. 2007:**58**(9):2339–2358. https://doi.org/10.1093/jxb/erm121

Hagege H, Klous P, Braem C, Splinter E, Dekker J, Cathala G, de Laat W, Forne T. Quantitative analysis of chromosome conformation capture assays (3C-qPCR). Nat Protoc. 2007:**2**(7):1722–1733. https://doi.org/10.1038/nprot.2007.243

Huang M, Zhang L, Zhou L, Wang M, Yung WS, Wang Z, Duan S, Xiao Z, Wang Q, Wang X, et al. An expedient survey and characterization of the soybean JAGGED 1 (GmJAG1) transcription factor binding preference in the soybean genome by modified ChIPmentation on soybean protoplasts. Genomics. 2021a:**113**(1):344–355. https://doi.org/10.1016/j.ygeno.2020.12.026

Huang M, Zhang L, Zhou L, Yung WS, Li MW, Lam HM. Genomic features of open chromatin regions (OCRs) in wild soybean and their effects on gene expressions. Genes (Basel). 2021b:**12**(5):640. https://doi.org/10.3390/genes12050640

Huang M, Zhang L, Zhou L, Yung WS, Wang Z, Xiao Z, Wang Q, Wang X, Li MW, Lam HM. Identification of the accessible chromatin regions in six tissues in the soybean. Genomics. 2022:**114**(3):110364. https://doi.org/10.1016/j.ygeno.2022.110364

Lu Z, Marand AP, Ricci WA, Ethridge CL, Zhang X, Schmitz RJ. The prevalence, evolution and chromatin signatures of plant regulatory elements. Nat Plants. 2019:**5**(12):1250–1259. https://doi.org/10.1038/s41477-019-0548-z

Sullivan AM, Arsovski AA, Lempe J, Bubb KL, Weirauch MT, Sabo PJ, Sandstrom R, Thurman RE, Neph S, Reynolds AP, et al. Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. Cell Rep. 2014:**8**(6):2015–2030. https://doi.org/10.1016/j.celrep.2014.08.019

Wang L, Jia G, Jiang X, Cao S, Chen ZJ, Song Q. Altered chromatin architecture and gene expression during polyploidization and domestication of soybean. Plant Cell. 2021:**33**(5):1430–1446. https://doi.org/10.1093/plcell/koab081

Weber B, Zicola J, Oka R, Stam M. Plant enhancers: a call for discovery. Trends Plant Sci. 2016:**21**(11):974–987. https://doi.org/10.1016/j.tplants.2016.07.013

Xiao Z, Wang Q, Li M-W, Huang M, Wang Z, Xie M, Varshney RK, Nguyen HT, Chan T-F, Lam H-M. Wildsoydb DataHub: a platform for accessing soybean multiomic datasets across multiple reference genomes. Plant Physiol. 2022:**190**(4):2099–2102. https://doi.org/10.1093/plphys/kiac419

Zhang L, Yung WS, Huang M. STARR-seq for high-throughput identification of plant enhancers. Trends Plant Sci. 2022:**27**(12):1296–1297. https://doi.org/10.1016/j.tplants.2022.08.008