

OPEN ACCESS

EDITED BY

Jelena Samardzic,
University of Belgrade, Serbia

REVIEWED BY

Yu Zheng,
Jiangnan University, China
Miriam Negussu,
University of Florence, Italy

*CORRESPONDENCE

Yufang Hu
✉ yufanghu163@163.com
Hua Yang
✉ yangh@lsbg.cn
Chang-En Tian
✉ changentian@aliyun.com

[†]These authors have contributed
equally to this work and share
first authorship

RECEIVED 07 April 2025

ACCEPTED 13 June 2025

PUBLISHED 30 June 2025

CITATION

Tianxiao L, Xiao H, Wang J, Zhang L,
Fan T, Huang M, Tian C-E, Yang H and
Hu Y (2025) Genome-wide characterization
of the GmLUX binding preferences and its
epigenic features in the soybean genome.
Front. Plant Sci. 16:1607224.
doi: 10.3389/fpls.2025.1607224

COPYRIGHT

© 2025 Tianxiao, Xiao, Wang, Zhang, Fan,
Huang, Tian, Yang and Hu. This is an open-
access article distributed under the terms of
the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Genome-wide characterization of the GmLUX binding preferences and its epigenic features in the soybean genome

Lv Tianxiao^{1†}, Hong Xiao^{1†}, Jingxing Wang^{2†}, Ling Zhang³,
Tian Fan¹, Mingkun Huang³, Chang-En Tian^{1*}, Hua Yang^{3*}
and Yufang Hu^{3*}

¹Guangdong Provincial Key Laboratory of Plant Adaptation and Molecular Design, Guangzhou Key Laboratory of Crop Gene Editing, Innovative Center of Molecular Genetics and Evolution, School of Life Sciences, Guangzhou University, Guangzhou Higher Education Mega Center, Guangzhou, China, ²Suihua Branch of the Heilongjiang Academy of Agricultural Sciences, Suihua, China, ³Jiangxi Provincial Key Laboratory of Ex Situ Plant Conservation and Utilization, Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, China

Transcription factors function in complex regulatory networks to regulate various biological and physiological processes. In soybean (*Glycine max*), GmLUX, as an important component of the evening complex, plays a critical role in the regulation of soybean flowering regulation. In this study, the genome-wide characterization and epigenetic features of GmLUX binding sites have been analyzed using high-throughput sequencing methods, such as ChIP-seq, Hi-C, histone modification and ATAC-seq. In addition, combined with molecular experiments, GmLUX was found to be able to directly regulate the *CO-like* gene by facilitating chromatin interactions, suggesting a new regulatory pathway of GmLUX in controlling flowering, which provided the important genomic resources for a further understanding of its regulatory mechanism.

KEYWORDS

glycine max, GmLUX, ChIP-seq, multiple omics, epigenetics

1 Introduction

Transcription factors (TFs) contribute significantly in regulating plant growth and development as well as the adaptation to the environment (e.g. stress responses, photosynthesis, specialized metabolite production) (Hrmova and Hussain, 2021; Strader et al., 2022; Dhatteval et al., 2024; Gao and Dubos, 2024). Plant TFs carrying sequence-specific DNA-binding domains, particularly recognize and bind to specific DNA sequences (TF binding sites, usually with a typical short conserved consensus motif) to regulate the expression of associated genes as an activators or repressors (Liu et al., 1999; Blanc-Mathieu et al., 2024). The interactions between TFs and DNA typically disrupt nucleosome stability and elevate chromatin accessibility, which is partially responsible for actively engaged *cis-*

regulatory elements (CREs) contained within accessible chromatin regions (ACRs) (Iwafuchi-Doi et al., 2016; Klemm et al., 2019; Ricci et al., 2019; Huang et al., 2022). ACRs with non-coding CREs are closely related to gene expression and TF-binding capacity (Klemm et al., 2019; Lu et al., 2019; Huang et al., 2022). Histone modifications are essential for regulating the chromatin accessibility, and flanking histone modifications indicate the transcriptional coregulators recruited to the ACRs (Klemm et al., 2019; Lu et al., 2019; Ricci et al., 2019; Zhou et al., 2021).

TFs play an important role in shaping plant development via regulating various downstream genes. As previously reported, soybean LUX transcription factor (GmLUX), a member of the evening complex, transcriptionally repressed the *E1* by binding to the LUX binding sites (LBS) of *E1s* promoters, which subsequently relieved the *E1s* suppression of two important *FT* genes (*FT2a* and *FT5a*) and promoted flowering (Lu et al., 2017; Bu et al., 2021). As a key TF, how GmLUX coordinated with other epigenetic signals to regulate downstream genes by interacting with genome-wide CREs remained elusive.

Here, in this study, we adopted the ChIP-seq method in combination with other high-throughput sequencing data, including Hi-C, histone modification and ATAC-seq, to genomically characterize the GmLUX (here referred to as Glyma.11G136600) binding sites and their epigenetic features in the soybean genome. Furthermore, by coupling these data and molecular experiments, we found that GmLUX could mediate the chromatin interaction to directly regulate a *CO-like* gene (Glyma.10G274300), which represents a new potential regulatory pathway of GmLUX. Together, these approaches provide important genomic resources for a comprehensive understanding of the regulatory mechanism of GmLUX.

2 Results

2.1 Genome-wide identification of GmLUX binding sites via ChIP-seq

To gain deeper insight into the GmLUX (Glyma.11G136600) regulatory network in soybean, two replicates of GmLUX ChIP-seq were generated from leaves of *GmLUX-flag* transgenic plants to characterize the binding preference genome-wide (Figure 1A; Supplementary Figure S1). Up to 55 M reads were obtained for the ChIP-seq and input libraries (Supplementary Table S1). After alignment to the soybean reference genome (Figure 1A), a promoter and downstream enrichment pattern of GmLUX was observed (Figure 1B), and a total of 4,373 GmLUX binding peaks were identified by the MACS2 software (Figure 1C; Supplementary Table S2). Additionally, these peaks tended to be located within 2 kb from the transcription start site (TSS) of their corresponding nearest genes, such as the promoter and exonic regions (Figures 1D, E). Furthermore, the gene ontology (GO) enrichment analysis of the GmLUX peak-associated genes (3,312) (Supplementary Table S2) revealed several enriched GO terms, such as nucleic acid binding transcription factor activity (GO:0001071), transcription factor

activity (GO:0003700) and sequence-specific DNA binding (GO:0043565), suggesting a robust transitional regulatory role of GmLUX (Figure 1F; Supplementary Table S3).

2.2 Construction of a putative GmLUX regulatory network

To further narrow down the potential genes directly targeted by GmLUX and constructed a GmLUX regulatory network, a comprehensive integration of GmLUX ChIP-seq, co-expression and motif scanning data was conducted, resulting in the identification of 33 key genes (Figures 2A, B; Supplementary Table S4). These genes, harboring a GmLUX binding motif in their promoters, were found to be co-expressed with GmLUX and associated with GmLUX enriched peaks, such as *Glyma.20G237200* and *Glyma.17G230500* (Figures 2C, D). Subsequently, these genes were then used to construct a putative GmLUX regulatory network with GmLUX as the central hub (Figure 2B). In addition, flower related genes (Supplementary Table S2) were identified in this network, such as *DOF* (Glyma.13G062500), *AP2/ERF* (Glyma.02G261700), *CO-like* (Glyma.10G274300) and *DnaJ* (Glyma.08G220000) (Figure 2B), which were reported to have potential roles in regulating flower and served as potential candidate genes for further investigation.

2.3 The epigenetic features of GmLUX binding peaks

The epigenetic states of chromatin are usually closely associated with TF binding sites, serving as an additional layer of gene regulation (Cuellar-Partida et al., 2012). Hence, three histone modifications were generated by ChIP-seq, including H3K9ac, H3K27ac and H3K4me1; moreover, ChIP-seq data for four additional histone modifications (H3K27me3, H3K4me3, H4K12ac and H3K14ac) were downloaded, along with ATAC-seq data, to comprehensively understand the epigenetic states around GmLUX binding peaks (Figure 3A; Supplementary Figures S2, S3). As shown in Figure 3A, a strong enrichment of ATAC-seq, H3K4me3 and H4K12ac was observed at the center of GmLUX binding peaks, accompanied by a relatively weaker signal of H3K27ac, H3K14ac and H3K9ac, while little enrichment of H3K27me3 and H3K4me1 signals was observed at these regions. Consistent with these observations, we identified over 2,000 GmLUX binding peaks with ATAC-seq, H3K4me3 and H4K12ac modifications, and a few of peaks with H3K27me3 and H3K4me1 modifications (Figure 3B). Since H3K27me3 modification is a repressive mark, in contrast to the remaining active epigenetic marks, a significant difference was observed between the expression of H3K27me3-modified GmLUX binding peaks associated genes and those genes with other epigenetic mark modified peaks (Figure 3C). These observations indicated that GmLUX binding peaks associated with the active epigenetic marks (e.g. ATAC, H3K4me3 and H4K12ac) (Supplementary Figure S5) to regulate downstream gene expression.

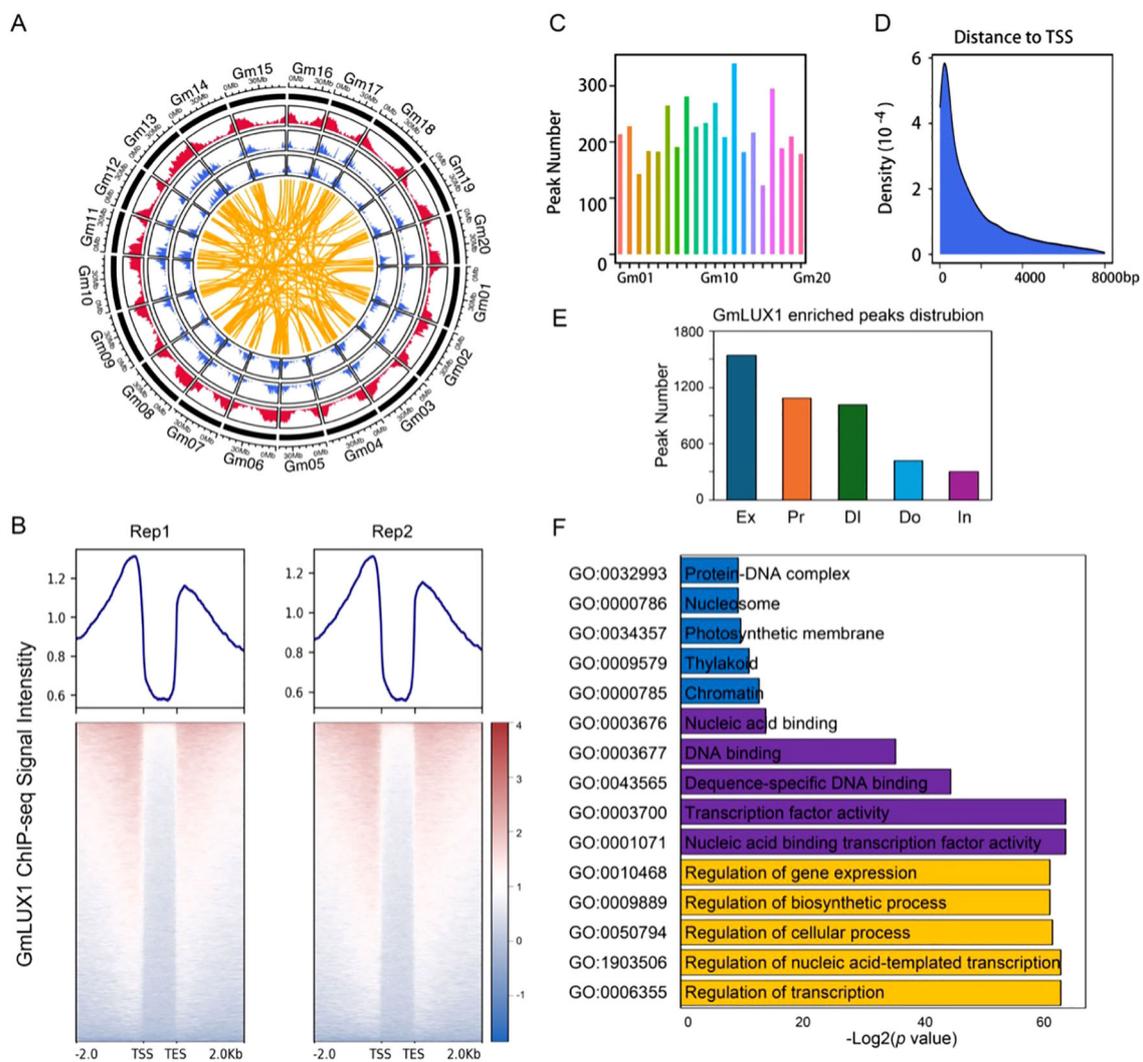


FIGURE 1
 Data summary of GmLUX ChIP-seq data. **(A)** A circos diagram showing the genomic distribution of GmLUX ChIP-seq peaks in the soybean genome. The red outer circle indicates the GmLUX associated genes, while the blue inner circles (two biological replicates) indicate the GmLUX enriched peaks. The yellow links indicate the homoeologous gene pairs targeted by GmLUX. **(B)** The GmLUX enriched profiles which have been normalized to input controls. Rep1 and Rep2 indicate two biological replicates. TSS, transcription start site. **(C)** Number of GmLUX enriched peaks in each chromosome. **(D)** The distance from GmLUX binding sites to transcription start site (TSS) of their nearest associated gene. **(E)** Genomic distribution of GmLUX enriched peaks in soybean genome. Ex, exon; Pr, Promoter; DI, Distal intergenic region; Do, Downstream region; In, Intron. **(F)** Gene ontology (GO) enrichment analysis of the genes associated with GmLUX enriched peaks.

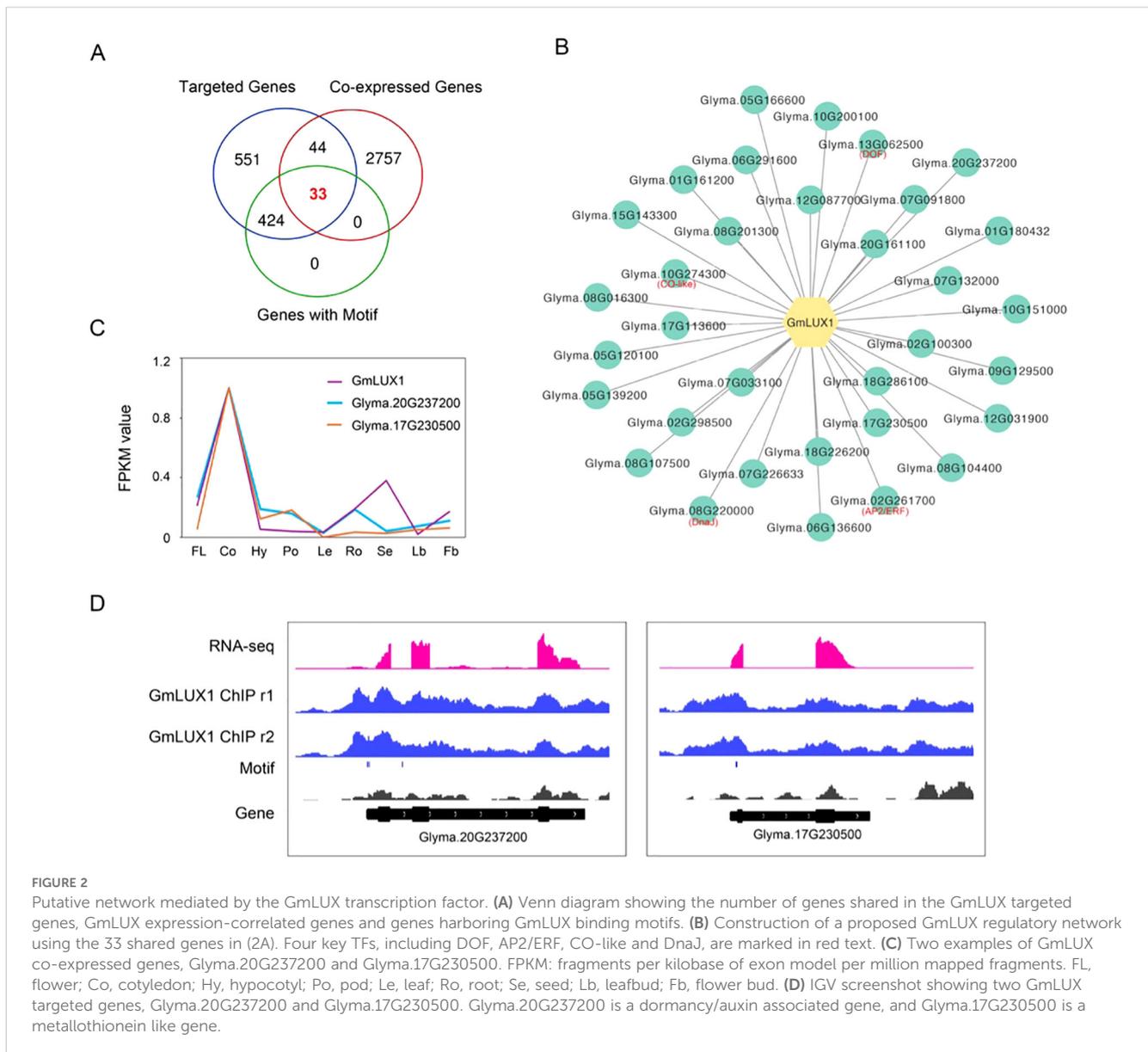
Furthermore, many GmLUX binding peaks with multiple modifications were observed, which could have a significant impact on the downstream gene expression (Figure 3D; Supplementary Table S5). For example, gene expression with ATAC and acetylation (Kac) modified peaks was significantly higher than that with ATAC alone, and similar situations were also found in the H3K4me3&Kac versus H3K4me3 group, as well as Kac versus H3K12ac group (Figure 3E).

In addition, the epigenetic modification enrichment pattern of the distal GmLUX binding peaks was distinct from the proximal peaks, except for the ATAC-seq signal (Supplementary Figure S4). As the distal ATAC-seq signal was served as an indicator of potential distal enhancer-like elements, this observation suggested

an enhancer role for distal GmLUX binding peaks, which required further investigation.

2.4 The impact of presence/absence of GmLUX binding peak on the expression of homologous genes

As a palaeopolyploid plant (Schmutz et al., 2010), soybean contains lots of duplicated genes in its genome. As previously reported, the gain and loss of *cis*-regulatory sequences (e.g. the ACR) could have a significant effect on the expression of homologous genes (hGenes) (Huang et al., 2021c; Fang et al., 2023).



We focused on 1,011 genes with a promoter GmLUX binding site, identified their corresponding hGenes, and checked whether their hGenes shared the GmLUX binding peaks with them. Among these 1,011 paired hGenes, only 224 hGenes (22.2%) were bound by the GmLUX, while 787 hGenes (77.8%) exhibited an absence of GmLUX binding (Supplementary Table S6). Furthermore, it was observed that those hGenes with both GmLUX binding peaks had similar epigenetic modifications and showed no significant impact on gene expression ($p = 0.673$) (Figures 4A, B; Supplementary Table S6). However, the expression of those hGenes without GmLUX binding peaks was significantly lower than their hGenes (Figures 4A, B; Supplementary Table S6). In addition, the absence of GmLUX binding peaks also caused obvious changes in the epigenetic features of hGenes (Figure 4B; Supplementary Table S6). These data suggested that the presence or absence of *cis*-regulatory sequences during gene duplication were important for maintaining the expression of

hGenes, which may have implications for the subsequent functionalization of these genes.

2.5 Molecular evidence for GmLUX-mediated promoter-enhancer interaction for gene expression

Enhancer elements played an important role in gene activation and could be located in the intergenic or intronic accessible regions (Zhu et al., 2015; Meng et al., 2021). In addition, as previously reported, TF could mediate the promoter-enhancer interaction to modulate gene expression in soybean (Huang et al., 2023). In this study, a total of 1,018 and 301 GmLUX binding peaks were observed located in the intergenic and intronic regions of the soybean genome, respectively (Supplementary Table S2). Of these,

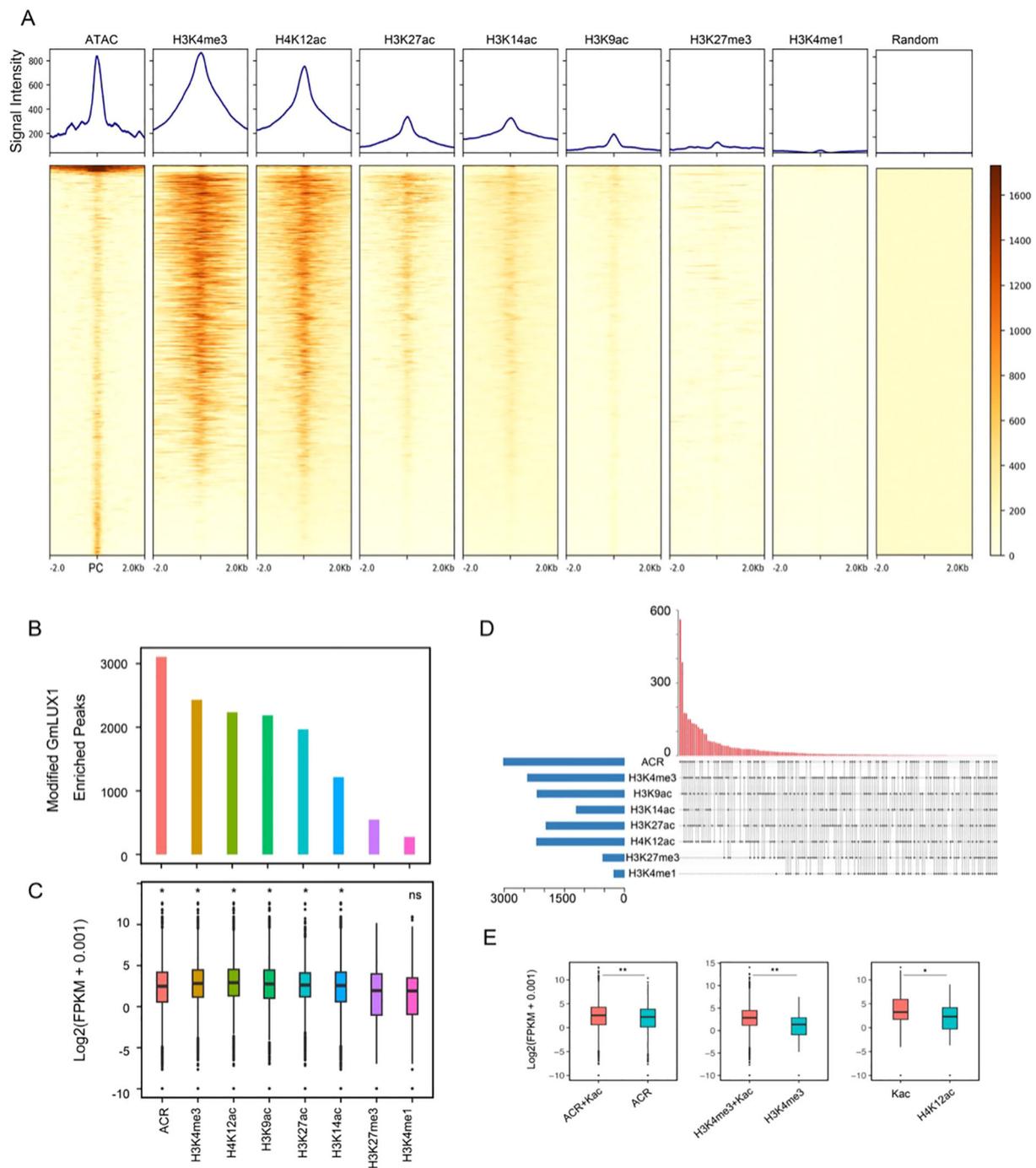
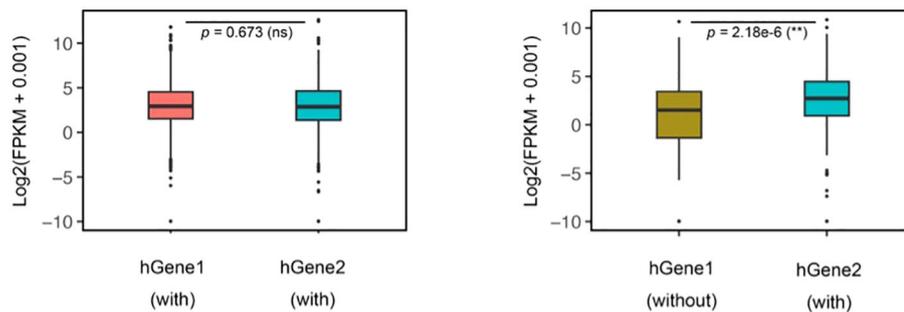


FIGURE 3

Epigenetic modifications at GmLUX binding peaks. **(A)** Enriched signals of ATAC-seq, H3K4me3, H4K12ac, H3K27ac, H3K14ac, H3K9ac, H3K27me3 and H3K4me1 ChIPed at the GmLUX peak center (PC), respectively. Random means the random genomic regions, which served as a control. **(B)** Number of GmLUX peaks associated with the distinct epigenetic modification signal. **(C)** Expression levels of genes associated with distinct modified GmLUX peaks. Asterisk (*) indicates the expression level is significantly higher than H3K27me3-modified GmLUX peaks associated genes ($p < 0.05$). Ns, not significant. P value was calculated by Wilcoxon test. ACR, accessible chromatin region. **(D)** Up-set Venn diagram showing the numbers of GmLUX enriched peak associated with multiple epigenetic modification signals. **(E)** The GmLUX targeted genes with multiple epigenetic modification signals are generally higher than those only with single modification signal. Two asterisks (**) indicate the p value less than 0.01, while one asterisk (*) means $0.01 < p < 0.05$. Kac indicates four histone acetylation. P value was calculated by Wilcoxon test.

A



B

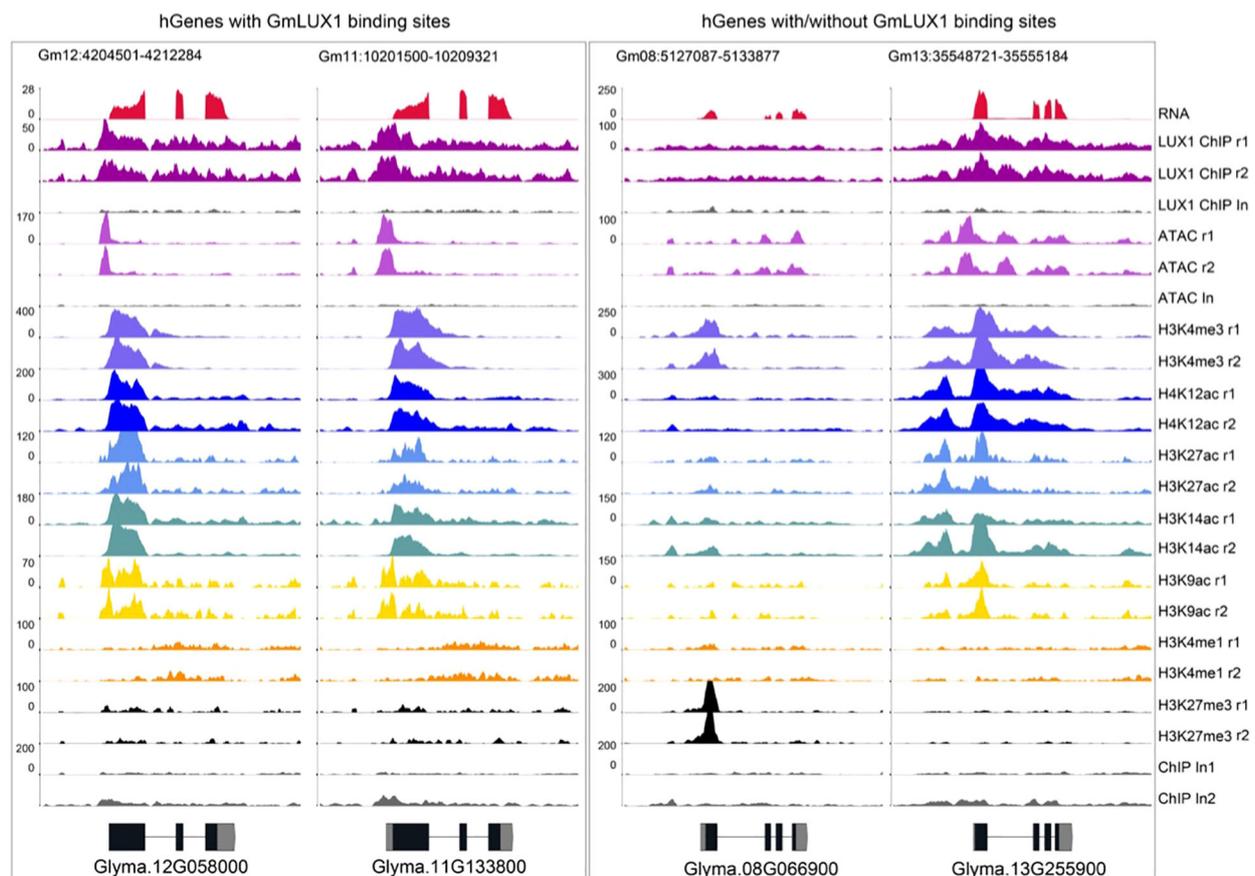
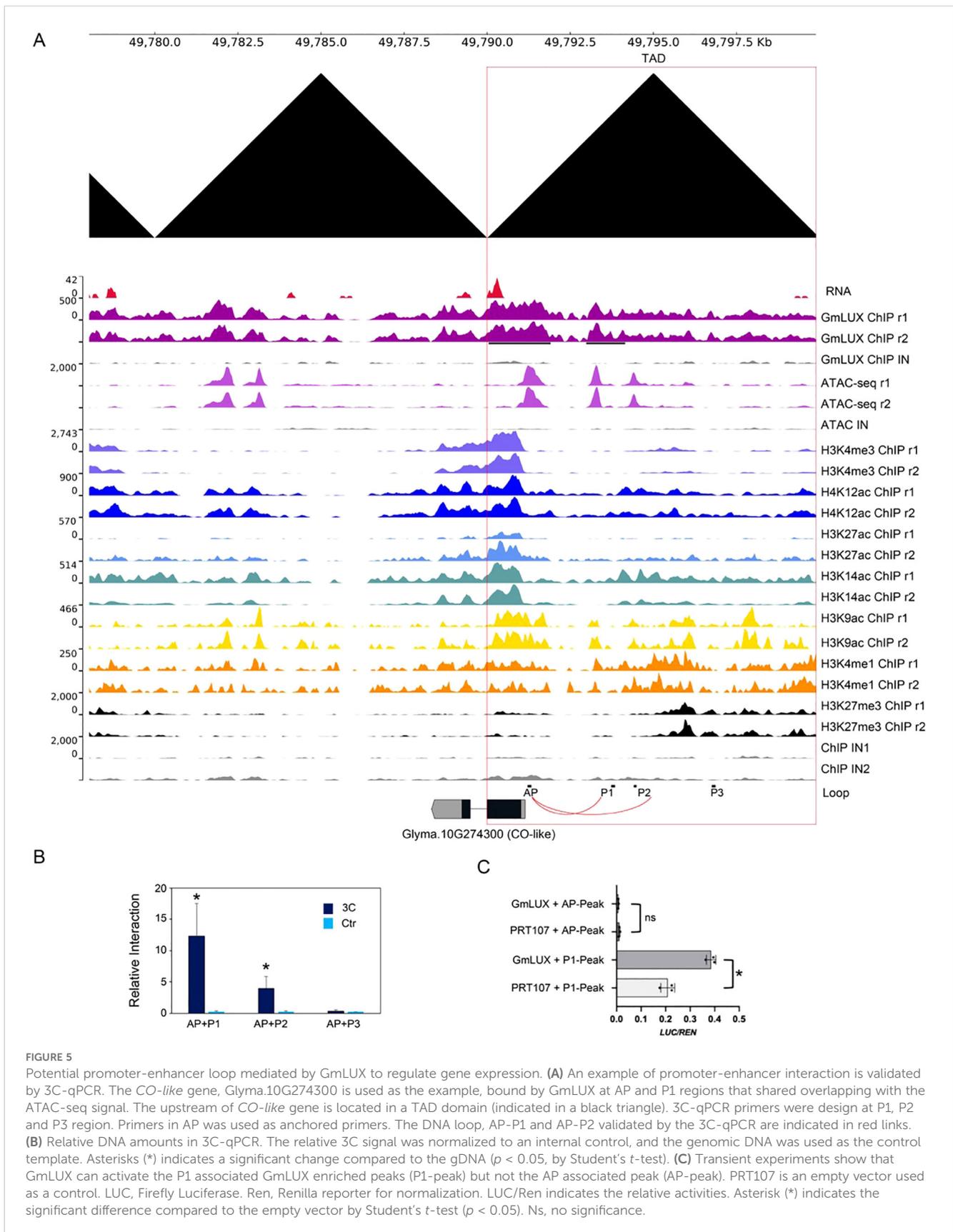


FIGURE 4

Potential effects of presence/absence of GmLUX binding region in homologous genes on gene expression. (A) Expression levels between homologous genes (hGenes) bound by GmLUX show no significant (ns), while the expression of genes associated with GmLUX is significantly higher than that of its homologous genes without GmLUX peaks. *P* value was calculated by the Wilcoxon test. (B) Two examples of hGenes with or without GmLUX binding peaks. Left panel, hGenes both with GmLUX have similar epigenetic modification signals; Right panel, distinct epigenetic modified signals are observed between hGenes with or without GmLUX binding peaks.

837 intergenic and 160 intronic peaks overlapped with ATAC-seq peaks. Furthermore, a total of 114 genes were associated with multiple GmLUX binding peaks (41.8% associated with promoter and intronic GmLUX peaks and 58.2% associated with promoter and intergenic peaks) (Supplementary Table S2). Since ACRs identified by ATAC-seq served as an indicator of enhancer-like elements, this observation indicated that GmLUX was able to mediate enhancer-like elements to regulate gene expression.

To further test our hypothesis, a flowering-related gene, *CO-like* (Glyma.10G274300) was selected as a candidate for further validation. This *CO-like* gene harbored multiple *cis*-regulatory regions located in its upstream region (three peaks of ATAC-seq and two peaks of GmLUX) (Figure 5A). Interestingly, *CO-like* was associated with several active histone modifications (e.g. H3K4me3, H4K12ac, H3K14ac etc.) and had relatively high expression in soybean (Figure 5). Moreover, the upstream region of this *CO-like*



gene located in a topologically associating domain (TAD, showed as a black triangle) region according to the Hi-C data, indicating a high frequency of chromatin interaction in these regions (Figure 5A). To confirm such an interaction, the 3C-PCR was performed using three primers (P1, P2 and P3) and an anchor primer (AP). It showed that the interaction signals of AP-P1 and AP-P2 were detected, except for that of AP-P3 (Figure 5B). Since AP and P1 were close to two GmLUX binding peaks, the data suggested that GmLUX mediated the AP-P1 interaction to regulate *CO-like* gene expression. Interestingly, in the transient experiments, only the P1 but not the AP *cis*-regulatory sequences had relatively high activities in protoplasts and could be activated by GmLUX overexpression (Figure 5C), indicating the distal binding region (e.g. P1) was necessary for GmLUX to regulate its target genes (e.g. the *CO-like* gene).

3 Discussion

Plant LUX is a SHAQYF-type GARP transcription factor containing a single MYB domain that can bind to the LBS motif (GATA/CCG) in target genes promoters (Hazen et al., 2005; Onai and Ishiura, 2005; Helfer et al., 2011; Zhang et al., 2019). In soybean, two LUX homologs were functionally redundant but together played critical roles in regulation of soybean flowering by directly binding to the LUX binding sites of *E1s* promoters (Lu et al., 2017; Bu et al., 2021; Lin et al., 2022). In this study, a multi-omics analysis has been adopted to identify a novel regulatory pathway of a GmLUX gene (Glyma.11G136600) that promotes soybean flowering.

In this study, the GmLUX ChIP-seq was performed using the robust, low-input requirement ChIPmentation method, which has been successfully validated in *Arabidopsis* (Lee and Bailey-Serres, 2019) and soybean (Huang et al., 2021b). Using this approach, the genome-wide distribution of GmLUX-specific binding sites has been determined, and a putative robust transitional regulatory network centered on GmLUX was also established. In this network, a *CO-like* gene (Glyma.10G274300) was identified, which has been reported to potentially regulate flowering (Wu et al., 2014; Cao et al., 2015), indicating the advantageous role of the multi-omics approach in mining the key regulatory factors involved in specific developmental stages.

Enhancer elements, essential *cis*-regulatory elements, can mediate the formation of chromatin loops to regulate gene expression (Li et al., 2019; Ricci et al., 2019; Huang et al., 2023). In this study, similar to the GmJAG1 (Huang et al., 2021b), we showed that GmLUX could mediate enhancer-like elements to regulate gene expression. This discovery is due to the relative ease of access to many high-throughput data (ChIP-seq, ATAC-seq and Hi-C etc.). For example, the high-throughput sequencing data, such as Hi-C, is necessary for the precise location of enhancer-like elements and the enhancer-mediated DNA loops on a genome-wide scale (Sanyal et al., 2012). Based on Hi-C analysis, a high frequency of sequence interactions has been observed between GmLUX and enhance-like elements of the *CO-like* gene which was associated with multiple active histone modifications and

contained several *cis*-regulatory regions located in its upstream region (Figure 5). Also, to genome-wide map the enhancer-promoter interaction, new technologies such as the STARR-seq (Arnold et al., 2013; Zhang et al., 2022), HiChIP (Mumbach et al., 2016) etc., can be used for further investigation. Given the importance of non-coding regulatory sequences in gene regulation (Huang et al., 2021a) and the increasing amount of high-throughput sequencing data, a comprehensive database such as the ENCODE project focused on these data in soybean, is important for further investigation into the function of these non-coding elements.

Taken together, a flowering-related gene, *CO-like* can be directly regulated by GmLUX binding to its promoter, according to a comprehensive analysis of ChIP-seq, histone modification, Hi-C, ATAC-seq and molecular experiments, suggesting a new potential GmLUX regulatory pathway in modulating soybean flowering.

4 Method and material

4.1 Plant materials and growth conditions

Glycine max (L.) Merr. cultivar Williams 82 (W82) was used as the wild type in this study, and transgenic W82 with 35S:*GmLUX-Flag* (Bu et al., 2021). Transgenic and wild-type seeds were sterilized with 75% ethanol and placed on sterilized vermiculite for germination, until their cotyledons were fully expanded. These seedlings were transferred to soil and grown in the greenhouse (16 h light/8 h dark, 25–28°C) to develop the healthy trifoliolate leaves, which were utilized for subsequent experiments.

4.2 ChIP-seq library construction and data processing

The construction of GmLUX ChIP-seq library was followed the previous study (Huang et al., 2021b), with minor modifications. Briefly, about 0.5 g of leaves from *GmLUX-flag* transgenic plants were utilized for nuclear isolation. The chromatin fragments were pull-down by the anti-Flag antibody (Sigma). The ChIPed and input DNA used for library construction were processed by Tn5 transposase (Vazyme). Histone ChIP-seq library construction was followed the method (Huang et al., 2021b; Zhang et al., 2023) and the anti-H3K27ac, anti-H3K9ac and anti-H3K4me1 antibodies (Millipore) were utilized in this study. All the ChIP-seq libraries were sequenced on the PE150 mode of the Illumina platform.

The raw data of ChIP-seq were trimmed by TrimGalore (<http://github.com/FelixKrueger/TrimGalore>) and the filtered reads were mapped to the soybean reference genome (V4, <https://phytozome.jgi.doe.gov>) by the Bowtie2 (Langmead and Salzberg, 2012). The high-quality mapped reads (MAPQ > 30) were extracted by SAMtools and used for peak calling by MACS2 (Zhang et al., 2008; Li et al., 2009). For broad peak calling, the parameters were set as '-trackline -extsize 147 -broad -q 0.01 -nomodel -buffer-size

500000', while the default parameters were used for narrow peak calling. The input library was used as a control. The profile of ChIP-seq is normalized to input controls by the BamCompare function in Deeptools (Ramírez et al., 2014). Enriched peaks annotation was performed via HOMER (<http://homer.ucsd.edu/homer/>). The visualization of ChIP-seq data was processed using the Deeptools (Ramírez et al., 2014), pyGenomtrack (Lopez-Delisle et al., 2021) or IGV (Thorvaldsdóttir et al., 2013) software.

4.3 RNA-seq data analysis

The RNA-seq data were obtained from a previous study (Huang et al., 2021b). Raw RNA-seq data were filtered by TrimGalore with the default parameters, and then mapped to the soybean reference genome using HISAT2 (<http://daehwankimlab.github.io/hisat2>). Gene expression levels were calculated by Cuffnorm (<http://cole-trapnell-lab.github.io/cufflinks/cuffnorm>) and represented by the fragments per kilobase per million mapped reads (FPKMs), which were utilized for subsequent analysis.

4.4 Hi-C data analysis

The leaf Hi-C data were downloaded from a previous report (Wang et al., 2021) and followed its analysis pipeline with minor modification. Briefly, the raw data was trimmed with the adaptors and were processed by the HiC-Pro (Servant et al., 2015) with default parameters. Seft-circle, dangling-end reads were removed, and the contacts were normalized by the iterative correction and eigenvector decomposition (ICE) method. Then the TAD domain was called by HiCexplore (Wolff et al., 2022) with default parameters and visualized by pyGenomtrack (Lopez-Delisle et al., 2021).

4.5 Motif analysis

The GmLUX motif (Motif ID: TFmatrixID_0354) position weight matrix (PWM) was downloaded from the PlantPAN database (Chow et al., 2019) and the FIMO from the MEME suite (Bailey et al., 2009) was used for motif scanning in the GmLUX peak region.

4.6 ATAC-seq data analysis

The soybean leaf ATAC-seq raw data were obtained from previous study (Huang et al., 2021b), and filtered by the TrimGalore and mapped to the reference genome via Bowtie2 (Langmead and Salzberg, 2012) with the parameters set as: -very-sensitive -X 1000. The mapped reads with a MAPQ value over 30 were used for peak calling via Genrich software on the ATAC mode (<https://github.com/jsh58/Genrich>). Visualization and annotation

of ATAC-seq data was performed similarly to the ChIP-seq mentioned above.

4.7 Measurement of transcriptional activity

To confirm the transcriptional activation of GmLUX, the full-length CDS of GmLUX was cloned into the pRT107 vector which was derived by the 35S promoter, while the two GmLUX enriched peak regions associated with the CO-like gene were cloned into the pGreenII-0800-LUC vectors and confirmed by Sanger sequencing. Then the validated plasmids of pRT107-GmLUX and pGreenII-0800-GmCO-like-Luc were co-transformed into *Arabidopsis* protoplasts, isolated from the leaves of 3-week-old Col-0 by the PEG/CaCl₂ method (Yoo et al., 2007). After 16 h of incubation at 22°C, the total proteins were extracted, and transcriptional activities were measured using the Dual-Luciferase Reporter Assay System (Promega, E1910) kit followed the manual instructions. The fluorescence value of firefly luciferase (LUC) and Renilla luciferase (Ren) was determined by Promega GloMax 20/20 system. The transcriptional activity of GmLUX against the CO-like associated peaks was evaluated by the ratio of LUC/Ren value. The experiments were performed in three biological replicates.

4.8 Chromosome conformation capture quantitative PCR

The 3C-qPCR was conducted according to the previous study (Hagege et al., 2007; Liu, 2017; Huang et al., 2023). Briefly, the isolated soybean leaf nuclei were digested with DnpII (NEB), and the sticky ends were filled with dWTP (W: A, T, G), biotin-14-dCTP (Sigma) and 40 U Klenow fragment (NEB). The blunt ends were then ligated with the T4 DNA ligase (NEB) following the manufacturer's protocol. The biotin-labeled 3C-DNA was extracted by MyOne™ Streptavidin C1 Dynabeads (Invitrogen) according to the manufacturer's protocol and used for qPCR using the TB Green Premix Ex II (Takara). The primers involved in 3C-qPCR analysis were listed in Supplementary Table S1.

5 Conclusions

In this study, the genomic characteristics of GmLUX (Glyma.11G136600) binding sites and their epigenetic features were revealed, using a modified ChIPmentation method, along with other high-throughput sequencing data, including Hi-C, histone modification ChIP-seq and ATAC-seq. Moreover, a new regulatory pathway of GmLUX has been identified, in which GmLUX facilitates the chromatin interaction to directly regulate a CO-like gene (Glyma.10G274300). This will provide important genomic resources for a comprehensive understanding of the regulatory mechanism of GmLUX.

Data availability statement

The data presented in the study are deposited in the NGDC (<https://ngdc.cnbc.ac.cn/>) repository, accession number PRJCA034927.

Author contributions

LT: Formal analysis, Funding acquisition, Writing – original draft, Investigation, Methodology. HX: Methodology, Formal analysis, Investigation, Writing – original draft. JW: Investigation, Methodology, Writing – original draft, Formal analysis, Software. LZ: Software, Formal analysis, Methodology, Writing – review & editing. TF: Software, Writing – review & editing, Writing – original draft, Methodology, Investigation. MH: Writing – review & editing, Data curation, Software, Formal analysis. CT: Data curation, Formal analysis, Writing – review & editing. HY: Writing – review & editing, Conceptualization, Formal analysis. YH: Writing – original draft, Conceptualization, Investigation, Formal analysis, Methodology.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was funded by the National Natural Science Foundation of China (32201800). This work was also supported by Natural Science Foundation of Guangdong Province (2023A1515011645).

Acknowledgments

We kindly thank the Prof. Fanjiang Kong for providing the transgenic soybean line (35S: GmLUX-Flag).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Arnold, C. D., Gerlach, D., Stelzer, C., Boryn, L.M., Rath, M., and Stark, A. (2013). Genome-wide quantitative enhancer activity maps identified by STARR-seq. *Science* 339, 1074–1077. doi: 10.1126/science.1232542
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. doi: 10.1093/nar/gkp335
- Blanc-Mathieu, R., Dumas, R., Turchi, L., Lucas, J., and Parcy, F. (2024). Plant-TFClass: a structural classification for plant transcription factors. *Trends Plant Sci.* 29, 40–51. doi: 10.1016/j.tplants.2023.06.023
- Bu, T., Lu, S., Wang, K., Dong, L., Li, S., Xie, Q., et al. (2021). A critical role of the soybean evening complex in the control of photoperiod sensitivity and adaptation. *Proc. Natl. Acad. Sci.* 118, e2010241118. doi: 10.1073/pnas.2010241118
- Cao, D., Li, Y., Lu, S., Wang, J., Nan, H., Li, X., et al. (2015). GmCOL1a and GmCOL1b function as flowering repressors in soybean under long-day conditions. *Plant Cell Physiol.* 56, 2409–2422. doi: 10.1093/pcp/pcv152
- Chow, C. N., Lee, T. Y., Hung, Y. C., Li, G. Z., Tseng, K. C., Liu, Y. H., et al. (2019). PlantPAN3.0: a new and updated resource for reconstructing transcriptional regulatory networks from ChIP-seq experiments in plants. *Nucleic Acids Res.* 47, D1155–D1163. doi: 10.1093/nar/gky1081

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1607224/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Pearson correlations between two replicates of GmLUX ChIP-seq. R1 and R2 indicate two biological replicates.

SUPPLEMENTARY FIGURE 2

Pearson correlations between two replicates of histone ChIP-seq data. R1 and R2 indicate two biological replicates. The H3K27me3 ChIP-seq data were downloaded from the previous study.

SUPPLEMENTARY FIGURE 3

Histone ChIP-seq enrichment analysis. H3K27ac, H3K4me1 and H3K9ac ChIP-seq signals are positively correlated with gene expression levels. TSS, transcription start site; TES, transcription end site; H, highly expressed genes (FPKM > 10); M, middle expressed genes (1 < FPKM ≤ 10); L, low expressed genes (FPKM ≤ 1); r1 and r2 indicate the two ChIP-seq replicates.

SUPPLEMENTARY FIGURE 4

Distinct enrichment pattern epigenetic modification between proximal and distal GmLUX enriched peaks. Proximal GmLUX peaks indicate those located near the gene region including promoter, exon, intron and downstream, while the distal peaks indicate the distance from peaks to nearest genes over 2 kb.

SUPPLEMENTARY FIGURE 5

Venn diagram showing the overlap of GmLUX bound peaks associated with ACR or H3K4me3/H3K12ac. P value was calculated by the hypergeometric test.

- Cuellar-Partida, G., Buske, F. A., McLeay, R. C., Whittington, T., Noble, W. S., and Bailey, T. L. (2012). Epigenetic priors for identifying active transcription factor binding sites. *Bioinformatics* 28, 56–62. doi: 10.1093/bioinformatics/btr614
- Dhatterwal, P., Sharma, N., and Prasad, M. (2024). Decoding the functionality of plant transcription factors: key factors and mechanisms. *J. Exp. Bot.* 75, 4745–4759. doi: 10.1093/jxb/erae231
- Fang, C., Yang, M., Tang, Y., Zhang, L., Zhao, H., Ni, H., et al. (2023). Dynamics of cis-regulatory sequences and transcriptional divergence of duplicated genes in soybean. *Proc. Natl. Acad. Sci.* 120, e2303836120. doi: 10.1073/pnas.2303836120
- Gao, F., and Dubos, C. (2024). The Arabidopsis bHLH transcription factor family. *Trends Plant Sci.* 29, 668–680. doi: 10.1016/j.tplants.2023.11.022
- Hagege, H., Klous, P., Braem, C., Splinter, E., Dekker, J., Cathala, G., et al. (2007). Quantitative analysis of chromosome conformation capture assays (3C-qPCR). *Nat. Protoc.* 2, 1722–1733. doi: 10.1038/nprot.2007.243
- Hazen, S. P., Schultz, T. F., Prunedo-Paz, J. L., Borevitz, J. O., Ecker, J. R., and Kay, S. A. (2005). LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. *Proc. Natl. Acad. Sci.* 102, 10387–10392. doi: 10.1073/pnas.0503029102
- Helfer, A., Nusinow, D. A., Chow, B. Y., Gehrke, A. R., Bulyk, M. L., and Kay, S. A. (2011). LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the Arabidopsis core clock. *Curr. Biol.* 21, 126–133. doi: 10.1016/j.cub.2010.12.021
- Hrmova, M., and Hussain, S. S. (2021). Plant transcription factors involved in drought and associated stresses. *Int. J. Mol. Sci.* 22, 5662. doi: 10.3390/ijms22115662
- Huang, M. K., Li, M. W., and Lam, H. M. (2021a). How noncoding open chromatin regions shape soybean domestication. *Trends Plant Sci.* 26, 876–878. doi: 10.1016/j.tplants.2021.06.008
- Huang, M. K., Zhang, L., Yung, W. S., Hu, Y. F., Wang, Z. L., Li, M. W., et al. (2023). Molecular evidence for enhancer-promoter interactions in light responses of soybean seedlings. *Plant Physiol.* 193, 2287–2291. doi: 10.1093/plphys/kiad487
- Huang, M. K., Zhang, L., Zhou, L. M., Wang, M. Z., Yung, W. S., Wang, Z. L., et al. (2021b). 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* 113, 344–355. doi: 10.1016/j.ygeno.2020.12.026
- Huang, M. K., Zhang, L., Zhou, L. M., Yung, W. S., Li, M. W., and Lam, H. M. (2021c). Genomic features of open chromatin regions (OCRs) in wild soybean and their effects on gene expressions. *Genes* 12, 640. doi: 10.3390/genes12050640
- Huang, M. K., Zhang, L., Zhou, L. M., Yung, W. S., Wang, Z. L., Xiao, Z. X., et al. (2022). Identification of the accessible chromatin regions in six tissues in the soybean. *Genomics* 114, 110364. doi: 10.1016/j.ygeno.2022.110364
- Iwafuchi-Doi, M., Donahue, G., Kakumanu, A., Watts, J. A., Mahony, S., Pugh, B. F., et al. (2016). The pioneer transcription factor foxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. *Mol. Cell* 62, 79–91. doi: 10.1016/j.molcel.2016.03.001
- Klemm, S. L., Shipony, Z., and Greenleaf, W. J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. doi: 10.1038/s41576-018-0089-8
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lee, T. A., and Bailey-Serres, J. (2019). Integrative analysis from the epigenome to transcriptome uncovers patterns of dominant nuclear regulation during transient stress. *Plant Cell* 31, 2573–2595. doi: 10.1105/tpc.19.00463
- Li, E., Liu, H., Huang, L., Zhang, X., Dong, X., Song, W., et al. (2019). Long-range interactions between proximal and distal regulatory regions in maize. *Nat. Commun.* 10, 2633. doi: 10.1038/s41467-019-10603-4
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Lin, X., Dong, L., Tang, Y., Li, H., Cheng, Q., Li, H., et al. (2022). Novel and multifaceted regulations of photoperiodic flowering by phytochrome A in soybean. *Proc. Natl. Acad. Sci.* 119, e2208708119. doi: 10.1073/pnas.2208708119
- Liu, C. (2017). “In situ hi-C library preparation for plants to study their three-dimensional chromatin interactions on a genome-wide scale,” in *Plant Gene Regulatory Networks: Methods and Protocols*. Eds. K. Kaufmann and B. Mueller-Roeber (Springer New York, New York, NY), 155–166.
- Liu, L., White, M. J., and MacRae, T. H. (1999). Transcription factors and their genes in higher plants. *Eur. J. Biochem.* 262, 247–257. doi: 10.1046/j.1432-1327.1999.00349.x
- Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., et al. (2021). pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* 37, 422–423. doi: 10.1093/bioinformatics/btaa692
- Lu, S., Zhao, X., Hu, Y., Liu, S., Nan, H., Li, X., et al. (2017). Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. *Nat. Genet.* 49, 773–779. doi: 10.1038/ng.3819
- Lu, Z., Marand, A. P., Ricci, W. A., Ethridge, C. L., Zhang, X., and Schmitz, R. J. (2019). The prevalence, evolution and chromatin signatures of plant regulatory elements. *Nat. Plants* 5, 1250–1259. doi: 10.1038/s41477-019-0548-z
- Meng, F., Zhao, H., Zhu, B., Zhang, T., Yang, M., Li, Y., et al. (2021). Genomic editing of intronic enhancers unveils their role in fine-tuning tissue-specific gene expression in Arabidopsis thaliana. *Plant Cell* 33, 1997–2014. doi: 10.1093/plcell/koab093
- Mumbach, M. R., Rubin, A. J., Flynn, R. A., Dai, C., Khavari, P. A., Greenleaf, W. J., et al. (2016). HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* 13, 919–922. doi: 10.1038/nmeth.3999
- Onai, K., and Ishiura, M. (2005). PHYTOCLOCK 1 encoding a novel GARP protein essential for the Arabidopsis circadian clock. *Genes Cells* 10, 963–972. doi: 10.1111/j.1365-2443.2005.00892.x
- Ramírez, F., Dündar, F., Diehl, S., Grüning, B. A., and Manke, T. (2014). deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 42, W187–W191. doi: 10.1093/nar/gku365
- Ricci, W. A., Lu, Z., Ji, L., Marand, A. P., Ethridge, C. L., Murphy, N. G., et al. (2019). Widespread long-range cis-regulatory elements in the maize genome. *Nat. Plants* 5, 1237–1249. doi: 10.1038/s41477-019-0547-0
- Sanyal, A., Lajoie, B. R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* 489, 109–113. doi: 10.1038/nature11279
- Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183. doi: 10.1038/nature08670
- Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C.-J., Vert, J.-P., et al. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 16, 259. doi: 10.1186/s13059-015-0831-x
- Strader, L., Weijers, D., and Wagner, D. (2022). Plant transcription factors — being in the right place with the right company. *Curr. Opin. Plant Biol.* 65, 102136. doi: 10.1016/j.pbi.2021.102136
- Thorvaldsdóttir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings Bioinf.* 14, 178–192. doi: 10.1093/bib/bbs017
- Wang, L., Jia, G., Jiang, X., Cao, S., Chen, Z. J., and Song, Q. (2021). Altered chromatin architecture and gene expression during polyploidization and domestication of soybean. *Plant Cell* 33, 1430–1446. doi: 10.1093/plcell/koab081
- Wolff, J., Backofen, R., and Grüning, B. (2022). Loop detection using Hi-C data with HiCExplorer. *GigaScience* 11. doi: 10.1093/gigascience/giac061
- Wu, F., Price, B. W., Haider, W., Seufferheld, G., Nelson, R., and Hanzawa, Y. (2014). Functional and evolutionary characterization of the CONSTANS gene family in short-day photoperiodic flowering in soybean. *PLoS One* 9, e85754. doi: 10.1371/journal.pone.0085754
- Yoo, S.-D., Cho, Y.-H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199
- Zhang, C., Gao, M., Seitz, N. C., Angel, W., Hallworth, A., Wiratan, L., et al. (2019). LUX ARRHYTHMO mediates crosstalk between the circadian clock and defense in Arabidopsis. *Nat. Commun.* 10, 2543. doi: 10.1038/s41467-019-10485-6
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoutte, J., Johnson, D. S., Bernstein, B. E., et al. (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* 9, R137. doi: 10.1186/gb-2008-9-9-r137
- Zhang, L., Yung, W.-S., Hu, Y., Wang, L., Sun, W., and Huang, M. (2023). Establishment of a convenient ChIP-seq protocol for identification of the histone modification regions in the medicinal plant *Andrographis paniculata*. *Med. Plant Biol.* 2. doi: 10.48130/MPB-2023-0006
- Zhang, L., Yung, W.-S., and Huang, M. (2022). STARR-seq for high-throughput identification of plant enhancers. *Trends Plant Sci.* 27, 1296–1297. doi: 10.1016/j.tplants.2022.08.008
- Zhou, C., Yuan, Z., Ma, X., Yang, H., Wang, P., Zheng, L., et al. (2021). Accessible chromatin regions and their functional interrelations with gene transcription and epigenetic modifications in sorghum genome. *Plant Commun.* 2, 100140. doi: 10.1016/j.xplc.2020.100140
- Zhu, B., Zhang, W., Zhang, T., Liu, B., and Jiang, J. (2015). Genome-wide prediction and validation of intergenic enhancers in Arabidopsis using open chromatin signatures. *Plant Cell* 27, 2415–2426. doi: 10.1105/tpc.15.00537