nature portfolio

Corresponding author(s): Yingxiang Wang & Cunyi Yang

Last updated by author(s): Jun 8, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The plants and flowers were photographed using a Canon E0S 90D. Immunofluorescence expriment were obtained using Zeiss Axio Imager A2 microscope. The subcellular localization of related genes was acquired by LSM-710 confocal microscope (Zeiss). The immunoblot analysis was acquired by eBlot Touch Imager. Illumina Hiseq X Ten platform was used for RNA-seq and ChIP-seq. CFX96 real-time PCR system (Bio-Rad) was used for RT-qPCR.
Data analysis	For BSA analysis: BWA 0.7.17-r1188, GATK 1.9, and SnpEff 3.1. MEGA 11 software was used in the phylogenetic tree construction. For RNA-seq analysis: HISAT2 2.1.0, StringTie2 2.2, FeatureCounts 2.0.3, DESeq2 package 1.34.0, GOATOOLS version 1.3.3, and KOBAS-I. For activity assay: LSM-710 confocal microscope (Zeiss) and ImageJ 1.8.0. For ChIP-seq analysis: Trim Galore 0.6.8, Picard 2.23.3, MACS2 2.2.9.1, ChIPseeker, deepTools2, bamCoverage, and Genome Viewer (IGV) 2.13.2. For haplotype analysis: BCFtools 1.14 and VCFtools 0.1.16 . For protein structure: Phyre2 and PyMoL 2.5. For statistical analysis: SPSS 19 and BEDTools 2.30.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The genome resequencing of HX3, ef1, wide type pool, and mutant pool, along with RNA-Seq and ChIP-seq data of HX3 and ef1, have been deposited in the Genome Sequence Archive (GSA) at the National Genomics Data Center (https://bigd.big.ac.cn/gsa/) under accession number CRA019775.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the phenotype statistic for the plants, at lest 10 samples for each genotype were randomly selected. For the field traits, we tested at least 6 plants of each accession. For RNA-seq and ChIP-seq, 2 independent repeats of every type materials. For RT-qPCR and ChIP-qPCR, 3 biological repeats were sampled. The sample sizes and material amount were determined based on experimental trials to allow for confident statistical analyses.
Data exclusions	No data were excluded from the analysis.
Replication	All analyses in this study were repeated more than two times depending on the different purposes of each experiment. All attempts at replication were successful.
Randomization	All candidate samples in this study including plants, bacterial strains, micrographs of cells and nucleus were randomly selected. All samples for biochemistry assays were also randomly selected in this study.
Blinding	All data in this study were collected from blinded to group allocation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Anti-H3K4me1 Polyclonal Antibody (Millipore, Cat. # 07-436; 1:2000 dilution); Anti-H3K4me2 Polyclonal Antibody (Millipore, Cat. # 07-030; 1:1000 dilution); Anti-H3K4me3 Polyclonal Antibody (Abcam, Cat. # ab8580; 1:2000 dilution); Anti-H3K9me2 Monoclonal Antibody (Abcam, Cat. # ab1220; 1:2000 dilution); Anti-H3K27me1 Polyclonal Antibody (Millipore, Cat. # 07-448; 1:2000 dilution); Anti-H3K36me3 Polyclonal Antibody (Abcam, Cat. # ab9050; 1:2000 dilution); Anti-H3 Polyclonal Antibody (Abcam, Cat. # ab1791; 1:4000 dilution); Anti-Actin Monoclonal Antibody (Abmart, Cat. # M20009M, 1:2000 dilution); Anti-Falg Monoclonal Antibody (GNI, Cat. # GNI4110-FG, 1:2000 dilution); GmLDL2 Rabbit Polyclonal Antibody was raised in rabbits and generated by GL Biochem, China (1:2000 dilution). Goat anti-Mouse IgG-HRP Secondary Antibody (GNI, Cat. # GNI9310-M, 1:2500 dilution); Goat anti-Rabbit IgG-HRP Secondary Antibody (GNI, Cat. # GNI9310-R, 1:2500 dilution); Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Flour 488 (Invitrogen, Cat. # A11008; 1:400 dilution); Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Flour 555 (Invitrogen, Cat. # A21428; 1:1000 dilution).

Validation

The validation statement of commercial antibodies can be found in the manufacturers' websites:
Anti-H3K4me1 Polyclonal Antibody (https://www.sigmaaldrich.cn/CN/zh/product/mm/07436);
Anti-H3K4me2 Polyclonal Antibody (https://www.sigmaaldrich.cn/CN/zh/product/mm/07030);
Anti-H3K4me3 Polyclonal Antibody (https://www.abcam.cn/products/primary-antibodies/histone-h3-tri-methyl-k4-antibody-chip-grade-ab8580.html);
Anti-H3K9me2 Monoclonal Antibody (https://www.abcam.cn/products/primary-antibodies/histone-h3-di-methyl-k9-antibody-mabcam-1220-chip-grade-ab1220.html);
Anti-H3K27me1 Polyclonal Antibody (https://www.sigmaaldrich.cn/CN/zh/product/mm/07448);
Anti-H3K36me3 Polyclonal Antibody (https://www.abcam.cn/products/primary-antibodies/histone-h3-tri-methyl-k36-antibody-chip-grade-ab9050.html);
Anti-H3 Polyclonal Antibody (https://www.abcam.cn/products/primary-antibodies/histone-h3-antibody-nuclear-marker-and-chip-grade-ab1791.html);
Anti-Actin Monoclonal Antibody (https://www.ab-mart.com.cn/page.aspx?node=%2059%20&id=%20985);
Anti-Falg Monoclonal Antibody (https://www.gnimission.com/resource/article/86);
Goat anti-Mouse IgG-HRP Secondary Antibody (https://www.gnimission.com/resource/article/5);
Goat anti-Rabbit IgG-HRP Secondary Antibody (https://www.gnimission.com/resource/article/2);
Goat anti-Rat IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Flour 488 (https://www.thermofisher.cn/cn/zh/antibody/product/Goat-anti-Rabbit-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-11008);
Goat anti-Rabbit IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Flour 555 (https://www.thermofisher.cn/cn/zh/antibody/product/Goat-anti-Rabbit-IgG-H-L-Cross-Adsorbed-Secondary-Antibody-Polyclonal/A-21428); GmLDL2 Rabbit Polyclonal Antibody has been validated in this study.

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
✘	☐	Public health
✘	☐	National security
✘	☐	Crops and/or livestock
✘	☐	Ecosystems
✘	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
✘	☐	Demonstrate how to render a vaccine ineffective
✘	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
✘	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
✘	☐	Increase transmissibility of a pathogen
✘	☐	Alter the host range of a pathogen
✘	☐	Enable evasion of diagnostic/detection modalities
✘	☐	Enable the weaponization of a biological agent or toxin
✘	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	The ef1, Gmldl2-2, Gmldl2-3 mutant, GmFER-OE and Gmldl2/fer double mutant trangenic seeds are stored in the seed bank of our laboratory; GmFER mutant: NJAU0800, NJAU1719, and NJAU1817 seeds come from Prof. Qingxin Song lab; The F2 seeds of HX3 × ef1 or ef1 × LNHM population, and 64 soybean accessions harboring different GmLDL2 haplotypesare stored in the seed bank of our laboratory.
Novel plant genotypes	The ef1 mutant was generated from soybean cultivar Huaxia 3 treated with 60Co γ-rays, and with a 1 bp deletion at the position The ef1 mutant was generated from soybean cultivar Huaxia 3 treated with 60Co γ-rays, and with a 1 bp deletion at the position number of 202nd in the first exon of GmLDL2; Gmldl2-2 and Gmldl2-3 were generated from Huaxia 3 by CRISPR/Cas9 system, Gmldl2-2 contains a 5 bp deletion at the position number of 28th in the first exon of GmLDL2, and Gmldl2-3 contains a 2 bp deletion at the position number of 618th bp in the first exon of GmLDL2, respectively. Gmldl2/fer double mutant were generated from Gmldl2 mutant by CRISPR/Cas9 system. Gmldl2/fer contains a 6 bp deletion at the position number of 463th in the first exon of GmFER.
Authentication	Genotyping of all mutant plants with primers (Supplementary Data 6).

ChIP-seq

Data deposition

✘ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	https://ngdc.cncb.ac.cn/gsa/
Files in database submission	06GH_1_1.fq 06GH_1_2.fq 06GH_3_1.fq 06GH_3_2.fq 06GM_2_1.fq 06GM_2_2.fq 06GM_3_1.fq 06GM_3_2.fq

K4M1H8_1.fq
K4M1H8_2.fq
K4M1H9_1.fq
K4M1H9_2.fq
K4M1M8_1.fq
K4M1M8_2.fq
K4M1M9_1.fq
K4M1M9_2.fq
K4M2H1_1.fq
K4M2H1_2.fq
K4M2H3_1.fq

Genome browser session
(e.g. UCSC)

N/A

Methodology

Replicates

Two independent biological replicates have been performed.

Sequencing depth

Illumina paired end sequencing on a Next 2000 System
ID Sample Total reads Mapping ratio Unique reads Unique mapping ratio Peak number (total reads)
06GH1 WT_GmLDL2_rep1 25,052,915 89.71% 5,965,626 23.81% 39,366
06GH3 WT_GmLDL2_rep2 12,092,030 91.98% 2,848,356 23.56% 13,769
06GM2 ef1_GmLDL2_rep1 21,640,747 94.72% 4,382,167 20.25% 31,146
06GM3 ef1_GmLDL2_rep2 20,363,928 96.27% 4,555,388 22.37% 24,358
K4M1H8 WT_H3K4me1_rep1 23,000,593 97.42% 8,848,091 38.47% 8,580
K4M1H9 WT_ H3K4me1_rep2 22,933,954 96.78% 6,676,473 29.11% 5,119
K4M1M8 ef1_ H3K4me1_rep1 22,465,705 96.97% 7,783,950 34.65% 9,968
K4M1M9 ef1_ H3K4me1_rep2 21,858,671 97.05% 7,985,956 36.53% 8,287
K4M2H1 WT_ H3K4me2_rep1 22,072,404 98.16% 9,956,289 45.11% 28,070
K4M2H3 WT_ H3K4me2_rep2 22,186,516 97.83% 9,995,199 45.05% 21,514
K4M2M1 ef1_ H3K4me2_rep1 21,470,369 97.73% 9,927,670 46.24% 24,514
K4M2M2 ef1_ H3K4me2_rep2 22,635,431 97.80% 9,871,356 43.61% 19,100
K4M3H1 WT_ H3K4me3_rep1 18,408,207 95.57% 6,518,539 35.41% 20,695
K4M3H2 WT_ H3K4me3_rep2 21,051,839 97.57% 7,987,484 37.94% 24,180
K4M3M1 ef1_ H3K4me3_rep1 24,195,696 97.16% 8,316,998 34.37% 61,694
K4M3M3 ef1_ H3K4me3_rep2 24,626,553 96.75% 8,990,273 36.51% 33,466
INPUT1 WT_input_rep1 21,439,436 45.85% 2,515,354 11.73%
INPUT3 WT_input_rep2 22,258,540 70.19% 4,147,068 18.63%
INPUT2 ef1_input_rep1 21,768,484 78.34% 4,090,465 18.79%
INPUT3 ef1_input_rep2 20,859,813 82.81% 4,535,238 21.74%

Antibodies

anti-H3K4me1 (Millipore, 07-436), anti-H3K4me2 (Millipore, 07-030), anti-H3K4me3 (Abcam, ab8580), and anti-GmLDL2.

Peak calling parameters

MACS2 callpeak --extsize 200 --nomodel -g 1e9 -B -q 0.05

Data quality

Duplicate reads were removed using Picard, and multiple mapped reads were filtered out strictly using the SAM file tag “XS:”. Unique mapped reads were then delivered for peak calling through MACS2 software version 2.2.9.1 with parameter --extsize 200 --nomodel -g 1e9 -B -q 0.05. Peaks were filtered based on FDR < 5% (q-value < 0.05) to control false discovery rates. The Peak numbers and the ratio of peaks with above 5-fold enrichment were below:
Sample Peak number above 5 fold enrichment ratio (above 5 fold enrichment)
WT_LDL2_rep1 39,366 10,226 25.98%
WT_LDL2_rep2 13,769 5,464 39.68%
ef1_LDL2_rep1 31,146 13,462 43.22%
ef1_LDL2_rep2 24,358 8,078 33.16%
WT_H3K4me1_rep1 8,580 2,057 23.97%
WT_ H3K4me1_rep2 5,119 1,876 36.65%
ef1_ H3K4me1_rep1 9,968 2,492 25.00%
ef1_ H3K4me1_rep2 8,287 1,966 23.72%
WT_ H3K4me2_rep1 28,070 4,997 17.80%
WT_ H3K4me2_rep2 21,514 3,375 15.69%
ef1_ H3K4me2_rep1 24,514 3,338 13.62%
ef1_ H3K4me2_rep2 19,100 2,732 14.30%
WT_ H3K4me3_rep1 20,695 10,541 50.94%
WT_ H3K4me3_rep2 24,180 8,605 35.59%
ef1_ H3K4me3_rep1 61,694 25,318 41.04%
ef1_ H3K4me3_rep2 33,466 14,238 42.54%

Software	Duplicated reads and low-mapping quality reads were identified and removed with SAMtools. Enriched intervals were identified by MACS version 2.2.9.1 with parameter.