Supplementary Information

A silkworm model reveals coordinated cellular communities and conserved Hippo regulation in silk production

Yan Ma[1], Zhiyong Zhang[2], Wenhui Zeng[1], Zhou Fang[1], Jie Hu[1], Yiyun Tang[1], Xin Yu[1], Lin Cheng[1], Zehui Ma[1], Qipeng Zeng[1], Dena Jiang[1], Xiao Li[1], Keshu Dong[1], Yumeng Zhu[1], Chen Zhao[1], Ying Zhou[1], Wei Tan[1], Yao Zhang[1], Zhonghuai Xiang[1], Hanfu Xu[1,*]

[1]State Key Laboratory of Resource Insects, College of Sericulture, Textile and Biomass Sciences, Southwest University, Chongqing 400715, China

[2]Beijing Junhejun Biotechnology Co., Ltd., Beijing 102200, China

***Corresponding author**

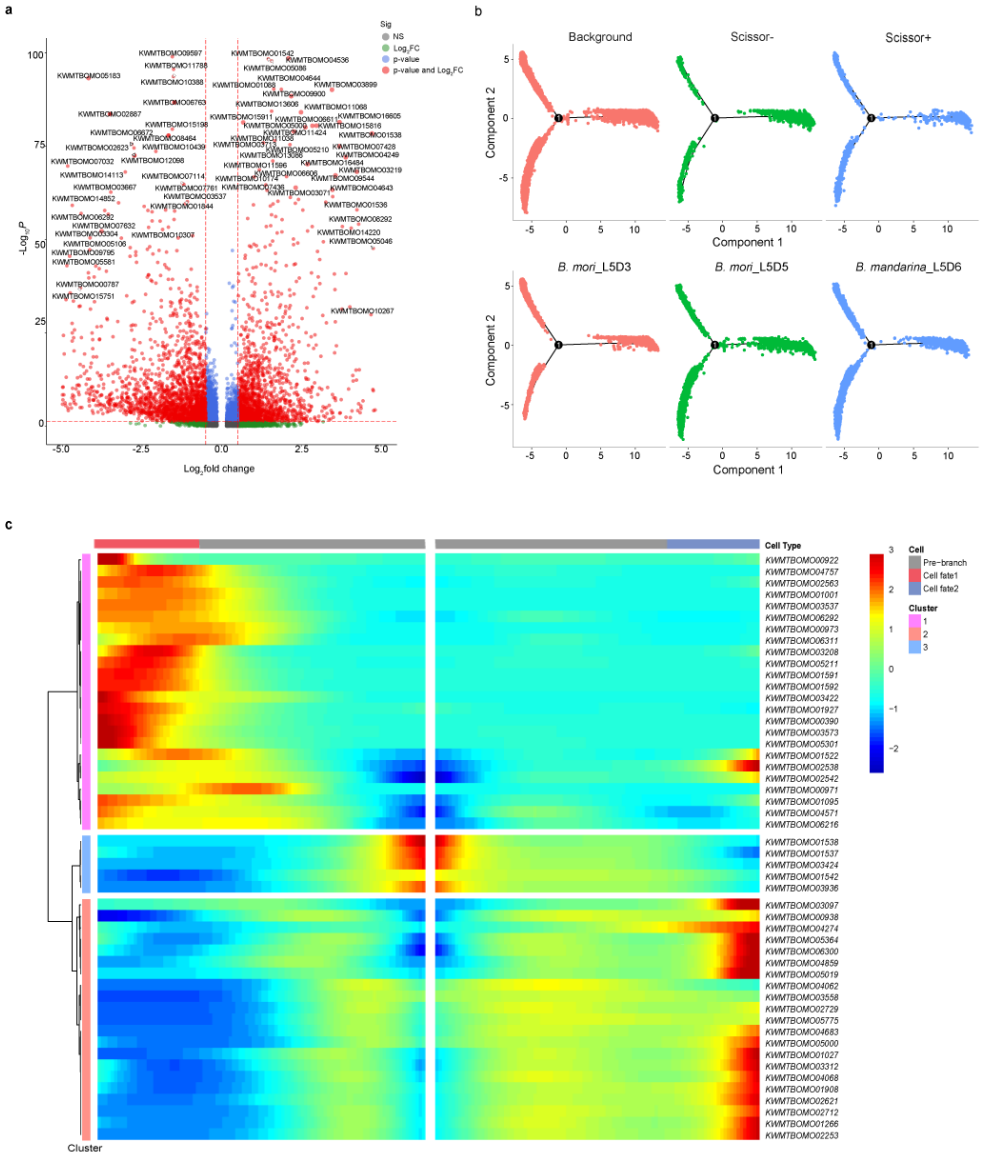
E-mail address: xuhf@swu.edu.cn

Tel/ Fax: +86 23 68254032.

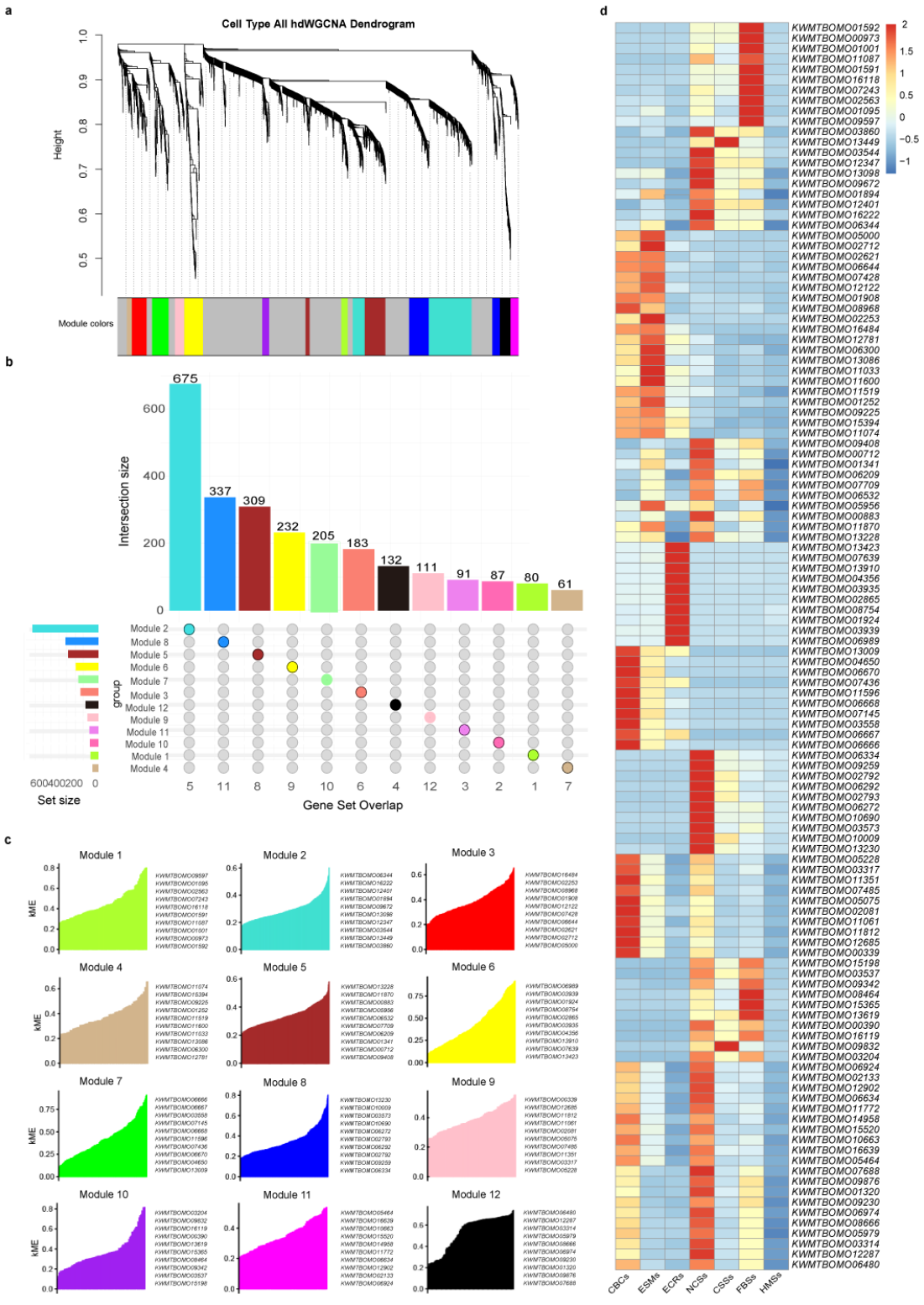
This PDF file includes:

Supplementary Figure

Supplementary Table

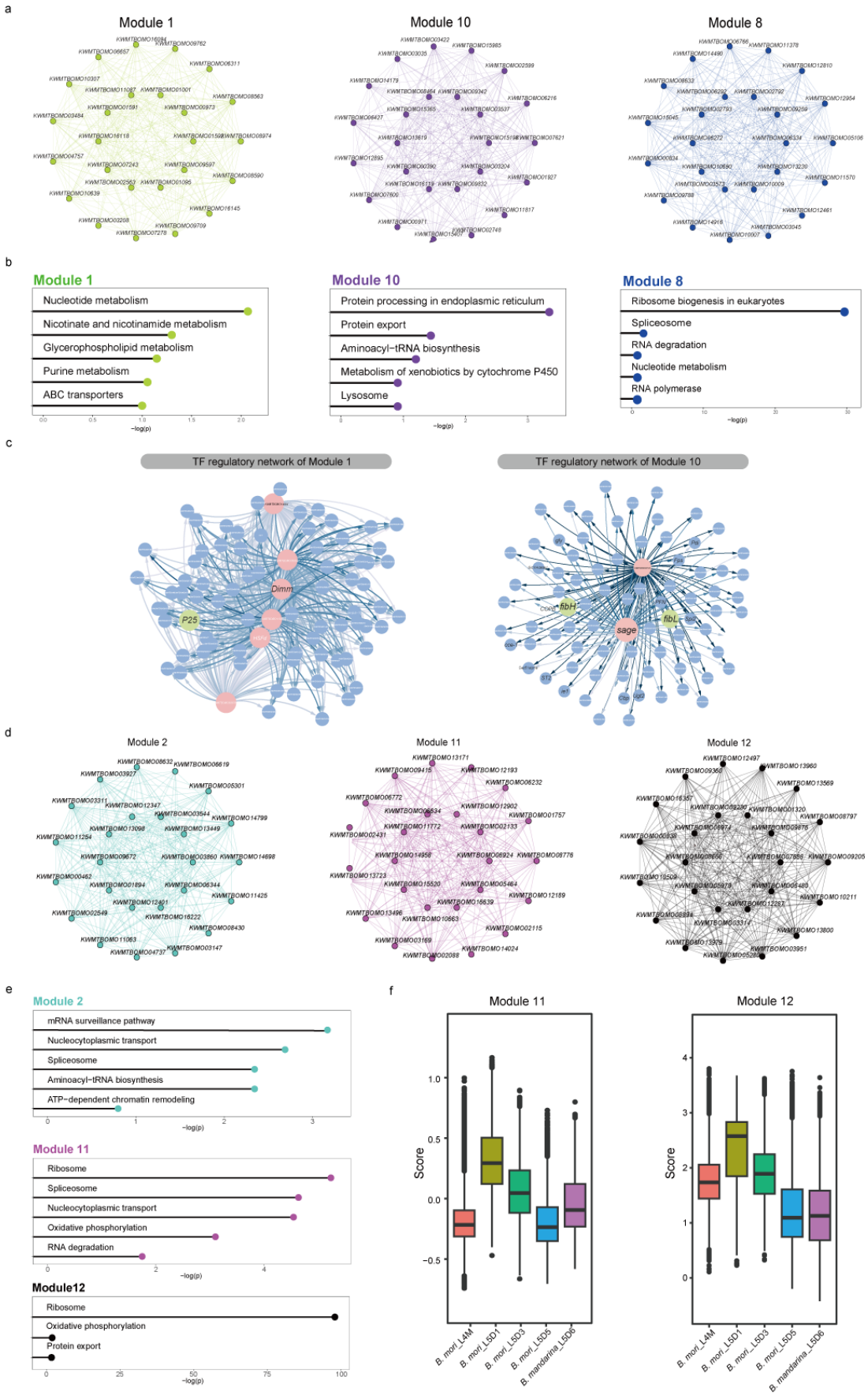


Supplementary Fig. 1. Differentially expressed genes (DEGs), distribution of Scissor-selected cells and dynamic changes in gene expression across pseudotime. (a) Volcano plot showing DEGs between the Scissor+ cells and the Scissor- cells. (b) Distribution of Scissor-selected cells along the pseudotime trajectory in *B. mori* (L5D3 and L5D5) and *B. mandarina* (L5D6). Cells are colored according to the developmental subpopulation (Scissor+, Scissor-, or background cells; top panel) and stage (bottom panel). (c) Heatmap illustrating dynamic changes in gene expression over pseudotime. Genes were grouped into three clusters based on their expression patterns, indicating covarying patterns across developmental progression. Cell fate 1 corresponds to cell state 1; cell fate 2 corresponds to cell state 3; and the prebranch state corresponds to cell state 2.



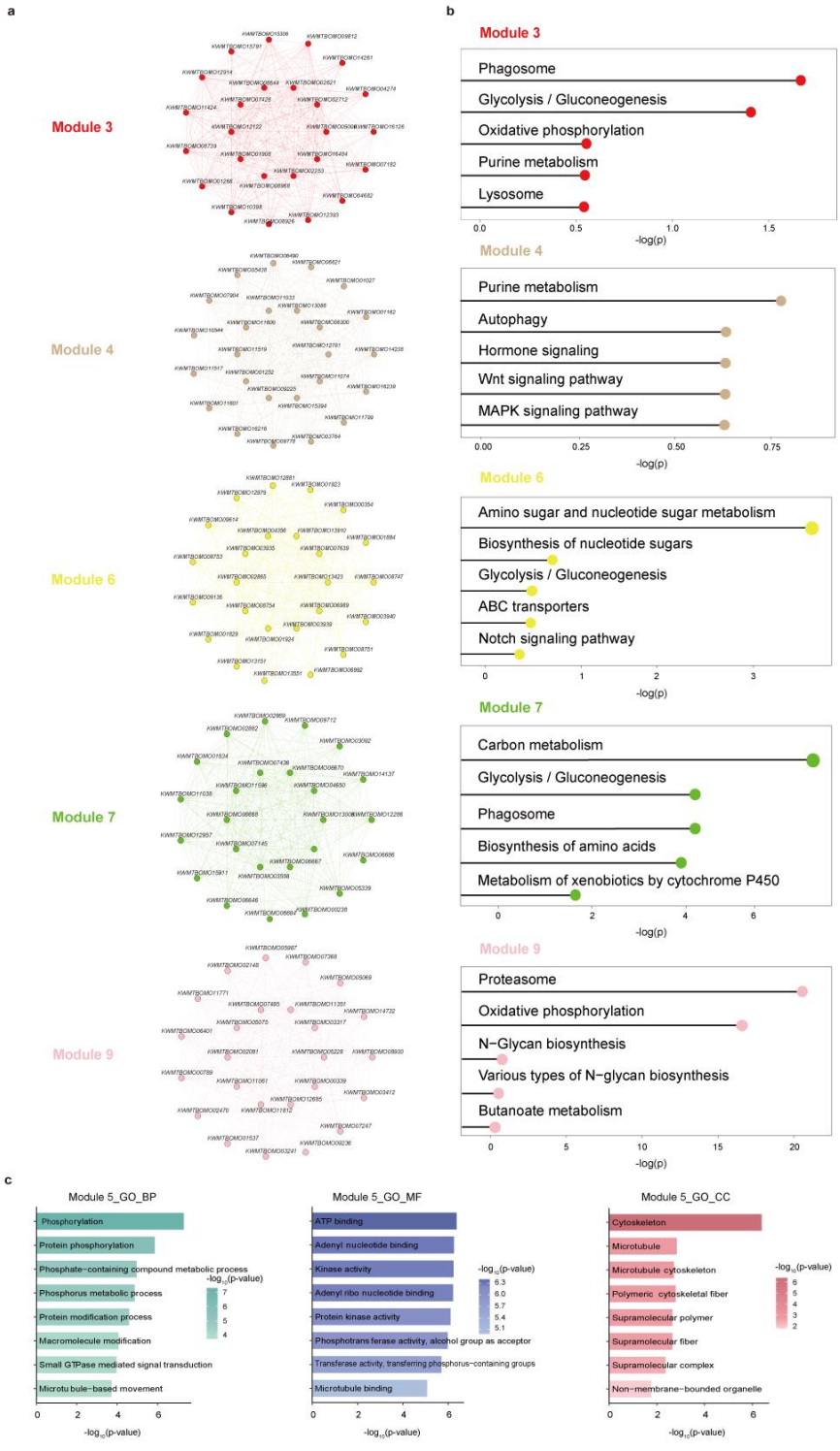
Supplementary Fig. 2. Identification of coexpression modules in the SG. (a) Dendrogram showing the hierarchical clustering of genes. Each leaf represents a gene; the color bars below indicate the module assignment. Genes not assigned to any module are grouped in the "gray" module. (b) Upset plot displaying set intersections

among coexpression modules. The bar chart indicates the total number of genes in each module. Colored dots represent genes present in the set; gray dots represent genes absent in the set. (c) kME plot of the top 10 module eigengenes for each module. (d) Heatmap of the normalized expression levels of the top 10 hub genes from each coexpression module across the seven cell types.



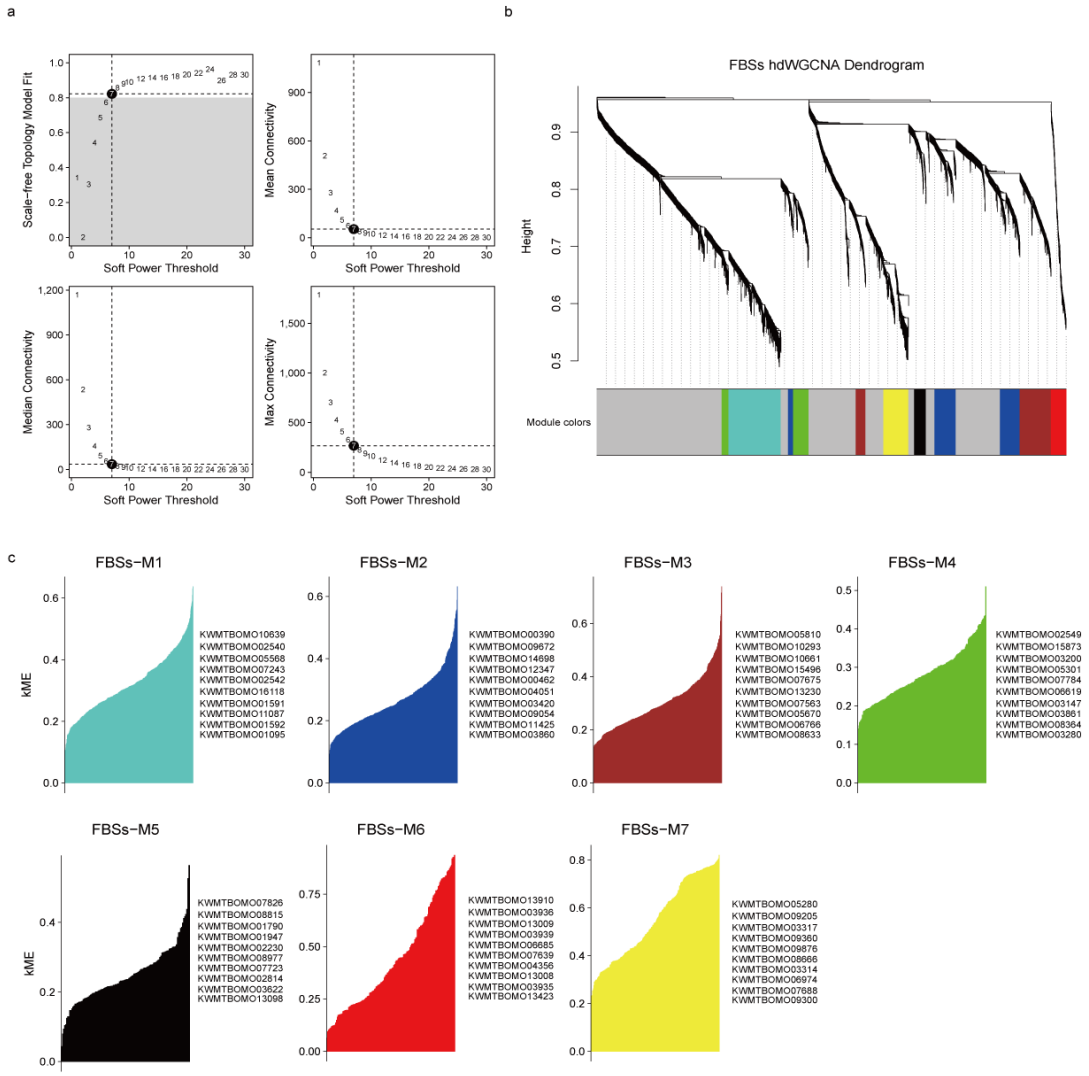
Supplementary Fig. 3. Analysis of coexpression modules functioning in silk

protein synthesis and support. (a) Top 25 hub genes with strong correlations in Modules 1, 10 and 8. The proximity of the node to the center indicates the correlation strength with module eigengenes. (b) Lollipop chart of significantly enriched KEGG terms for Modules 1, 10 and 8. The bar length corresponds to $-\log_{10}$(P.adj). Length represents the $-\log_{10}$ (P.adj), with a longer lollipop indicating more significantly enriched terms. (c) Transcriptional regulatory networks for Modules 1 and 10 inferred using GENIE 3. Pink nodes represent transcription factors (TFs); green nodes represent silk protein genes; and blue nodes represent target genes. The arrow color reflects the strength of regulatory interactions, with deeper colored arrows indicating higher weights. (d) The top 25 hub genes with strong correlations in Modules 2, 11 and 12. The proximity of the node to the center indicates the correlation strength with module eigengenes. (e) Lollipop chart of significantly enriched KEGG terms for Modules 2, 11 and 12. Length represents the $-\log_{10}$(P.adj), with a longer lollipop indicating more significantly enriched terms. (f) Distribution of module eigengene scores for Modules 11 and 12 across samples from *B. mori* (L4M, L5D1, L5D3 and L5D5) and *B. mandarina* (L5D6). Box colors represent different samples.

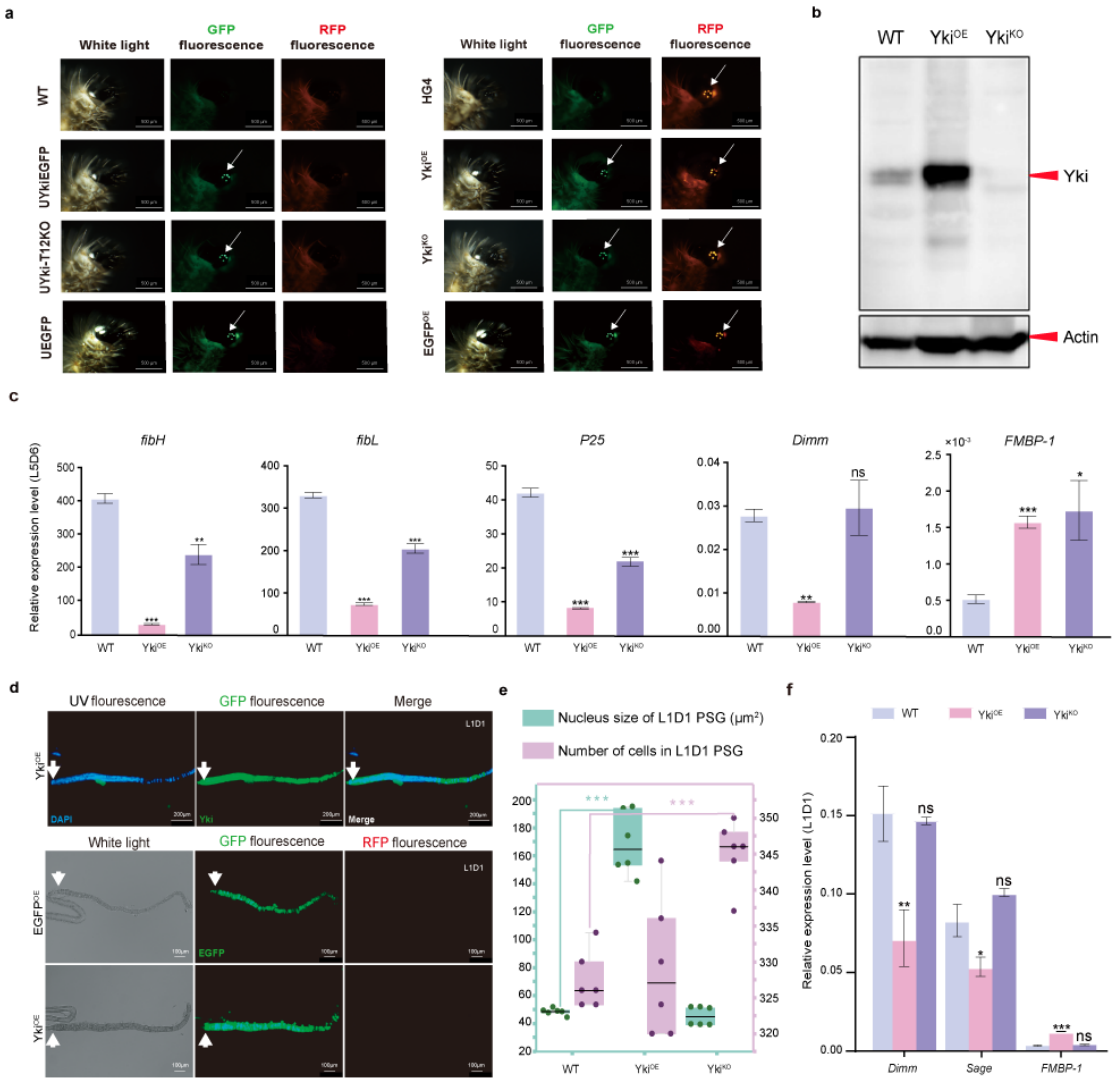


Supplementary Fig. 4. Analysis of coexpression modules functioning in silk protein processing and Module 5. (a) The top 25 hub genes with strong correlations in Modules 3, 4, 6, 7 and 9. The proximity of the node to the center indicates the correlation strength with module eigengenes. (b) Lollipop chart of significantly enriched KEGG terms for Modules 3, 4, 6, 7 and 9. Length represents the $-\log_{10}$

(P.adj), with a longer lollipop indicating more significantly enriched terms. (c) GO enrichment analysis of Module 5. The bar length and color intensity correspond to the statistical significance of enrichment $-\log_{10}$ (P.adj).

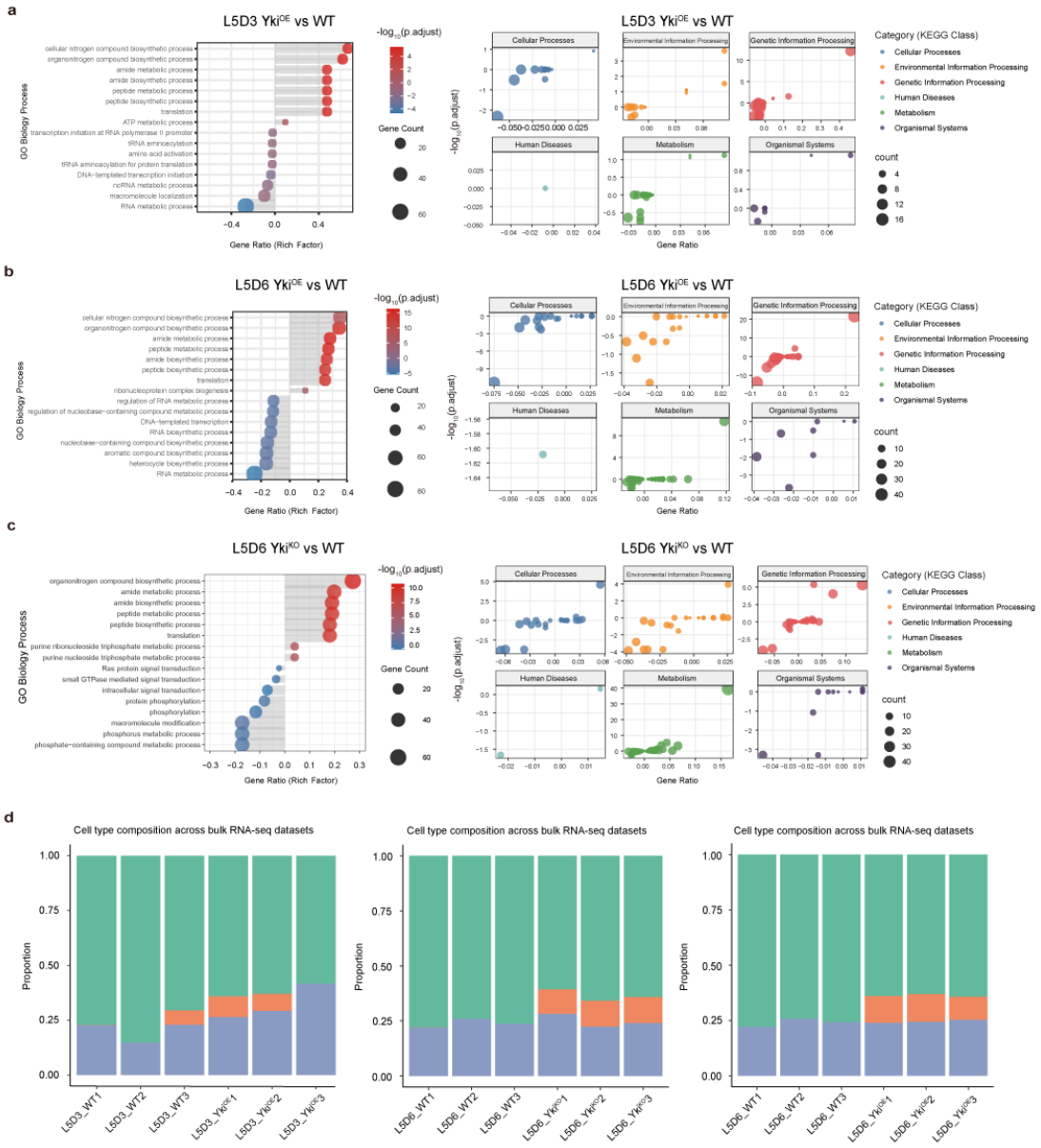


Supplementary Fig. 5. Identification of coexpression modules of FBSs. (a) Scale-free topology fit and soft-thresholding analysis of the coexpression network in FBSs. The optimal soft-thresholding power was determined to be 7. (b) Dendrogram showing the hierarchical clustering of genes. Each leaf represents a gene; the color bars below indicate the module assignment. Genes not assigned to any module are grouped in the "gray" module. (c) kME plot of the top 10 module eigengenes for each module, ranked by kME. A higher kME indicates greater functional centrality within the module.



Supplementary Fig. 6. Construction of *Yki* overexpression and knockout mutants and molecular profiling. (a) Transgenic silkworms, including a Gal4 line (HG4), UAS lines (UYkiEGFP, UYki-T12KO, and UEGFP), and Gal4/UAS lines (Yki^{OE}, Yki^{KO}, and EGFP^{OE}), were obtained by screening for GFP and/or RFP fluorescence in the ocelli of newly hatched larvae. WT and UEGFP^{OE} were used as negative and positive controls (similarly hereafter). The arrows (white) show the position of the larval eyes. Scale bars, 500 μm. (b) Levels of the Yki protein in the L5D6 PSGs of WT, Yki^{OE} and Yki^{KO} were detected using Western blotting. Actin served as the control. (c) The relative expression levels of fibroin genes and the regulators *Dimm* and *FMBP-1* in the L5D6 PSGs of WT, Yki^{OE} and Yki^{KO} were measured using qRT–PCR. (d) PSGs were dissected from L1D1 Yki^{OE} larvae for use in the immunostaining analysis. Scale bars: 200 μm (upper panel). Phenotypes of the L1D1 PSGs in EGFP^{OE}

and Yki^{OE} larvae. GFP fluorescence shows the distribution of *Yki* in the PSG. Scale bars, 100 μm (lower panel). The arrowhead (white) shows the boundary of the PSG and the MSG. (e) Quantification of cell numbers and nuclear size of the 1L1D PSGs in WT, Yki^{OE}, and Yki^{KO} larvae. The data are presented as the means ± S.D; n=6 for each strain. (f) The relative expression levels of SG transcription factor (*Dimm*, *Sage* and *FMBP-1*) in the L1D1 PSGs of WT, Yki^{OE} and Yki^{KO}.



Supplementary Fig. 7. Functional enrichment of DEGs and analysis of the proportions of different cell types in *Yki*-modified strains. (a–c) Bar plot of GO biological process terms and bubble plot of KEGG pathways (based on bulk RNA-seq) enriched in DEGs identified in the L5D3 Yki^{OE}, L5D6 Yki^{OE} and L5D6 Yki^{KO} larvae. (d) Proportions of FBSs, HMSs, and NCSs across L5D3 Yki^{OE}, L5D6 Yki^{OE}, and L5D6 Yki^{KO} larvae compared with WT larvae. Three biological replicates were performed for each sample.

Supplementary Table 1. Primer sequences used for the qRT–PCR analysis

Name	Forward primer sequences (5'- 3')	Reverse primer sequences (5'- 3')
CycE	GTCCACCCCACACTCTAATAAA	TCAGCCCAAGACAATCCAG
CDK2	TACCATCAACAGCCTTACGGGA	GGTAGTGGTCCCTTCGTCAGAT
E2F1	TACAGCAGACCGTCCAGTT	CCGCCGCTATGTTCAAAT
Yki	TCGGACGGTGAACAGAAA	TGTCGCATACGAAGAGGC
Fzr	CACGGCTACTCACAGAACCA	CTTGTGTGAGGGCGTTTTG
CycA	TTCAAGATGAGAACAGGACGG	CTCACATTTTCAGCAGCATTCAC
MCM3	CAAGAGGAGTTTTTTGTGGCAT	GACCACCTTAGGACGAACCAAT
MCM6	GCTGGTGACCGTTATGATTTTA	CTTTATGCCTTCCATTTGACC
MCM8	TAATAGAGCAAAAACGGTGGTC	CCTGTGCCGATTTCTTTGTT
MCM9	TGGAGCACTGGTTCTGTCGG	TGAATCGTAGTTCCCCTTAG
MCM10	CAATGATTACGGTTCACAAG	ACTTTAGGAGGACTGAACACA
ATG3	ATGACAAGTATTACCAAACTCCGA	CGCTCTCCGTTACCGTCTCTA
ATG5	ACCCTTGAAATGGCACTATCC	ATCTTTGTTTGGGCAATGAA
ATG6	ATGGGTAATACGGGCTTTTG	TGTGGCTCAGATTTGTCCTC
ATG7	CGTCAGAAGAGGTCGGTACTC	CGACTCTTTTCGGATGCTG
ATG12	CAATCAGTGGCTGGAATGGA	TCCATAATCCATCCAATAGGCTT
fibH	CAGGGGATACGGACAAGGT	TTCACACAAGGCAGTGCTCT
fibL	GGAGGTGGAAGAATCTATGAC	TGTAGGCAGCGATGTTGT
P25	GGGTCTGCCCATCTTCCAC	CTCGCCAGCCAGTTCCTCT
Dimm	CACCGAATCTCCTGACCAA	CATCACTTCCGCTACCACTAT
FMBP-1	TGGAAATCATCTGTCGGTGA	CTTCGTCTACGCGCTTTAGG
Sage	ATTACGAGCCCAAGAGGAT	CCACTACGGTGTCACGAAC
Hr3	CAGGGTTCTTGGACGCAGACTTC	AGTCCAACCACATCTCCTCGTAG
Ftz-f1	GCAAAGTGCTGGACCAAAAT	GTCAAACTTCTGCCCGTTGT
Hr4	CCAGACCAGAACTGCGTACA	TAATTTCACAGCCACCGTCA

Supplementary Table 2. Antibodies used for western blotting analysis

Name of antibody	Form and species of antibody	Production company
Yki	Polyclonal antibody, rabbit source	Zoonbio Biotechnology
Actin	Mouse mAb	CST, #3700