Supplementary Materials for

Cytoplasmic U5 snRNA promotes the heterotrophic-to-autotrophic transition via selective translational activation

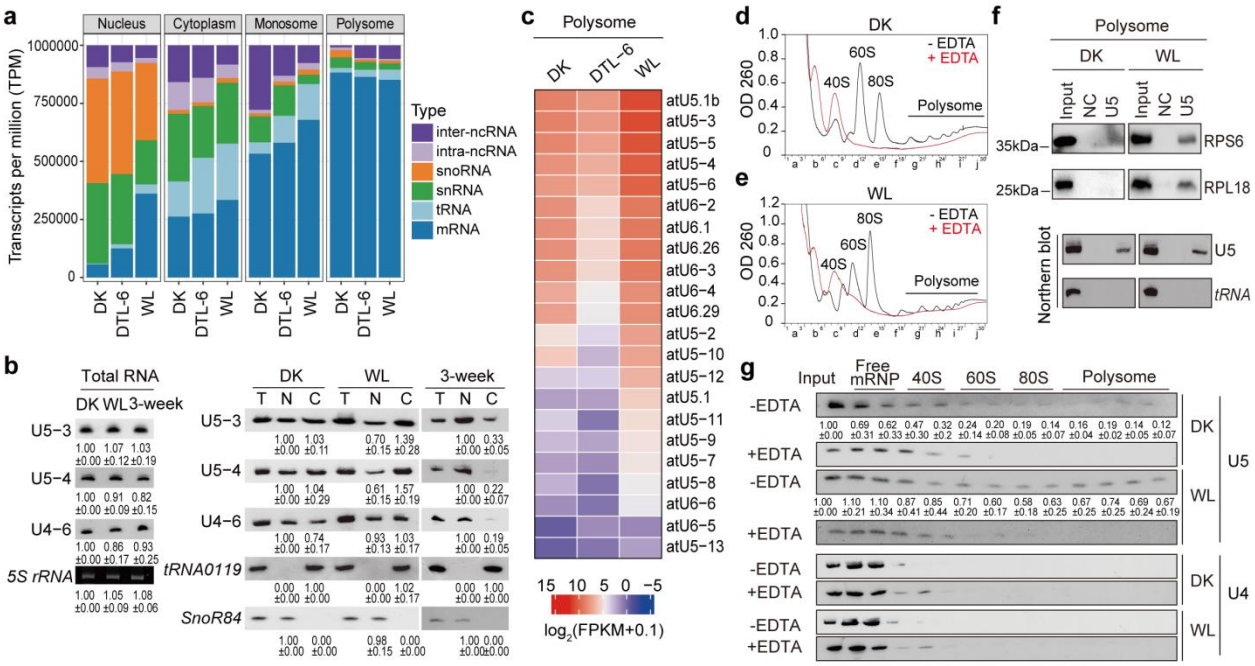
Yuntong Jiao *et al*.

Corresponding author(s): zhudanmeng@pku.edu.cn (D.Z.)

The PDF file includes:

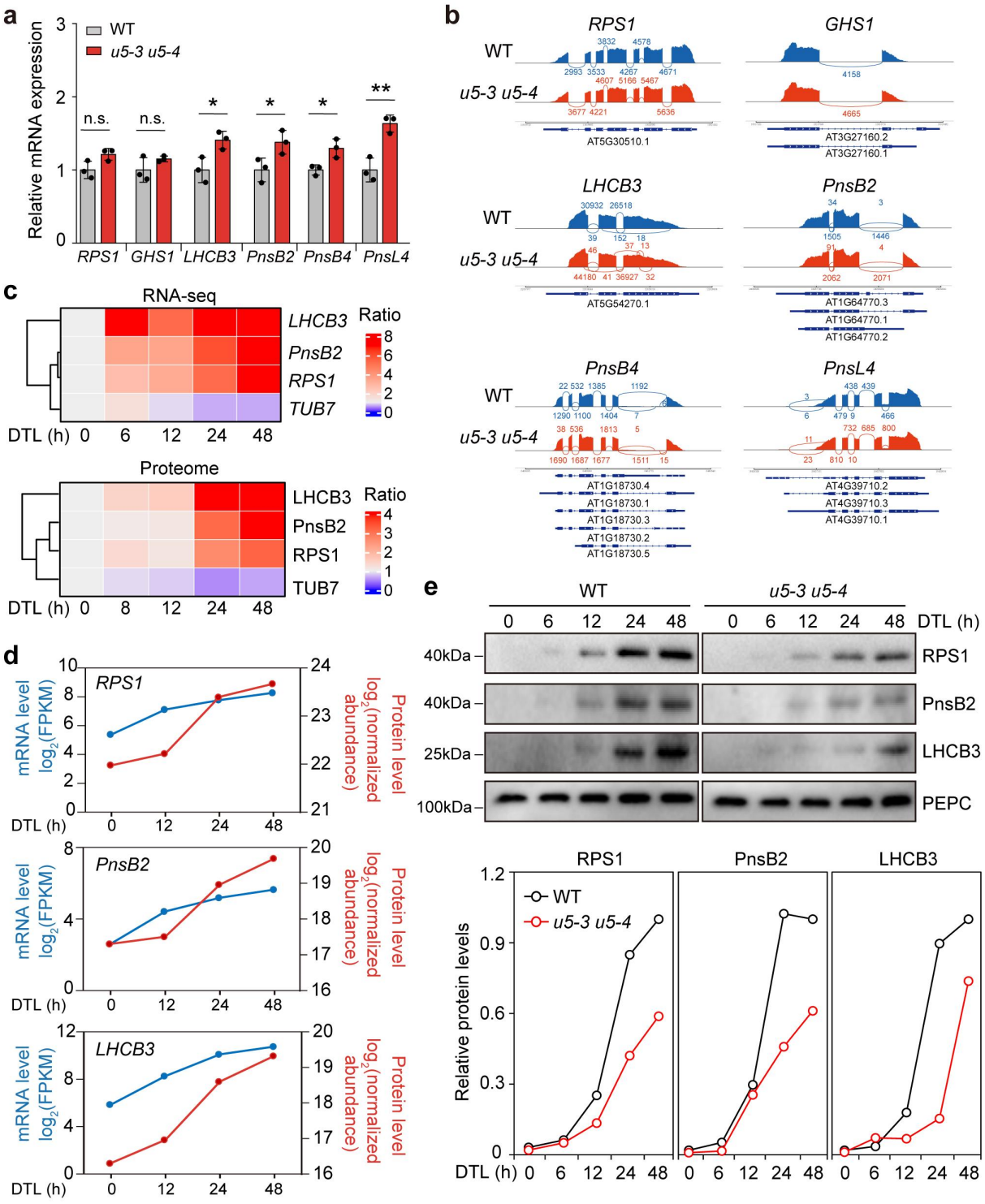
Supplementary Figs. S1-S13

Supplementary Tables S1-S9



Supplementary Fig. S1 Global RNA profiling reveals light-induced enrichment of U5 snRNAs in the cytoplasm.

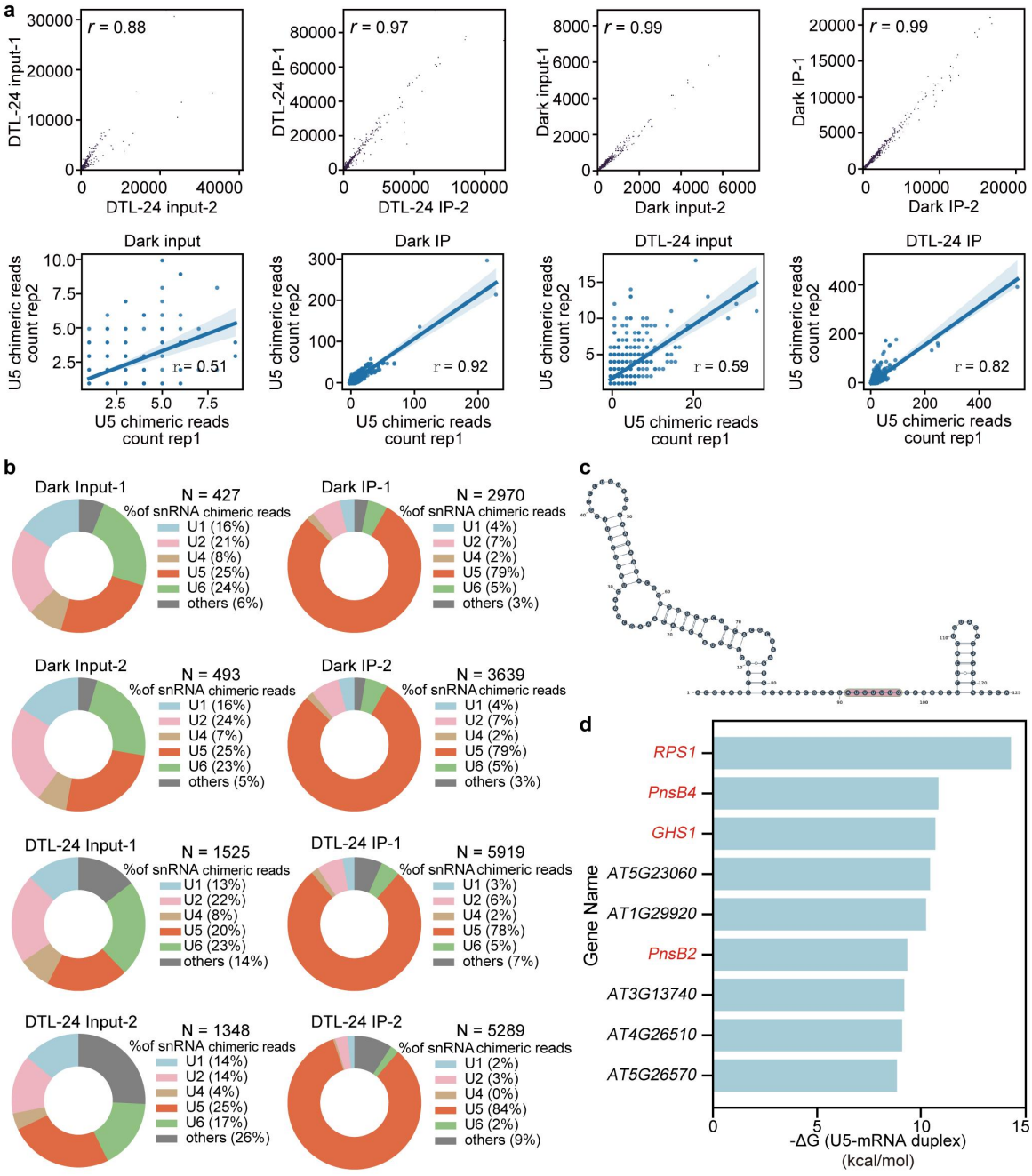
a RNA composition of each subcellular fraction. Stacked bars showing the proportion of total transcripts per million (TPM) contributed by six RNA classes in the nuclear, ribosome-free cytoplasmic, monosomal and polysomal fractions from dark-grown seedlings (DK), dark-grown seedlings exposed to light for 6 hours (DTL-6), and white-light-grown seedlings (WL). inter-ncRNA, intergenic noncoding RNA; intra-ncRNA, intragenic noncoding RNA; snoRNA, small nucleolar RNA; snRNA, small nuclear RNA; tRNA, transfer RNA; mRNA, messenger RNA. **b** Northern blots showing the higher abundance of U5 snRNAs in the cytoplasm of WL-grown seedlings. Total: total RNA; N: nucleus; C: cytoplasm. **c** Heatmaps showing the abundance of U5 and U6 snRNAs in polysome fractions under the indicated light conditions. Values are presented as $\log_2$ (FPKM + 0.1). **d-e** Polysome profiles of WT seedlings grown under DK (**d**) and WL (**e**) conditions. Cell lysates were treated with or without EDTA. **f** RNA pull-down assays showing the interaction of U5 snRNA with ribosomal proteins (RPS6 and RPL18) in the polysome fractions of WL seedlings. Non-biotin-labeled U5 snRNA probes were used as a negative control (NC). Data represent one of three biological replicates. **g** Northern blots of each fraction from polysome profiling with non-treated (-EDTA) and EDTA-treated (+EDTA) cell lysates of WT seedlings grown under DK or WL. The numbers below the images in panels b and f represent the relative expression levels of each snRNA, quantified with SD from three replicates.



Supplementary Fig. S2 U5 snRNAs promote the timely up-regulation of RPS1, LHCB3 and PnsB2.

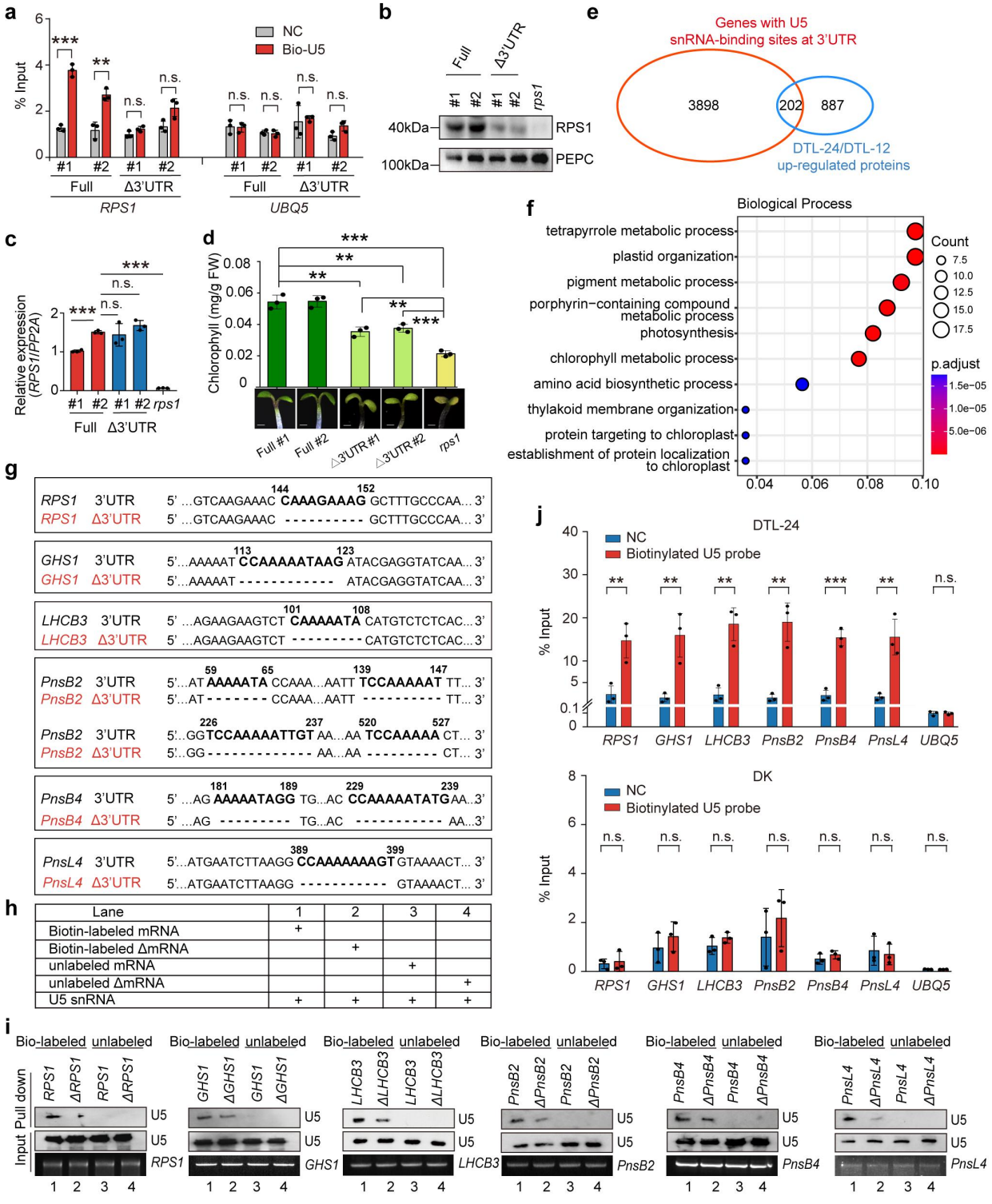
a Relative expression of representative chloroplast-related genes in the WT and *u5-3 u5-4* double mutant grown under DTL-24. Data are presented as mean ± SD, n = 3. $^{**}P < 0.01$, and $^{*}P < 0.05$

by two-tailed Student's *t*-test; n.s., not significant. **b** Sashimi plots showing that the frequencies of alternative splicing of genes in (**a**) remain unchanged in the *u5-3 u5-4* double mutant compared to WT. **c** mRNA (upper) and protein expression (lower) levels of indicated genes in WT seedlings under indicated time points. Heatmap was generated after normalization of the expression levels of each mRNA or protein relative to the levels at the first time point (DTL-0). **d** Comparison of the mRNA and protein levels of RPS1, PnsB2 and LHCB3 in WT seedlings during the dark-to-light transition for indicated time points. FPKM values were $\log_2$-transformed directly, without adding a pseudocount. **e** Western blots (upper) and corresponding band intensities (lower) showing protein levels of RPS1, PnsB2 and LHCB3 in WT and *u5-3 u5-4* during the dark-to-light transition for indicated time points.



Supplementary Fig. S3 U5 snRNA targets mRNA through the Sm site.

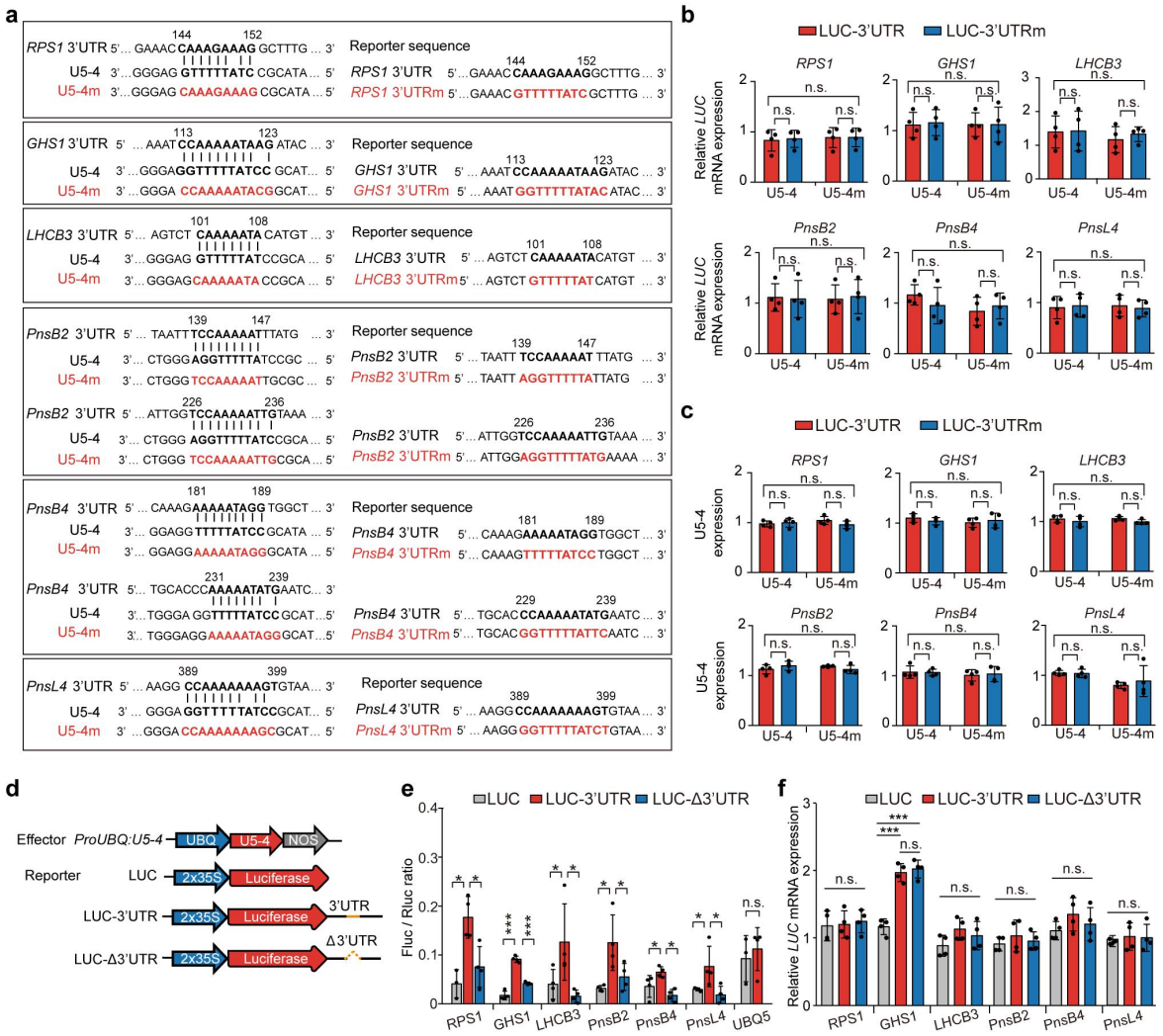
a Correlations of total KARR-seq reads (upper) and U5 snRNA-mRNA chimeric reads (lower) between replicates in input (without U5 ASO enrichment) and IP (with U5 ASO enrichment) samples under DK and DTL-24. **b** Distribution of snRNA types in input and enriched chimeric reads from U5 snRNA KARR-seq in seedlings grown under DK and DTL-24. **c** Predicted structure of U5 snRNA using VARNA, highlighting the Sm site element (red). **d** Free energies of RNA duplexes formed between U5 snRNA Sm site and the mRNA targets.



Supplementary Fig. S4 mRNAs interact with U5 snRNAs Sm site via UBE at 3'UTR.

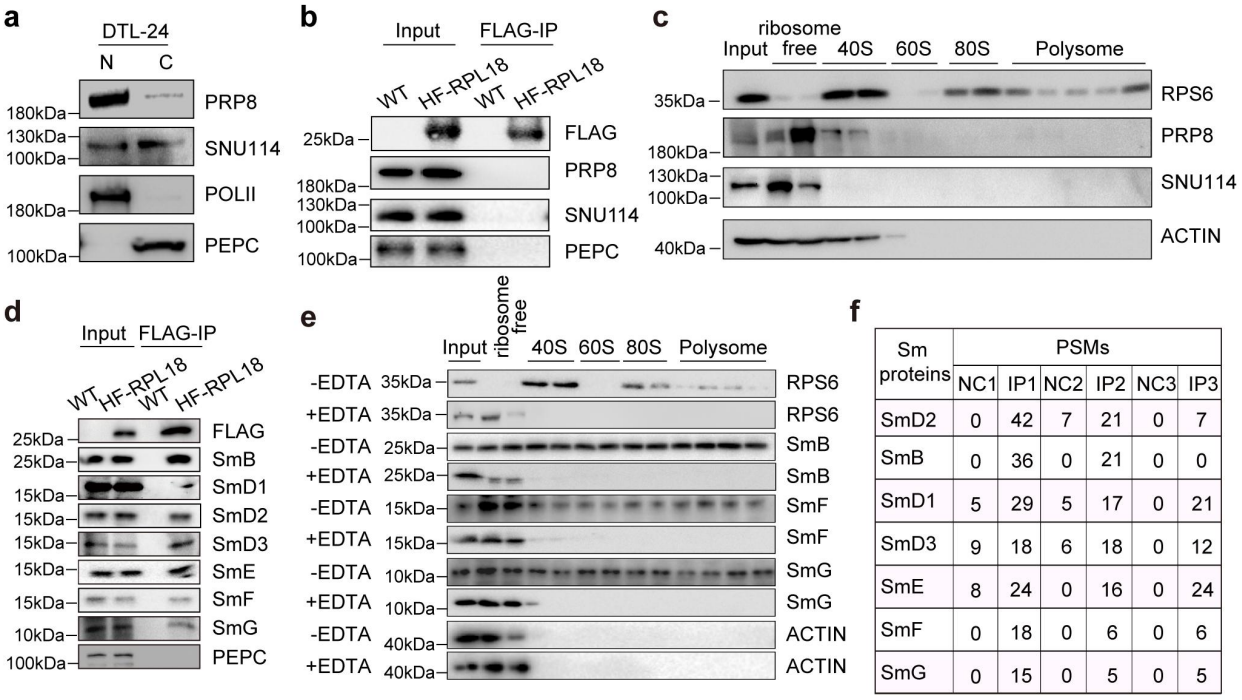
a *RPS1* mRNA was not enriched in endogenous U5 snRNA pull-down assays in transgenic lines expressing *RPS1* without UBE in the 3'UTR. Non-biotin-labeled U5 snRNA probes were used as a negative control (NC). Full: *ProRPS1:RPS1-3'UTR/rps1*; Δ 3'UTR: *ProRPS1:RPS1-Δ*

3'UTR/rps1. Values are mean ± SD, n = 3. **b-c** Protein (**b**) and mRNA (**c**) levels of RPS1 in transgenic lines grown under DTL-24. Values are mean ± SD, n = 3. **d** Chlorophyll contents (upper) and phenotypes (lower) of two independent transgenic seedlings as indicated. Scale bar = 0.5 mm. **e** Venn diagram showing the overlap among genes with up-regulated protein levels under DTL-24 compared with DTL-12 (blue), genes with U5 snRNA-binding sites in the 3'UTR (red). **f** GO analysis of the overlapping genes shared in (**e**). **g** Schematic of the sequence position of UBE deletion in 3'UTR of each mRNA. **h** Schematic of the prey and bait RNAs in the in vitro pull-down assays. Biotin-labeled mRNA, biotinylated mRNA with UBE in the 3'UTR; Biotin-labeled ΔmRNA, biotinylated mRNA with UBE-deleted 3'UTR; unlabeled mRNA, nonbiotinylated mRNA. **i** In vitro pull-down assays demonstrating the requirement of UBEs at the 3'UTRs for interaction of U5 snRNA with *RPS1*, *GHS1*, *LHCB3*, *PnsB2*, *PnsB4* and *PnsL4*. Results shown are representative of three different experiments. **j** The representative mRNAs were enriched in endogenous U5 snRNA pull-down assays at DTL-24 (upper), but not in DK (lower). *UBQ5* was used as the negative control. Data are presented as mean ± SD, n = 3. ***P<0.001, **P<0.01, and *P<0.05 by two-tailed Student's *t*-test; n.s., not significant.



Supplementary Fig. S5 U5 snRNAs facilitate translation via Sm site-UBE interaction.

a UBEs (bold in black) at the 3'UTRs of indicated genes, WT (black) and mutated (red) 3'UTR dual-luciferase reporters. WT (black) and mutated (red) U5-4 snRNA effectors. **b** Results of RT-qPCR analysis showing that the mutation of the UBE in the 3'UTR had no effect on the levels of reporter mRNAs. Values are mean ± SD, n = 4. **c** The levels of U5-4 determined by RT-qPCR. Values are mean ± SD, n = 4. **d** Schematic representation of effector and reporter constructs used in the dual-luciferase reporter assays. The orange line in the 3'UTR indicates the UBE, and the dotted line indicates a deleted UBE. **e** Translation of dual-luciferase reporters with WT or UBE-deleted 3'UTRs of the indicated genes. The *UBQ5* 3'UTR reporter served as a negative control. Values are mean ± SD, n = 4. **f** Results of RT-qPCR analysis showing that the deletion of the UBE in the 3'UTR had no effect on the levels of reporter mRNAs. Values are mean ± SD, n ≥ 3. *P* values are from two-tailed Student's *t*-test. ***$P<0.001$, **$P<0.01$, *$P<0.05$, n.s., not significant.

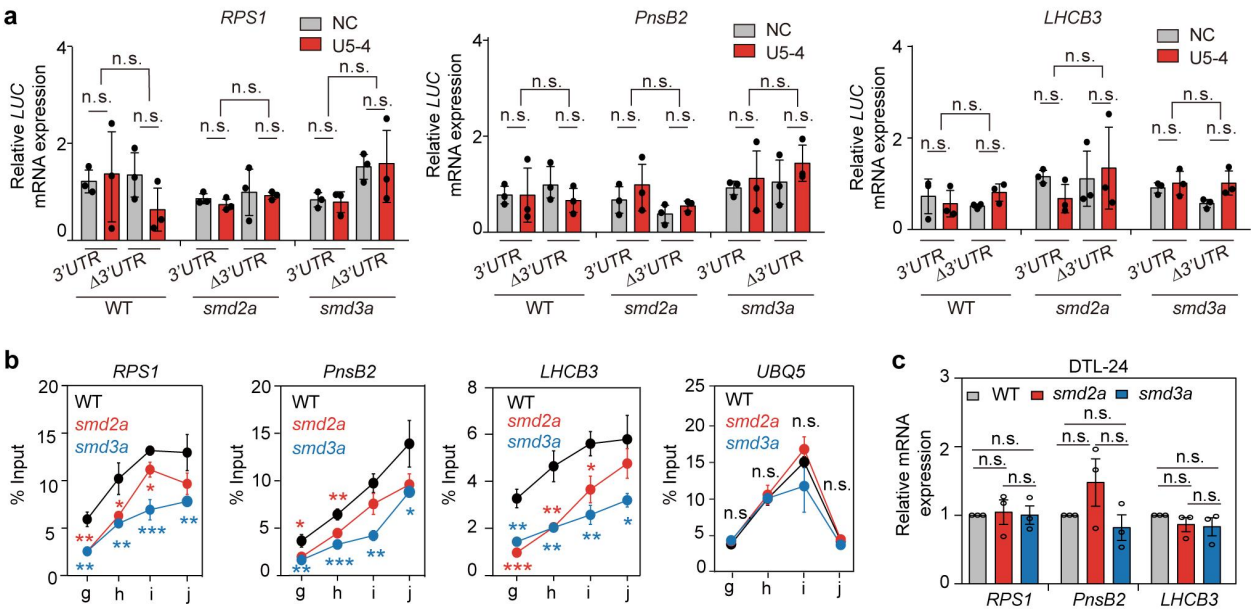


f

Sm proteins	PSMs					
	NC1	IP1	NC2	IP2	NC3	IP3
SmD2	0	42	7	21	0	7
SmB	0	36	0	21	0	0
SmD1	5	29	5	17	0	21
SmD3	9	18	6	18	0	12
SmE	8	24	0	16	0	24
SmF	0	18	0	6	0	6
SmG	0	15	0	5	0	5

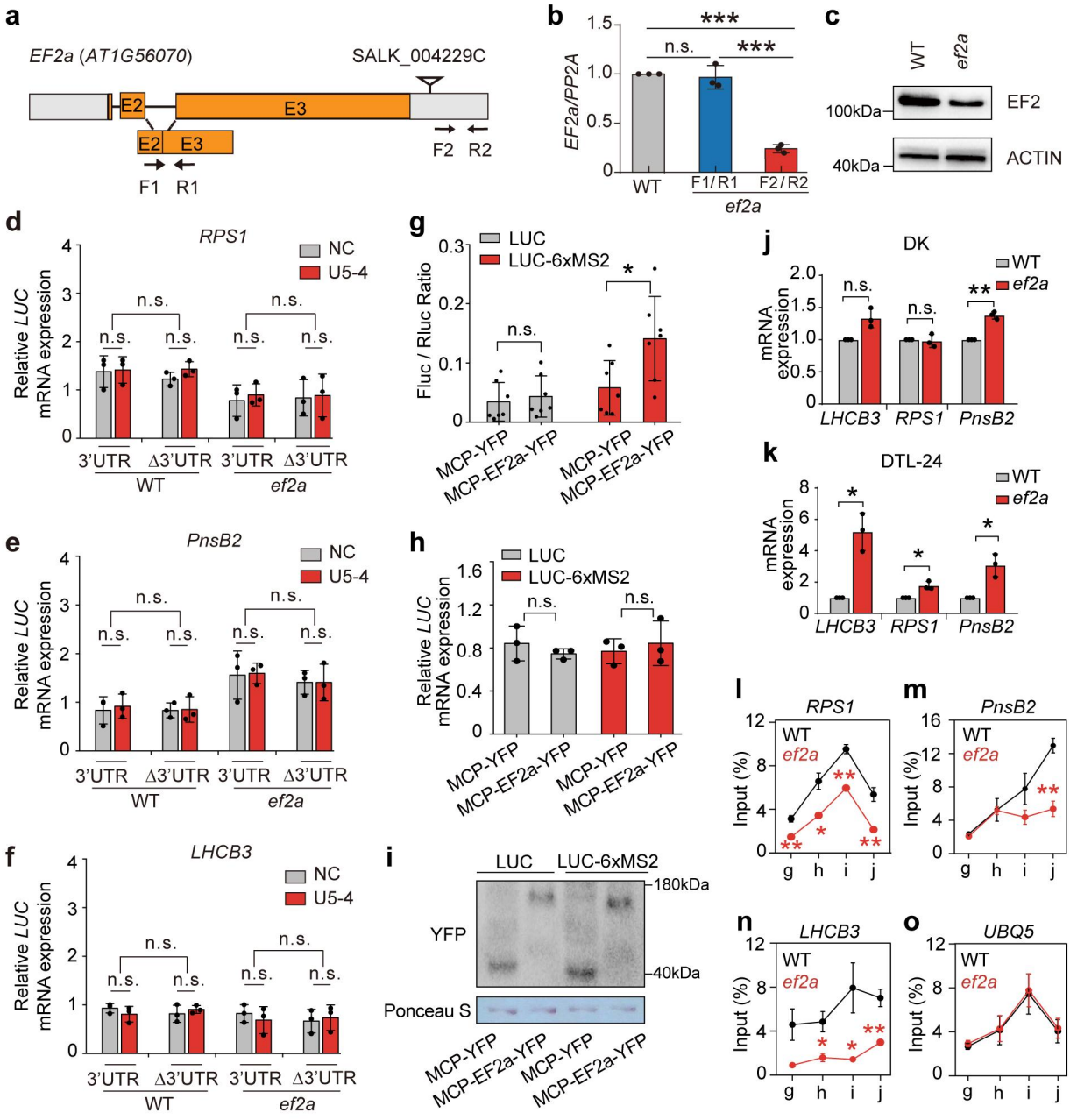
Supplementary Fig. S6 Sm proteins are associated with polysomes.

a Western blots showing the subcellular accumulation of PRP8 and SNU114 in seedlings grown under DTL-24. POLII and PEPC were used as markers for the nucleus and cytoplasm, respectively. **b** Co-IP assays showing RPL18 does not associate with PRP8 and SNU114 in transgenic seedlings expressing HIS-FLAG-tagged RPL18 (HF-RPL18) grown under DTL-24. **c** Western blots of each fraction from polysome profiling of WT seedlings grown under DTL-24 showing PRP8 and SNU114 do not associate with polysomes. **d** Co-IP assays showing the association of HIS-FLAG-tagged RPL18 (HF-RPL18) with each of the seven Sm proteins in transgenic seedlings grown under DTL-24. **e** Western blots of each fraction from polysome profiling of WT seedlings grown under DTL-24 showing the association of Sm proteins to the polysomes, as exemplified by SmB, SmF, and SmG. **f** List of Sm proteins in polysome fractions identified by IP-MS. The number of peptide spectrum matches (PSMs) corresponding to each protein is indicated. Sm proteins were co-immunoprecipitated by anti-HA beads from *Pro35S:SmFa-3HA*/WT transgenic plants. NC, WT plants were used as the negative control.



Supplementary Fig. S7 Polysome-associated Sm proteins are required for U5 snRNA-promoted mRNA translation.

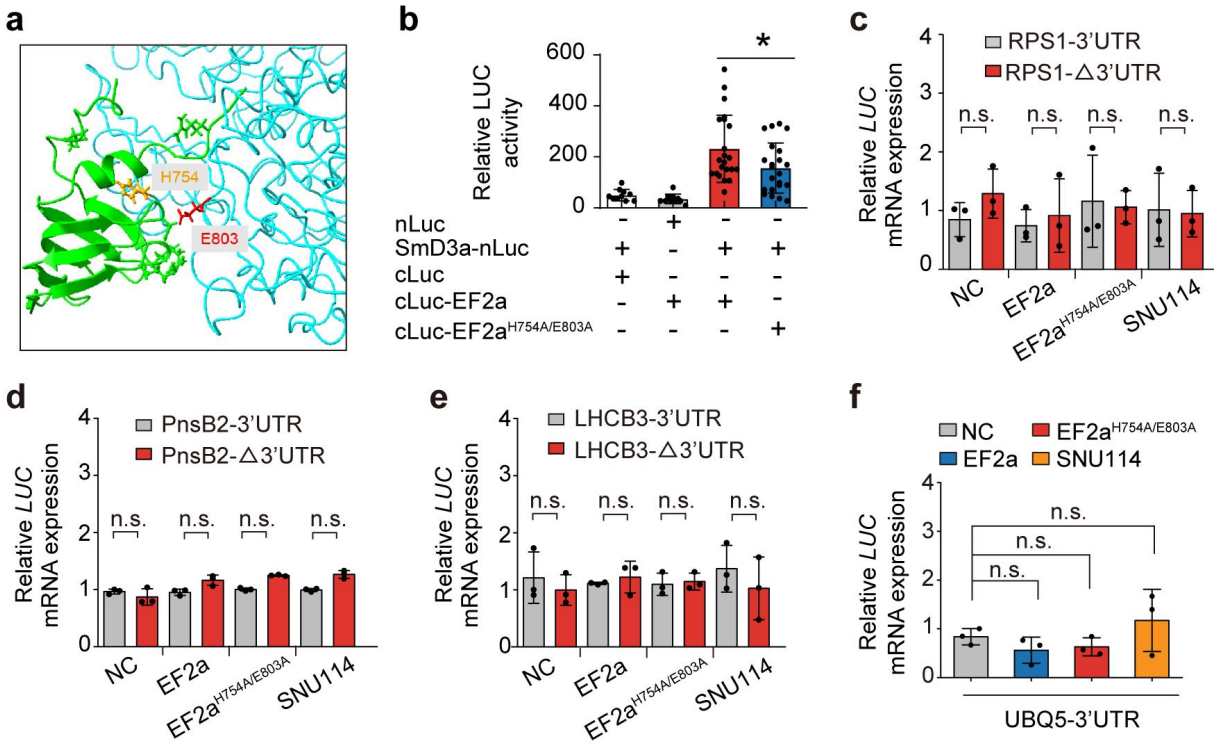
a RT-qPCRs showing similar mRNAs levels of *RPS1*, *PnsB2* and *LHCB3* reporters with WT and deleted-UBE in the 3'UTR in the protoplasts of WT, *smd2a* and *smd3a*. **b** RT-qPCR showing the decreased polysome association of *RPS1*, *PnsB2*, and *LHCB3* mRNAs in *smd2a* and *smd3a* mutants. *UBQ5*: negative control. **c** RT-qPCR showing the expression levels of *RPS1*, *PnsB2*, and *LHCB3* mRNAs in WT, *smd2a* and *smd3a* seedlings grown under DTL-24. Values are mean ± SD, n = 3. Data were analyzed by two-way ANOVA (**a**), and two-tailed Student's *t*-test (**b**, **c**). ***P<0.001, **P<0.01 and *P<0.05; n.s., not significant.



Supplementary Fig. S8 EF2a is required for U5 snRNA-mediated translation activation.

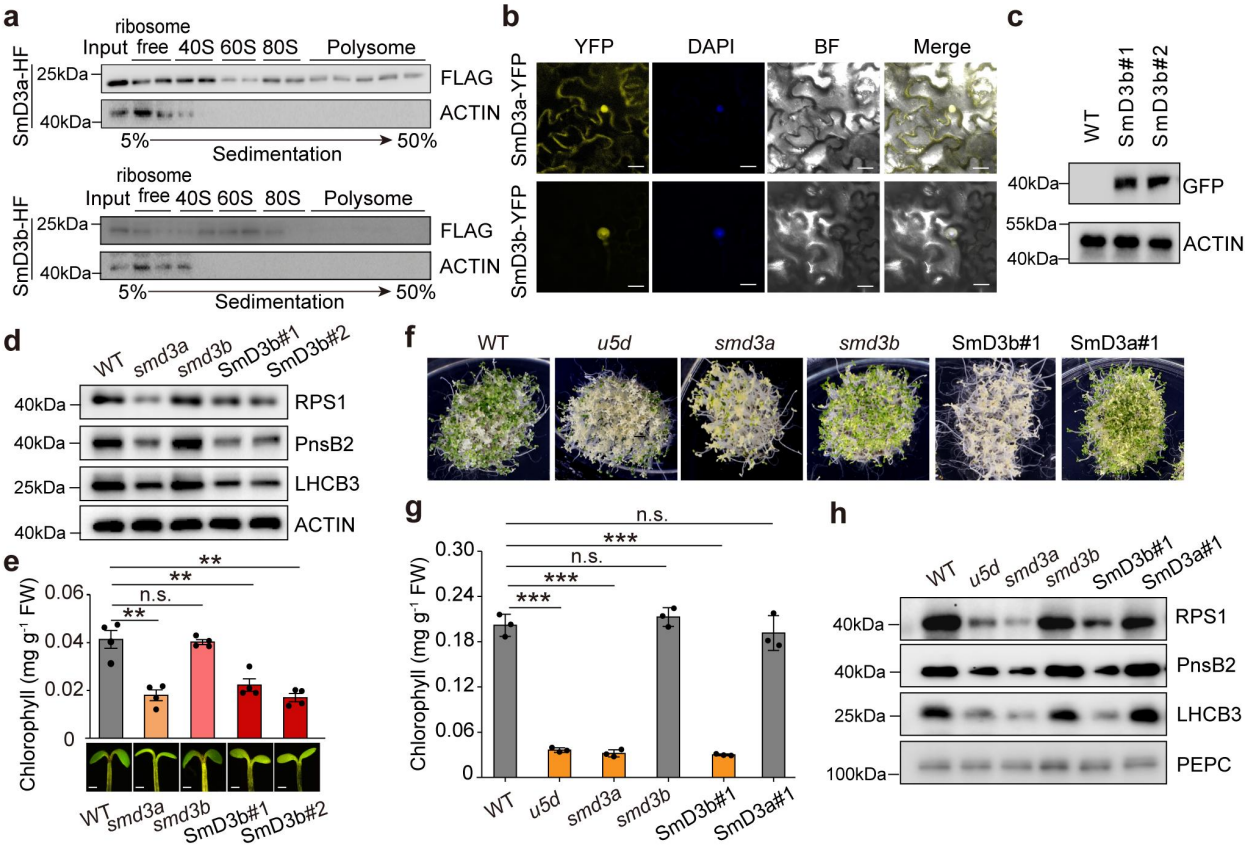
a Position of the T-DNA insertions in EF2a. Exons are indicated by orange rectangles, introns by black lines, and T-DNA insertion by the triangle. Arrows indicate primers used for RT-qPCR. **b-c** RT-qPCR (**b**) and western blot (**c**) were used to determine the expression levels of EF2a transcripts and protein levels in WT and the *ef2a* mutant. **d-f** RT-qPCRs showing similar mRNAs levels of *RPS1* (**d**), *PnsB2* (**e**) and *LHCB3* (**f**) reporters with WT and deleted-UBE in the 3'UTR in the protoplasts of WT and *ef2a*. **g** MS2-tethering reporter assay for the expression of EF2a in regulating mRNA translation. Values are mean ± SD, n = 7. **h-i** mRNA (**h**) and protein (**i**) levels

of MCP-YFP and MCP-EF2a-YFP expressed in the *N. benthamiana* leaves. Values are mean ± SD, n = 3. Each dot represents a biological replicate. **j-k** RT-qPCR showing the expression levels of *RPS1*, *PnsB2*, and *LHCB3* mRNAs in WT and *ef2a* seedlings grown under DK (**j**) and DTL-24 (**k**). **l-o** RT-qPCR showing decreased polysome association of mRNAs in *ef2a*. Data were analyzed by two-way ANOVA (**d-f**), and two-tailed Student's *t*-test (**b**, **g-h**, **j-o**). ***$P<0.001$; **$P<0.01$; *$P<0.05$; n.s., not significant.



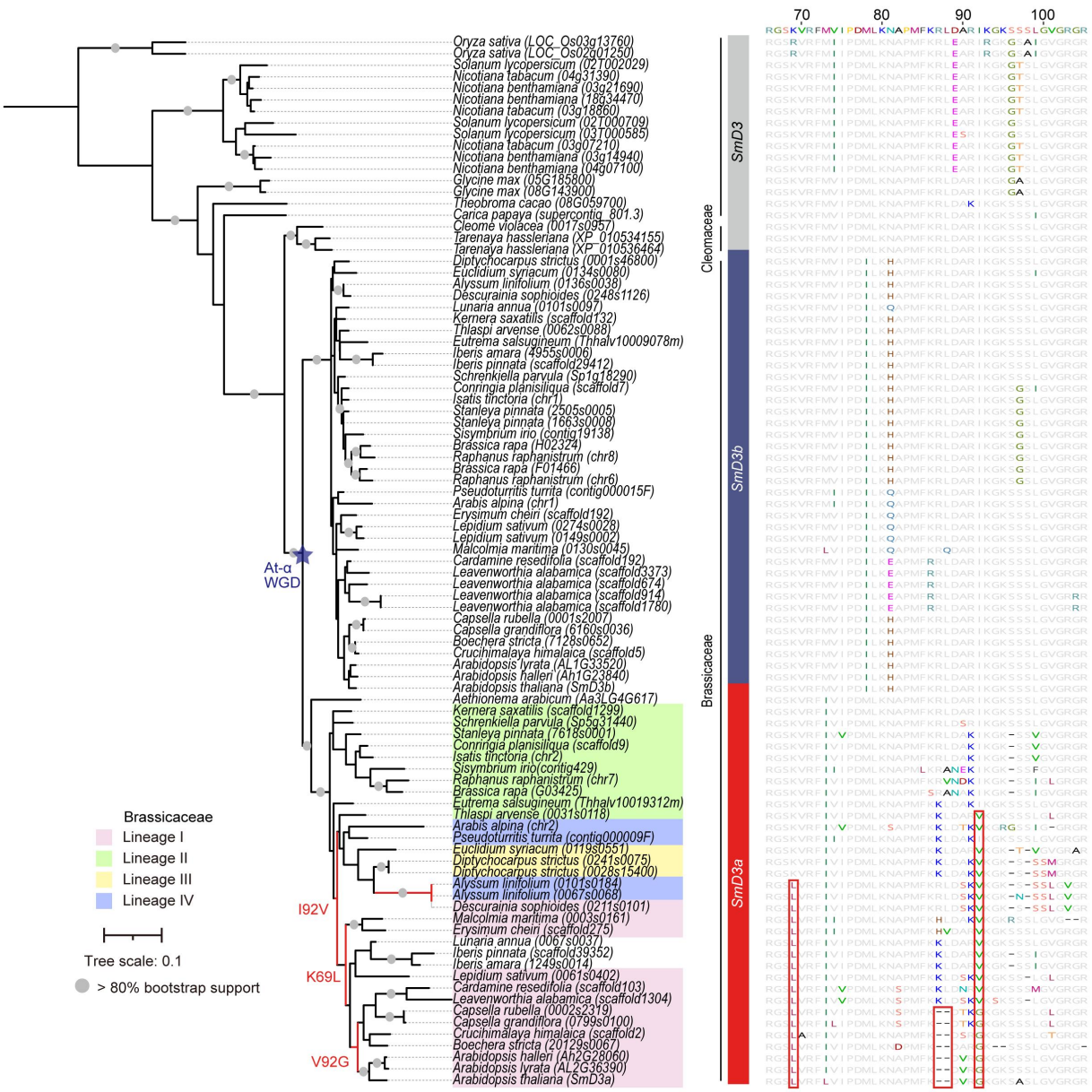
Supplementary Fig. S9 H754 and E803 of EF2a are key sites for the interaction between EF2a and SmD3a.

a Predicted interaction surface between EF2a and SmD3a. EF2a is represented in cyan, while SmD3a is shown in green. H754 and E803 are highlighted in yellow and red, respectively. **b** LCI assay showing that H754 and E803 of EF2a mediate the interaction between EF2a and SmD3a. Values are mean ± SD, n ≥ 10. **c-f** RT-qPCR assays showing the effects of EF2a, EF2a$^{H754A/E803A}$ and SNU114 on the expression of *RPS1* (**c**), *PnsB2* (**d**), *LHCB3* (**e**), and *UBQ5* (**f**) reporter mRNAs. Values are mean ± SD, n = 3. Data were performed by two-tailed Student's *t*-test. $^{**}P<0.01$; $^{*}P<0.05$; n.s., not significant.

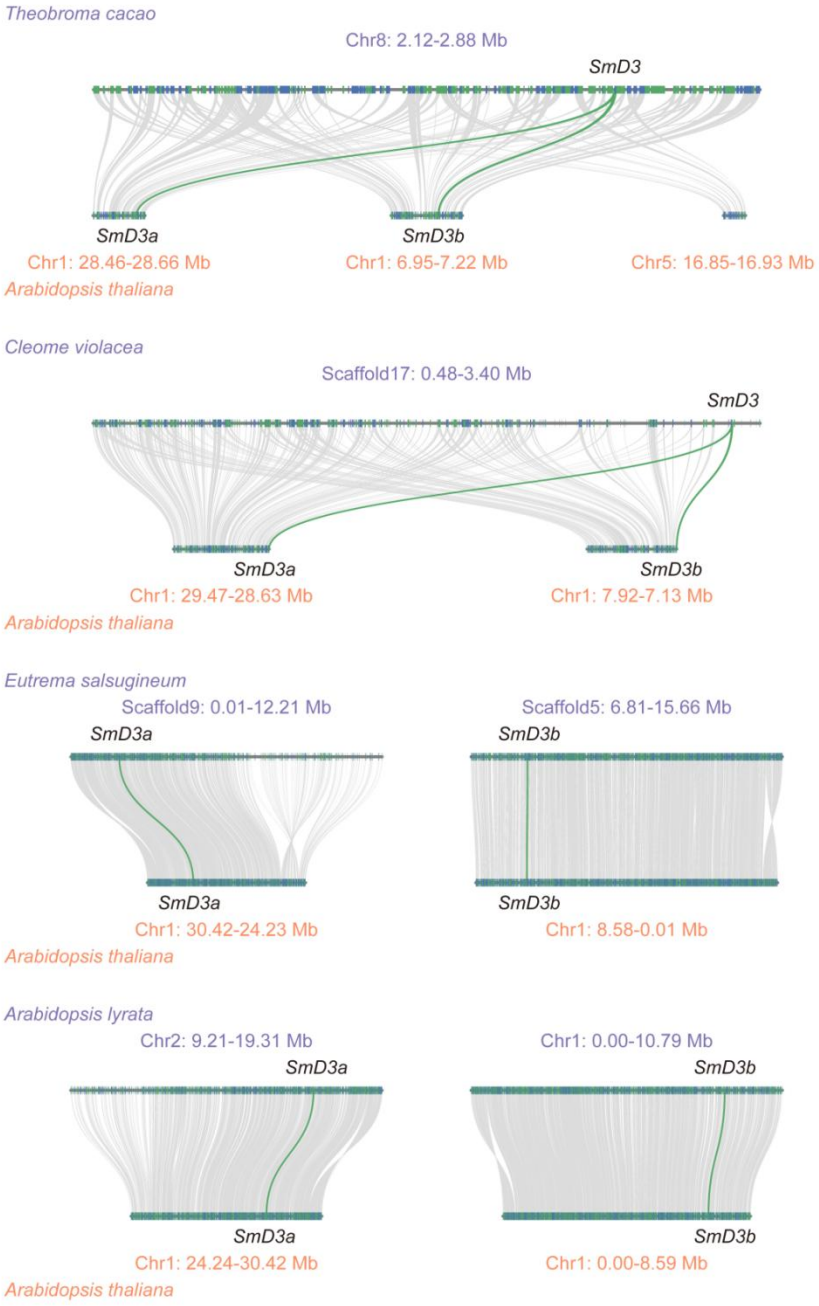


Supplementary Fig. S10 Homologous SmD3a and SmD3b proteins differ in polysome association and function.

a Western blot of each fraction from polysome profiling of *Pro35S:SmD3a-HIS-FLAG*/WT (SmD3a-HF) and *Pro35S:SmD3b-HIS-FLAG*/WT (SmD3b-HF) transgenic seedlings grown under DTL-24 showing SmD3a (*upper*), but not SmD3b (*lower*), associates with polysomes. **b** Confocal microscopy showing the subcellular localization of SmD3a-YFP and SmD3b-YFP transiently expressed in *N. benthamiana* leaves. Scale bar = 20 μm. **c** Western blots showing the expression of YFP-tagged SmD3b in two independent *ProSmD3a:SmD3b-YFP/smd3a* (SmD3b#1 and SmD3b#2) transgenic seedlings. **d-e** Protein levels of U5 targets (**d**) and chlorophyll contents (**e**) of WT, *smd3a*, *smd3b*, SmD3b#1 and SmD3b#2 transgenic seedlings grown under DTL-24. Scale bar = 0.5 mm. Statistical analysis was performed using two-tailed Student's *t*-test. $^{**}P<0.01$, n.s., not significant. **f-h** Phenotype (**f**), chlorophyll contents (**g**), and protein levels of U5 targets (**h**) in the WT, *u5-3 u5-4* (*u5d*), *smd3a*, *smd3b*, *ProSmD3a:SmD3b-YFP/smd3a* (SmD3b#1), and *ProSmD3a:SmD3a-YFP/smd3a* (SmD3a#1) seedlings following heat stress.

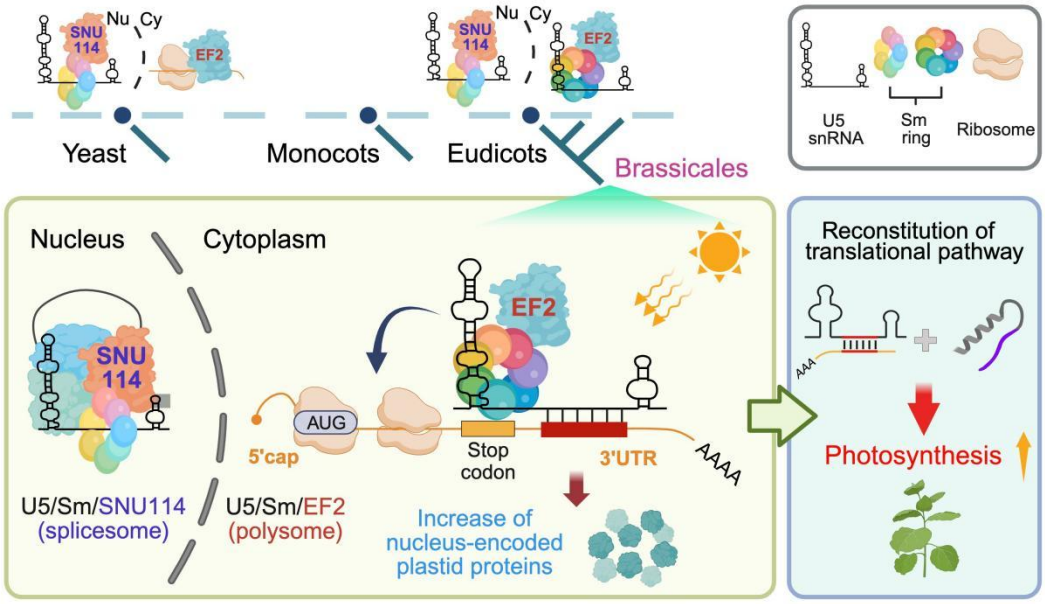


Supplementary Fig. S11 Phylogenetic relationship of SmD3 with more Brassicaceae species. The maximum-likelihood phylogenetic tree is reconstructed based on the coding sequences. The branch length reflects the nucleotide substitution rate. The circles denote bootstrap support higher than 80%. The orthologs of *A. thaliana* SmD3a and SmD3b in Brassicaceae are clustered separately, which are labeled in red and blue, respectively. Brassicaceae SmD3a genes are colored according to the species lineage. Multiple sequence alignment results indicate that *SmD3a*-specific mutations gradually emerged after gene duplication, including I92V, K69L and V92G substitutions, along with the deletion of the amino acids K87 and L88.



Supplementary Fig. S12 Synteny relationship of the SmD3 gene.

The plots show the local synteny blocks containing the SmD3 gene between *Arabidopsis thaliana* and *Theobroma cacao*, *Cleome violacea*, *Eutrema salsugineum*, and *Arabidopsis lyrata*. The SmD3 gene is highlighted in green. In *T. cacao* and *C. violacea*, lacking At-α WGD, both *A. thaliana* SmD3a and SmD3b are best matched to a single genomic region. In contrast, a one-to-one relationship is observed in Brassicaceae species (*E. salsugineum* and *A. lyrata*) that underwent At-α WGD, indicating that SmD3a and SmD3b originated in the basal Brassicaceae and derived from the At-α WGD event.



Supplementary Fig. S13 A working model of U5 snRNA/Sm protein-mediated translational activation.

Upon light illumination, U5 snRNA associates with polysomes to selectively enhance the production of nuclear-encoded plastid proteins. This occurs through base-pairing with the 3′ UTRs of target mRNAs and interaction with heptameric Sm proteins containing SmD3, a cytoplasmic paralog derived from a Brassicaceae-specific gene duplication, revealing a light-activated pathway that repurposes spliceosomal RNA and its cognate protein partners for coordinated plastid proteome remodeling. AI-assisted engineering of SmD3b, combined with optimized U5 snRNA, enhances chloroplast protein accumulation and the maximum photochemical efficiency of PSII in tobacco.

Supplementary Tables

Supplementary Table S1. Expression information in subcellular RNA-seq.

Supplementary Table S2. Genome coordinates of all snRNAs in *Arabidopsis thaliana*.

Supplementary Table S3. List of differentially expressed genes (DEGs) regulated by U5.

Supplementary Table S4. Translation efficiency (TE) and protein levels in *u5-3 u5-4* mutant.

Supplementary Table S5. Constitutive and alternative splicing events detected in *u5-3 u5-4* mutant.

Supplementary Table S6. KARR-seq results for U5 snRNA and mRNA chimeric reads.

Supplementary Table S7. List of genes with U5 snRNA-binding sites in the 3'UTR.

Supplementary Table S8. Polysome-associated interactors of SmD3a and SmFa identified by IP-MS.

Supplementary Table S9. Primers used in this study.