iScience, Volume *23*

Supplemental Information

Spatiotemporal Resolved Leaf Angle Establishment Improves Rice Grain Yield via Controlling Population Density

Rongna Wang, Chang Liu, Qinzhong Li, Zhina Chen, Shiyong Sun, and Xuelu Wang

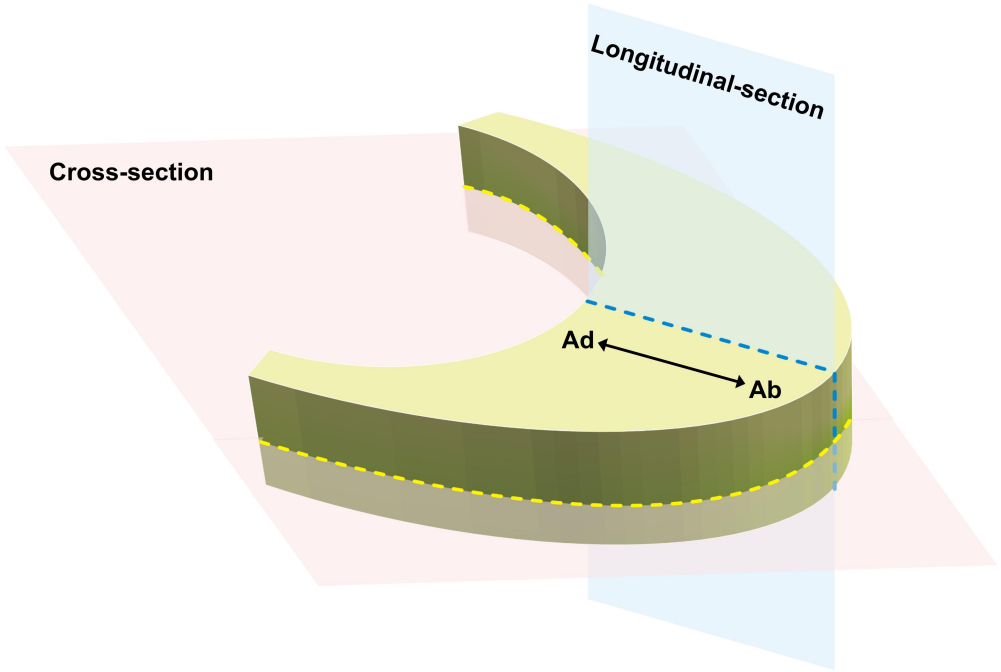

Figure S1. 3D schematics diagram of the positions of the paraffin sections produced from rice LJs, Related to Figure 1.

The cross section is indicated in pink, and the longitudinal-section is indicated in blue. Ad and Ab represent the adaxial and abaxial sides of the LJ, respectively.

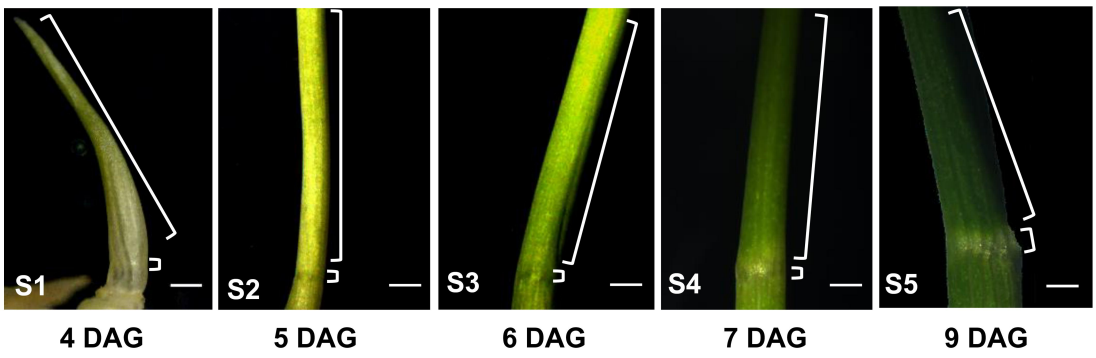

Figure S2. The regions of the first complete leaf collected for RNA-seq analysis at five stages of development, Related to Figure 2 and Figure 3.

The blade regions used for RNA-seq analysis are adjacent to LJs; the LJ regions include the LJs. DAG, days after germination. Scale bars, 500 μm.

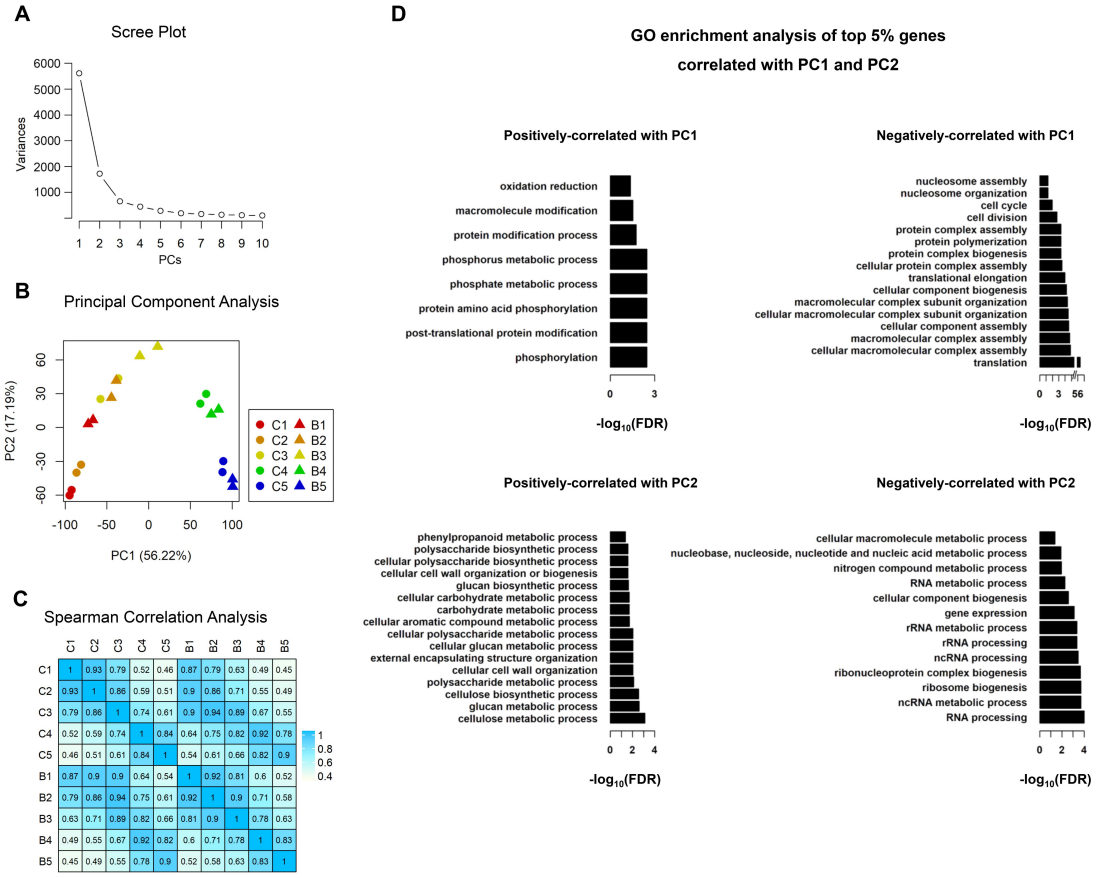

Figure S3. Correlation among the different samples and replicates used for mRNA sequencing, Related to Figure 2.

A, Scree plot for PCA.

B, Principal component analysis (PCA) of 20 libraries from LJs and blades at five stages of development (2 biological replicates per sample). The biological replicates are indicated by the same color and shape; circle represents the LJ, triangle represents the blade; C1 to C5 indicate LJ samples at five stages of development, B1 to B5 indicate blade samples at five stages of development. The proportions of variance represented by PC1 and PC2 are indicated in brackets.

C, GO enrichment analysis (FDR ≤ 0.05) of the top 5% of genes positively or negatively correlated to PC1 and PC2.

D, Spearman's correlation coefficients (SCCs) analysis of the transcriptomes from ten samples using the transformed value $\log_2$(FPKM+1) of genes with more than 1 FPKM in at least one sample.

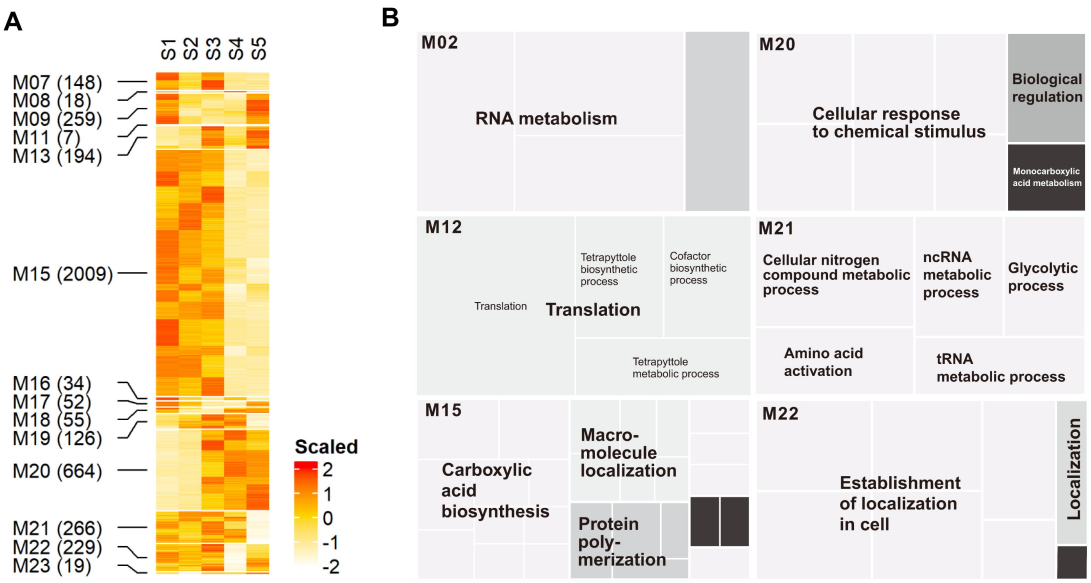

Figure S4. The remained clusters of DEGs and the GO enrichment analysis, Related to Figure 2.

A, Remained clusters of DEGs in LJs. The expression values were normalized using the Z-score method; the number of genes in each cluster is indicated in brackets.

B, Biological processes revealed by GO enrichment analysis (FDR ≤ 0.05) in the remained clusters. The related GO terms are displayed in similar colors; aggregate size indicates the significance of the GO term.

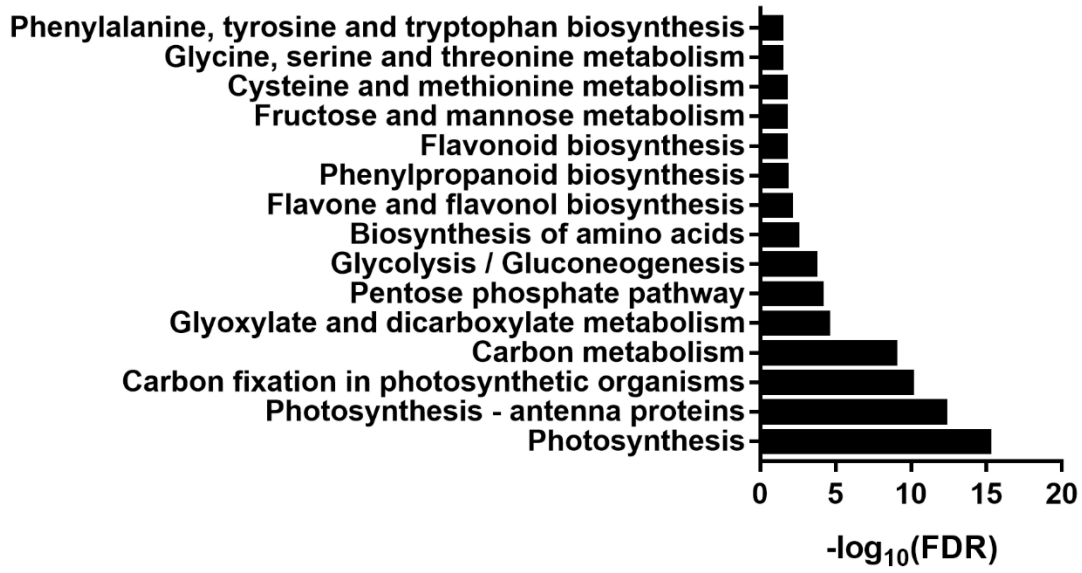

Figure S5. Enriched functional processes in cluster M04 determined by KEGG pathway analysis (FDR ≤ 0.05) , Related to Figure 2.

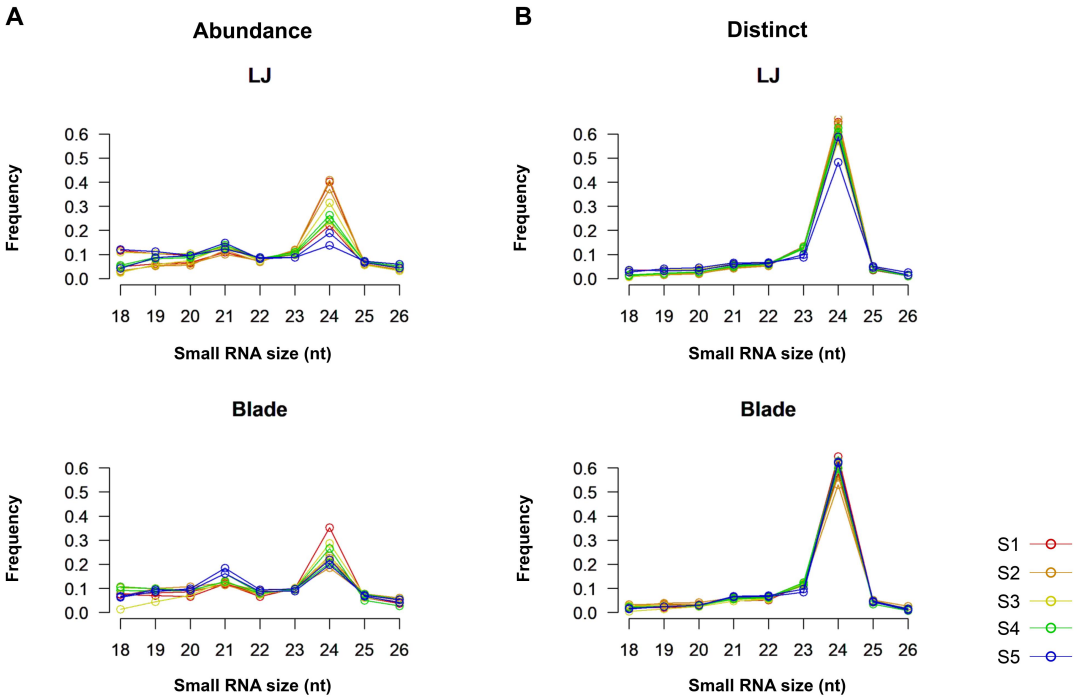

Figure S6. Size distribution of small RNAs in libraries from LJs and blades at five stages of development, Related to Figure 3.

The sizes of small RNAs are plotted versus frequency (percentage) among total sequences (A) and distinct sequences (B). The distinct sequences represent the small RNA species in each library after removing the duplicates from total sequences.

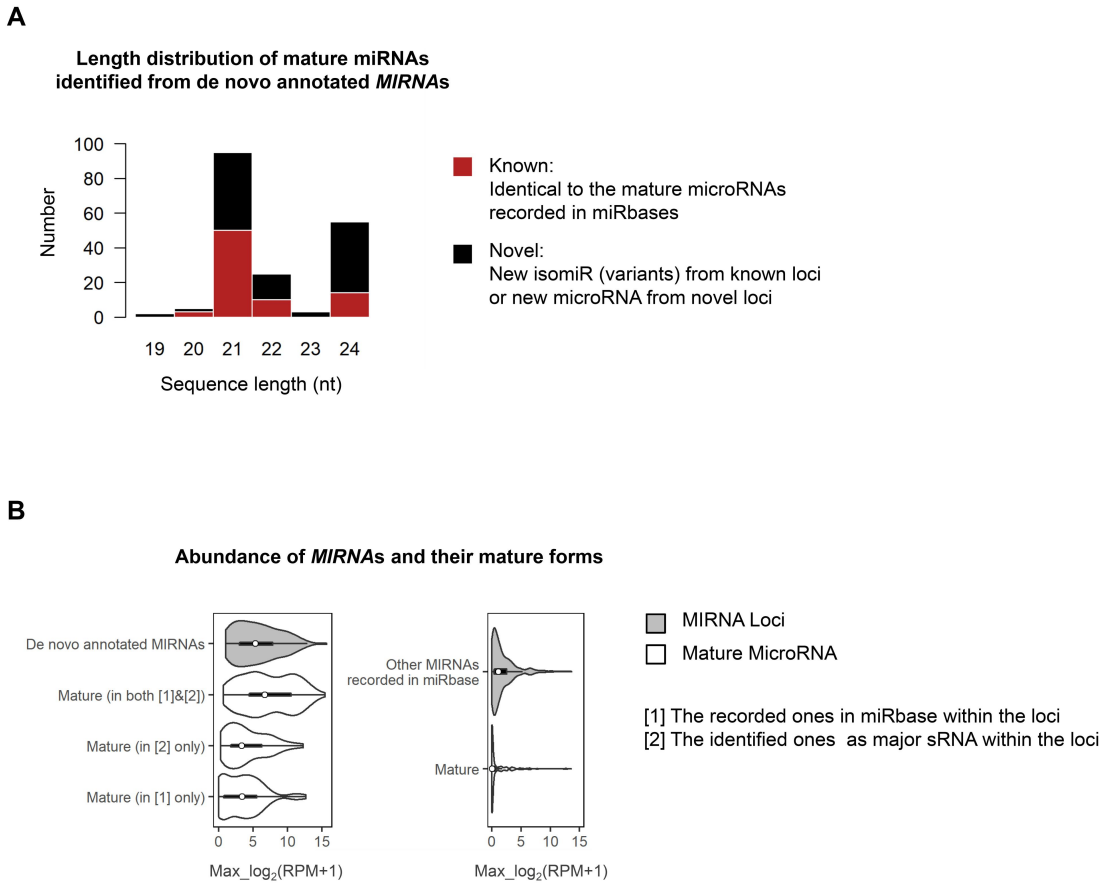

Figure S7. Characters and abundance of mature miRNAs, Related to Figure 3.
A, Characters of the mature miRNAs identified from de novo annotated *MIRNAs*.
B, Abundance of *MIRNAs* and their mature miRNAs identified by de novo annotation and miRBase.

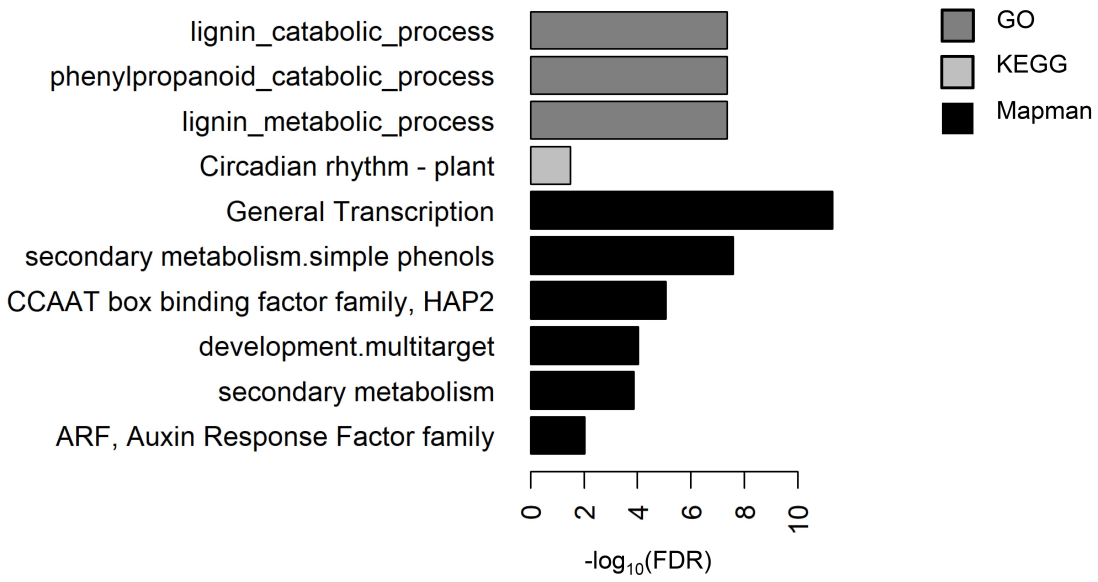

Figure S8. Enriched functional processes (GO/KEGG/MAPMAN; FDR ≤ 0.05) for predicted target genes by stage-specific miRNAs that are differentially expressed in LJs during development (expectation ≤ 2.5), Related to Figure 3.

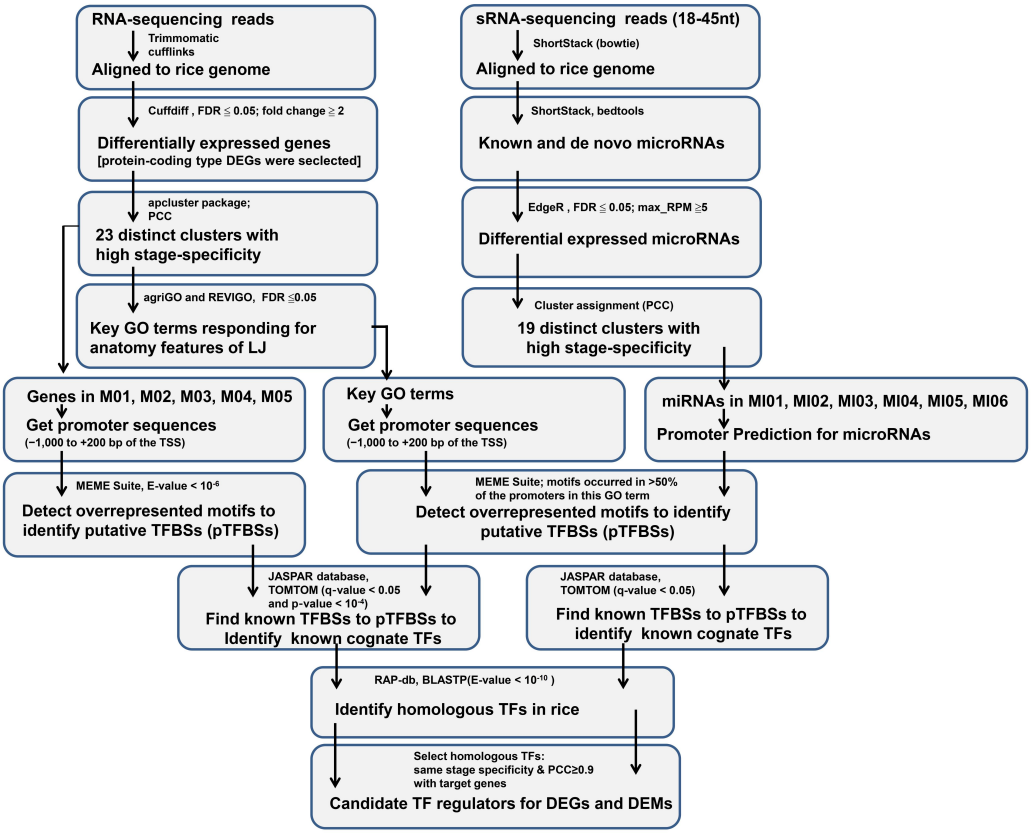

Figure S9. The detailed pipeline used to predict TFBSs and candidate TFs for LJ development in rice, Related to Figure 4.

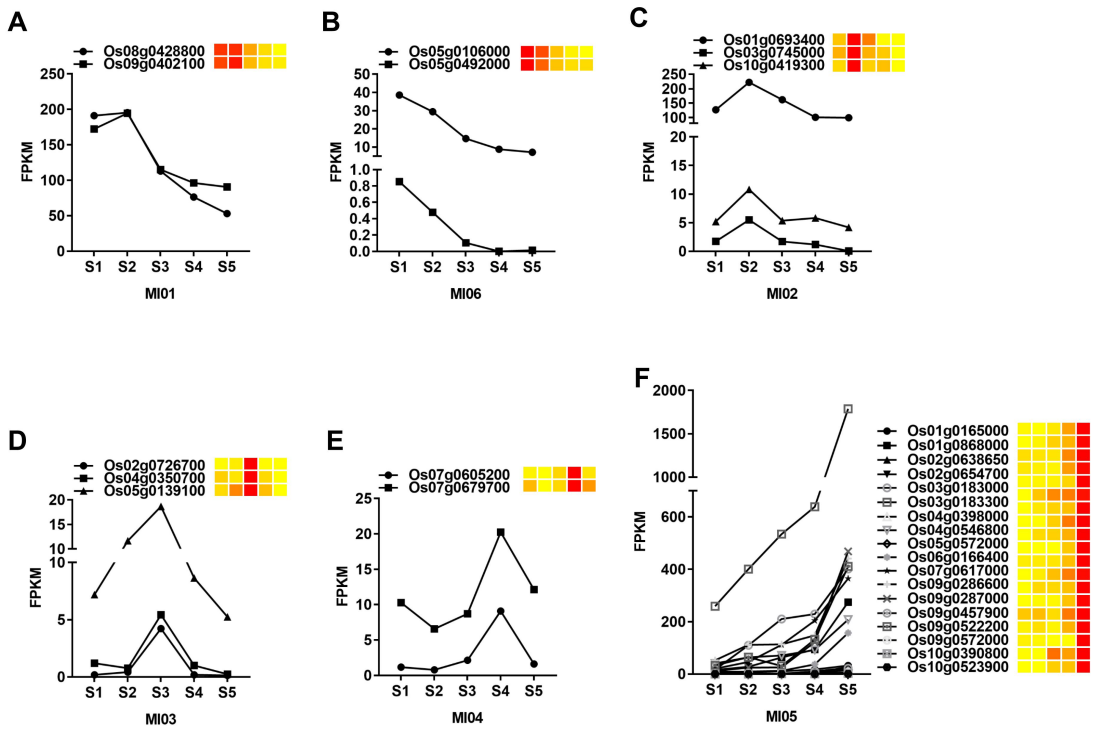

Figure S10. Identification of key TFs determining the expression patterns of *MIRNAs* in LJs, Related to Figure 4.

The FPKM values and the normalized expression values at five stages of development are shown. Heatmaps represent the normalized expression values determined using the Z-score method.

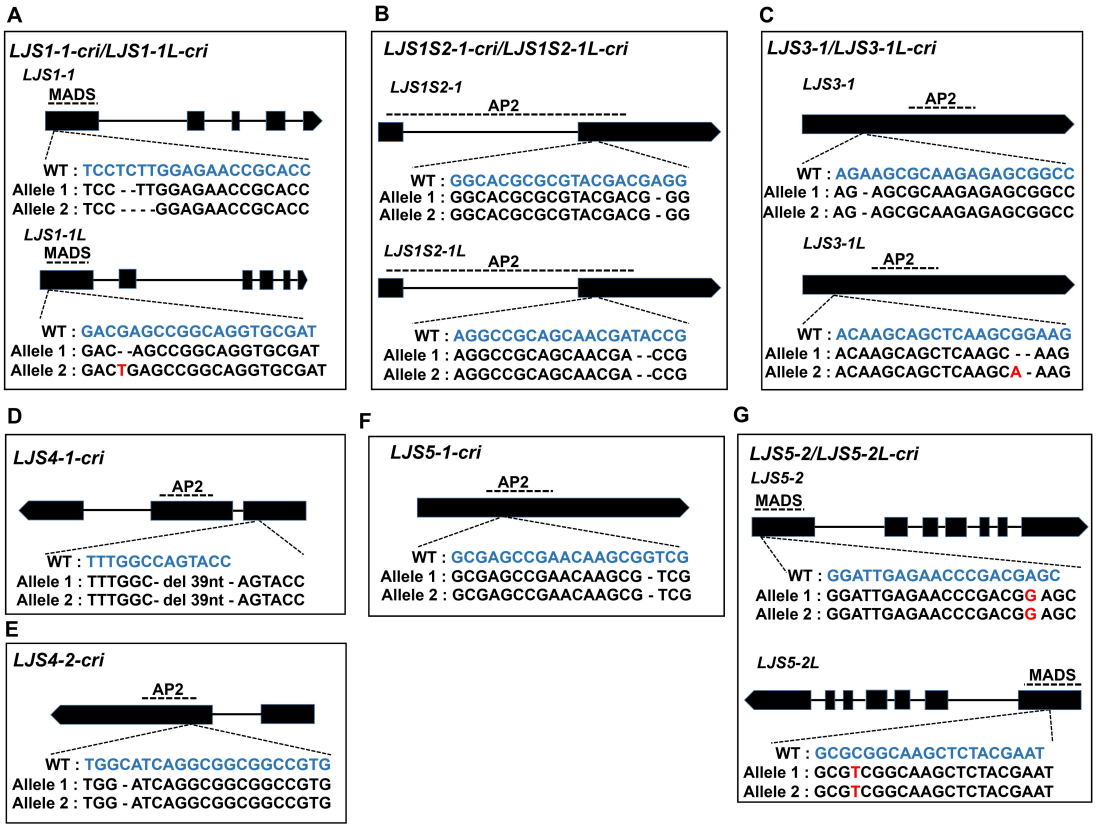

Figure S11. Sequencing results for the knockout lines of the candidate TFs generated by CRISPR/Cas9, Related to Figure 4.

The gRNA-targeting sites are indicated in blue. Insertions are highlighted in red and deletions are indicated by dashes.

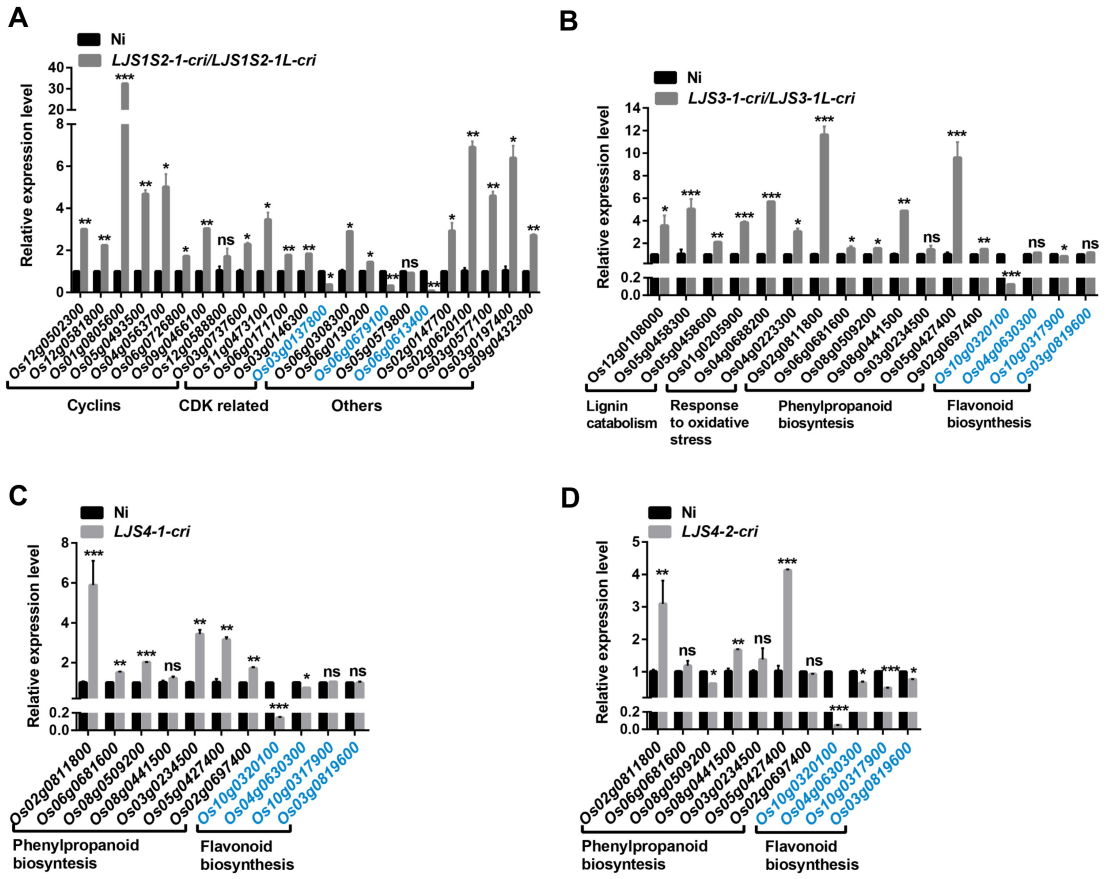

Figure S12. Relative expression levels of putative target genes in specific GO terms, Related to Figure 5.

A, Relative expression levels of putative target genes from the cell cycle term in M06, which were detected in the LJs of *LJS1S2-1-cri/LJS1S2-1L-cri* and Ni at S1, respectively. Gene IDs in black represent genes with positive roles in the cell cycle, whereas gene IDs in blue represent genes with negative roles in the cell cycle.

B, Relative expression levels of putative target genes in the GO terms lignin catabolism and response to oxidative stress in M03, and phenylpropanoid biosynthesis and flavonoid biosynthesis in M04, which were detected in the LJs of *LJS3-1-cri*/*LJS3-1L-cri* and wild-type Ni. Gene IDs in black represent genes with positive roles in lignin biosynthesis, whereas gene IDs in blue represent genes with negative roles in lignin biosynthesis.

C,**D**, Relative expression levels of putative target genes in the GO terms phenylpropanoid biosynthesis and flavonoid biosynthesis in M04, which werre detected in the LJs of *LJS4-1-cri*, *LJS4-2-cri*, and wild-type Ni at S4. Gene IDs in black represent genes with positive roles in lignin biosynthesis, whereas gene IDs in blue represent genes with negative roles in lignin biosynthesis.

The expression levels in Ni were defined as 1. Data are means ± SD. n=3. Significance was tested using Student's *t*-test with $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, and ns indicates no significance.

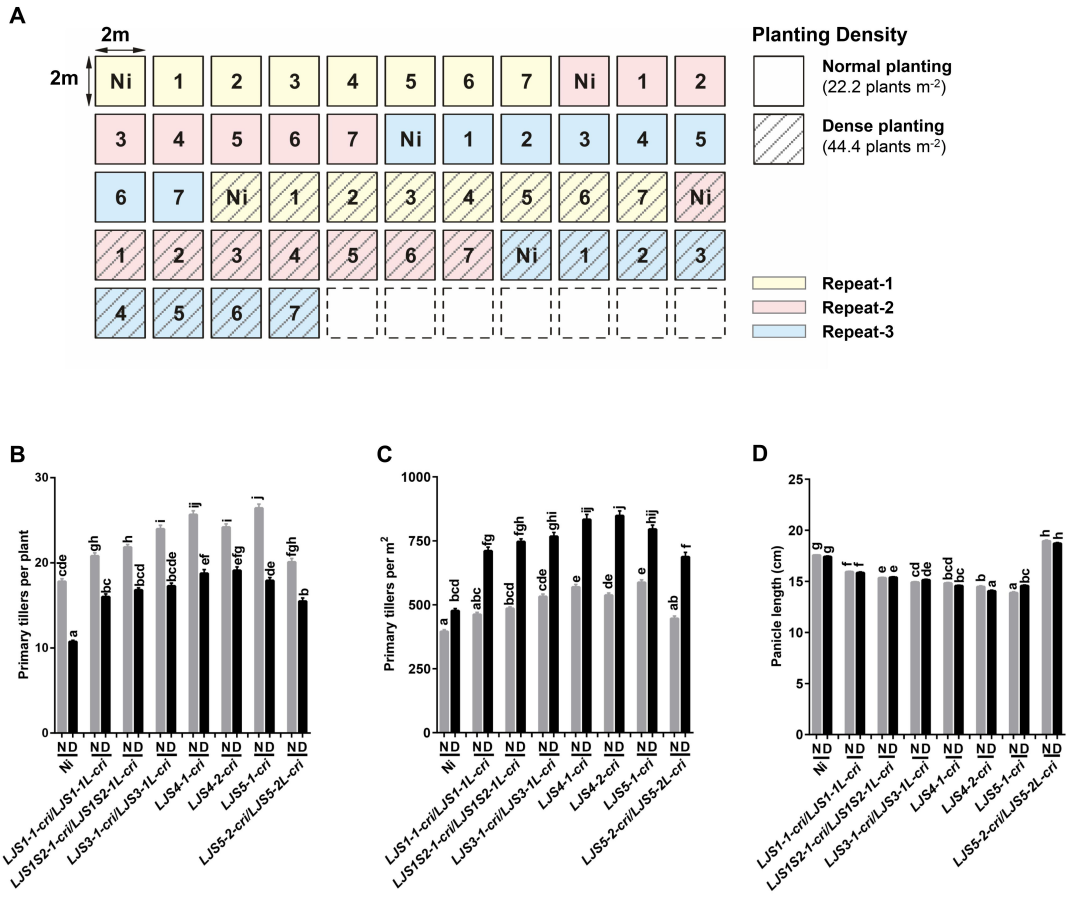

Figure S13. Tiller number and panicle length under the different planting densities, Related to Figure 6.

A, Sketch map of the paddy trial in Wuhan, China in 2019 used to compared knockout lines with erect leaves and wild-type plants (Ni) under two planting densities with three replicates. Knockout lines 1-7 are 1-*LJS1-1cri/LJS1-1L-cri*, 2-*LJS1S2-1-cri/LJS1S2-1L-cri*, 3-*LJS3-1-cri/LJS3-1L-cri*, 4-*LJS4-1-cri*, 5-*LJS4-2-cri*, 6-*LJS5-1-cri*, and 7-*LJS5-2-cri/LJS5-2L-cri*.

B, Primary tiller number per plant under different planting densities. n = 40.

C, Primary tiller number per square meter under different planting densities.

D, Panicle length under different planting densities. n = 70.

N: normal planting (22.2 plants m^{-2}); D: dense planting (44.4 plants m^{-2}). Data are the means ± SE. Lowercase letters indicate significant differences at the level of $P < 0.05$ within a parameter (Tukey's Honest Significant Difference test).

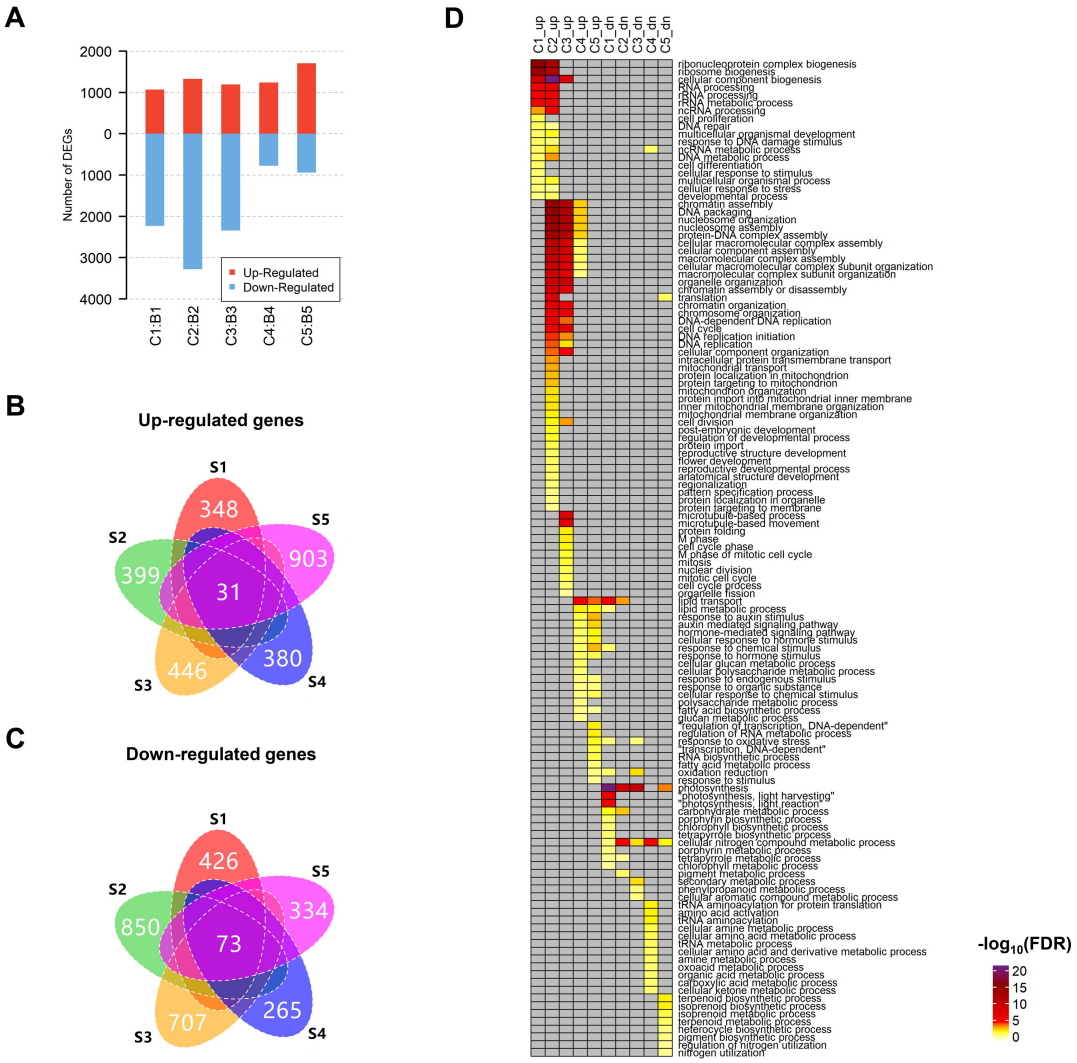

Figure S14. Identification of genes with organ-specificity at five stages of development, Related to Figure 2.

A, Number of organ-specific genes at each stage (FDR ≤ 0.05 and fold change ≥ 2).

B, Up-regulated genes in LJs compared to blades at each stage.

C, Down-regulated genes in LJs compared to blades at each stage.

D, Enriched GO terms in the biological processes category (FDR ≤ 0.05) for the organ-specific genes at each stage. Gray box indicates that a GO term is not enriched in a specific gene set.

Transparent Methods

Plant growth conditions and sample collection

The rice variety used in this study was *japonica* (*Oryza sativa*) cultivar, Nipponbare (Ni). The seeds were soaked in distilled water at room temperature for 1 day and imbibed for 2 days in water at 37°C in a constant temperature incubator (referred to as the first and second day after germination, respectively). The water was changed twice daily. The germinated seeds were transferred to 96-well plates with cut bottoms and grown in water under a 16 h (light, 28°C)/8 h (dark, 25°C) cycle. LJ and blade tissue form the first complete leaves of seedlings were separately cut and collected on the 4^{th}, 5^{th}, 6^{th}, 7^{th}, and 9^{th} day after germination, the length of LJ were less than 1 mm, and the length of blade was less than 10 mm. In addition, the leaf materials inside the outer leaf were removed respectively. Approximately 150 seedlings at each stage were dissected to isolate RNA for LJs and leaf blades. The collected tissues were immediately frozen in liquid nitrogen and subjected to total RNA extraction with TRIzol (Invitrogen). Two biological replicates per sample were produced to construct RNA-seq and small RNA-seq libraries.

Paraffin sectioning

LJ and blade tissues were fixed in FAA solution (50% ethanol, 5% acetic acid, and 10% formaldehyde in water), vacuum infiltrated for one hour, dehydrated with a graded ethanol series (from 50% ethanol to absolute ethanol), and embedded in paraffin. Microtome sections (7 um) were cut with a paraffin-slicing machine (Leica) and affixed to egg white-glycerin-coated slides. The sections were de-paraffinized in xylene, dehydrated through a graded ethanol series, and stained with Safranin O/Fast Green, and observed under a light microscope (Zeiss, Germany).

RNA library preparation, sequencing, and data processing

RNA-seq libraries were constructed using a NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (NEB, USA) and sequenced on the Illumina HiSeq 4000 platform to generate 150 bp paired-end reads. After removing low-quality reads, the clean reads

were trimmed with Trimmomatic-0.36 and then mapped to the rice reference genome rice IRGSP-1.0 with TopHat v2.1.1 (Kim et al., 2013). The uniquely mapped reads were used to calculate FPKM (fragments per kilobase of transcript per million reads) values for the genes with Cufflinks v2.2.1 (Trapnell et al., 2012) to estimate their expression levels in different samples. The rice genome and gene annotation file was downloaded from Ensembl Plants (http://plants.ensembl.org/index.html). Principal component analysis (PCA) and Spearman's correlation coefficients (SCCs) analysis based on gene expression profiles were performed using the prcomp and cor functions in R software.

Identification of DEGs and DEMs, clustering, and functional enrichment analysis

Differentially expressed genes (DEGs) in each pair of samples were identified using Cuffdiff (Cufflinks v2.2.1, FDR ≤ 0.05 and Fold change ≥ 2) and merged into three gene sets based on stage-level differences between LJ samples only or blade samples only and the organ-level differences between LJs and blades at the same stage. The stage-level DEGs for LJs or blades with annotated coding sequences were used to perform clustering to obtain predominant gene sets with distinct stage specificity to uncover active functional modules during organ development in subsequent analysis. The DEGs were assigned into small clusters by the affinity propagation-clustering algorithm (Bodenhofer et al., 2011) using the R package apcluster with default parameters. Given that many of these clusters still exhibited similar stage specificity, we further combined them into 23 gene sets based on their similarity to the expected expression pattern with the highest Pearson correlation coefficient in 5 LJ samples (PCC5). The expected patterns were determined by permuting five elements with values of “1” or “0” in a numeric array, which represent single stage or stages-group specificity. The DEGs between LJs and blades at each stage were defined as organ-specific genes.

The differentially expressed microRNAs (DEMs) for each sample pair were identified with edgeR (FDR ≤ 0.05; maximum expression value ≥ 5RPM). The DEMs between LJ samples were directly clustered into 19 stage-specific modules, based on their similarity to the expected expression pattern with the highest PCC, as mentioned

above.

Gene ontology (GO) enrichment analysis (FDR ≤ 0.05) of particular gene sets was performed using the web tool agriGO (http://bioinfo.cau.edu.cn/agriGO/index.php), and the GO terms were summarized with REVIGO (Supek et al., 2011). The KEGG enrichment analysis (FDR ≤ 0.05) for particular gene sets was performed with the R package "clusterProfiler", and the pathway annotations of rice genes were downloaded from KEGG (https://www.genome.jp/kegg/).

Small RNA library preparation, sequencing, and miRNA identification

The total RNAs were sent to BGI (Shenzhen, China) for small RNA library construction and high-throughput sequencing on the BGISEQ-500 platform. Adaptor sequences, contamination, and low-quality reads were removed from the raw reads. Mapping of the clean data to the reference genome and de novo *MIRNA* identification were performed with ShortStack (v 3.8.2; parameter set: --dicermin 18 --dicermax 26 --foldsize 400) for each sample and merged *MIRNA* loci were obtained with bedtools v2.17.0. The identified *MIRNA*s (microRNA loci) were compared to the annotated *MIRNA*s in miRBase (v21; osa.gff3). The major sRNA species with the highest abundance identified from each *MIRNA* locus were considered to be the mature miRNAs in different samples, which were also compared with he mature miRNAs in miRBase. A locus with a mature miRNA whose sequence was similar to that of a known miRNA annotated in miRBase was considered to be a paralogous member of that known family (Meyers et al., 2008). A mature miRNA whose sequence differed from the known sequence at the same locus was considered to be a variant or isomiR.

Target prediction and differentially expressed miRNAs

Target prediction of the mature miRNAs was performed with psRNATarget (http://plantgrn.noble.org/psRNATarget/; expectation ≤ 2.5), and differentially expressed miRNAs between samples were detected with edgeR (FDR ≤ 0.05; maximum RPM ≥ 5). The miRNAs identified in comparisons between C1-C5 were considered to be stage-specific miRNAs that function in LJ development. The stage-specific miRNAs were directly assigned to 19 clusters based on their expression dynamics with the highest similarity to the expected expression patterns, which

represent single stage or stages-group specific miRNAs, as mentioned above.

Identification of cis-motifs and candidate cognate TFs

For identification of cis-elements, first, the gene sets including all the genes in clusters M01 to M05 were used. Second, we also used genes sets enriched in GO terms which were the main developmental features of LJ in different stages, including lignin catabolism, lipid metabolism, lipid transport, response to oxidative stress in M03, carbohydrate catabolism in M04, and the cell cycle in M06. To predict TFs which potentially regulated specific gene sets, first, the overrepresented motifs in promoters of genes in the specific gene set were detected using the MEME program (Bailey and Elkan, 1994): for predicting motifs in gene sets M01-M05, the top 10 motifs with E-value $\leq 10^{-6}$ were selected as putative transcription factor binding sites (pTFBSs); for predicting motifs in gene sets of GO terms, the top 10 motifs that occurred in $\geq$ 50% of the promoters were selected as pTFBSs. Then, the pTFBSs were compared to known cis-motifs (JASPAR CORE database) using the TOMTOM program (Gupta et al., 2007) to identify the significantly similar binding sites (q-value $\leq$ 0.05 and p-value $\leq 10^{-4}$) that could be recognized by known cognate TFs. Third, the rice homologs of these cognate TFs were identified by BLASTp analysis (e-value $\leq 10^{-10}$; the protein sequences of cognate TFs and rice homologs are downloaded from UniProt and RAP-DB, respectively). Finally, the candidate TFs were identified with two characters: first, they belonged to the same TF family with the cognate TFs, which is predicted by plantTFDB (Jin et al., 2017; http://planttfdb.cbi.pku.edu.cn/); second, they exhibited the similar stage specificity with their predicted targets (PCC5 $\geq$ 0.9). Notably, the TF-targeted genes are the genes from the corresponding gene set whose promoters contain the TF-specific pTFBS from above MEME analysis. We have showed the detailed pipeline in Supplementary Figure 9.

The identified *MIRNA*s (microRNA loci) were compared with the annotations of protein-coding genes (https://rapdb.dna.affrc.go.jp/) to uncover intragenic and intergenic *MIRNA*s (Cui et al., 2009). Given that the intragenic *MIRNA*s were potentially transcribed as a part of their overlapping genes (Lee et al., 2004), their promoter regions were considered to be the same as those of these overlapping genes. To predict the promoters of intergenic *MIRNA*s, the specified upstream regions were

extracted as previously described, which were considered to include TSSs driving transcription (Bailey and Elkan, 1994). The method used to predict TFs regulating miRNA expression was almost the same as that used to predict the regulators of genes from the enriched GO terms mentioned above, except that a threshold of q-value ≤ 0.05 was used (TOMTOM) in the pTFBS comparisons to retain enough results for the clusters we were interested in.

Vector construction and plant transformation

To generate the CRISPR/Cas9 knockout mutants, the pCXUN-CAS9 vector (He et al., 2017) was used. The CRISPR target sequences for selected TF genes were obtained from CRISPR-P 2.0 (http://crispr.hzau.edu.cn/CRISPR2/). To genotype homozygous mutants generated by CRISPR/Cas9, less than 800 bp DNA fragments covering the CRISPR target sequences were amplified and sequenced.

RNA extraction and RT-qPCR

Total RNA was extracted from the samples using the TRIzol method (DP432, Tiangen). First-strand cDNA was synthesized using a PrimeScript RT reagent kit (TaKaRa). For RT-qPCR, SYBRMaster Mix for PCR (Invitrogen) was used. *OsUBQ* (*Os03g0234200*) was used as an internal control to normalize the data. Three biological repeats were performed. The primers are listed in Table S7.

Paddy trials and yield evaluation under different planting densities

Paddy trials were conducted in Wuhan in 2019. The KO and wild-type rice plants were grown in a paddy field with row spacing and plant spacing in a row at a distance of 30 × 15 cm (22.2 plants per m^2) or 15 × 15 cm (44.4 plants per m^2). Each treatment was replicated three times in randomized blocks. Each plot was 2 × 2 m. The number of tillers per treatment was counted in the field from 20 plants excluding marginal plants. Fifty plants were harvested from each plot excluding marginal plants. Tillers with fertile panicles (more than 5 fertile seeds) and panicle length were determined from approximately 30 plants. After harvest, the plants were dried for 7 days, and 1,000- grain weight and grain yield per plant were measured from approximately 30 plants. Grain yield was converted to value per hectare. Standard statistical procedures

were used to analyze the data using SPSS 17.0 (Softonic International, Barcelona, Spain). Tukey's Honest Significant Difference test was used for multiple comparisons at the $P < 0.05$ level.

Supplemental References

Bailey, T.L. and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28-36.

Bodenhofer, U., Kothmeier, A., and Hochreiter, S. (2011). APCluster: an R package for affinity propagation clustering. Bioinformatics *27*, 2463-2464.

Cui, X., Xu, S., Mu, D., and Yang, Z. (2009). Genomic analysis of rice microRNA promoters and clusters. Gene *431*, 61-66.

Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L., and Noble, W. S. Quantifying similarity between motifs. Genome Biol. *8*, R14 (2007).

He, Y., Zhang, T., Yang, N., Xu, M., Yan, L., Wang. L., Wang, R., and Zhao, Y. (2017). Self-cleaving ribozymes enable the production of guide RNAs from unlimited choices of promoters for CRISPR/Cas9 mediated genome editing. J. Genet. Genomics *44*, 469-472.

Jin, J., Tian, F., Yang, D., Meng, Y., Kong, L., Luo, J., and Gao, G. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Res. *45*, D1040-D1045.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Lee, Y., Kim, M., Han, J., Yeom, K., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. EMBO J. *23*, 4051-4060.

Meyers, B.C., Axtell, M.J., Bartel, B., Bartel, D.P., Baulcombe, D., John L. Bowman,

J.L., Cao, X., Carrington, J.C., Chen, X., Pamela J. Green, P.J., et al. (2008). Criteria for annotation of plant MicroRNAs. Plant Cell *20*, 3186-3190.

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat. Protoc. *7*, 562-578.