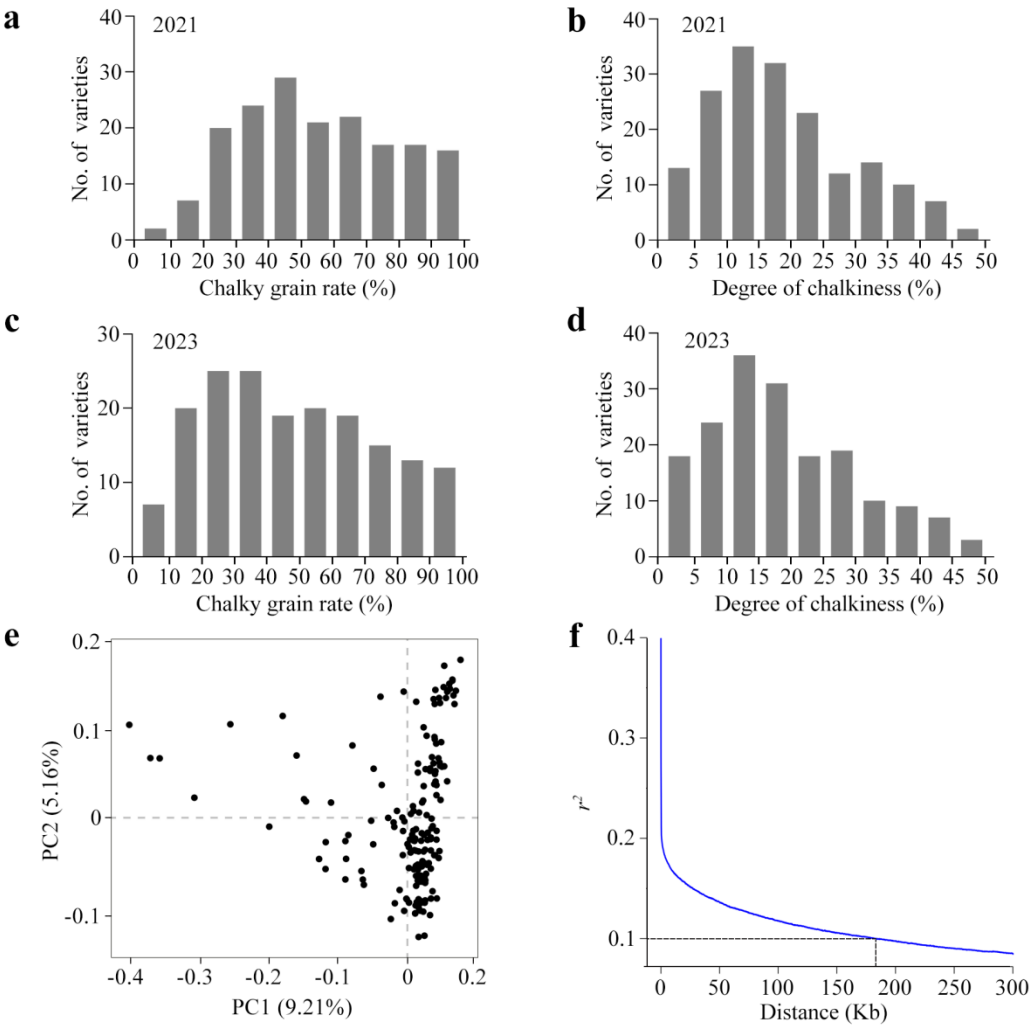
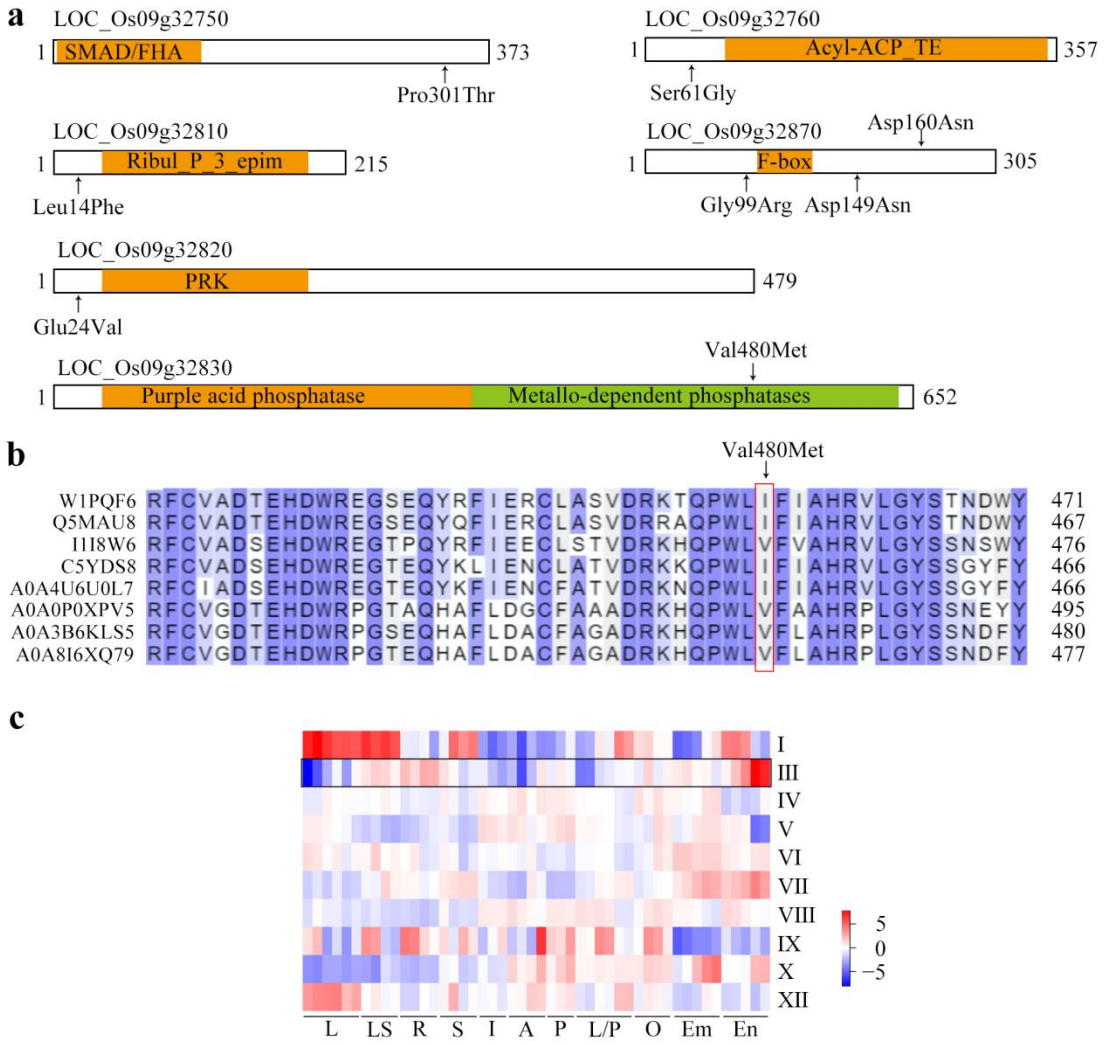
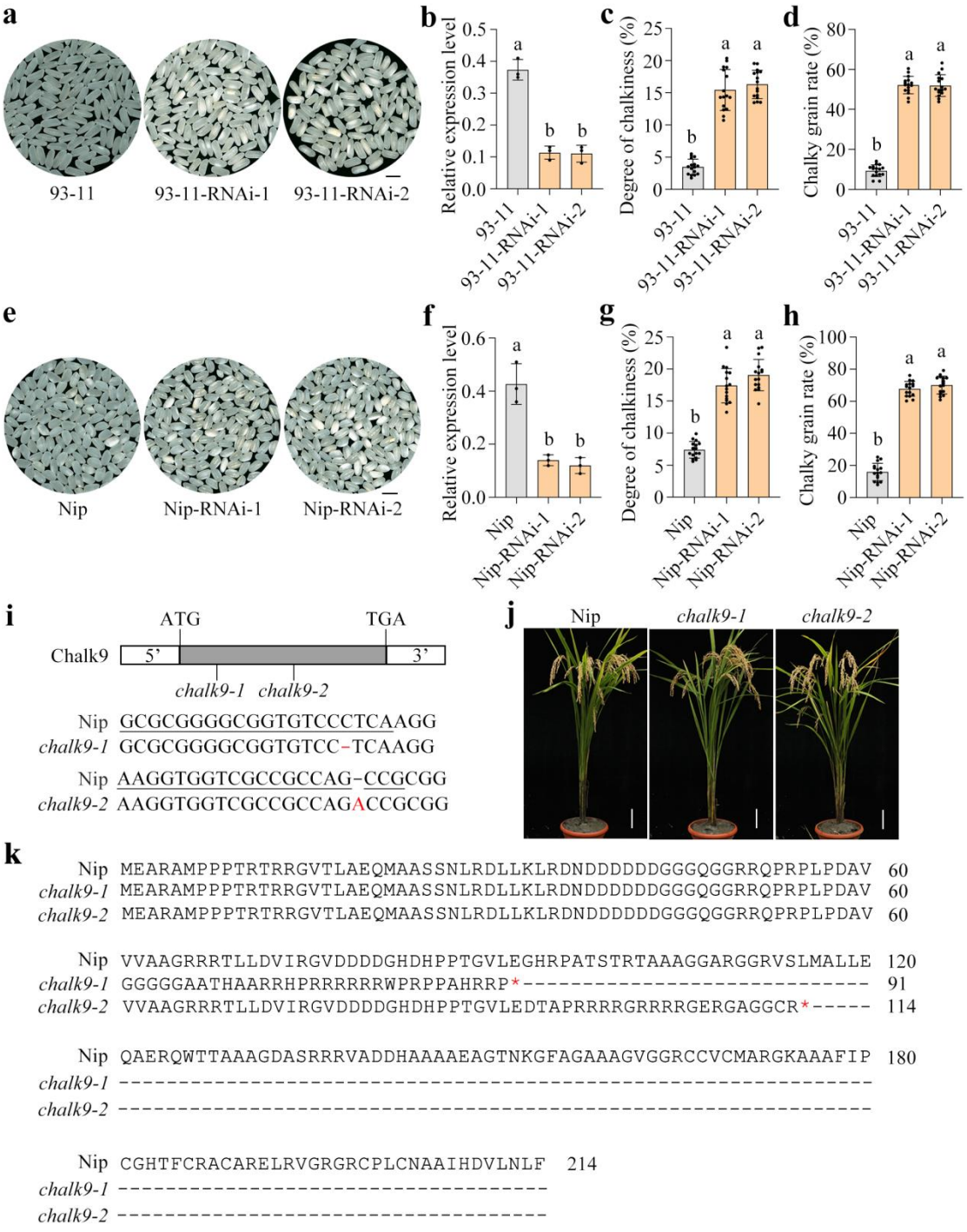


Supplementary Fig. 1 Variations of chalky grain rate and degree of chalkiness in 175 *indica* varieties. a-d, Frequency distributions of chalky grain rate (**a** and **c**) and degree of chalkiness (**b** and **d**) in two years. **e**, PCA for the 175 indica rice varieties based on whole-genome sequencing data. PC1 and PC2 indicate score of principal components 1 and 2, respectively. Values in parentheses indicate percentage of variance in the data explained by each principal component. **f**, Genome-wide average LD decay in the 175 varieties. LD was calculated as the squared Pearson's correlation coefficient (r^2).



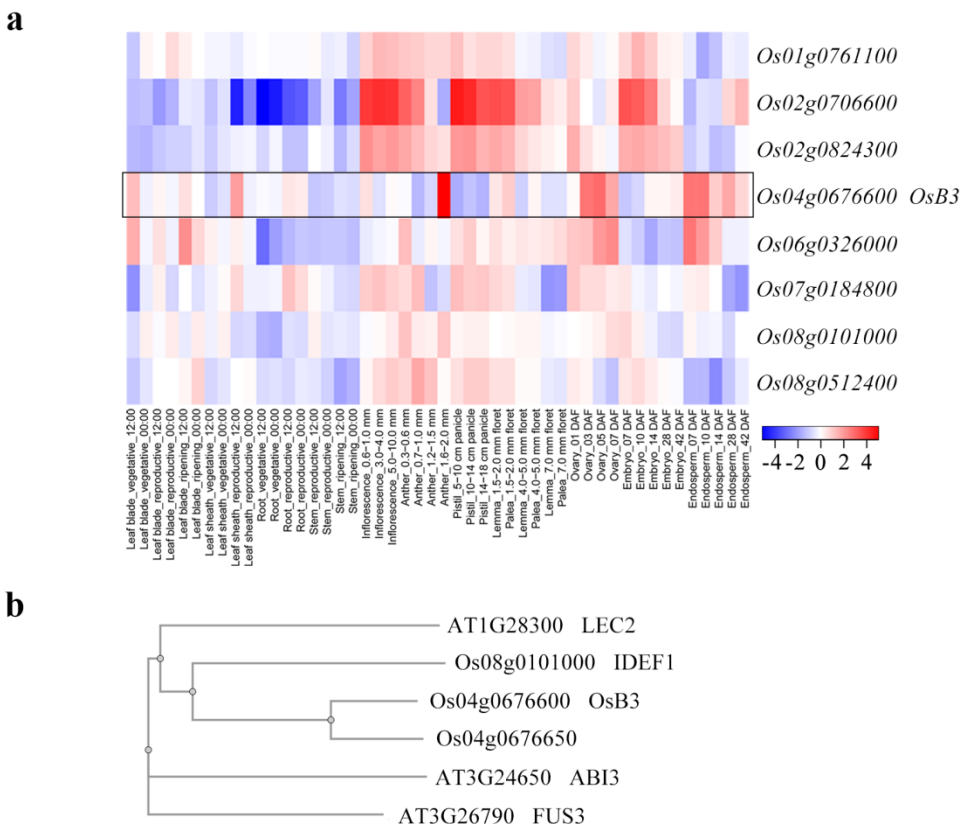
Supplementary Fig. 2 Functional importance estimation of SNPs located in the coding region. a, Schematic presentation of the candidate proteins with the amino acid substitution and the corresponding functional domain. **b**, The amino acid sequences in the region of the amino acid substitution site were compared using ClustalW. The analysis included sequences from *Oryza sativa* (A0A0P0XPV5/LOC_Os09g32830) and its orthologs in *Amborella trichopoda* (W1PQF6), *Arabidopsis thaliana* (Q5MAU8), *Brachypodium distachyon* (I1I8W6), *Sorghum bicolor* (C5YDS8), *Setaria viridis* (A0A4U6U0L7), *Triticum aestivum* (A0A3B6KLS5), and *Hordeum vulgare subsp. vulgare* (A0A8I6XQ79) were compared by ClustalW. The residues surrounded by red box correspond to allelic variants found in the 175 rice varieties. **c**, Expression

analysis of the candidate genes from GWAS in various tissues. The result of two genes (II and XI) was not found in RiceXPro database. L, leaf blade; LS, leaf sheath; R, root; S, stem; I, inflorescence; A, anther; P, pistil; L/P, lemma/palea; O, ovary; Em, embryo; En, endosperm.

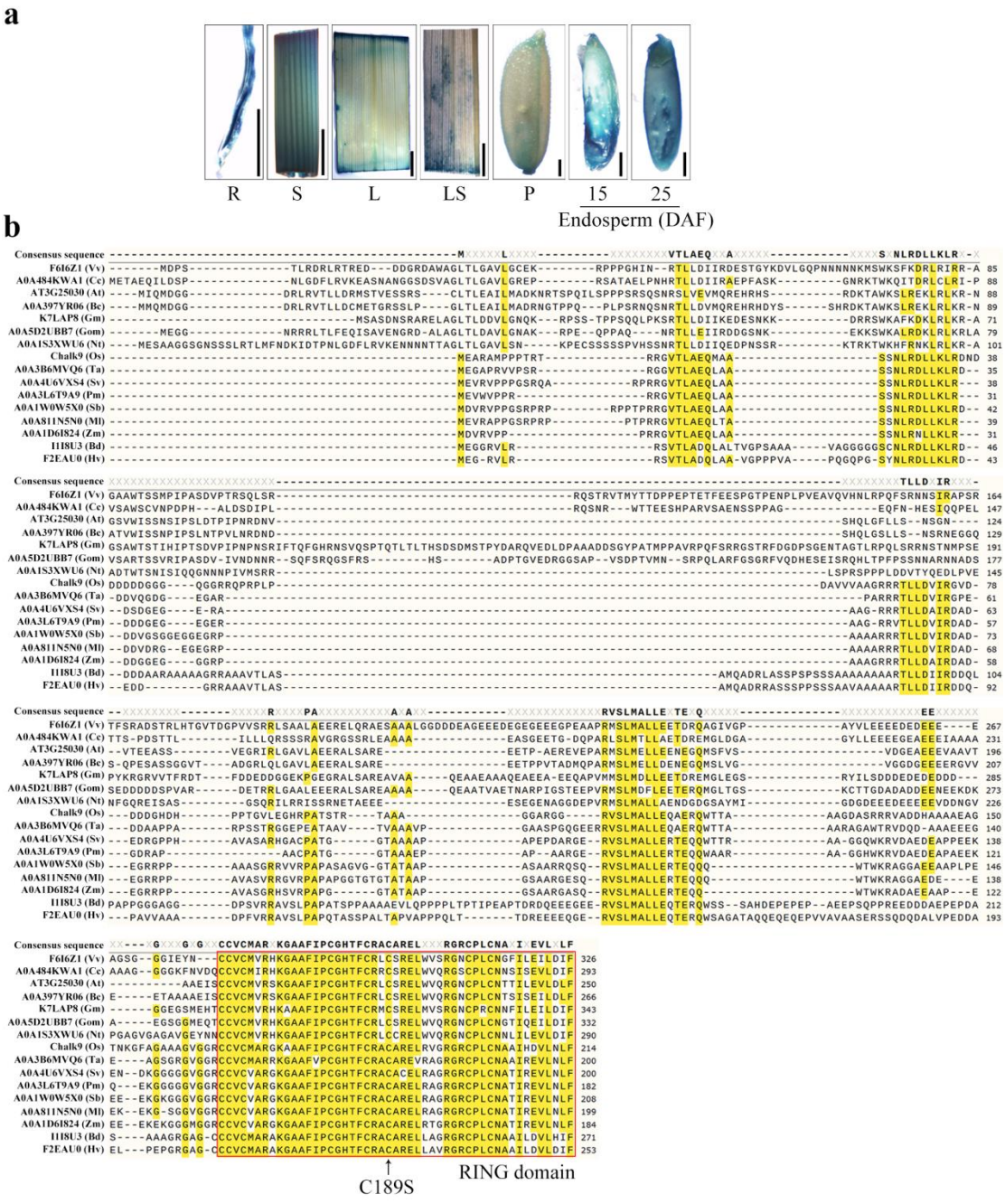


Supplementary Fig. 3 Identification of *Chalk9* RNAi and knockout plants. a, Grain chalkiness in 93-11, 93-11-RNAi-1, and 93-11-RNAi-2 plants. Scale bar: 8 mm. **b**, Relative expression level of *Chalk9* in 93-11, 93-11-RNAi-1, and 93-11-RNAi-2 plants. Data show means ± SD (n = 3 biological replicates). **c**,**d**, Degree of chalkiness (**c**) and chalky grain rate (**d**) in 93-11, 93-11-RNAi-1, and 93-11-RNAi-2 plants. Data show

means ± SD (*n* = 16 plants). **e**, Grain chalkiness in Nip, Nip-RNAi-1, and Nip-RNAi-2 plants. Scale bar: 5 mm. **f**, Relative expression level of *Chalk9* in Nip, Nip-RNAi-1, and Nip-RNAi-2 plants. Data show means ± SD (*n* = 3 biological replicates). **g**,**h**, Degree of chalkiness (**g**) and chalky grain rate (**h**) in Nip, Nip-RNAi-1, and Nip-RNAi-2 plants. Data show means ± SD (*n* = 16 plants). **i**, Targeted mutagenesis of *Chalk9*. The target mutated sites are indicated on the gene structure of *Chalk9*. Gray box indicates the exon of *Chalk9* gene. The 20-bp target sequence for CRISPR/Cas9-mediated editing is underlined. **j**, Plant morphology of Nip, *chalk9-1*, and *chalk9-2* plants at the mature stage. Scale bar: 10 cm. **k**, Comparison of amino acid sequences encoded by *Chalk9* gene in the wild-type Nip and the truncated sequences of Chalk9 protein in *Chalk9* knockout mutants. Red asterisk indicates the termination codon. In **b**-**d** and **f**-**h**, different letters indicate significant differences ($P < 0.05$, one-way ANOVA with Tukey's multiple comparison test); for *P* values, see Source Data.

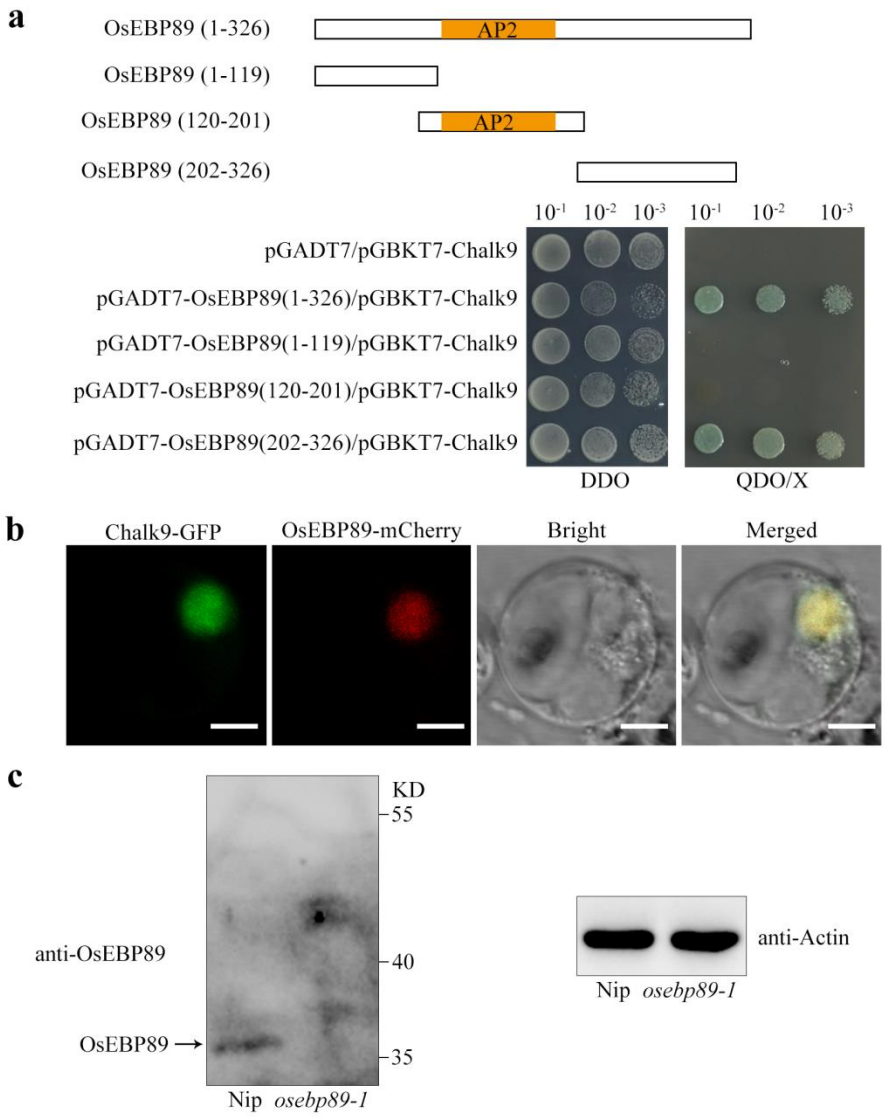


Supplementary Fig. 4 Identification of the candidate genes in transcriptional factors analysis. a, Expression pattern of the candidate genes from transcriptional factors analysis in various organs and tissues. **b**, Phylogenetic analysis of ABI3, FUS3 and LEC2 in *Arabidopsis* and their orthologues of B3 domain transcription factor in rice. The neighbour-joining tree was constructed by Clustal Omega tool.



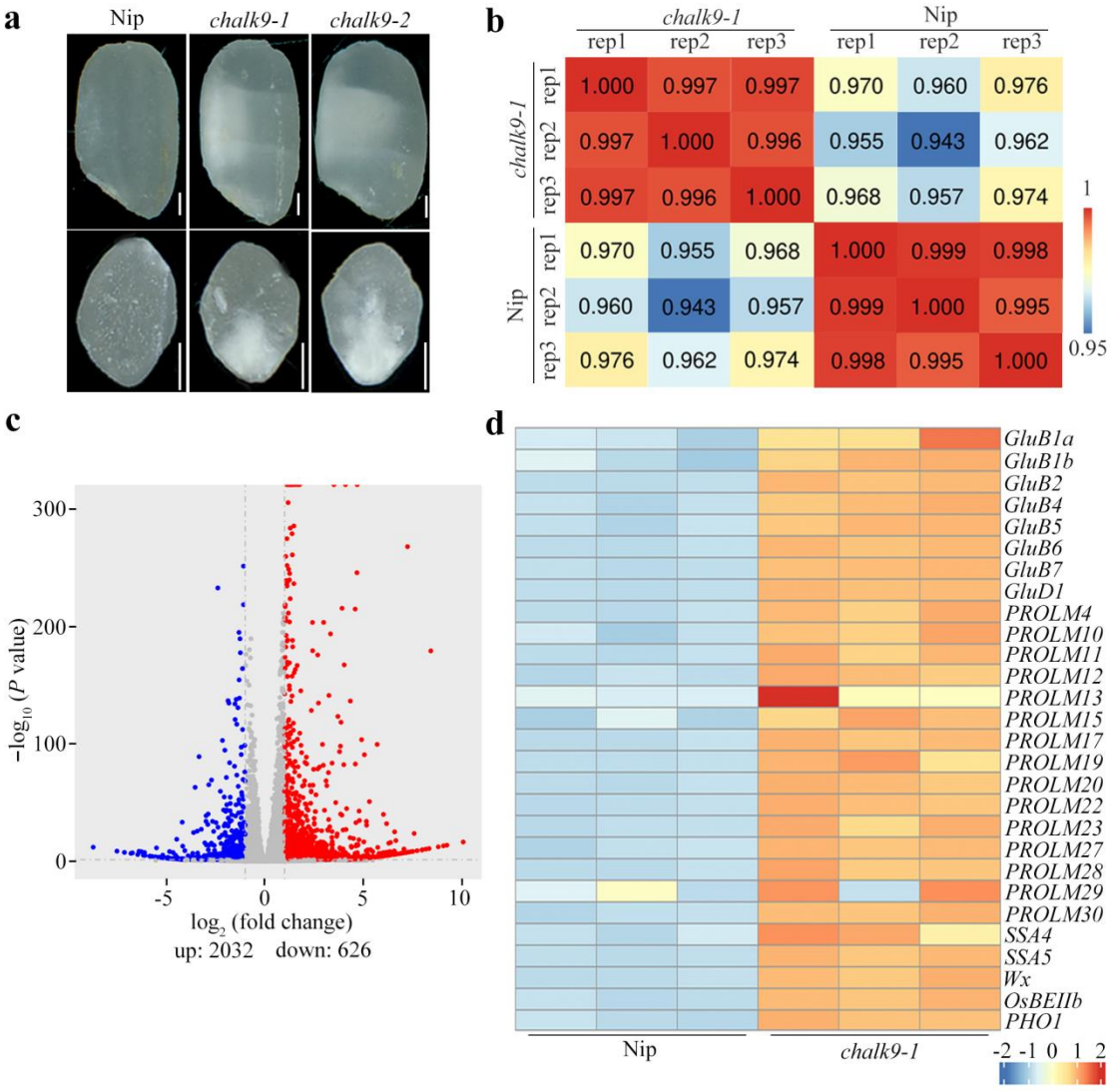
Supplementary Fig. 5 The amino acid sequence of RING domain in Chalk9 is highly conserved in plants. a, *Chalk9* promoter activity was monitored using *Chalk9pro::GUS* transgenic plants. Histochemical GUS staining was performed on root (R), stem (S), leaf (L), leaf sheath (LS), panicle (P), and longitudinally sectioned seed. DAF, days after flowering. Scale bar: 1 mm. **b**, The full-length sequences of Chalk9 orthologs from various plants were subjected to multiple sequence alignment. The red

box indicates RING domain, which is conserved among higher plants. Arrow indicates the site mutated in the RING finger domain of Chalk9. Multiple sequence alignment was performed using ClustalW. The following abbreviations are used: Os: *Oryza sativa*, Cc: *Cuscuta campestris*, Bc: *Brassica campestris*, Sv: *Setaria viridis*, Ta: *Triticum aestivum*, Pm: *Panicum miliaceum*, Sb: *Sorghum bicolor*, Ml: *Miscanthus lutarioriparius*, Bd: *Brachypodium distachyon*, Vv: *Vitis vinifera*, At: *Arabidopsis thaliana*, Gm: *Glycine max*, Gom: *Gossypium mustelinum*, Nt: *Nicotiana tabacum*, Zm: *Zea mays*, Hv: *Hordeum vulgare*.

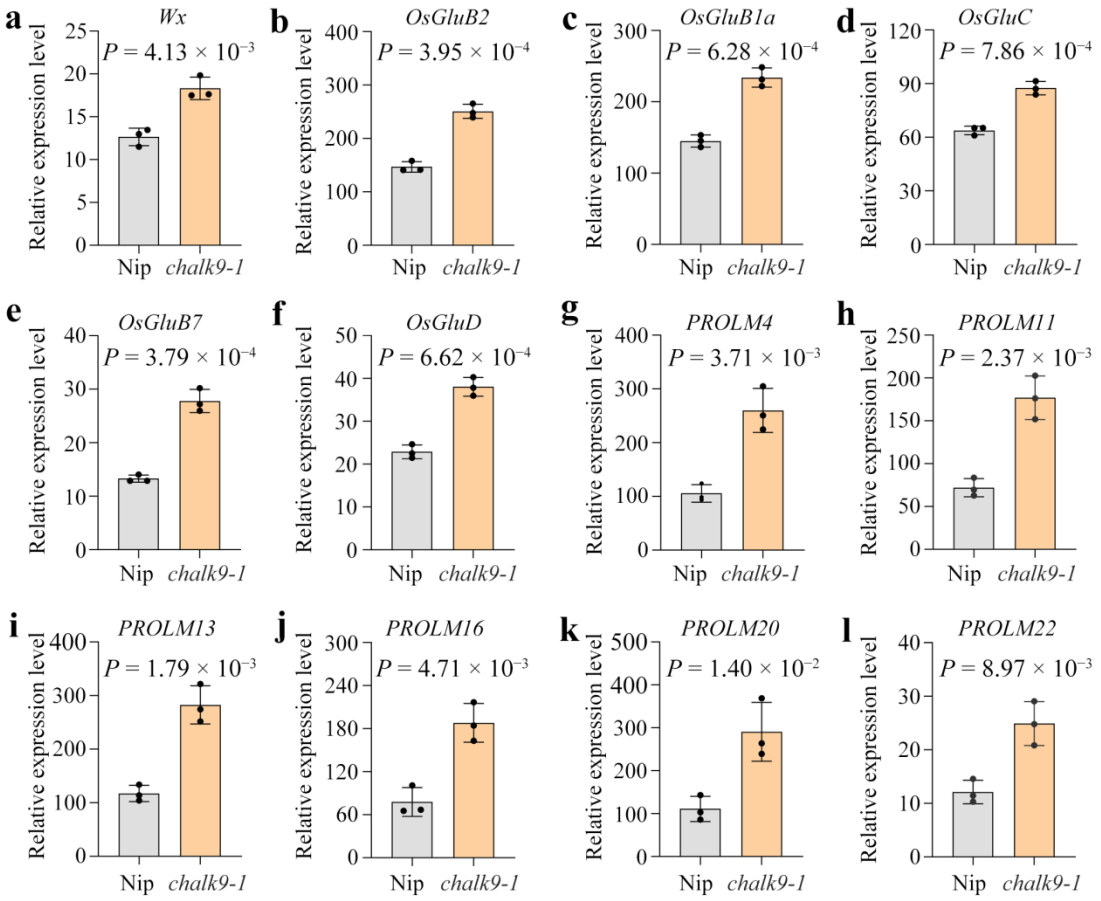


Supplementary Fig. 6 Functional analysis of OsEBP89 and identification of the anti-OsEBP89 antibody. a, Schematic diagram showing the full-length and three truncated fusions of OsEBP89 to the Gal4 activation domain (pGADT7), along with the yeast two-hybrid assay results involving the indicated OsEBP89 constructs and Chalk9. The truncated forms include OsEBP89 (1-119), containing the N-terminal domain, OsEBP89 (120-201), containing the AP2 domain, and OsEBP89 (202-326), containing the C-terminal region. **b**, Protoplasts were co-transfected with OsEBP89-mCherry and Chalk9-GFP constructs, followed by incubation in light for 12 hours prior to confocal microscopy imaging. Bright indicates the bright field image. Scale bars: 5 μm. **c**, Immunoblot analysis of total rice proteins isolated from wild-type Nip and

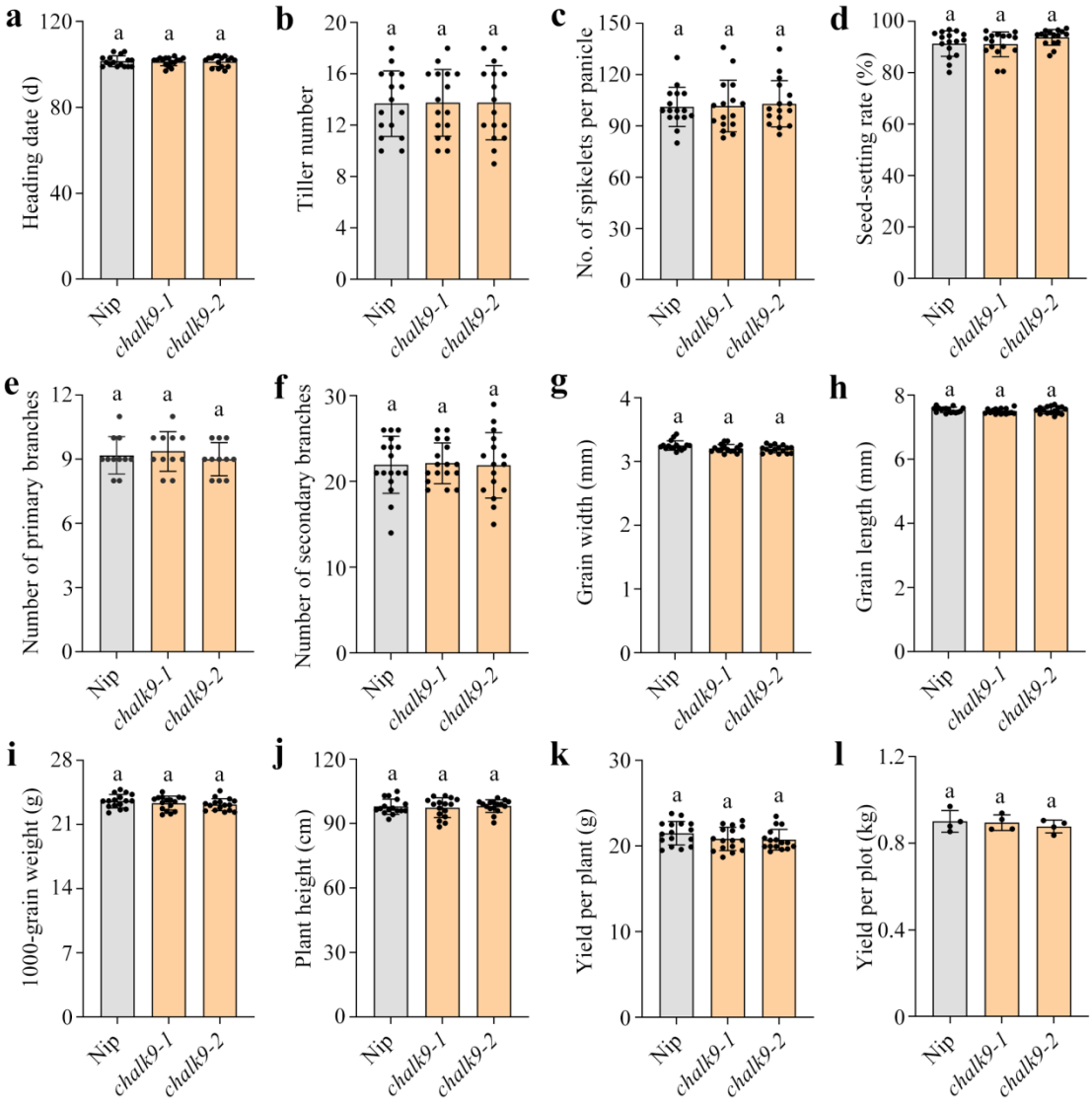
93 *osebp89-1* plants, based on equal fresh weight, using the anti-OsEBP89 antibody. Actin
94 was used as loading control.
95



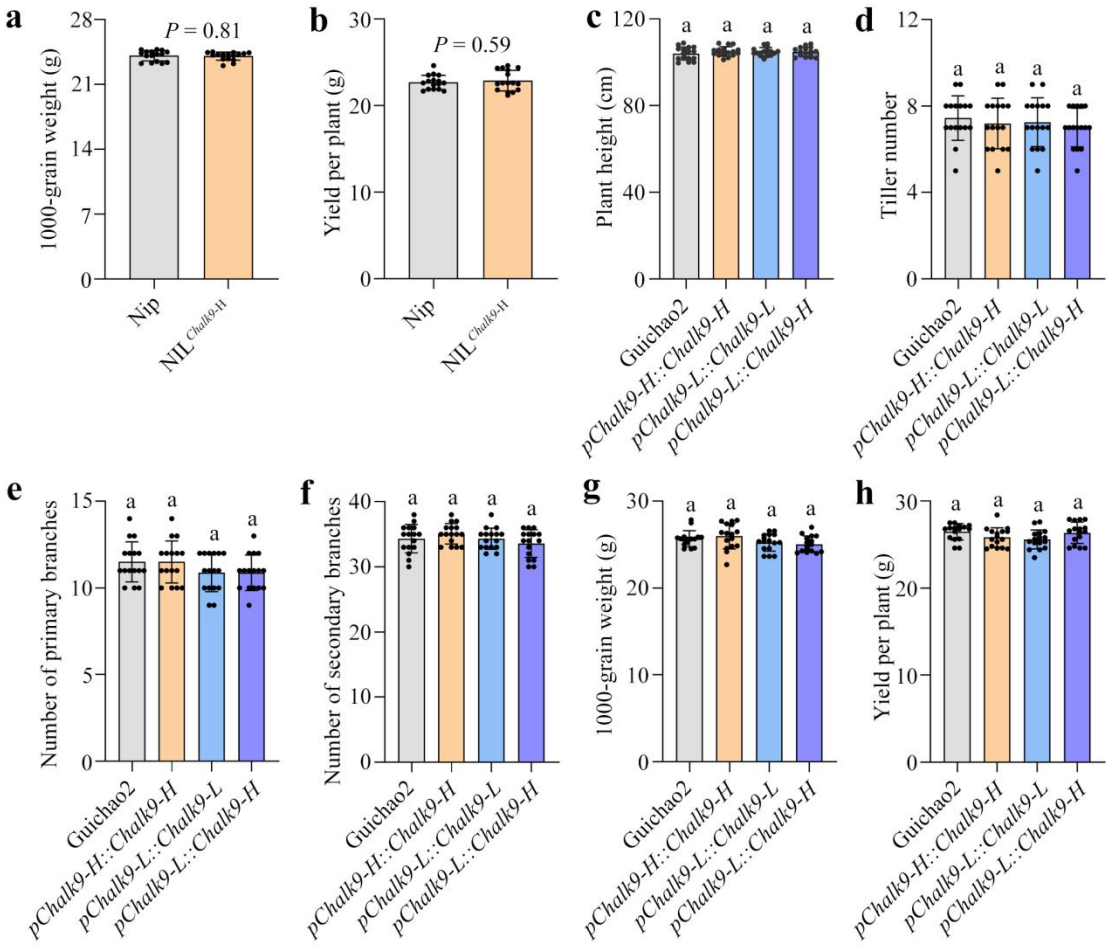
Supplementary Fig. 7 Transcript levels of storage substance-related genes in seeds of Nip and *chalk9-1* plants from RNA-seq data. a, The observation of appearance (upper) and transverse section (down) of mature grains from Nip, *chalk9*-1, and *chalk9-2* plants. Scale bars: 1 mm. **b**, Pearson's correlation coefficients between replicates of Nip and *chalk9-1* plants. **c**, Volcano plots of gene expression levels in *chalk9-1* relative to Nip. Upregulated and downregulated differentially expressed genes (DEGs) based on absolute fold-change ($|\mathrm{Log_2(FC)}| > 1$) and *P*-value ($P < 0.05$) are represented as blue and red dots, respectively. **d**, Heatmap of normalized FPKM values for known starch- and storage protein-related DEGs in the Nip and *chalk9-1* plants.



Supplementary Fig. 8 Transcript levels of storage substance-related genes in the seeds form Nip and *chalk9-1* plants. a-l, Expression analysis of *Wx* (**a**), *OsGluB2* (**b**), *OsGluB1* (**c**), *OsGluC* (**d**), *OsGluB7* (**e**), *OsGluD* (**f**), *PROLM4* (**g**), *PROLM11* (**h**), *PROLM13* (**i**), *PROLM16* (**j**), *PROLM20* (**k**), and *PROLM22* (**l**) in Nip and *chalk9-1*. Total RNA isolated from endosperms of Nip and *chalk9-1* was used for qRT-PCR. Data are means ± SD ($n = 3$ biological replicates). Statistical analysis was performed by two-tailed Student's *t*-test.



Supplementary Fig. 9 Agronomic traits for *chalk9* mutants. a-l, Comparisons of heading date (**a**), tiller number (**b**), No. of spikelet per panicle (**c**), seed-setting rate (**d**), number of primary branch (**e**), number of secondary branch (**f**), grain width (**g**), grain length (**h**), 1000-grain weight (**i**), plant height (**j**), yield per plant (**k**), and yield per plot (**l**) in Nip, *chalk9-1*, and *chalk9-2* plants. In **a**-**k**, Data show means ± SD (n = 16 plants). In **l**, Data show means ± SD (n = 4 plots; plot = 10 plants × 4 rows). Different letters indicate significant differences ($P < 0.05$, one-way ANOVA with Tukey's multiple comparison test); for P values, see Source Data.



Supplementary Fig. 10 Agronomic traits for near-isogenic lines and transgenic plants. a,**b**, Comparisons of 1000-grain weight (**a**) and yield per plant (**b**) between Nip and NIL$^{Chalk9\text{-H}}$ plants. **c**-**h**, Comparisons of plant height (**c**), tiller number (**d**), the number of primary branch (**e**), the number of secondary branch (**f**), 1000-grain weight (**g**), and yield per plant (**h**) in Guichao2 and transgenic plants. Data show means ± SD (n = 16 plants). In **a** and **b**, statistical analysis was performed by two-tailed Student's t-test. In **c**-**h**, different letters indicate significant differences ($P < 0.05$, one-way ANOVA with Tukey's multiple comparison test); for P values, see Source Data.