

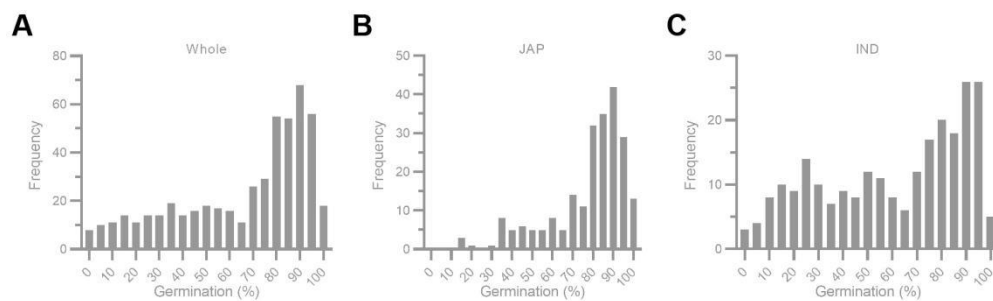
1
2
3
4
5 ***Sdr4* dominates pre-harvest sprouting and facilitates adaptation**
6 **to local climatic condition in Asia cultivated rice**

7
8 Bo Zhao^{1†}, Hui Zhang^{1,2†}, Tianxiao Chen¹, Ling Ding¹, Liying Zhang¹, Xiali Ding¹, Jun
9 Zhang¹, Qian Qian^{1,3*}, Yong Xiang^{1*}

10
11 This file includes Figure S1-S13. Table S1-S9 are merged as a separate excel file and
12 is available at Journal of Integrative Plant Biology.

13
14 **SUPPORTING INFORMATION**

15



16

17

Figure S1. Distribution of germination rate (%) for different rice subpopulations

18

(A–C) Histograms showing the frequency distribution of GR for freshly harvested seeds from the whole population **(A)**, JAP subpopulation **(B)** and IND subpopulation **(C)**, respectively.

19

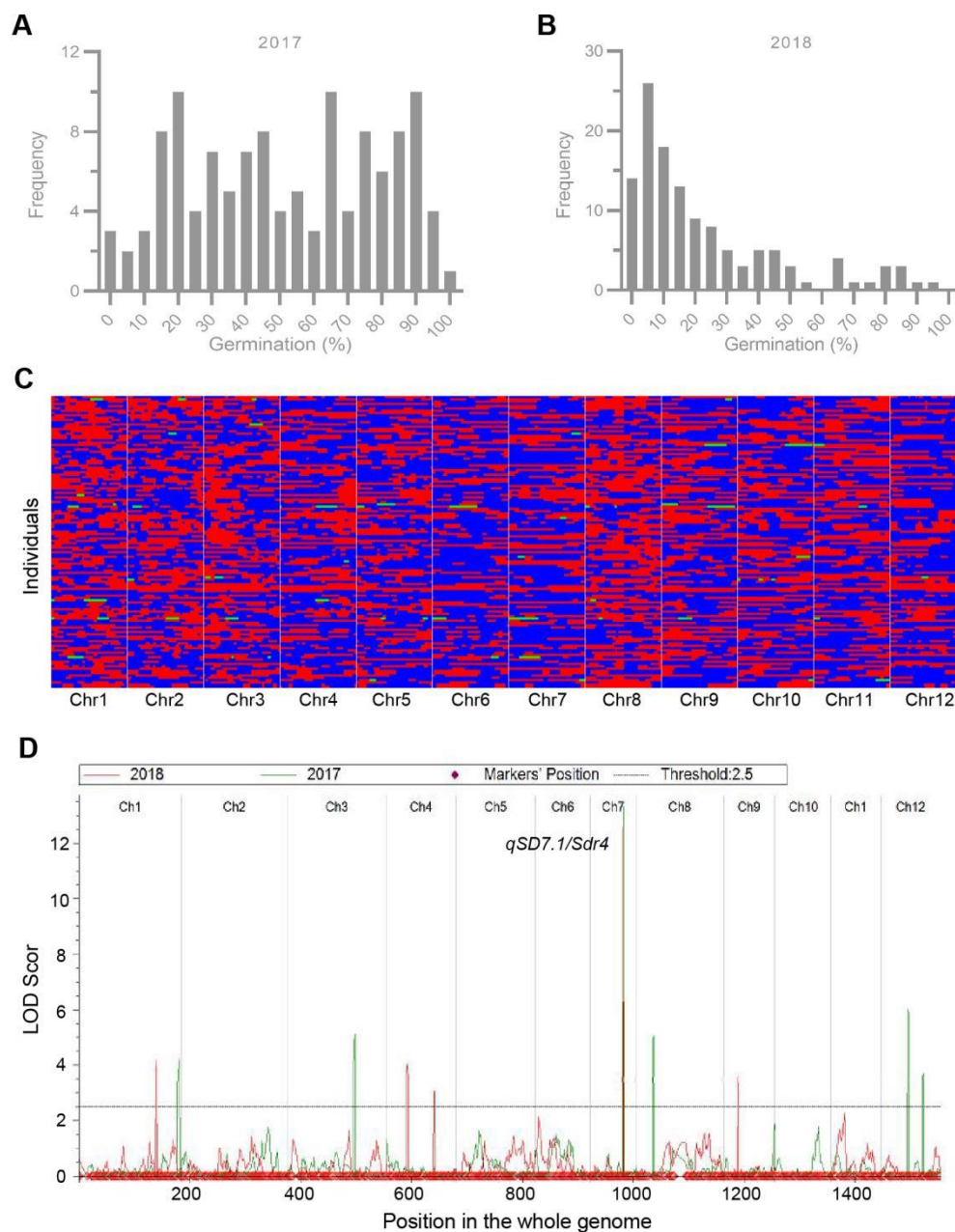


Figure S2. QTL analysis of germination rate (%) of fresh seeds in a RIL population

(A, B) Frequency distribution of germination rate (%) of newly harvested seeds for the RIL population in 2017 **(A)** and 2018 **(B)**. **(C)** Recombination bin map of 134 RILs. Blue represents the PA64s genotype, red represents the 9311 genotype, and green represents the heterozygous genotype. **(D)** QTL analysis of germination rates (%) of newly harvested seeds.

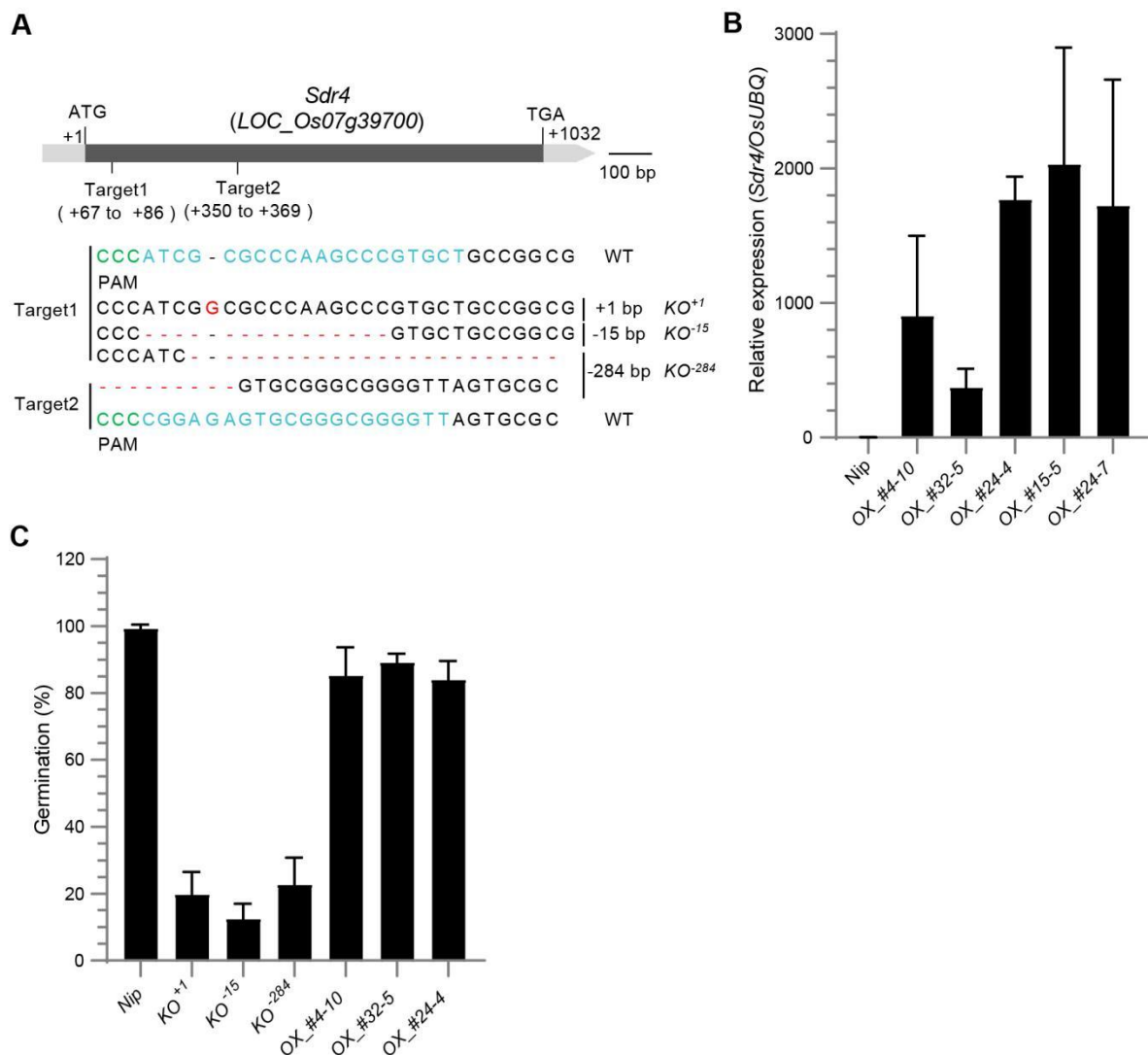


Figure S3. Generation of *Sdr4* mutants and overexpression lines

(A) CRISPR/Cas9-mediated targeted mutagenesis of *Sdr4*. Exons, black box; UTRs, gray boxes. The positions of the two gRNAs are shown below the gene structure. *KO*⁺¹, line with a 1-bp insertion; *KO*⁻¹⁵, line with a 15-bp deletion; *KO*⁻²⁸⁴, line with a 284-bp deletion. **(B)** qRT-PCR analysis of *Sdr4* expression in young leaves of Nipponbare (Nip) and T₂ *Sdr4* overexpression lines (*Sdr4*-OX). Rice *UBIQUITIN* served as an internal control. Data are presented as means $\pm$ SD ($n = 2-3$). **(C)** Germination rates (%) of wild-type (Nipponbare (Nip)), *sdr4*-KO and *Sdr4*-OX seeds after 1 month of storage at room temperature in Shenzhen. Data are presented as means $\pm$ SD ($n = 3-5$).

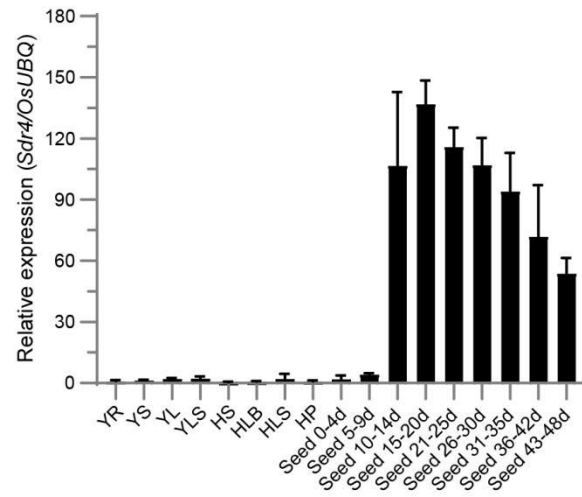


Figure S4. Relative expression of *Sdr4* in various wild-type (Nipponbare) tissues examined by qRT-PCR

Rice *UBIQUITIN* was used as an internal control and young seedling root was used as a reference sample. Data are presented as means $\pm$ SD ($n = 3$). YR, young seedling root; YC, young seedling stem; YL, young seedling leaf; YLS, young seedling leaf sheath; HS, heading stage stem; HLB, heading stage leaf blade; HLS, heading stage leaf sheath; HP, heading stage panicle, Seed 0–4d, 0–4 days-after-pollination (DAP) seeds.

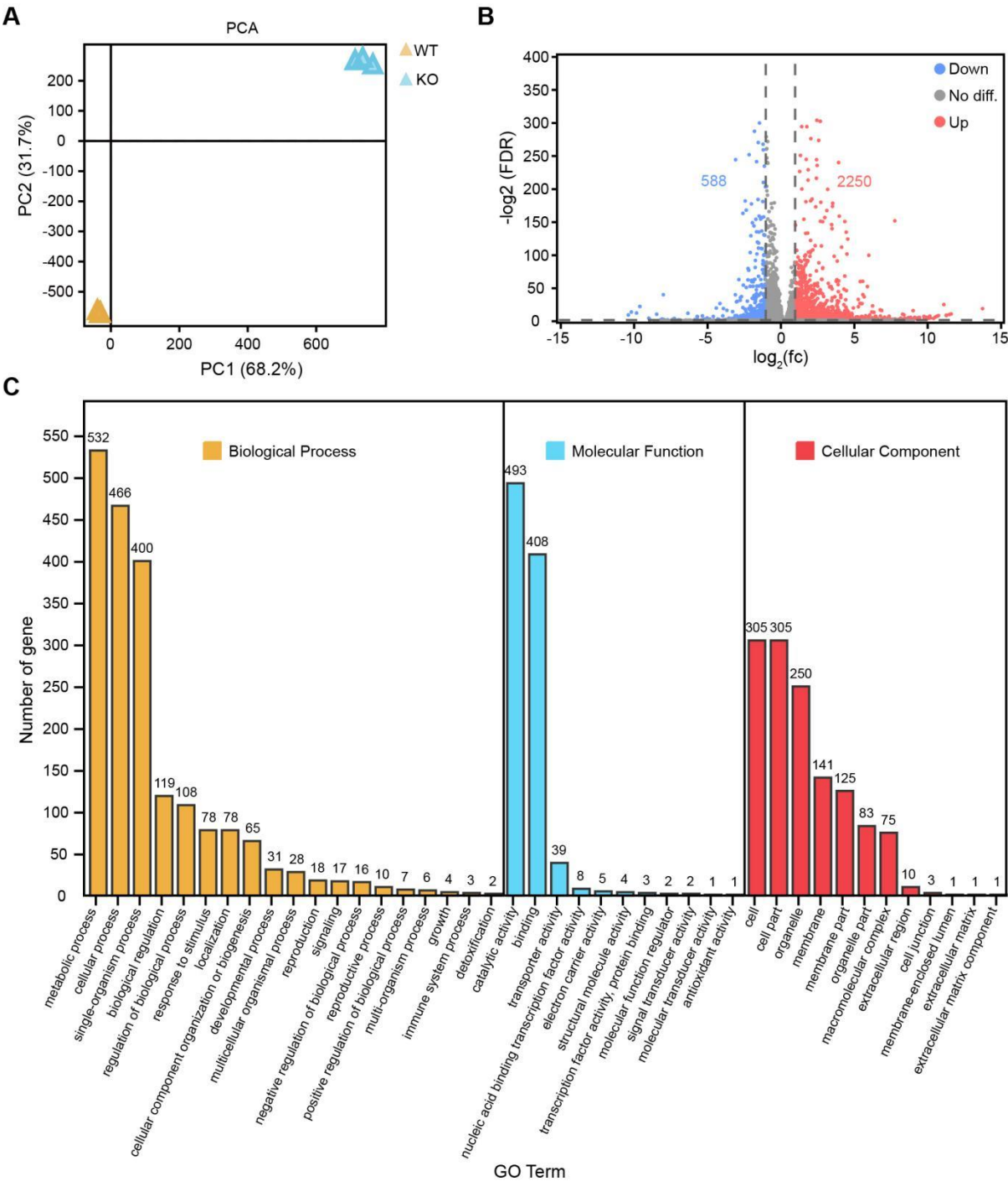
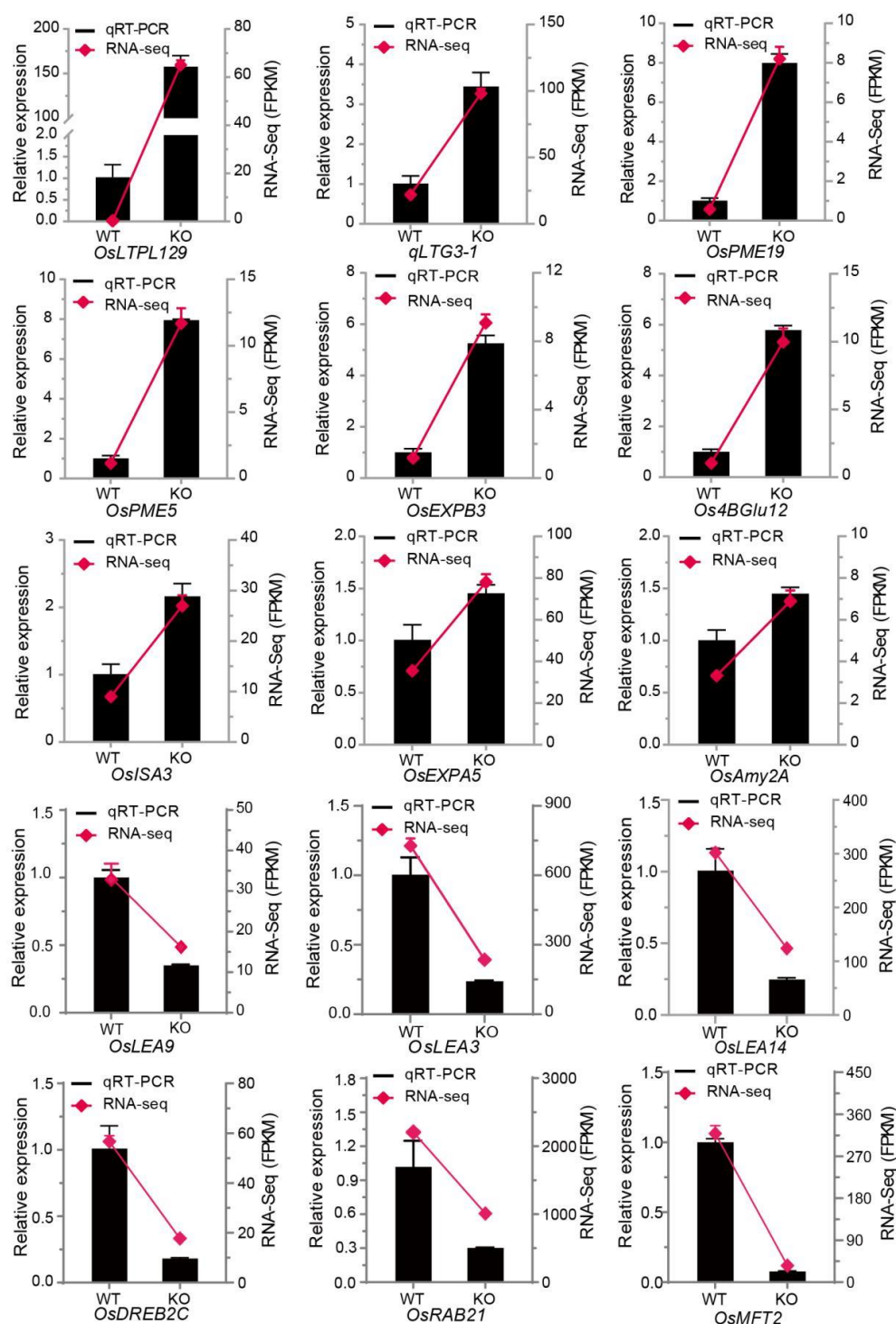


Figure S5. Differentially expressed genes (DEGs) analysis in 21-DAP seeds of the *sdr4-KO* mutant and wild type

(A) PCA (principal component analysis) of DEGs. **(B)** Volcano map of DEGs. **(C)** Gene Ontology analysis of DEGs.



58
 59 **Figure S6.** qRT-PCR verification of DEGs between the *sdr4*-KO mutant and wild type in RNA-
 60 seq
 61 Data are presented as means $\pm$ SD (n = 3). Rice *UBIQUITIN* was used as an internal control.

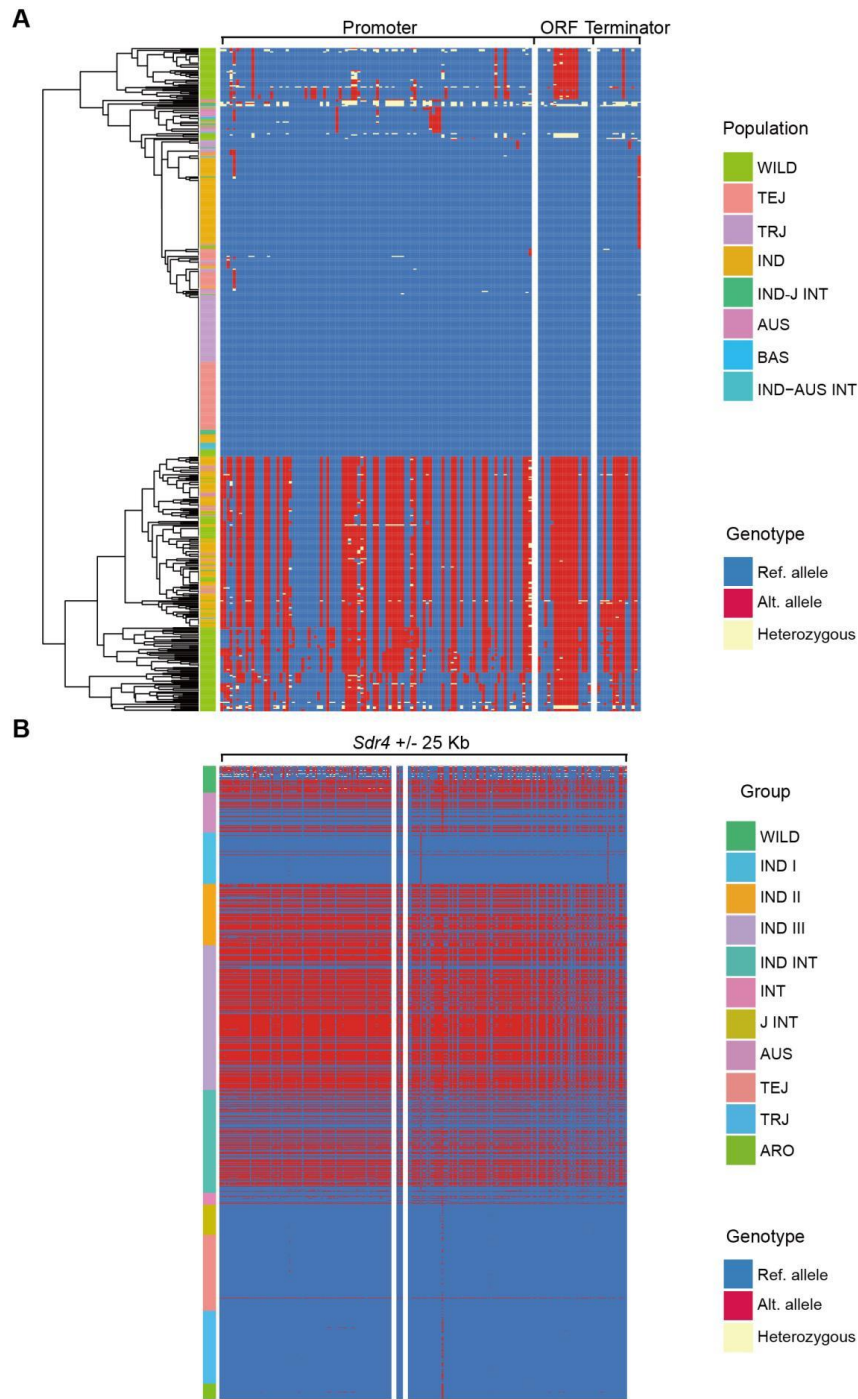
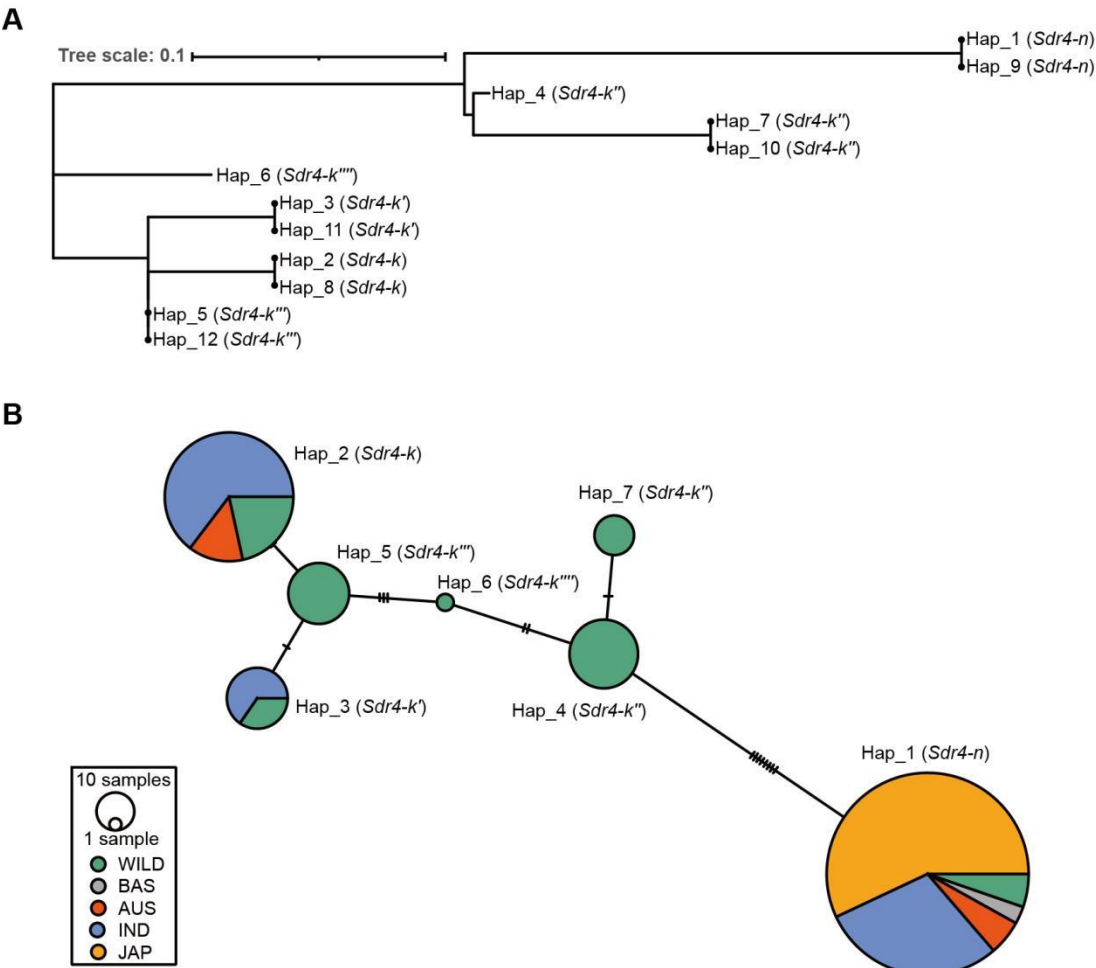


Figure S7. Heatmaps of nucleotide polymorphisms across the regions harboring *Sdr4*

(A) Heatmap of nucleotide polymorphisms and phylogenetic tree of 404 QTN Library and 168 WILD accessions basing on the *Sdr4* genomic region including the 3-kb promoter, 1032-bp ORF, and 2-kb 3'UTR. **(B)** Heat maps of nucleotide polymorphisms of *Sdr4* +/- 25 kb region with 4726 RiceVarMap and 168 WILD accessions. The region between the two white solid lines in **(A, B)** indicates the *Sdr4* ORF region.



70

71

72

73

74

75

76

Figure S8. Phylogenetic relationships among different *Sdr4* haplotypes based on the variations in the *Sdr4* genome sequence including the 3-kb promoter, 1032-bp ORF, and 2-kb 3'UTR
(A) Phylogenetic tree of different *Sdr4* haplotypes constructed using the neighbor-joining method with MEGA X. **(B)** Minimum-spanning tree constructed basing on different *Sdr4* haplotype sequences using PopART software. Each haplotype group is represented by a circle; circle size represents the number of samples within the haplotype.

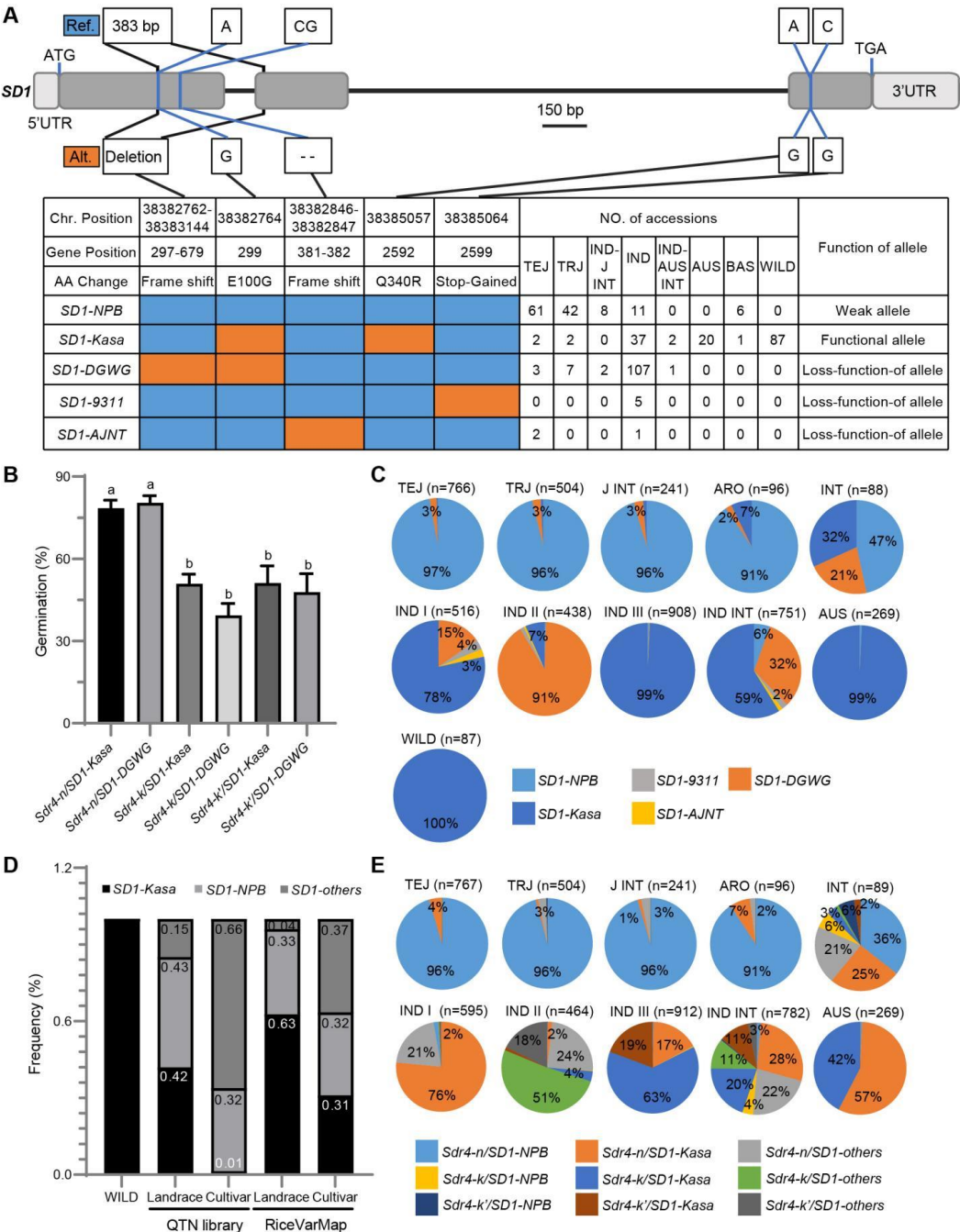


Figure S9. Haplotype analysis of *SD1* across the open reading frame (ORF) regions in cultivated and WILD accessions

(A) Five haplotypes identified based on 13 sequence variations among 313 cultivated (QTN library) accessions and 87 wild (WILD) accessions. Dark blue indicates the reference (Nipponbare) allele sequence; orange indicates the alternative sequence. **(B)** Comparison in seed dormancy of different combinations of haplotypes between *SD1* and *Sdr4* in cultivated

rice. Different letters above the columns indicate significant differences ($P < 0.05$) in a pairwise comparison using Tukey's multiple-comparison test. Data are presented as means $\pm$ SEM. **(C)** Frequencies of different *SD1* haplotypes among ten subgroups of the 4577 RiceVarMap cultivated accessions and 87 WILD accessions. **(D)** Proportions of three haplotypes (*SD1-Kasa*, *SD1-NPB*, *SD1-others*) among WILD accessions, landraces, and cultivars from the QTN library and RiceVarMap accession panels. **(E)** Frequencies of different combinations of haplotypes between *SD1* and *Sdr4* in cultivated rice among ten subgroups of the 4577 RiceVarMap cultivated accessions.

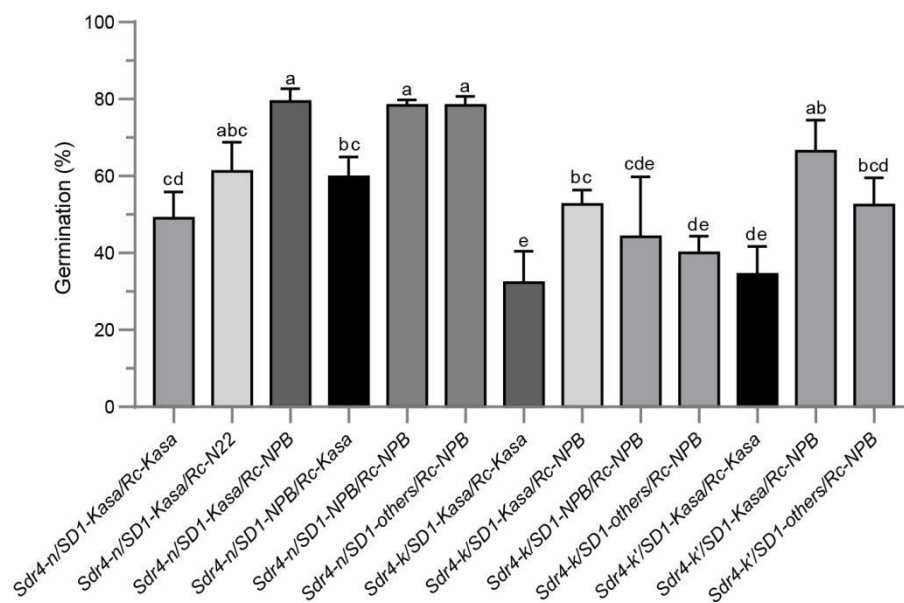


Figure S10. Joint haplotype analysis of *Sdr4*, *SD1* and *Rc* genes in the GWAS germplasm panel.

Different letters above the columns indicate significant differences ($P < 0.05$) in multiple comparisons using Fisher's LSD test. Data are presented as means $\pm$ SEM.

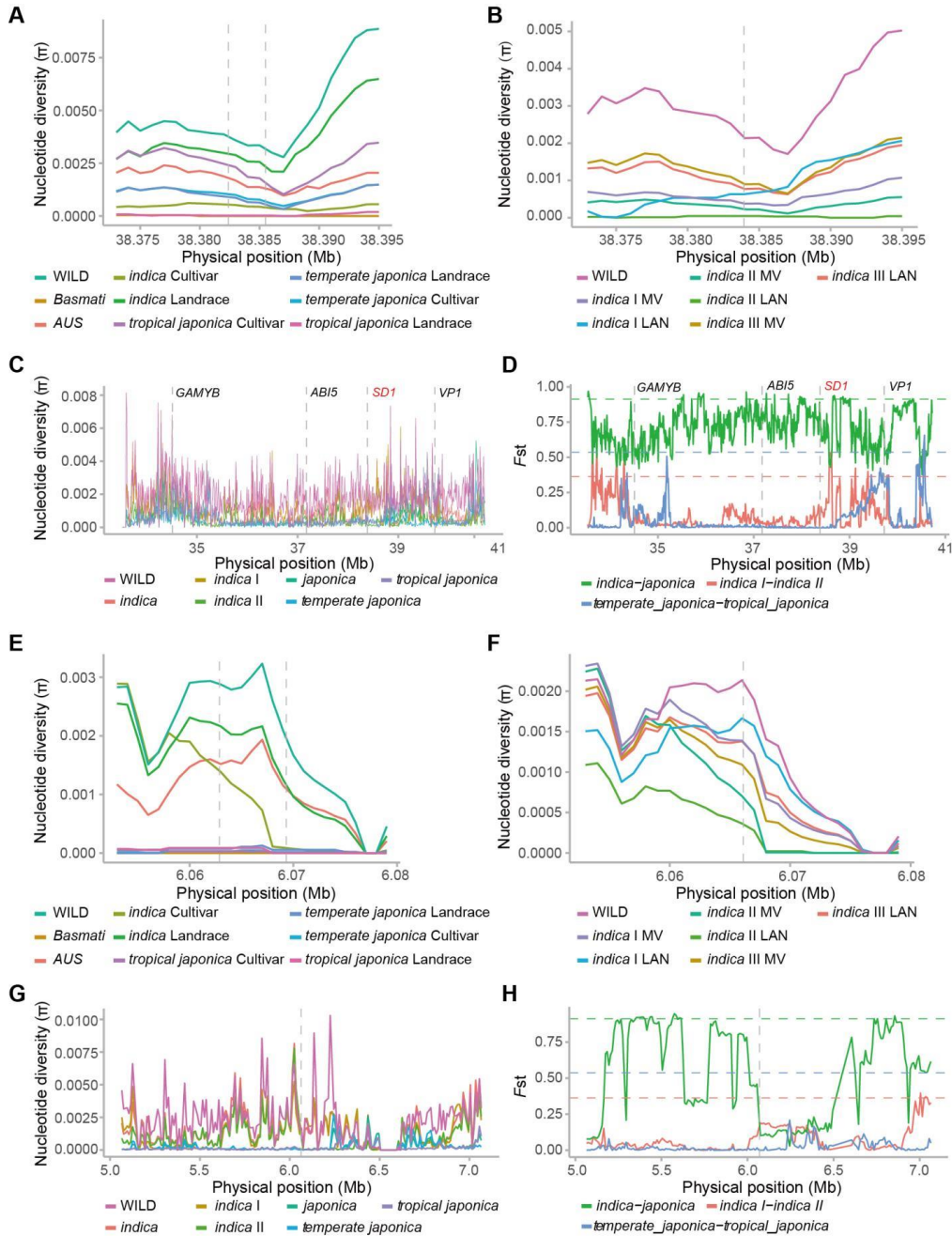


Figure S11. Distribution of nucleotide diversity (π) and population differentiation index (F_{st}) across the regions harboring *SD1* and *Rc*

(A, E) Nucleotide diversity of different subgroups in the 404 QTN library accessions and 168 WILD accessions across the *SD1* (A) and *Rc* (E) ORFs (the regions between two dashed lines) ± 10 kb regions. (B, F) Nucleotide diversity of three *indica* subgroups in the 3776 RiceVarMap accessions and 168 WILD accessions across the *SD1* (B) and *Rc* (F) ORFs ± 10 kb regions. (C, G) Nucleotide diversity of different subgroups in the 4726 RiceVarMap accessions and 168 wild

107 accessions across the *SD1* (**C**) and *Rc* (**G**) regions. (**D**, **H**) Population differentiation (F_{st}) across
108 the *SD1* (**D**) and *Rc* (**H**) regions between the different subgroups. The dashed horizontal lines in
109 different colors represent the threshold (95% quantile of the genome-wide value) for the
110 corresponding comparative groups. The dashed vertical lines in (**A–H**) represent the position of
111 *SD1* (**A–D**) and *Rc* (**E–H**) ORF regions.

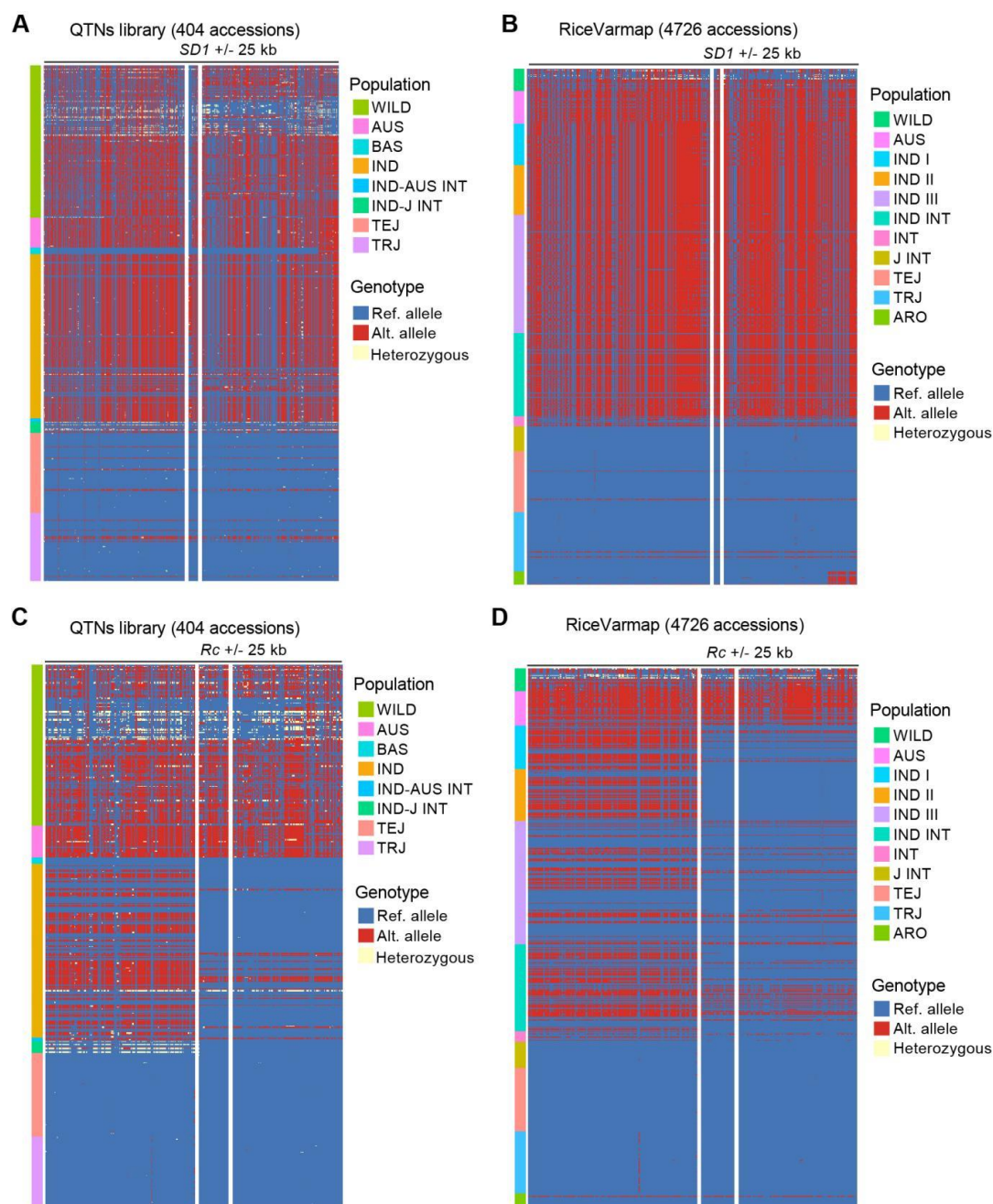


Figure S12. Heatmaps of nucleotide polymorphisms across the regions harboring *SD1* and *Rc* (A, C) Heatmap of nucleotide polymorphisms of 404 QTN Library and 168 WILD accessions across the *SD1* $\pm$ 25 kb region (A) and *Rc* $\pm$ 25 kb region (C). (B, D) Heat maps of nucleotide polymorphisms of 4726 RiceVarMap and 168 WILD accessions across the *SD1* $\pm$ 25 kb region (B) and *Rc* $\pm$ 25 kb (D). The region between the two white solid lines in (A–D) represent the position of *SD1* (A, B) and *Rc* (C, D) ORF regions.

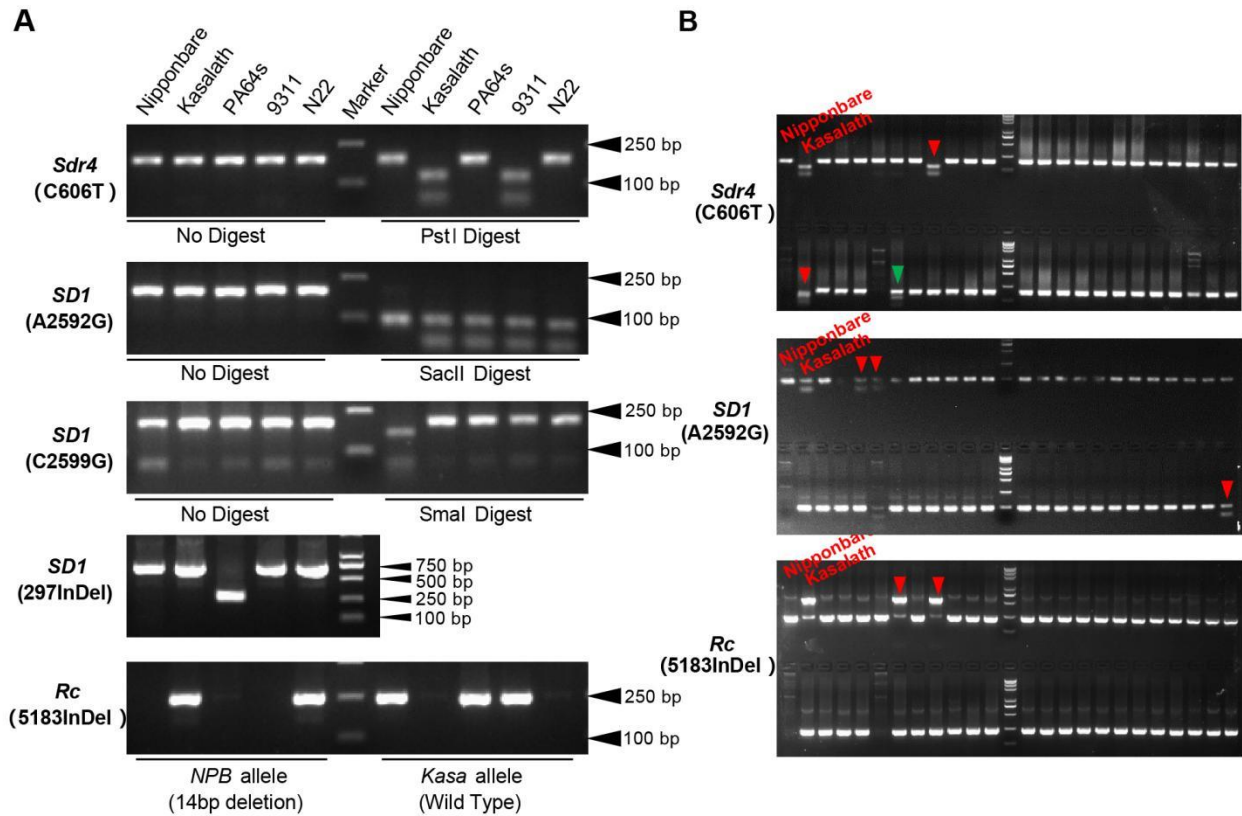


Figure S13. Functional marker development based on the variations in the exons of *Sdr4*, *SD1* and *Rc* alleles

(A) Five accessions, Nipponbare (JAP accession), Kasalath (AUS accession), 9311 (IND accession), PA64s (intermediate accession) and N22 (AUS accession), were used to test the polymorphisms of the functional markers for the *Sdr4*, *SD1*, and *Rc* loci. **(B)** The genotypes of the *Sdr4*, *SD1*, and *Rc* loci of 44 CSSLs derived from a cross between Nipponbare and Kasalath were determined using the corresponding functional markers. The red arrows indicate the genotypes of the introgression lines from the donor parent, Kasalath, for *Sdr4*, *SD1*, and *Rc*. The green arrow indicates the heterozygous genotype.

Table S1. Samples information, haplotypes of *Sdr4*, *SD1* and *Rc* loci and GR phenotype for 499 GWAS accessions from 3K Project

Table S2. Samples information and haplotypes of *Sdr4*, *SD1* and *Rc* loci for 4726 RiceVarMap accessions

Table S3. Germination rate (%) of fresh seeds in the whole population and four subpopulations

Table S4. Germination rate (%) of the freshly harvested seeds in the ten subgroups of cultivated rice

139 **Table S5.** QTL for seed dormancy detected in RILs in 2017 and 2018 in Shenzhen (2017_SZ
140 and 2018_SZ)
141 **Table S6.** List of differentially expressed genes between sdr4-KO mutant and WT
142 **Table S7.** Samples information and haplotypes of Sdr4, SD1 and Rc loci for 404 QTN library
143 accessions and 168 WILD accessions
144 **Table S8.** GR trait and haplotypes of *Sdr4*, *SD1* and *Rc* loci for 46 CSSLs of NK set
145 **Table S9.** List of primers used in this study