

Description of Additional Supplementary Files

File Name: Supplementary Data 1

Description: Chalky grain rate and degree of chalkiness of 175 *indica* accessions in two years.

File Name: Supplementary Data 2

Description: Annotation of significant SNPs associated with grain chalkiness in the candidate region. *P* values were determined using a two sided Wald test and assessed after Bonferroni correction for multiple comparisons.

File Name: Supplementary Data 3

Description: Significant SNPs located in coding or regulatory regions.

File Name: Supplementary Data 4

Description: Annotation of candidate genes on Chromosome 9 associated with the grain chalkiness by MSU Rice Genome Annotation Project Release 7.

File Name: Supplementary Data 5

Description: Missense SNPs were identified in only six genes. *P* values were determined using a two-sided Wald test and assessed after Bonferroni correction for multiple comparisons.

File Name: Supplementary Data 6

Description: All variations of *Chalk9* in 149 *indica* accessions were identified by re-sequencing based on PCR amplification. v1 to v28 indicates the name of the variations. 0, no insertion. NA, not detected.

File Name: Supplementary Data 7

Description: Major haplotypes of *Chalk9* were identified from significant variations in *indica* accessions. v4, v5, v10, v12, v14, and v15 indicate the significant variants.

File Name: Supplementary Data 8

Description: Prediction of transcription factors binding to the 64-bp sequence in *Chalk9*.

File Name: Supplementary Data 9

Description: 2,658 differentially expressed genes generated from *chalk9-1* vs Nip. FDR, false discovery rate.

File Name: Supplementary Data 10

Description: Expression analysis of starch and SSP genes generated from *chalk9-1* vs Nip by RNA-seq.

File Name: Supplementary Data 11

Description: Putative cis-regulatory elements identified in the promoters of genes involved in starch and storage protein biosynthesis.

File Name: Supplementary Data 12

Description: The SNPs at the *OsEBP89* gene region define one major haplotype in indica varieties. The position corresponds to the coordinates in the chromosome 3 of Nipponbare genome.

File Name: Supplementary Data 13

Description: The SNPs at the *OsEBP89* gene region from sequencing data of 4,726 rice accessions. The position corresponds to the coordinates in the chromosome 3 of Nipponbare genome. The 34 SNPs with MAF greater than or equal to 1% define 4 haplotypes.

File Name: Supplementary Data 14

Description: The SNPs at the *Chalk9* promoter region in 3K Rice Genomes Project. The position corresponds to the coordinates on chromosome 9 of the Nipponbare genome. The 37 SNPs with MAF greater than or equal to 1% define 9 haplotypes. The SNPs at position 19521397 (marked in red), which are highly correlated with the causal indel, were used to assign the haplotypes to the Haplotype-H group (shaded in red) and Haplotype-L group (shaded in green).

File Name: Supplementary Data 15

Description: Primers (5'-3') used in this study.

File Name: Supplementary Data 16

Description: Genomic sequences of *Chalk9* from wild and cultivated rice.

File Name: Supplementary Data 17

Description: Genomic variations of *OsEBP89* in cultivated rice varieties.