PNAS
www.pnas.org

Supplementary Information for

Ef-cd locus shortens rice maturity duration without yield penalty

Jun Fang, Fantao Zhang, Hongru Wang, Wei Wang, Fei Zhao, Zijuan Li, Changhui Sun, Faming Chen, Fan Xu, Shuoqi Chang, Liang Wu, Qingyun Bu, Pingrong Wang, Jiankun Xie, Fan Chen, Xuehui Huang, Yijing Zhang, Xinguang Zhu, Bin Han, Xiaojian Deng, and Chengcai Chu

Jun Fang and Fantao Zhang contributed equally to this work.

Corresponding author: Chengcai Chu, Xiaojian Deng, and Jun Fang

Email: ccchu@genetics.ac.cn, xjdeng@sicau.edu.cn, or fangjun@iga.ac.cn.

This PDF file includes:

Supplementary text

Figs. S1 to S11

www.pnas.org/cgi/doi/10.1073/pnas.1815030116

Supplementary Text

SI Materials and Methods

Material preparation and growth conditions

During all stages of developing early-maturing NILs using six late-maturing recurrent parents and early-maturing donor line 6442S-7 (*SI Appendix*, Fig. S1*A*), the genotypes of offspring plants were analyzed for *Ef-cd* locus using markers RM231, RM22, C515, and InDel1 (1, Dataset S4), and the homozygotes were selected for continuous backcrosses and self-pollinations until the NILs of *Ef-cd* became genetically stable. Two recurrent parents of SH881 and MH63, and their two early-maturing NILs D248 and D330 were selected for phenotyping the major agronomic traits in 2009 of Beijing. Two T-DNA insertion mutant lines of *Ef-cd* (*ef-cd-1*: PFG_1D-05526.R and *ef-cd-2*: PFG_2A-10486.R) and one T-DNA insertion mutant line of *OsSOC1* (PFG_3A-05349.L) were ordered from the Postech rice mutant database (http://signal.salk.edu/cgi-bin/RiceGE) (2, 3).

Large-scale field tests for the *Ef-cd* NIL hybrids E-SY63 (ZS97A/D330), E-II-You838 (E-II-32A/FH838) and their corresponding hybrids Shanyou63 (ZS97A/MH63) and II-You838 (II-32A/FH838) were performed in a rice paddy with a randomized complete block design during the regular rice cultivation season in 2015 at four experimental stations: the Institute of Genetics and Developmental Biology, Chinese Academy of Sciences (39^{o}54’ N, Beijing City), the Jiaxing Academy of Agricultural Sciences (30^{o}75’ N, Jiaxing, Zhejiang Province), the Sichuan Agricultural University (30^{o}42’ N, Chengdu, Sichuan Province), and the Fujian Academy of Agricultural Sciences (26^{o}08’ N, Fuzhou, Fujian Province). The plant spacing was 26 cm (row space) and 17 cm (plot space). Field management followed the normal agricultural practices for applied fertilizer (per hectare) as follows: 50 kg N, 60 kg P, and 95 kg K as basic fertilizer; 90 kg N at the tilling stage; and 30 kg N at the heading stage. Three replicates were used for plot yield assays. For short-term hydroponic culture, rice seedlings were grown in modified Kimura B solution in a growth chamber with a 12-h-light (30°C)/12-h-dark (28°C) photoperiod, ~200 μmol m^{-2} s^{-2} photon density, and ~70% humidity.

The data of 1,439 elite hybrid rice cultivars were obtained from the National Center for Gene Research, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (4).

Fine mapping of *Ef-cd*

NIL (*Ef-cd*) D248 was crossed with SH881, and the resultant F_1 plants were self-pollinated to result in F_2 population in Lingshui (18°22' N, Hainan Province) in the winter of 2007. 4,800 late-maturing individuals were strictly selected from the F_2 population for fine mapping of *Ef-cd.* Molecular markers were developed from the sequences of the *Ef-cd*-flanking regions of SH881 and 6442S-7 genomic DNA for the mapping, which were provided in Dataset S4.

Vector construction and plant transformation

To generate the specific *Ef-cd* RNAi vector, a 168-bp fragment (not overlapped with *Os03g0122600*) of the 5' region of *Os03g0122500* was amplified from the full-length cDNA by using the primers listed in Dataset S4 and sequentially cloned into the pUCRNAi vector with *Xho*I/*Bgl*II and *Sal*I/*Bam*HI sites. Finally, the full stem-loop fragment was cloned into pCAMBIA2301. The resultant construct was transformed into Nipponbare (*Oryza sativa* spp. *japonica*) using the *Agrobacterium*-mediated transformation method (5).

The *Os03g0122500-Cas9* mutant plants were generated using CRISPR/Cas9 by the Biogle company (Hangzhou, China). We designed two sgRNAs targeting *Os03g0122500* at different locations: sg1213: 5'-GTCGATTCGGTGCGTGATAA-3', and sg1214: 5'-GTCATATACAACAGGGGTAC-3'. The two sgRNAs were created in the BGK032-DSG vector containing Cas9, which was introduced into *A. tumefaciens* strain EHA105 and transformed into Nipponbare. To examine the function of CRISPR/Cas9 *in vivo*, genomic DNA was extracted from the transformants and primer pairs flanking the designed target site were used for PCR amplification.

RNA isolation and quantitative RT-PCR

Total RNA was extracted from the fresh leaf blades of 40 d seedlings using TRIzol reagent (Invitrogen, USA). One microgram of total RNA was used to synthesize first-stand cDNA with the ReverTra Ace qPCR RT Master Mix (Toyobo, Japan),

according to the manufacturer's instructions. PCR amplification was performed with the SYBR Green Real-Time PCR Master Mix reagent (Toyobo, Japan) using a real-time PCR detection system, according to the manufacturer's instructions (CFX96, Bio-Rad, USA). The rice *OsActin1* gene was used as the internal control, and all primers used for the analysis of the flowering genes are listed in Dataset S4.

3' rapid amplification of cDNA ends (RACE)

3' RACE was used to confirm the polyadenylation sites and alternative splicing events in this study. Total RNA was extracted from the fresh leaf blades of 40 d seedlings using TRIzol reagent (Invitrogen, USA). First-strand cDNA synthesis was performed using FirstChoice® RLM-RACE Kit (Ambion, Cat. no. AM1700) according to the manufacturer's manual. And then, two rounds of PCR were performed. For the first round, the 3'-RACE outer primer and gene-specific primer 1/3 were used; and then, the 3'-RACE inner primer and gene-specific primer 2/4 were used for the second round PCR. Relevant primers are listed in Dataset S4.

LUC assays

The 2.0-kb promoter region of *Ef-cd* in D248 and SH881 was amplified and inserted into the destination vector pGWB535 to generate the fusion construct with LUC by using an in-fusion kit (Clontech, USA). Luciferase activity was measured using the Dual-Luciferase reporter kit (Promega, USA) on the TD-20/20 Luminometer (Turner Designs, USA).

Labeling with 15N-nitrate or 15N-ammonium for determination of 15N accumulation

The ^{15}N accumulation experiments were performed according to methods described previously (6). ^{15}N-nitrate or ^{15}N-ammonium labeling was performed with ^{15}N-labeled KNO_3 (98 atom % ^{15}N, Sigma-Aldrich) or ^{15}N-labeled NH_4Cl (98 atom % ^{15}N, Sigma-Aldrich), respectively. For ^{15}N-nitrate accumulation experiment, rice seedlings were cultured in the Kimura B solution for 10 days. Next, the seedlings were pretreated with the Kimura B solution for 2 hours and then transferred to modified Kimura B solution containing 5 mM ^{15}N-KNO_3 for 3 hours. At the end of labeling, shoots were collected and dried at 70°C. Finally, the samples were ground and the ^{15}N content was evaluated by an isotope ratio mass spectrometer with an elemental

analyzer (Thermo Finnigan Delta Plus XP; Flash EA 1112). For ^{15}N-ammonium accumulation experiment, the treatment was conducted as above except that 5 mM ^{15}N-KNO_3 was replaced with 1 mM ^{15}N-NH_4Cl and 1 mM KNO_3.

Photosynthesis related parameter measurements

The maximum light-saturated rate of leaf photosynthetic CO_2 uptake rate (A_{sat}) was measured with a portable infrared gas analyzer system LI-6400XT (LI-COR, Lincoln, NE, USA) for flag leaves at the grain filling stage. The photosynthetic photon flux density (PPFD) used was 1,600 μmol photo $m^{-2}s^{-1}$ and were measured after leaves were acclimated in the leaf chamber for about 5 min. We used 5 replicates. The leaf temperatures and ambient CO_2 concentration used during the measurement were 25~28^{o}C and 385 ppm respectively. The chlorophyll content (SPAD value) was determined according to methods described previously (7).

Chromatin immunoprecipitation (ChIP) assay

ChIP assays were performed as previously described (8). Briefly, fresh leaf blades of 40 d seedlings were harvested for fixation. Chromatin was isolated and sonicated to generate DNA fragment with an average size of about 500 bp. The solubilized chromatin was immunoprecipitated by Protein A beads (Millipore) with anti-H3K36me3 antibody (Abcam, ab9050), and the co-immunoprecipitated DNA was recovered and analyzed by ChIP-seq and quantitative PCR. More than 10 ng ChIP DNA each sample was used for Illumina library generation following the manufacturer's instructions (Illumina, http://www.illumina.com/). Library construction and deep sequencing were performed by Genergy Biotechnology Co. Ltd. (Shanghai, China) using Illumina HiSeq 2000 following the manufacturer's instructions (Illumina). Relative enrichment was calculated by normalizing the amount of a target DNA fragment against the respective input DNA samples.

Strand-specific RNA sequencing and ChIP-seq data analysis

Raw sequencing reads were trimmed by quality using Trimmomatic (9) and the cleaned reads were mapped to the *Os*-Nipponbare-Reference-IRGSP-1.0 (10) using BWA 0.7.12-r1039 (11) for DNA sequencing and STAR 2.5.3.a (12) for strand-specific RNA sequencing (ssRNA-seq), both with default settings. The returned alignments were stringently filtered so as to remove duplicates, ambiguously mapped

reads and read pairs with conflicting alignments. We used StringTie (13) to predict transcript structures based on the mapped reads in different samples. Integrative Genomics Viewer (IGV) (14) was used for illustrating the genomic tracks and Sashimi plots, and we use Counts Per Million (CPM) mapped reads to normalize the number of reads per bin.

Analysis of differentially expressed genes

For differentially expressed genes (DEGs) analysis, ssRNA-seq reads were aligned to rice reference genome IRGSP-1.0 by STAR (12) with "--quantMode GeneCounts" option to report the read counts per gene. Then DEGs were determined using EBSeq (15) with the read counts. Genes with the adjusted *p*-value below 0.05 were considered as DEGs. GO enrichment analysis was used 'agriGO' online toolkit (http://systemsbiology.cau.edu.cn/agriGOv2/index.php) (16).

Accession numbers

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (17) and are accessible through GEO Series accession number GSE133746 (strand-specific RNA-seq data) and GSE134021 (ChIP-seq data).

References

1. X. J. Deng *et al*., Identification and gene mapping of completely dominant earliness in rice. *Agric. Sci. China* **1**, 11–18 (2002).
2. J. S. Jeon *et al*., T-DNA insertional mutagenesis for functional genomics in rice. *Plant J.* **22**, 561–570 (2000).
3. D. H. Jeong *et al*., T-DNA insertional mutagenesis for activation tagging in rice. *Plant Physiol.* **130**, 1636–1644 (2002).
4. X. H. Huang *et al*., Genomic analysis of hybrid rice varieties reveals numerous superior alleles that contribute to heterosis. *Nat. Commun.* **6**, 6258 (2015).
5. C. H. Sun *et al*., The histone methyltransferase SDG724 mediates H3K36me2/3 deposition at *MADS50* and *RFT1* and promotes flowering in rice. *Plant Cell* **24**, 3235–3247 (2012).
6. W. Wang *et al*., Expression of the nitrate transporter gene *OsNRT1.1A*/*OsNPF6.3*

confers high yield and early maturation in rice. *Plant Cell* **30**, 638–651 (2018).

7. F. M. Gao *et al*., Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross Zhou 8425B/Chinese spring. *Front Plant Sci.* **6**, 1099 (2015).
8. H. Wang *et al*., *Arabidopsis* flower and embryo developmental genes are repressed in seedlings by different combinations of polycomb group proteins in association with distinct sets of *cis*-regulatory elements. *PLoS Genet.* **12**, e1005771 (2016).
9. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
10. H. Sakai *et al*., Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* **54**, e6 (2013).
11. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
12. A. Dobin *et al*., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
13. M. Pertea *et al*., StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
14. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
15. N. Leng *et al*., EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**, 1035–1043 (2013).
16. T. Tian *et al*., agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**, W122–W129 (2017).
17. R. Edgar, M. Domrachev, A. E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).

A

Recurrent parents	NILs	Combination	Generation
Shuhui881	D248	6442S-7/SH881	BC_5F_{11}
Minghui63	D330	6442S-7/MH63	BC_9F_{12}
Shuhui527	D488	6442S-7/SH527	BC_9F_{12}
Mianhui725	D439	6442S-7/MH725	$BC_{13}F_{10}$
Yihui1577	D374	6442S-7/YH1577	$BC_{10}F_9$
II-32B	B3115	6442S-7/II-32B	BC_6F_{14}

B

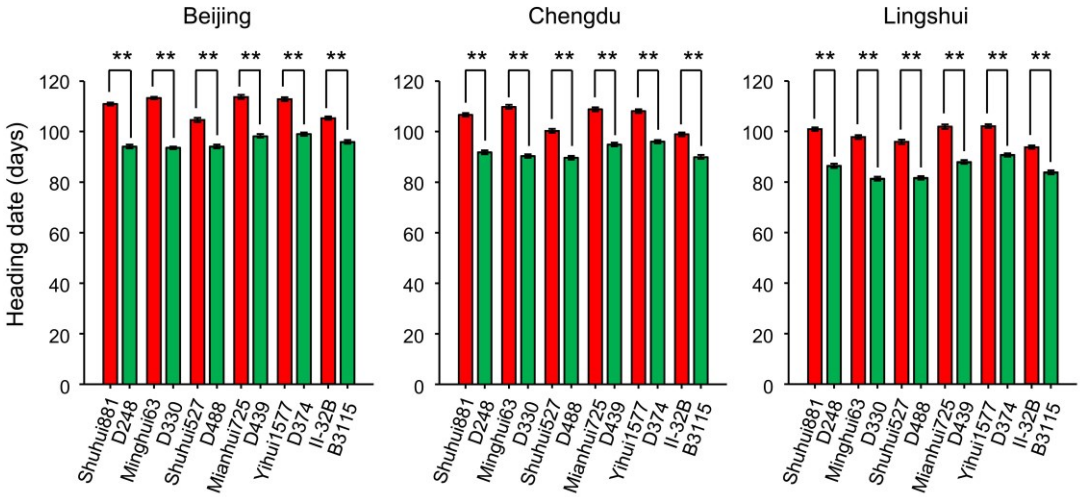

Fig. S1. Near-isogenic lines (NILs) and comparison of heading days with their corresponding recurrent parents. (*A*) Origin and generation of the NILs. (*B*) Comparison of heading days in pairs under three different locations: Beijing (39°54' N, Beijing City, sowing date was May 4, 2009), Chengdu (30°42' N, Sichuan Province, sowing date was May 8, 2009) and Lingshui (18°22' N, Hainan Province, sowing date was Nov 15, 2009). The *p* value was calculated using Student's *t*-test. ** for significant at $p < 0.01$.

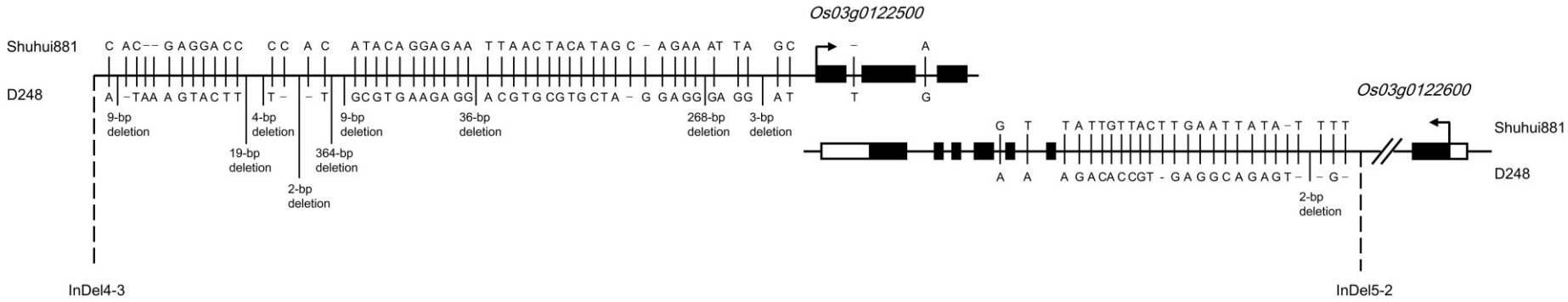

Fig. S2. Sequence polymorphisms in the entire mapping interval between the recurrent parent Shuhui881 and early-maturing NIL D248.

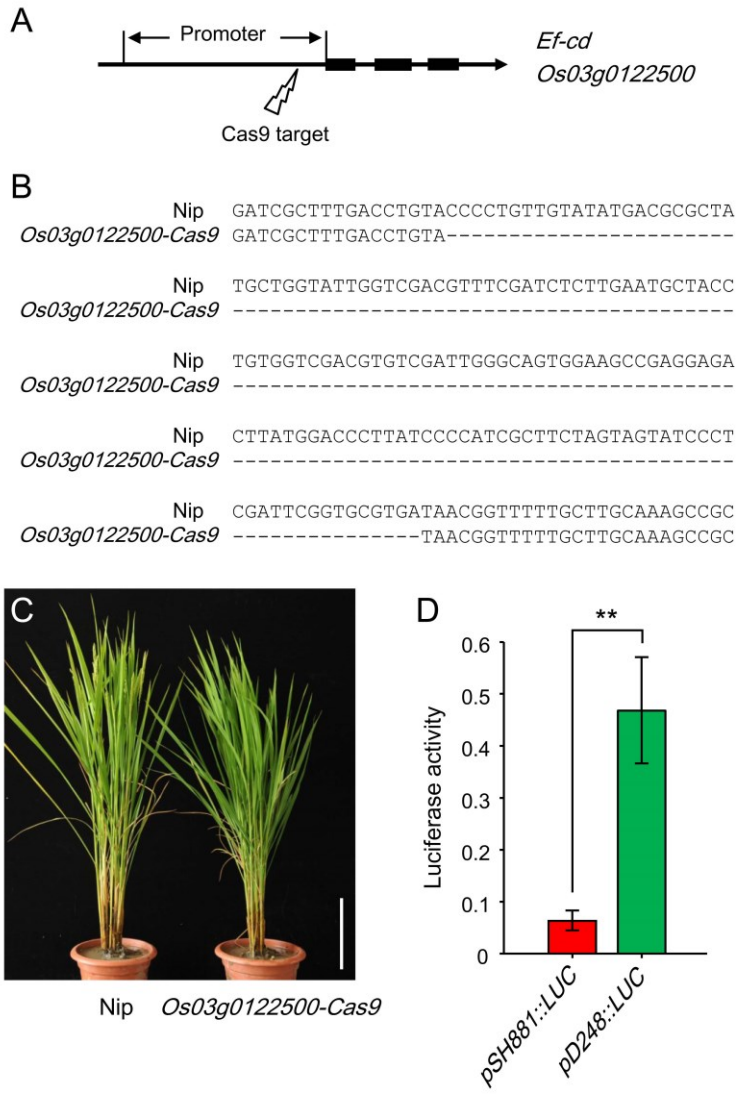

Fig. S3. Comparative analysis of promoter activity of *Ef-cd*. (*A*) Genomic structure of *Ef-cd* (*Os03g0122500*). The target site of Cas9 was located on the promoter region of *Os03g0122500*. (*B*) Sequencing results of target site from the wild type Nipponbare and the mutant line *Os03g0122500-Cas9*. A 158-bp deletion was occurred in the promoter region of *Os03g0122500*. (*C*) *Os03g0122500-Cas9* mutant line headed later than the wild type Nipponbare. The plants were planted in Beijing City (39°54' N). The sowing date was May 8, 2015. Bar = 20 cm. (*D*) Comparison of the luciferase activity in the protoplast cells between NIL *pD248::LUC* and its recurrent parent *pSH881::LUC*. The *p* value was calculated using Student's *t*-test. ** for significant at $p < 0.01$.

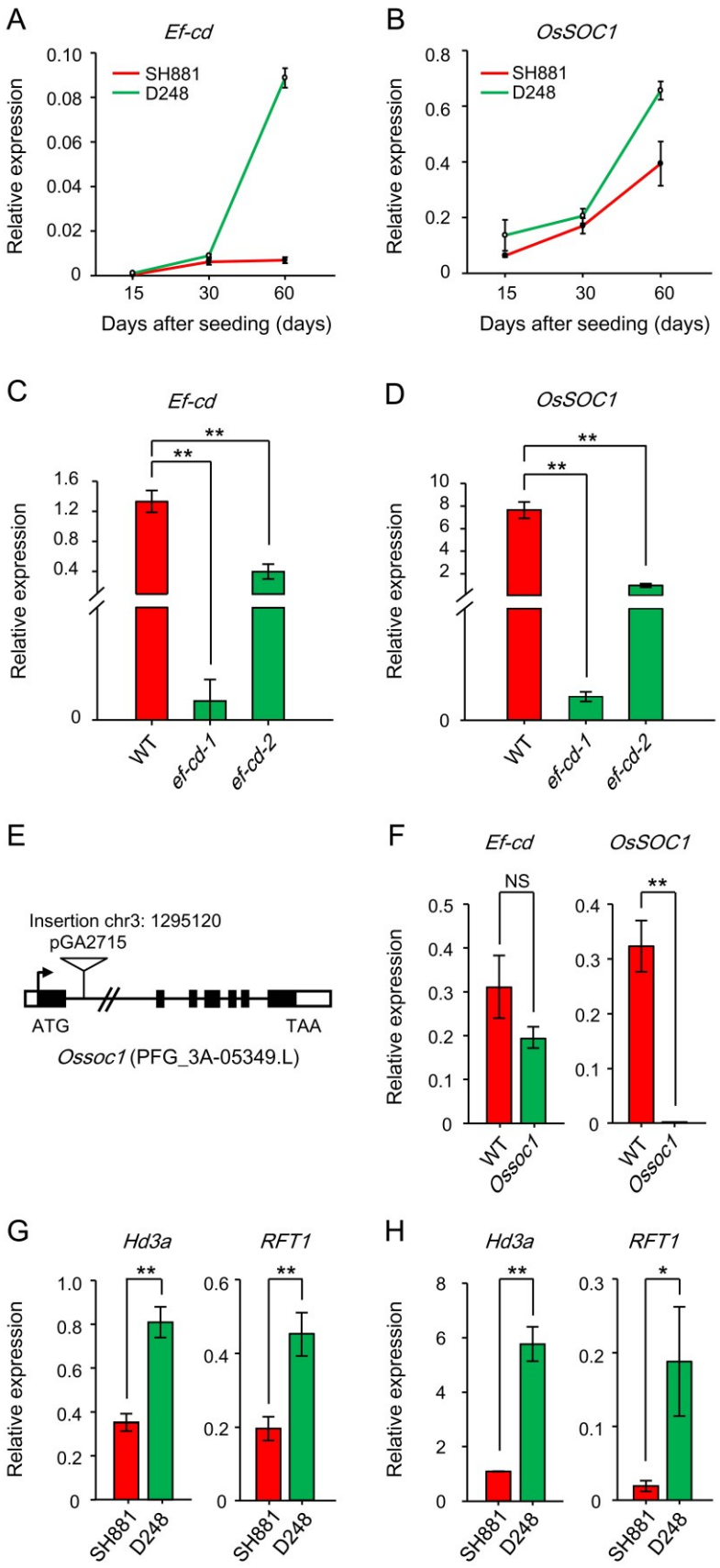

Fig. S4. Comparison of gene expression. (*A*) *Ef-cd*: D248 had greater expression than SH881 since the 30th day after seeding. (*B*) *OsSOC1*: D248 had greater expression

than SH881 all the way through. (*C*) The expression levels of *Ef-cd* were significantly reduced in *ef-cd* mutant plants. (*D*) The expression levels of *OsSOC1* gene were significantly reduced in *ef-cd* mutant plants. (*E*) Schematic of the *Ossoc1* mutant carrying a T-DNA insertion in the first intron. The triangle represents the T-DNA insertion. The type of the binary vector used for constructing the T-DNA insertion mutant was shown on the triangle. (*F*) *Ef-cd* and *OsSOC1* in wild-type Dongjin and its *Ossoc1* mutant plant (PFG_3A-05349.L), the expression level of *Ef-cd* was not significantly reduced in *Ossoc1* mutant plant. (*G*) *Hd3a* and *RFT1* in short-day condition: D248 had greater expression than SH881. (*H*) *Hd3a* and *RFT1* in long-day condition: D248 had greater expression than SH881. The *p* value was calculated using Student's *t*-test. * for significant at $p < 0.05$; ** for significant at $p < 0.01$; NS for not significant.

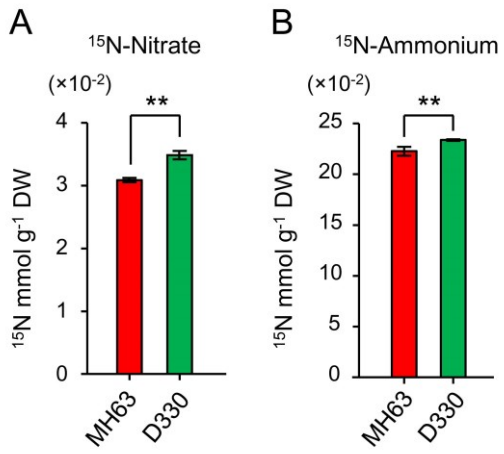

Fig. S5. ^{15}N accumulation assays in the shoots of MH63 and its early-maturing NIL D330 plants labeled with ^{15}N-nitrate or ^{15}N-ammonium. (*A*) The acquisition of nitrate was significantly increased in D330 than MH63. (*B*) The acquisition of ammonium was significantly increased in D330 than MH63. The *p* value was calculated using Student's *t*-test. ** for significant at $p < 0.01$.

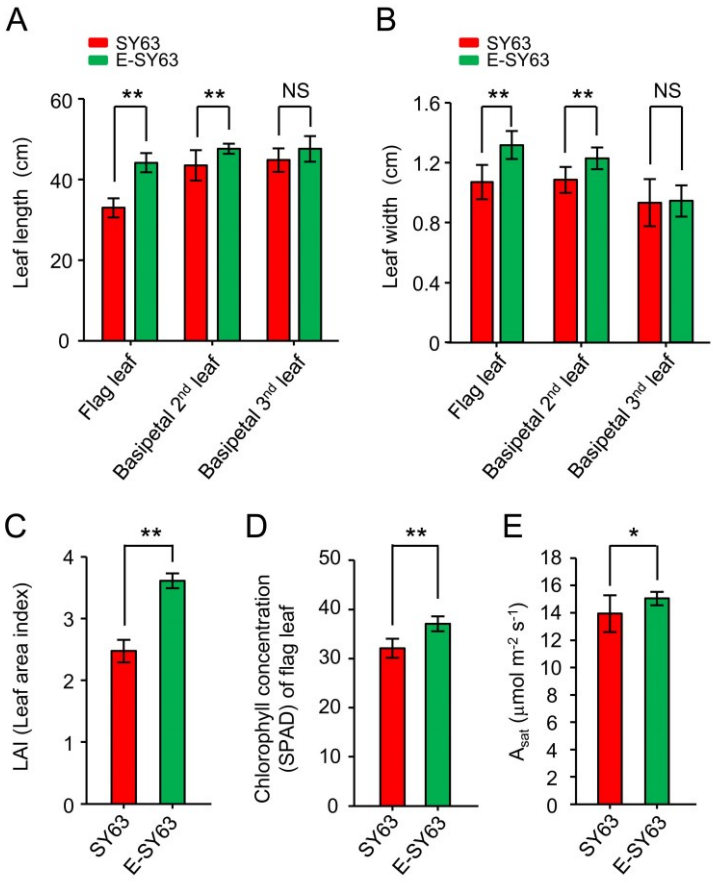

Fig. S6. Comparison of the physiology traits of E-SY63 and SY63 in Changsha City (28°12'N, Hunan Province, China, sowing date was May 30, 2015). (*A*) Length of upper three leaves. (*B*) Width of upper three leaves. (*C*) The leaf area index at the grain filling stage. (*D*) The chlorophyll concentration of flag leaf at the grain filling stage. (*E*) The light-saturated photosynthetic rate (A_{sat}) of flag leaf at the grain filling stage. The *p* value was calculated using Student's *t*-test. * for significant at $p < 0.05$; ** for significant at $p < 0.01$; NS for not significant.

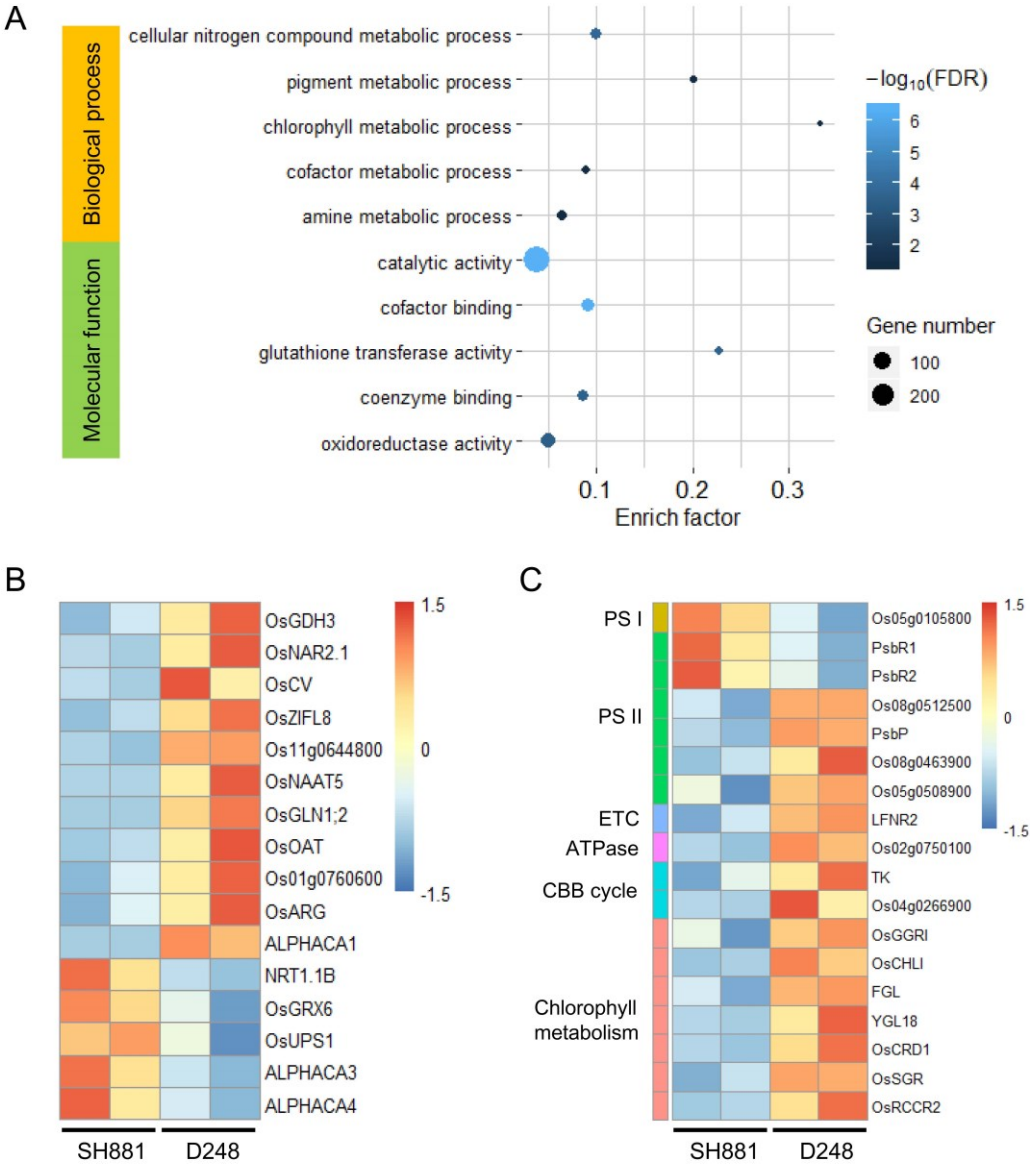

Fig. S7. Analysis of differentially expressed genes between the early-maturing NIL D248 and the recurrent parent SH881. (*A*) Top five enriched GO terms in each category of up-regulated genes in D248. (*B*) Differentially expressed genes related to nitrogen metabolism between SH881 and D248. (*C*) Differentially expressed genes of photosynthesis and chlorophyll metabolism between SH881 and D248. Abbreviation: PSI: photosystem I; PSII: photosystem II; ETC: electron transport chain; CBB cycle: Calvin-Benson-Bassham cycle.

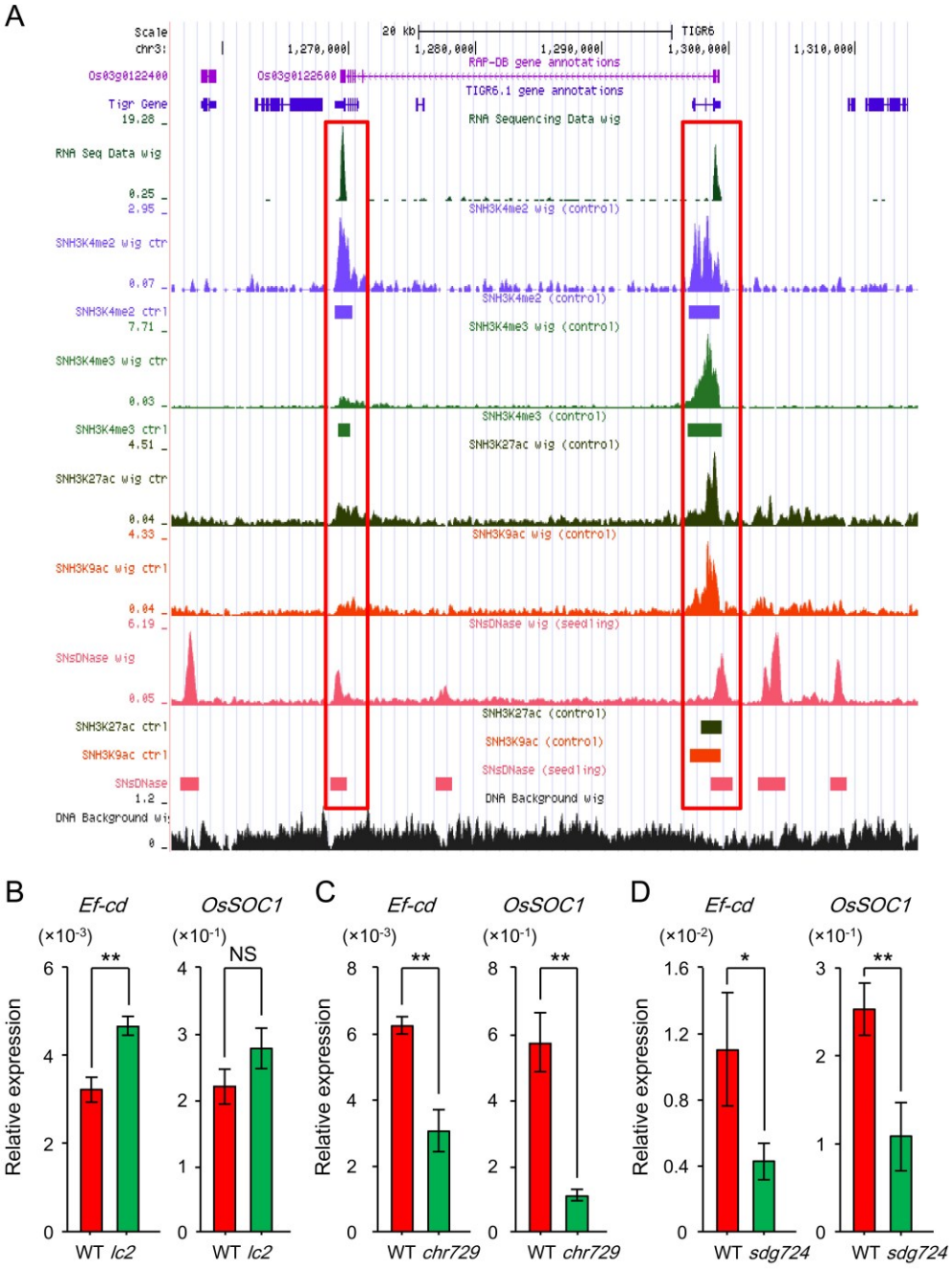

Fig. S8. Analysis of chromatin modification levels and mRNA expression levels. (*A*) The *Ef-cd-OsSOC1* locus had distinct chromatin modifications. Red blocks indicated the distinct regions. Reference website: http://structuralbiology.cau.edu.cn/cgi-bin/hgTracks. (*B-D*) Pairwise comparison of the expression levels of *Ef-cd* and *OsSOC1* in three mutants *lc2*, *chr729* and *sdg724* with their wild types (WT). The *p* value was calculated using Student's *t*-test. * for significant at $p < 0.05$; ** for significant at $p < 0.01$; NS for not significant.

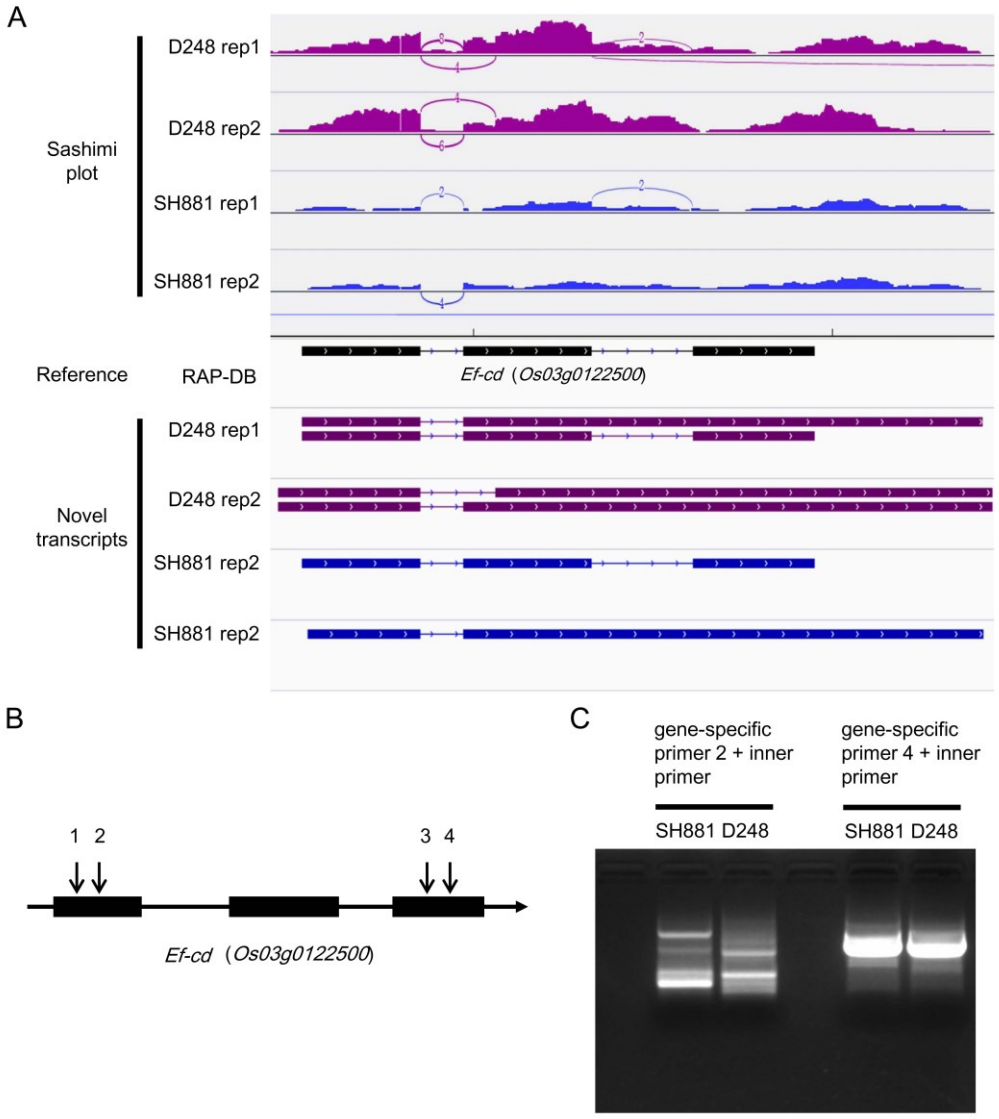

Fig. S9. Transcript structure analysis of *Ef-cd*. (*A*) *Ef-cd* locus alternative splicing events in D248 (2 replication) and SH881 (2 replication) are displayed in the IGV Sashimi-plot. The solid line indicated the number of associated exon-exon junction spanning reads. Reference *Ef-cd* transcript structure (black) are shown in the middle track. The structure of *Ef-cd* transcript that StringTie assembles from ssRNA-Seq data are shown below the reference track. (*B*) The position of the gene-specific primers designed for 3'-RACE. Relevant primers are listed in Dataset S4. (*C*) Other isoforms of *Ef-cd* with polyadenylated tails were detected by 3'-RACE.

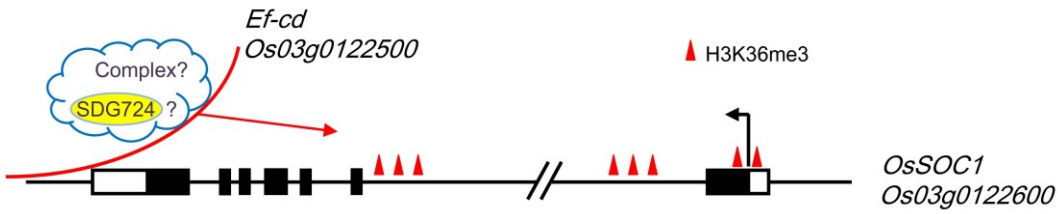

Fig. S10. Schematic illustration of the regulation of *Ef-cd*.

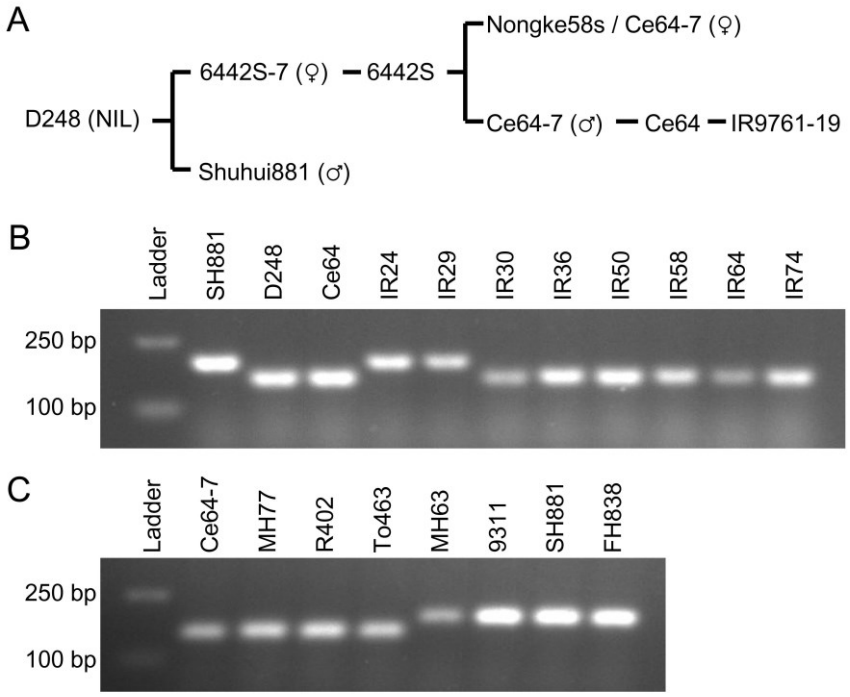

Fig. S11. Pedigree analysis of D248 and popularity of *Ef-cd* test in rice cultivars. (*A*) The pedigree information of D248. (*B*) Analysis of the popularity of *Ef-cd* in rice cultivars using 122500InDel-1 marker (closely linked to the 36-bp insertion/deletion locus). Early-maturing cultivars Ce64, IR30, IR36, IR50, IR58, IR64, and IR74 all contained the *Ef-cd* locus, but absent in late-maturing cultivars IR24 and IR29. (*C*) Analysis of the popularity of *Ef-cd* in rice cultivars using the 122500InDel-1 marker. *Ef-cd* locus is present in the Chinese leading early-maturing restorer cultivars such as Ce64-7, Minghui77 (MH77), R402, and To463, but absent in the Chinese leading late-maturing restorer cultivars such as Minghui63 (MH63), 9311, Shuhui881 (SH881), and Fuhui838 (FH838).