Supplemental Information

Genome-wide targeted mutagenesis in rice using CRISPR/Cas9 system

Contents

Supplemental Figure 1. Summary of CRISPR/Cas9 targets in rice genome.
Supplemental Figure 2. Example of a synthesized oligonucleotides of sgRNAs in RGKO#1.
Supplemental Figure 3. Colony PCR result for identification of CRISPR/Cas9 plasmid construction.
Supplemental Figure 4. Example of an amplicon for NGS-based genotyping.
Supplemental Table 1. Detail information of sgRNA library for rice.
Supplemental Table 2. Subgroups of sgRNA library for rice.
Supplemental Table 3. Primers for amplification of subgroups of sgRNA library.
Supplemental Table 4. Sequencing result of RGKO plasmid libraries.
Supplemental Table 5. Primers for NGS-based genotyping.
Supplemental Table 6. Examples of NGS-based genotyping result (960 samples).
Supplemental Table 7. Primers for plasmid construction and identification.
Supplemental Materials and Methods
Supplemental References
Supplemental Figures

Supplemental Figure 1

Supplemental Figure 1. Summary of CRISPR/Cas9 targets in rice genome. Targets comprising 5'-N20NGG-3' or 5'-N20NAG-3' were computationally derived and analyzed. sgRNAs with NGG pam were selected to construct the RGKO libraries.

Supplemental Figure 2

Supplemental Figure 2. Example of a synthesized oligonucleotides of sgRNAs in RGKO#1. The 20-bp sgRNA is shown in red and two Bsal sites are marked in blue. Primers for amplification of this oligonucleotides were underlined.
Supplemental Figure 3

Supplemental Figure 3. Colony PCR result for identification of CRISPR/Cas9 plasmid construction. Vectors with (ccdB+) or without (ccdB-) ccdB gene were used for ligation.

Supplemental Figure 4

Supplemental Figure 4. Example of an amplicon for NGS-based genotyping. The 20-bp sgRNA is shown in red. Primers were underlined and barcodes are shown in blue.
Supplemental Tables (Attached in Excel file)

Supplemental methods

**Vector construction**

Cas9 amplified from pX260 (Cong et al., 2013) were linked to the maize ubiquitin promoter (UBI) in an intermediate plasmid, and then this expression cassette was inserted into binary vector pCAMBIA1300 (Cambia, Australia) which contains the HPT (hygromycin B phosphotransferase) gene. The originally existed BsaI site in the pCAMBIA1300 backbone was removed using point mutation kit (Transgen, China). A fragment comprised of a OsU6 promoter (Feng et al., 2013) and a gRNA derived from pX260 (Cong et al., 2013) was inserted into this vector using In-fusion cloning kit (Takara, Japan) to produce the vector BGK03 (Figure 1A). Vector BGK032 was constructed by inserting the negative selection marker gene ccdB (flanked by two BsaI sites) at the sgRNA site of BGK03. *E. coli* strain DB3.1 was used for maintaining this binary vector.

**Genome-wide sgRNA design for RGKO library**

The rice genome sequence and its GTF file (MSU7) containing the exonic coordinates of all protein coding genes were obtained. Targets comprising 5’-N20NGG-3’ or 5’-N20NAG-3’ were computationally derived and analyzed. All sgRNAs that against 5’ constitutive coding exons were selected. Filtering was performed as follows. First, to minimize off-target cleavages, only sgRNAs that satisfy stringent conditions were chosen: from position 8, the 5’-N12NGG-3’ of each sgRNA must contain at least one mismatch with all other targets in the whole genome. Second, GC content of each sgRNA must be in the range of 25% to 80%. Third, sgRNAs that are positioned at least 60 bp away from the translation initiation site and in the first half of coding sequences were collected. Finally, up to 3 sgRNAs were chosen for each gene, prioritizing sgRNAs with fewer predicted off-target sites and higher predicted targeting efficiency. A total of 88,541 sgRNA sequences was chosen to produce the sgRNA library (Supplementary Table 1).
**RGKO library construction**

According the designed sgRNAs, a 109-mer oligo pool was purchased from Synbio Technologies Inc (Supplementary Table 1). The single-stranded oligo pool was converted to double-stranded DNA by PCR using KOD Neo (Toyobo) with 100 fmol of the oligo as template (Supplementary Table 1). PCR were conducted using the following conditions: 95 °C for 60 s, 10 cycles of 95 °C for 20 s, 60 °C for 30 s and 68 °C for 15 s, and the final extension, 68 °C for 2 min. The PCR products were purified with the Gel purification kit (Tiangen), digested with BsaI and separated by PAGE. The 29-bp fragment was excised from the PAGE gel. Ligation was performed using 5 ng of the 29-bp fragment and 100 ng of BsaI-digested BGK032 with T4 DNA ligase (NEB). Ten ligation reactions were combined, ethanol precipitated and dissolved with 30 μl water. Ten electroporations were carried out using 2 μl of the purified ligation product and 25 μl of electro-competent cells (Takara) per reaction according to the manufacturer’s instruction. The electroporated cells were combined and plated onto fifty agar plates (15-cm, LB + Kanamycin). Colonies were identified using Sanger sequencing individually in advance. Then bacteria from all plates (100× library complexity) were scraped off and combined. The plasmid DNA was purified with Plasmid Maxi kit (Tiangen) to produce the RGKO libraries. In order to verify these plasmid pools, PCR product amplified from RGKO libraries using primer (HGP-F0 and HGP-R0) were sequenced using NGS.

**Rice transformation using RGKO library**

The RGKO libraries were introduced into A. tumefaciens strain EHA105 using electroporation. Agrobacterium colonies were directly resuspended for rice transformation. Transformation of rice was performed as described previously (Nishimura et al., 2006). Rice cultivar ZH11 were used for all transformation.

**Detection of genome modifications**

To identify the sgRNA of regenerated seedlings, genomic DNA from approximately 0.1 g of leaf tissue was extracted and the PCR product amplified using primers (PUV3-R and gRNA-R5) was identified using Sanger sequencing. In order to assess the mutagenesis frequency, the targeted loci of sgRNAs were amplified for Sanger sequencing and then analyzed using Dsdecode (Liu et al., 2015).

**NGS-based high-throughput genotyping**

For primer design, 96 unique barcodes (6 bp) were computationally derived with at least two mismatches
with all other barcodes. PCR primers tailed with 6-bp additional nucleotides (Supplementary Table 5) were then verified using software Oligo7. 20-μl PCR reactions using 10 μl PCR mixture (Bioteke), 4 μl each of upper and lower primers (2 μM) and 2 μl DNA, were performed in the following conditions: 95 °C for 3 min, 32 cycles of 95 °C for 20 s, 55 °C for 30 s and 72 °C for 15 s, and the final extension, 72 °C for 2 min. Then, all PCR solutions were mixed together for NGS sequencing. sgRNA(s) of each mutant was computationally distinguished from NGS data according to its barcodes.

Supplemental references