Supplementary Data

Synthesis-dependent repair of Cpf1-induced double-strand DNA breaks enables targeted gene replacement in rice

Shaoya Li¹, Jingying Li¹, Jiahui Zhang¹, Wenming Du¹, Jindong Fu¹, Suhas Sutar¹, Yunde Zhao^{2, 3, *} and Lanqin Xia^{1, *}

Contents

Fig. S1 The map of plasmids used in this study

- (A) pCXUN-LbCpf1-OsU3-RCR1-RCR2-left-armed-DRT
- (**B**) pCXUN-LbCpf1-OsU3-RCR1-RCR2-armed-DRT

Fig. S2 Sequences of the DRTs

- (A) DRT with only left homology arm
- (B) DRT with two homology arms
- Table S1 The primer sets used in this study

Table S2 Characterization of regenerated rice plants in T₀ generation

Table S3 Analysis of potential off-target effects



pCXUN-LbCpf1-OsU3-RGR1-RGR2-left-armed-DRT



pCXUN-LbCpf1-OsU3-RGR1-RGR2-armed-DRT

Fig. S1. A schematic show of the vectors used in this study

- (A) pCXUN-LbCpf1-OsU3-RGR1-RGR2-left-armed-DRT
- (B) The map of pCXUN-LbCpf1-OsU3-RGR1-RGR2-armed-DRT

B

A	A Target 1 (23 bp)	Left arm (97 bp)							
	AACCAACATTTQGGTATGGTGGTGCAATGGGAGGA	TTCCTCATGAACATTCAGGAGCTGGCATTGATCCGCATTGAGAACCTCCCTGTGAAGGTGATGG							
	<mark>ŦĠŦŦĠĂĂĊĊĂĂĊĊĂĂĊĊŦġĠĠċĂŦĠĠŦċĠŦċĊĂġ<mark>ŦŀĠ</mark>ĠĂ<u>ĠĠĠĂ</u>ŦĂĠĠŦŦŦŦĂĊĂĂĠĠĊĠĊĂŦĂĠĠĠĊĠĊĂŦĂĊĊŦĠĠĠĊĂĂĊĊĊĠĠĂĂŦĠŦĠĂĠĂĠĊĠĂĠĂŦĂŦĊĊĂĠĂŦŦŦ</mark>								
	Mutated target 1 Mutated G to T: W-L	Mutated <i>Eco</i> RV							
	GTGACTATTGCTAAGGGGTTCAATATTCCTGCAGTCCGTGTAACAAAGAAGA	GTGAAGTCCGTGCCGCCATCAAGAAGATGCTCGAGACTCCAGGGCCATACTTGTTG <mark>GAcATC</mark> ATC							
	GTCCCGCACCAGGAGCATGTGCTGCCTATGATCCCA <mark>A(T</mark> GGGGGGCGCATTCAA Mutated G to T: S-I	AGGACATGATCCTGGATGGTGATGGCAGGACTGTGTATTAATCTATAATCTGTATGTTGGCAAAGC							
	ACCAGCCCGGCCTATG <mark>TcTGACgTGAATGACtCATAAAGAGTG<mark>TTTGACCTGA</mark></mark>	ATGACCCATAAAGAGTG GTATGCCTA							
	Mutated target 2	Target 2 (23 bp)							
ן ה -									
R	Target 1 (23 bp) Left arm (97 bp)								
AACCAACATTTCCCCCCCCCCCCCCCCCCCCCCCCCCC									
							Mutated target 1 Mutated G to T: W-L Mut		
GTGACTATTGCTAAGGGGTTCAATATTCCTGCAGTCCGTGTAACAAAGAAGAGGGGGAGGTCCGTGCCGCCATCAAGAAGATGCTCGAGACTCCAGGGCCATACTTGTTGGGA									
	GTCCCGCACCAGGAGCATGTGCTGCCTATGATCCCAAITGGGGGGGGGG								
	ACCAGCCCGGCCTATCTCTGACgTGAATGACtCATAAAGAGTGGTATGCCTATG	ATGTTTGTATGTGCTCTATCAATAACTAAGGTGTCAACTATGAACCATATGCTCTTCTGTTTTACTT							
	Mutated target 2	Right arm (121 bp)							
	GTTTGATGTGCTTGGCATGGTAATCCTAATTAGCTTCCTGCTG	NTGACCCATAAAGAGTGGTATGCCTA							
	Та	rget 2 (23 bp)							

Fig. S2. Sequences of the DRT with only left homology arm and the DRT with two homology arms

(A) DRT with only left homology arm; (B) DRT with two homology arms. The sequences shadowed in yellow are derived from the wild-type ALS gene sequence, sequences of two targets are underlined, the PAM and mutated PAM sequences are boxed, the left and right arms are shadowed in blue, the mutated *Eco*RV target site are shadowed in gray, the mutated amino acids are shadowed in green, and the mutated synonymous changes in the donor are in red lower case.

Table S1. The primer sets used in this study

Primer	Sequence (5' to 3')	Use of PCR products	Annealing (°C)	Fragment size (bp)
RGR1F2 RGR-common-R	TAATTTCTACTAAGTGTAGATGGTATGGTGGTGCAATGGGAGGAGGCCGGCATGGTCCCAGCCT GTCCCATTCGCCATGCCGAAGC	Construct vector for RGR1	58	112
RGRF1 RGR-common-R	AAATTACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTAATTTCTACTAAGTGTAGAT GTCCCATTCGCCATGCCGAAGC	Construct vector for RGR1and RGR2	58	155
RGR2-F2 RGR-common-R	TAATTTCTACTAAGTGTAGATACCTGAATGACCCATAAAGAGTGGGCCGGCATGGTCCCAGCCT GTCCCATTCGCCATGCCGAAGC	Construct vector for RGR2	58	112
OsU3F OsU3R OsU3-RGR1R	AAGGAATCTTTAAACATACGAACAGATCA TGCCACGGATCATCTGCAC CACGGACTCATCAGTAATTTTGCCACGGATCATCTGCAC	Construct vector for OsU3-RGR1-RGR2	56	404
RGR-Common-F RGR1-10 random-R	AAATTACTGATGAGTCCGTGAGGAC GTAATTTGGTCGTATTCGTCCCATTCGCCATGCCG	Construct vector for OsU3-RGR1-RGR2	58	172
RGR2-10 random-F Sac-RGR2R	CGAATGGGACGAATACGACCAAATTACTGATGAGTCCGTGAGGAC CCCTTTCGCCAGGGGTACCGGTCCCATTCGCCATGCCGAAGC	Construct vector for OsU3-RGR1-RGR2	58	195
Sac-OsU3F Sac-RGR2R	GATTACGAATTCGAGCTCAAGGAATCTTTAAACATACGAACAGATCA CCCTTTCGCCAGGGGTACCGGTCCCATTCGCCATGCCGAAGC	Construct vector for OsU3-RGR1-RGR2	58	739
Pme-donorF Pme-donorR	GTTTCCCGCCTTCAGTTTAAACGTAAAACGACGGCCAGT GTCAAACACTGATAGTTTAAACCAGGAAACAGCTATGAC	Construct vector for donor	56	873

donor-armLF donor-armLR1	AACCAACATTTGGGTATGGTG CATAGGCCGGGCTGGT	Construct vector for donor (with only left arm)	60	486
donor-armLF donor-armLR2	AACCAACATTTGGGTATGGTG CACTCTTTATGGGTCATTCAGGTCAAACATAGGCCGGGCTGGT	Construct vector for donor (with only left arm)	58	522
pme-donor-armLF pme-donor-armLR	GTTTCCCGCCTTCAGTTTAAACAACCAACATTTGGGTATGGTG GTCAAACACTGATAGTTTAAACCACTCTTTATGGGTCATTCAGGT	Construct vector for donor (with only left arm)	58	566
ALSTestF T2MR	GGCAATCATCGCTACTGGTG CACTCTTTATGAGTCATTCACGTCAG	Detect edited ALS gene in calli	58	648
ALSTestF ALSTestR	GGCAATCATCGCTACTGGTG ATGCCTACAGAAAACAACACACTACA	Detect edited ALS gene in plants	58	803
hptF hptR	GAGGGCGTGGATATGTCCTG ATTGACCGATTCCTTGCGGT	Detect <i>hptII</i> gene in plants	58	306
U3F U3R	AAGGAATCTTTAAACATACGAACAGATC TTTGGATCCTCTAGAGATTATGTGG	Detect crRNA sequence in plants	57	509
LbCpf1F LbCpf1R	GCCTGTTGCACAAGGATTAGACTTC GGAGCAGAGCTGTTCATGAGG	Detect LbCpf1 transgene in plants	58	684
ALS-OFF1-F ALS-OFF1-R	GCTGCTCTGTTGTCTGGGTC ACGAGCAACAGCCAGAATCA	Detect mutations on putative off-target site of <i>ALS</i> target 1	59	800

ALS-OFF2-F ALS-OFF2-R	GGTTGCAAAGTCACATGCGTG TCCGGAGATTCCTGACGTGAG	Detect mutations of putative off-target site of <i>ALS</i> target 1	on of S	59	818
ALS-OFF3-F ALS-OFF3-R	GGTTTACCTTTCTAGCGCACTTC CGATGTTCTTAGGCGCACC	Detect mutations of putative off-target site of <i>ALS</i> target 1	on of s	59	658
ALS-OFF4-F ALS-OFF4-R	CACCCATAAGGAATGGGTCTTC GCTCTTGATTGGGCCGGC	Detect mutations of putative off-target site of <i>ALS</i> target 2	on of S	57	999
ALS-OFF5-F ALS-OFF5-R	CTACCTAGGGCAGCTGCTG CAGAGCCATCCCAGACTTTAC	Detect mutations of putative off-target site of <i>ALS</i> target 2	on of :	59	524

Table S2.Characterization and genotyping of regenerated plants in T₀ generation

Vector	NO. of calli for transform	NO. of resistant calli	NO. of individual rice plants recovered	Line	Genotype	
	152	15	71	171-2	Heterozygous HR	
pCXUN-LbCpf1-OsU3-RG				171-9	chimeric HR	
R1-RGR2-left-armed-DRT				170-1 172-1	Partial HR	
	164	20	94	169-2	chimeric HR	
				169-3		
				168-2	Partial HR	
pCXUN-LbCpf1-OsU3-RG				168-4		
R1-RGR2-armed-DRT				168-5		
				169-5		
				169-6		
				169-1	Random indels	

Target	Name of putative off-target site	Putative off-target locus	Sequence of the putative off-target site	No. of m ismatching b ases	No. of plants sequenced	No. of plants with mutations
	ALS1-OFF1	chr02:4712511-4712537	<u>TTTA</u> G <mark>AA</mark> AT T GT T GTGCAA <mark>G</mark> GGGAGGA	5	12	12
Target 1	ALS1-OFF2	chr05:26016305-26016331	<u>TTTG</u> GGTA <mark>A</mark> GAT <mark>C</mark> ATGCAAAGGGAGAA	6	12	12
	ALS1-OFF3	chr03:33738572-33738598	<u>TTTG</u> ATTTTGGTGGTG <mark>G</mark> AATGG <mark>A</mark> AGAA	6	12	12
Torrect 2	ALS2-OFF1	chr08:4284947-4284973	TTTCACCTAAATAACCCATTAGGAGTA	5	12	12
rarget 2	ALS2-OFF2	chr07:584165-584191	TTTGCCCTGGATGGCCCTTTAAGAGGG	6	12	12

Note: The PAM motif is underlined; mismatching bases are shown in red.