

Targeted C•G-to-T•A base editing with TALE cytosine deaminases in plants

Dingbo Zhang, Vanessa Pries, Jens Boch*

Leibniz University Hannover, Institute of Plant Genetics, Herrenhäuser Str. 2,
30419 Hannover, Germany

Additional file 1

Figure S1. Schematic of the GUS^{G537} cytosine base editing reporter.

Figure S2. Editing efficiency of TALE-DdCBEs with fusion of Rad51.

Figure S3. DddA variants showed enhanced editing in the GUS^{G537} reporter.

Figure S4. TALE-DdCBE11-induced C•G-to-T•A editing in the rice chloroplast genome.

Figure S5. Schematic illustration for assembling TALE-DdCBEs.

Table S1. Inheritance of mutations in T1 lines.

Table S2. Analyzing potential off-target editing of TALE-DdCBE11 in T0 *N. benthamiana* plants.

Table S3. Plasmids used in this study.

Table S4. TALE binding sequences and corresponding RVDs.

Table S5. Oligos used in this study.

Table S6. Alignment of amplicon sequencing results.

Supplemental Sequences. Amino acid sequences of TALE-DdCBE architectures.

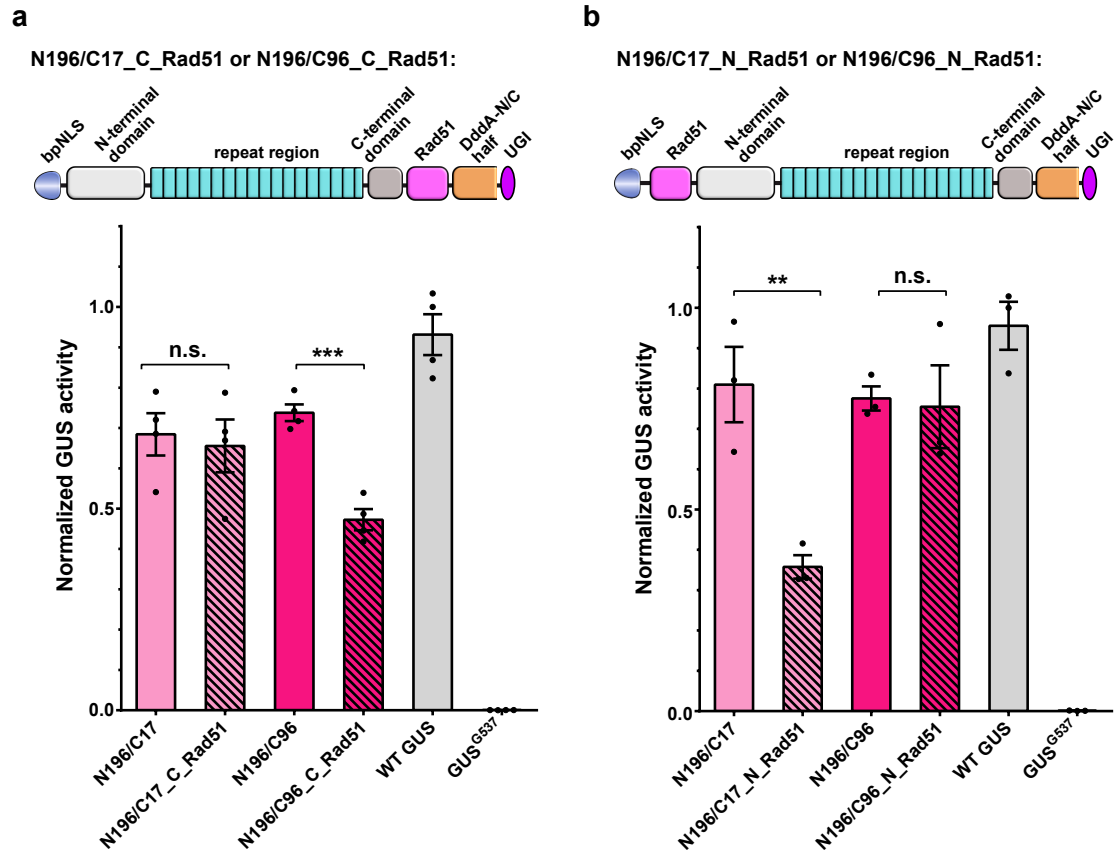


Figure S2. Editing efficiency of TALE-DdCBEs with fusion of Rad51.

(a) C•G-to-T•A editing efficiencies of TALE-DdCBEs with Rad51 inserted between TALE C-terminal domain and DddA halves in GUS^{G537} reporter, n = 4.

(b) C•G-to-T•A editing efficiencies of TALE-DdCBEs with Rad51 inserted between bpNLS and TALE N-terminal domain in GUS^{G537} reporter, n = 3.

bpNLS: bipartite nuclear localization sequence. UGI: uracil glycosylase inhibitor. GUS activities were measured and normalized to 2x35S::GUS (WT GUS, positive control). Values and error bars indicate the mean $\pm$ SEM, ***p < 0.001, **p < 0.01, n.s.: not significant, using Student's two-tailed unpaired t-test.

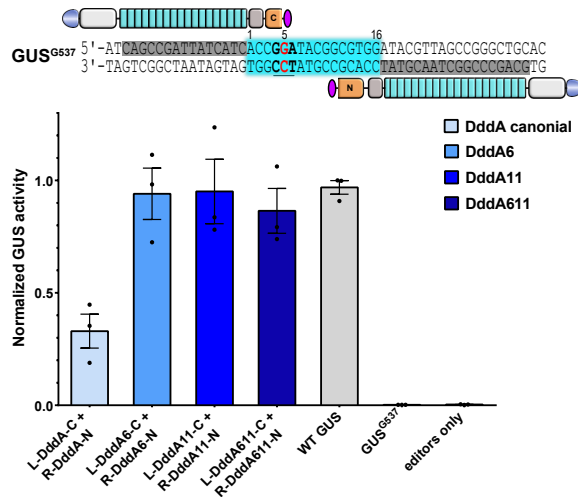


Figure S3. DddA variants showed enhanced editing in the GUS^{G537} reporter.

C•G-to-T•A editing efficiencies of TALE-DdCBEs harboring canonical DddA or DddA6 or DddA11 or DddA611 in GUS^{G537} reporter. Top: TALE-DdCBEs using left TALE-DddA-C / right TALE-DddA-N architecture, targeted cytosine located at C5 within the 16-nt spacer. Bottom: values and error bars indicate the mean $\pm$ SEM, n=3.

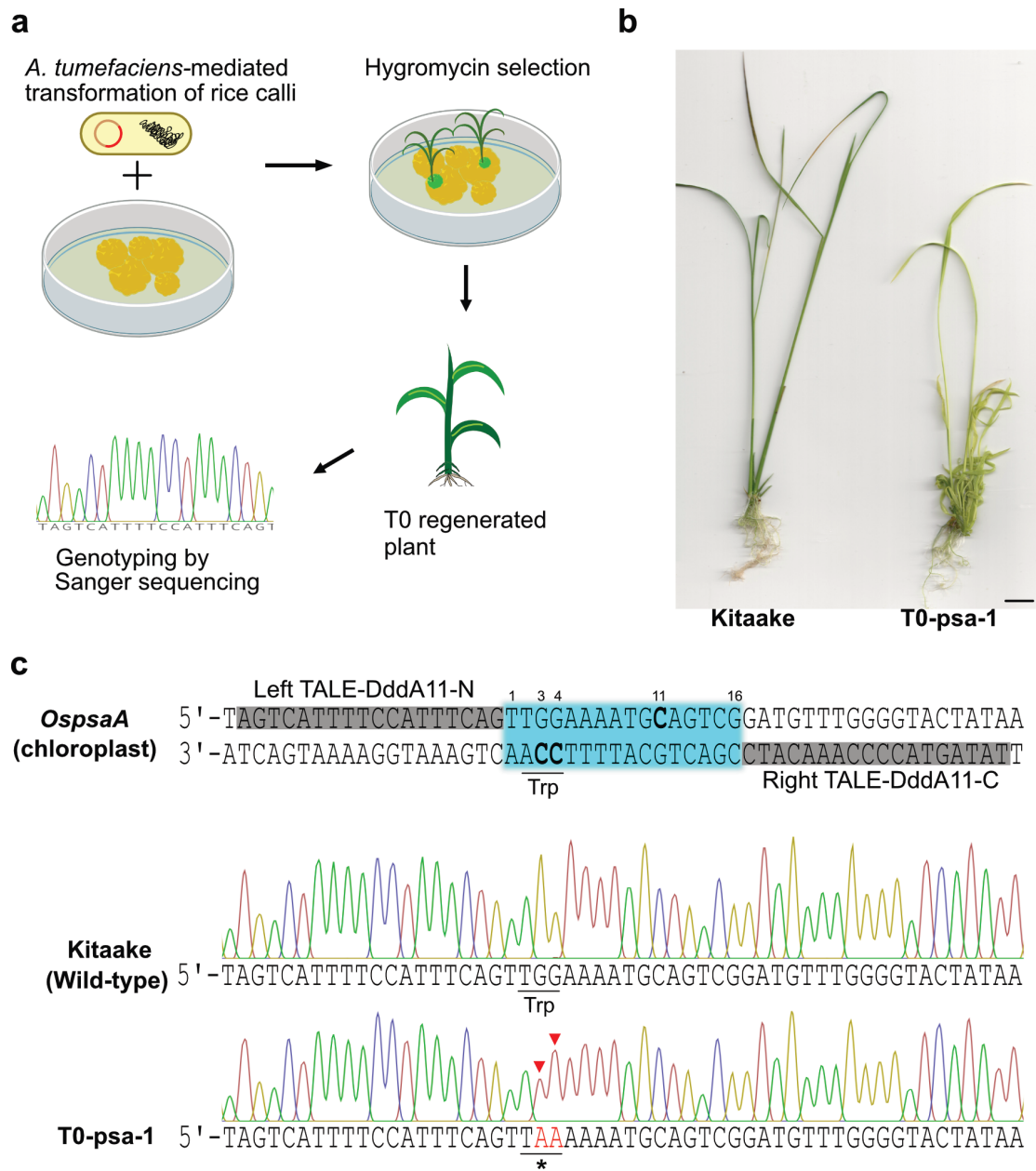


Figure S4. TALE-DdCBE11-induced C•G-to-T•A editing in the rice chloroplast genome. (a) Schematic C•G-to-T•A editing mediated by TALE-DdCBE11 in rice plants. **(b)** Phenotypes of the *OspsaA* mutant line T0-psa-1 and a wild-type Kitaake plant. Bar = 1 cm. **(c)** Genotypes of line T0-psa-1 and wild-type Kitaake plants. A pair of TALE-DddA11 was targeting the rice *OspsaA* chloroplast gene. TALE-binding sites are in gray background and the spacer region is in cyan background. Sanger sequencing chromatograms of wild-type and regenerated T0 plants are shown. Base conversions are indicated in red and marked by red triangles. The tryptophan (Trp) codon TGG was converted to a TAA stop codon (*).

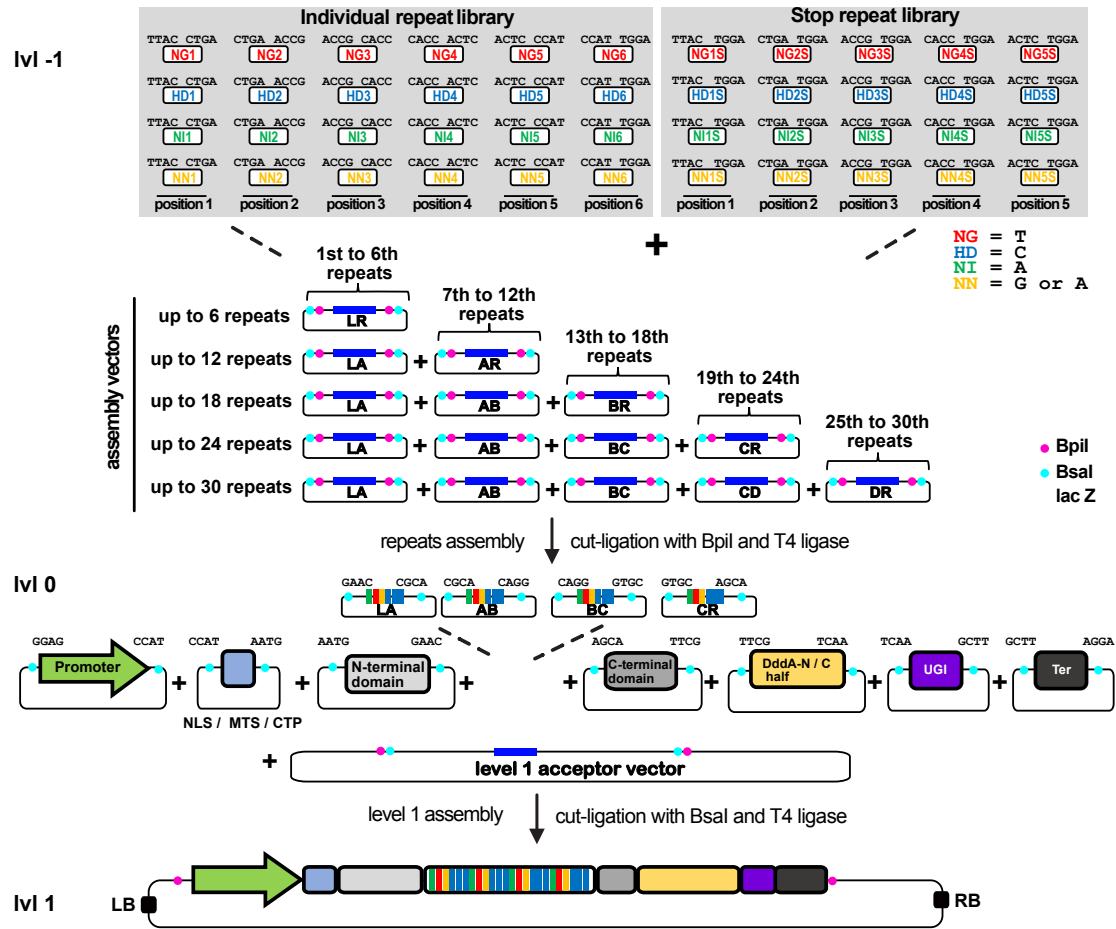


Figure S5. Schematic illustration for assembling TALE-DdCBEs.

Golden-Gate assembly of TALE-DdCBEs based on the MoClo design. Corresponding to the target sequence (with a thymine at the 5' end), up to six single repeats are ligated into the corresponding receiving plasmids. Next, those multi-repeat modules are combined with promoter, NLS or MTS or CTP, N-terminal domain, C-terminal domain, DddA half, terminator (Ter) and level 1 acceptor vector to yield the level 1 TALE-DdCBEs transcription unit. Level 1 vectors are used for transient transformation (protoplasts assay and infiltration of *N. benthamiana* leaves). Four-base overhang sequences are shown on the modules. NLS: Nuclear localization signal. MTS: Mitochondrial targeting signal. CTP: chloroplast transition peptide.

Table S1. Inheritance of mutations in T1 *N. benthamiana* lines.

Target gene	T0 lines	T0 genotypes [†]	Number of T1 plants	Number, and genotypes of T1 plants [†]
<i>NbSuRB</i>	T0-6	C6->T (He)	4	4, C6->T (He)
	T0-7	C6->T (He)	2	1, C6->T (He); 1, C6->T (Ho)
	T0-9	C6->G (He)	4	2, C6->G (He); 2, C6->G (Ho)
	T0-10	C6->T (He), C7->T (He)	2	1, C6->T (Ho) & C7->T (Ho) 1, C6->T (Ho) & C7->T (He)

†: He: Heterozygous. Ho: Homozygous.

Table S2. Analyzing potential off-target editing of TALE-DdCBE11 in T0 *N. benthamiana* plants.

Target sites	Left TALE binding sequence (5'-3') [†]	Spacer (5'-3')	Right TALE binding sequence (5'-3') [†]	Off-target editing (line No., editing)
On-target (<i>NbSuRB</i>)	TGTCATGCTGGGTC	AGATTGGAAGTT	TCCAACCTAAAGAGG	-
Off-target site 1	TCTCATGCTTCTTC	CTTCAAGTT	TCCAACCTAAAAGG	T0-3, not detected T0-6, not detected T0-10, not detected
Off-target site 2	TATCATGCAAGAGTT	TGATTGCA	TCCAACCTAAAGAAG	T0-3, not detected T0-6, not detected T0-10, not detected
Off-target site 3	TGTCATGCTGGGTC	AAACCTCGACCT TTCC	CACAACACAAAGACG	T0-3, not detected T0-6, not detected T0-10, not detected
Off-target site 4	TATCATGCTGGGTC	AACCTTTATGAC CTTTCC	CACAACATAAAGATG	T0-3, not detected T0-6, not detected T0-10, not detected
Off-target site 5	TATCATGATGAGTT	TTCCATCA	TCCAACCTAAAAG	T0-3, not detected T0-6, not detected T0-10, not detected

[†]: Mismatched sequences are highlighted in red.

Table S3. Plasmids used in this study.

Name	MoClo level	Content
L0_35s_Pro5U	level 0	2X35S promoter with GGAG/CCAT overhang
L0_ZmUbi_Pro5U	level 0	Maize ubiquitin promoter with GGAG/CCAT overhang
L0_Noster	level 0	Nos terminator with GCTT/AGGA overhang
L0_35s_ter	level 0	35S terminator with GCTT/AGGA overhang
L0_bpNLS	level 0	bpNLS with CCAT/AATG overhang
L0_TND_288	level 0	TALE N-Terminal domain (288 aa) with AATG/GAAC overhang
L0_TND_196	level 0	TALE N-Terminal domain (196 aa) with AATG/GAAC overhang
L0_TND_135	level 0	TALE N-Terminal domain (135 aa) with AATG/GAAC overhang
L0_TCD_17	level 0	TALE C-Terminal domain (17 aa) with AGCA/TTCG overhang
L0_TCD_28	level 0	TALE C-Terminal domain (28 aa) with AGCA/TTCG overhang
L0_TCD_63	level 0	TALE C-Terminal domain (63 aa) with AGCA/TTCG overhang
L0_TCD_96	level 0	TALE C-Terminal domain (96 aa) with AGCA/TTCG overhang
L0_TCD_247	level 0	TALE C-Terminal domain (247 aa) with AGCA/TTCG overhang
L0_UGI	level 0	UGI-NLS with TCAA/GCTT overhang
L0_CTP	level 0	Chloroplast transition peptide with CCAT/AATG overhang
L0_DddA_N	level 0	DddA-N half with TTCG/TCAA overhang
L0_DddA_C	level 0	DddA-C half with TTCG/TCAA overhang
L0_DddA6_N	level 0	DddA6-N half with TTCG/TCAA overhang
L0_DddA6_C	level 0	DddA6-C half with TTCG/TCAA overhang
L0_DddA11_N	level 0	DddA11-N half with TTCG/TCAA overhang
L0_DddA11_C	level 0	DddA11-C half with TTCG/TCAA overhang
L0_DddA611_N	level 0	DddA611-N half with TTCG/TCAA overhang
L0_DddA611_C	level 0	DddA611-C half with TTCG/TCAA overhang
L0_DddA6 ^{K1389A} _N	level 0	DddA6 ^{K1389A} -N half with TTCG/TCAA overhang
L0_DddA6 ^{K1389A} _C	level 0	DddA6 ^{K1389A} -C half with TTCG/TCAA overhang
L0_DddA6 ^{T1391A} _N	level 0	DddA6 ^{T1391A} -N half with TTCG/TCAA overhang
L0_DddA6 ^{T1391A} _C	level 0	DddA6 ^{T1391A} -C half with TTCG/TCAA overhang
L0_DddA11 ^{K1389A} _N	level 0	DddA11 ^{K1389A} -N half with TTCG/TCAA overhang
L0_DddA11 ^{K1389A} _C	level 0	DddA11 ^{K1389A} -C half with TTCG/TCAA overhang

L0_DddA11 ^{T1391A} _N	level 0	DddA11 ^{T1391A} -N half with TTCG/TCAA overhang
L0_DddA11 ^{T1391A} _C	level 0	DddA11 ^{T1391A} -C half with TTCG/TCAA overhang
L0_DddA_N_UGI	level 0	DddA-N half-UGI-NLS fusion with TGGC/AGGA overhang
L0_DddA_C_UGI	level 0	DddA-C half-UGI-NLS fusion with TGGC/AGGA overhang
L0_NLS_Rad51_N	level 0	bpNLS-Rad51 fusion with CCAT/AATG overhang
L0_Rad51_C	level 0	Rad51 domain with TTCG/TGGC overhang
pICH47732_DZ	level 1 acceptor	MoClo level 1 acceptor in position 1 with GGAG/AGGA overhang
pICH47742_DZ	level 1 acceptor	MoClo level 1 acceptor in position 2 with GGAG/AGGA overhang
pGUS_WT	level M	Wild-type GUS expression plasmid containing a p19 silencing suppressor
pGUS_537	level M	GUS ^{G537} expression plasmid containing a p19 silencing suppressor

For TALE repeats see Geißler et al., 2011

Table S4. TALE binding sequences and corresponding RVDs.

Gene	Binding sequence [†] (5'-3')	RVDs
GUS ^{G537} -left- TALE1	T CAGCCGATTATC	HD NI NN HD HD NN NI NG NG NI NG HD
GUS ^{G537} -left- TALE2	T CAGCCGATTATCA	HD NI NN HD HD NN NI NG NG NI NG HD NI
GUS ^{G537} -left- TALE3	T CAGCCGATTATCAT	HD NI NN HD HD NN NI NG NG NI NG HD NI NG
GUS ^{G537} -left- TALE4	T CAGCCGATTATCATC	HD NI NN HD HD NN NI NG NG NI NG HD NI NG HD
GUS ^{G537} -left- TALE5	T CAGCCGATTATCATCA	HD NI NN HD HD NN NI NG NG NI NG HD NI NG HD NI
GUS ^{G537} -left- TALE6	T CAGCCGATTATCATCAC	HD NI NN HD HD NN NI NG NG NI NG HD NI NG HD NI HD
GUS ^{G537} -left- TALE7	T CAGCCGATTATCATCACC	HD NI NN HD HD NN NI NG NG NI NG HD NI NG HD NI HD HD
GUS ^{G537} -left- TALE8	T CAGCCGATTATCATCACCG	HD NI NN HD HD NN NI NG NG NI NG HD NI NG HD NI HD HD NN
GUS ^{G537} - right TALE1	T AACGTATCCACG	NI NI HD NN NG NI NG HD HD NI HD NN
GUS ^{G537} - right TALE2	T AACGTATCCACGC	NI NI HD NN NG NI NG HD HD NI HD NN HD
GUS ^{G537} - right TALE3	T AACGTATCCACGCC	NI NI HD NN NG NI NG HD HD NI HD NN HD HD
GUS ^{G537} - right TALE4	T AACGTATCCACGCCG	NI NI HD NN NG NI NG HD HD NI HD NN HD HD NN
GUS ^{G537} - right TALE5	T AACGTATCCACGCCGT	NI NI HD NN NG NI NG HD HD NI HD NN HD HD NN NG
GUS ^{G537} - right TALE6	T AACGTATCCACGCCGTA	NI NI HD NN NG NI NG HD HD NI HD NN HD HD NN NG NI
GUS ^{G537} - right TALE7	T AACGTATCCACGCCGTAT	NI NI HD NN NG NI NG HD HD NI HD NN HD HD NN NG NI NG
GUS ^{G537} - right TALE8	T GCAGCCCGGCTAACGTAT	NN HD NI NN HD HD NN NN HD NG NI NI HD NN NG NI NG
Nb-T1- left TALE	T GCATAGCGCATTA	NN HD NI NG NI NN HD NN HD NI NG NG NI NI
Nb-T1- right TALE	T CACGAATGAGGTTG	HD NI HD NN NI NI NG NN NI NN NN NG NG NN
OsALS-T1- left TALE	T GCTTGCATTTGGCGTGCG	NN HD NG NG NN HD NI NG NG NG NN NN HD NN NG NN HD NN

OsALS-T1-right TALE	TGCAAAAGCCTCAATTTTCCT	NN HD NI NI NI NI NN HD HD NG HD NI NI NG NG NG NG HD HD HD NG
OsALS-T2-left TALE	TGGGTATGGTTGTGCAATG	NN NN NN NG NI NG NN NN NG NG NN NG NN HD NI NI NG NN
OsALS-T2-right TALE	TATGCGCCCTATTTGCCTT	NI NG NN HD NN HD HD HD NG NI NG NG NG NN HD HD NG NG
OsPDS-left TALE	TCCTGGTAAATAGGTAGT	NG HD HD NG NN NN NG NI NI NI NG NI NN NN NG NI NN NG
OsPDS-right TALE	TGCTCAAAGCCAGCAGTCA	NN HD NG HD NI NI NI NN HD HD NI NN HD NI NN NG HD NI
OspsaA-left TALE	TAGTCATTTTCATTTTCAG	NI NN NG HD NI NG NG NG NG HD HD NI NG NG NG HD NI NN
OspsaA-right TALE	TATAGTACCCCAAACATC	NG HI NG NI NN NG NI HD HD HD HD NI NI NI HD NI NG HD
NbSuRB-left TALE	TGTCATGCTGGGTC	NN NG HD NI NG NN HD NG NN NN NN NG HD
NbSuRB-right TALE	TCCAACCTAAAGAGG	HD HD NI NI HD HD NG NI NI NI NN NI NN NN
unspecific-TALE	TCACGAATGAGGTTG	HD NI HD NN NI NI NG NN NI NN NN NG NG NN

Table S5. Oligos used in this study.

Name	Sequence 5' - 3'	Application
NB1-T1-AP-F1	GATATTTATGTTTGTTC	Amplify NB-T1
NB1_R	CCTAGGACTCGAGAGAAAGCTGCAAATC	
OsALS-T1-AP-F1	GTGATGATCCGTTGTCCC	Amplify OsALS-T1
OsALS-T1-AP-R1	GCAAGCTTAACATCTGCG	
OsALS-T2-AP-F1	TGCAGGAATATTGAACCC	Amplify OsALS-T2
OsALS-T2-AP-R1	TATTGATGGGGATGGTAG	
OsPDS-AP-F1	CGTATCTGTATTCCGTAG	Amplify <i>OsPDS</i>
OsPDS-AP-R1	ACCTCGTTTGCCTCCTC	
PsA-F	GCTCCTGGTGCAACAACAAG	Amplify <i>OspsaA</i>
PsA-R	AGGCCAGACAAAATGAGCA	
PsA-AP-F	TCTTGGTTTTTCGCTTCCCTT	Amplify <i>OspsaA</i> for NGS
PsA-AP-R	AATTACTTGGGATGCCTGTG	
NbSURB-F	GCAGAGCTAGTTCTAGGCGTC	Amplify <i>NbSURB</i>
NbSURB-R	GGTTCTGGTTTATTGATAGAGCTGC	
NB-OT1-F	TTTTCAGCCGTGGAGGTCTC	Amplify <i>NbSURB</i> predicted off-target site 1
NB-OT1-R	TTGCCAGTGACGATCGAACA	
NB-OT2-F	ATTGAGGCCACCACCATGTT	Amplify <i>NbSURB</i> predicted off-target site 2
NB-OT2-R	TTCCCTCAAGCTCCGTCAAC	
NB-OT3-F	GGGAAACGCACCCATAGTCA	Amplify <i>NbSURB</i> predicted off-target site 3
NB-OT3-R	AAGAGTGAGGGCAATGGACG	
NB-OT4-F	ATGTCGGGGCAAGAAAGGAG	Amplify <i>NbSURB</i> predicted off-target site 4
NB-OT4-R	TGCTCGCCCATTAGTGTGTT	
NB-OT5-F	GAGGCTGCCACCATGTTTTG	Amplify <i>NbSURB</i> predicted off-target site 5
NB-OT5-R	GGTCGCACATGTCAAGCATC	

Table S6. Alignment of amplicon sequencing results.

Aligned result of amplicon sequencing results			
	Aligned sequences		
<i>OsPDS</i> reference	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	Reads	Reads (%)
<i>OsPDS</i> (DddA) Repetition 1	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	4825	100
<i>OsPDS</i> (DddA) Repetition 2	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	14709	99.117
<i>OsPDS</i> (DddA) Repetition 3	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	14356	99.363
<i>OsPDS</i> (DddA11) Repetition 1	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	3420	95.265
	CCTGGTAAATAGGTAGTAGGAGAGATGTTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	57	1.588
	CCTGGTAAATAGGTAGTAGGAATAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	7	0.195
	CCTGGTAAATAGGTAGTAGAATAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	6	0.167
	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAATGACTGCTGGCTTTGAGCATGCGGCAT	4	0.111
	CCTGGTAAATAGGTAGTAGGAGAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	3	0.084
<i>OsPDS</i> (DddA11) Repetition 2	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	10636	97.221
	CCTGGTAAATAGGTAGTAGGAGAGATGTTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	49	0.448
	CCTGGTAAATAGGTAGTAGGAATAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	8	0.073
	CCTGGTAAATAGGTAGTAGGAGAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	4	0.037
	CCTGGTAAATAGGTAGTAGAATAATGTTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	1	0.009
	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAATGACTGCTGGCTTTGAGCATGCGGCAT	1	0.009
<i>OsPDS</i> (DddA11) Repetition 3	CCTGGTAAATAGGTAGTAGGAGAGATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	12194	95.895
	CCTGGTAAATAGGTAGTAGGAGAGATGTTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	170	1.337
	CCTGGTAAATAGGTAGTAGGAATAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	34	0.267
	CCTGGTAAATAGGTAGTAGAATAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	13	0.102
	CCTGGTAAATAGGTAGTAGGAGAATGCTGAGATGACTGCTGGCTTTGAGCATGCGGCAT	10	0.079

	CCTGGTAAATAGGTAGTAGGAGAGATGCTGA A ATGACTGCTGGCTTTGAGCATGCGGCAT	1	0.008
OsALS-T1 reference	TGTT GCTTGCATTGGCGTGCG TTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	Reads	Reads (%)
OsALS-T1 (DddA)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	5043	93.216
Repetition 1	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	108	1.996
OsALS-T1 (DddA)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	5174	93.141
Repetition 2	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	158	2.844
OsALS-T1 (DddA)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	7477	92.663
Repetition 3	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	275	3.408
OsALS-T1 (DddA11)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	4067	95.291
Repetition 1	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	124	2.905
OsALS-T1 (DddA11)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	14381	95.937
Repetition 2	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	352	2.348
OsALS-T1 (DddA11)	TGTTGCTTGCATTGGCGTGCGGTTTGATGATCGTGTGACAGGGAAAATTGAGGCTTTTG	15817	94.741
Repetition 3	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT T GTGTGACAGGGAAAATTGAGGCTTTTG	569	3.408
	TGTTGCTTGCATTGGCGTGCGGTTTGATGAT A GTGTGACAGGGAAAATTGAGGCTTTTG	24	0.144
OsALS-T2 reference	GGTATGGTTGTGCAATGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	Reads	Reads (%)
OsALS-T2 (DddA)	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	4079	99.585
OsALS-T2 (DddA)	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	6819	99.810
OsALS-T2 (DddA)	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	7443	99.612
Repetition 3	GGTATGGTTGTGCAATGGGAGGATA A GTTTTACAAGGCAAATAGGGCGCATACATACTTG	2	0.0003
OsALS-T2 (DddA11)	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	5252	93.336
Repetition 1	GGTATGGTTGTGCAATGGGAGGATA A ATTTTACAAGGCAAATAGGGCGCATACATACTTG	233	4.141
	GGTATGGTTGTGCAATGGGAG A AATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	13	0.231
	GGTATGGTTGTGCAATGGGAGGATA A GTTTTACAAGGCAAATAGGGCGCATACATACTTG	9	0.160

OsALS-T2 (DddA11) Repetition 2	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	18009	97.7104
	GGTATGGTTGTGCAATGGGAGGATAAATTTTACAAGGCAAATAGGGCGCATACATACTTG	116	0.629
	GGTATGGTTGTGCAATGGGAGAAATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	3	0.0163
	GGTATGGTTGTGCAATGGGAGGATATGTTTACAAGGCAAATAGGGCGCATACATACTTG	2	0.011
OsALS-T2 (DddA11) Repetition 3	GGTATGGTTGTGCAATGGGAGGATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	20117	95.143
	GGTATGGTTGTGCAATGGGAGGATAAATTTTACAAGGCAAATAGGGCGCATACATACTTG	602	2.847
	GGTATGGTTGTGCAATGGGAGAAATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	22	0.104
	GGTATGGTTGTGCAATGGGAGTATAGGTTTACAAGGCAAATAGGGCGCATACATACTTG	7	0.0331
	GGTATGGTTGTGCAATGGGAGGATAGTTTACAAGGCAAATAGGGCGCATACATACTTG	2	0.009
	GGTATGGTTGTGCAATGGGAGGATAGATTTTACAAGGCAAATAGGGCGCATACATACTTG	2	0.009
	GGTATGGTTGTGCAATGGGAGGATAAATTTTACAAGGCAAATAGGTCGCATACATACTTG	2	0.009
NB-T1 reference	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	Reads	Reads (%)
NB-T1 (DddA) Repetition 1	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	16230	95.297
	AGTTGCATAGCGCATTAATTGCGTGATTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	386	2.266
	AGTTGCATAGCGCATTAATTGCGTGATATGCAACCTCATTCGTGAAAAAGATCTTCTCCA	16	0.094
	AGTTGCATAGCGCATTAATTGCGTAATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	11	0.065
NB-T1 (DddA) Repetition 2	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	8557	95.036
	AGTTGCATAGCGCATTAATTGCGTGATTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	228	2.532
	AGTTGCATAGCGCATTAATTGCGTAATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	1	0.011
NB-T1 (DddA) Repetition 3	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	13801	94.748
	AGTTGCATAGCGCATTAATTGCGTGATTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	355	2.437
NB-T1 (DddA6) Repetition 1	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	13071	95.256
	AGTTGCATAGCGCATTAATTGCGTGATTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	355	2.587
	AGTTGCATAGCGCATTAATTGCGTAATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	9	0.066

NB-T1 (DddA6) Repetition 2	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	14799	95.207
	AGTTGCATAGCGCATTAATTGCGGTATTTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	496	3.191
	AGTTGCATAGCGCATTAATTGCGTAAATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	4	0.026
NB-T1 (DddA6) Repetition 3	AGTTGCATAGCGCATTAATTGCGTGATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	8837	95.032
	AGTTGCATAGCGCATTAATTGCGGTATTTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	322	3.463
	AGTTGCATAGCGCATTAATTGCGTAAATCTGCAACCTCATTCGTGAAAAAGATCTTCTCCA	3	0.032

Supplemental Sequences. Amino acid sequences of TALE-DdCBE architectures.

TALE-DdCBE (from N- to C-terminus):

bpNLS-N-terminal domain-TALE repeats-C-terminal domain-GS linker-DddA-N/C half-SGGS linker-UGI-NLS

bpNLS:

MKRTADGSEFESPCKKRKV

N-terminal domain

N288:

MDPIRSRTPSPARELLSGPQPDGVQPTADRGVSPAGGPLDGLPARRTMSRTRLPSPPAPSPAFA
ADSFSDLRLQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARP
PRAKPAPRRRAAQPSDASPAQVDLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVA
LSQHPAALGTAVVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQL
LKIAKRGGVTAVEAVHAWRNALTGAPLN

N196:

MAHHTEAATGEWDEVQSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAQVDLRT
LGYSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTAVVKYQDMIAALPEATH
EAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN

N135:

DLRTLGYSSQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTAVVKYQDMIAALP
EATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHAWRNALT
GAPLN

C-terminal domain

C17:

SIVAQLSRPDPSLAALT

C28:

SIVAQLSRPDPSLAALTNDHLVALACLG

C47:

SIVAQLSRPDPSLAALTNDHLVALACLGGRPALDAVKKGLPHAPALI

C63:

SIVAQLSRPDPSLAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVA

C96:

SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADH
AQVVRVLGFFQCHSHPAQAFDDAMTQFGMS

C247:

SIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPERTSHRVADH
AQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGLTPPASQRWDRILO
ASGMKRAKPSPTSTQTPDQASLHAFADSLERDLAPSPMHEGDQTRASSRKRSDRAVTGPSAQ
QSFEVVRVPEQRDALHLPLLSWGVKRPRTTRIGLLDPGTPMDADLVASSTVVW

DddA-N:

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFSSGGPTPYPNYANAGHVEGQSALFM
RDNGISEGLVFHNNPEGTCGFCVNMTETLLPENAKMTVVPPEG

DddA-C:

AIPVKRGATGETKVFTGNSNSPKSPTKGGCSGGS

DddA6-N:

GSYALGPYQISAPQLPAYNGRTVGTFYYVNDAGGLESKVFISGGPTPYPNYANAGHVEGQSALFM
RDNGISEGLVFHNNPEGTCGFCVNMIETLLPENAKMTVVPPEG

DddA6-C / DddA6^{K1389A}-C / DddA6^{T1391A}-C:

AIPVKRGATGETKVFIGNSNSPKSPTKGGCSGGS

DddA6^{K1389A}-N:

GSYALGPYQISAPQLPAYNGRTVGTFYYVNDAGGLESKVFISGGPTPYPNYANAGHVEGQSALFM
RDNGISEGLVFHNNPEGTCGFCVNMIETLLPENAAMTVVPPEG

DddA6^{T1391A}-N:

GSYALGPYQISAPQLPAYNGRTVGTFYYVNDAGGLESKVFISGGPTPYPNYANAGHVEGQSALFM
RDNGISEGLVFHNNPEGTCGFCVNMIETLLPENAKMAVVPPEG

DddA11-N:

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFISGGPTPYPNYVSAGHVEGQSALFM
RDNGISEGLVFHNNPKGTCGFCVNMIETLLPENAKMTVVPPEG

DddA11-C / DddA11^{K1389A}-C / DddA11^{T1391A}-C:

AIPVKRGATGETKVFIGNSNSPKSPTKGGCSGGS

DddA11^{K1389A}-N:

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFISGGPTPYPNYVSAGHVEGQSALFM
RDNGISEGLVFHNNPKGTCGFCVNMIETLLPENAAMTVVPPEG

DddA11^{T1391A}-N:

GSYALGPYQISAPQLPAYNGQTVGTFYYVNDAGGLESKVFISGGPTPYPNYVSAGHVEGQSALFM
RDNGISEGLVFHNNPKGTCGFCVNMIETLLPENAKMAVVPPEG

UGI-NLS:

TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYK
PWALVIQDSNGENKIKMLSGGSPKKKRKV

Chloroplast transition peptide:

MAPSVMASSATTVAPFQGLKSTAGMPVARRSGNSSFGNVSNNGGRIRCMQVWPIEGIKKFETLSYL
PPLGNSSFGNVSNNGGRIRC