

BIOLOGICAL SCIENCES
Plant Biology

Precise translocation and duplication using CRISPR-Cas9 in *Arabidopsis thaliana*

Authors: Peter G. Lynagh¹, Soichi Inagaki^{1,2}, Kirk R. Amundson¹, Mohan P.A. Marimuthu¹, Brett Randolph Pike¹, Isabelle M. Henry¹, Ek Han Tan^{1,3} and Luca Comai¹

UC Davis Genome Center and Department of Plant Biology, Davis, California¹
National Institute of Genetics, Mishima, Japan²
School of Biology and Ecology, University of Maine, Orono, Maine³

Corresponding Author: Luca Comai

Author Contributions:

P.G.L., S.I., K.R.A., B.R.P. and E.H.T. developed genetic constructs. M.P.A.M. designed and conducted the fluorescent microscopy experiment and localized RPS5A-tD Tomato expression. P.G.L., S.I., I.M.H. and E.H.T. genotyped plants. P.G.L., S.I. and E.H.T. phenotyped plants. P.G.L. conducted the translocation and duplication experiments. P.G.L., I.M.H. and L.C. designed most of the research, analyzed data and wrote the paper.

Keywords: Plant, Arabidopsis, CRISPR, Genome, Recombination, NHEJ, Translocation, Duplication

Supplementary Information

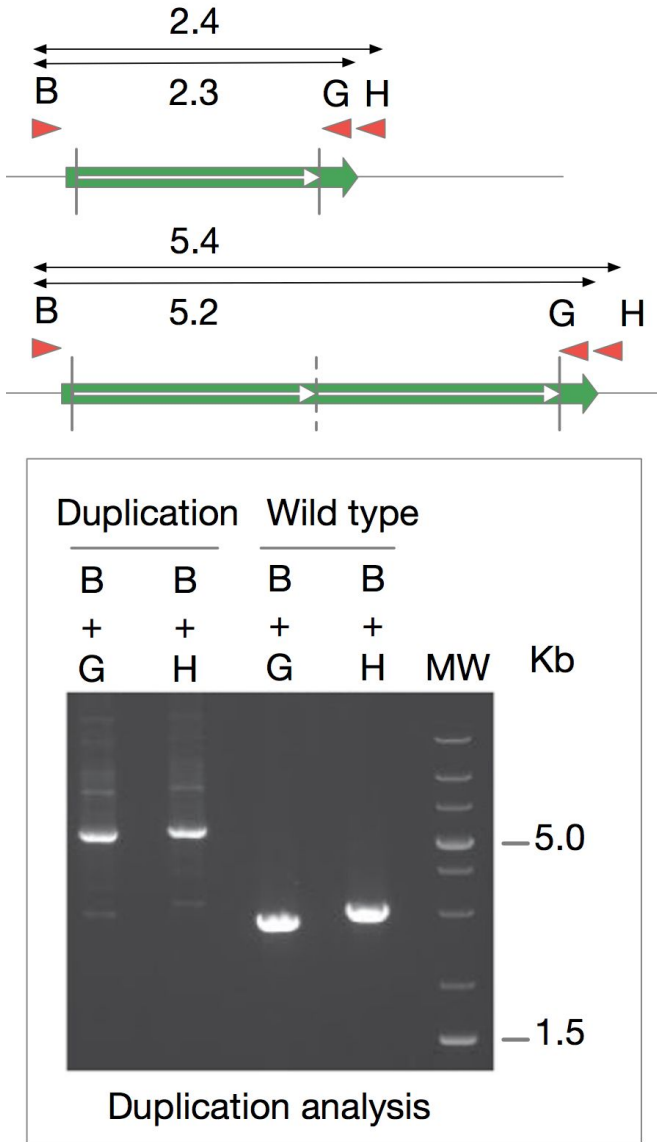


Figure S1. PCR for CH1 tandem duplication was performed using two combinations of primers (B + G and B + H). Using leaf DNA from T2 plant #9 from T1 plant #24 as a template yields bands that suggest a duplication of the targeted segment.



Figure S2: Continued:

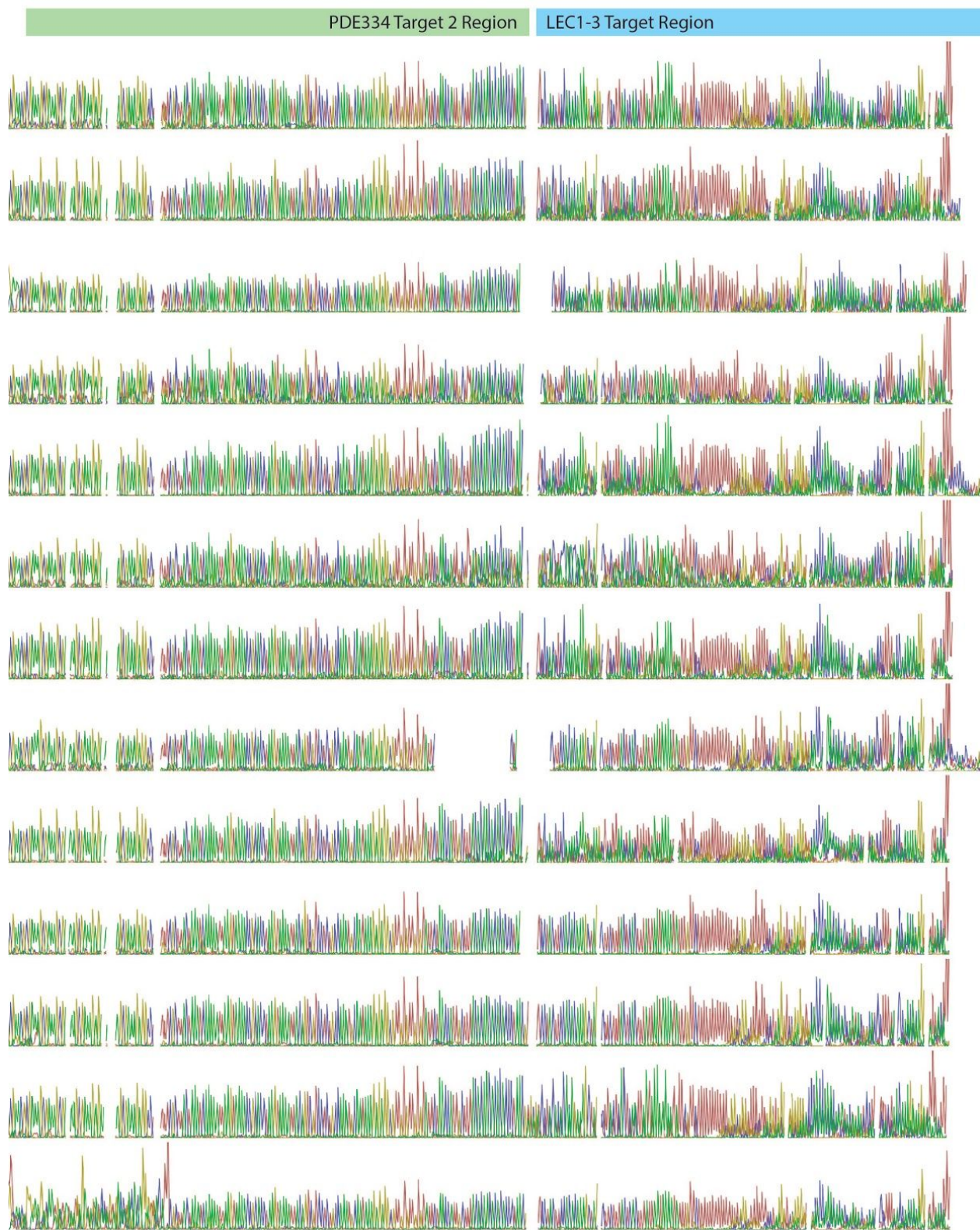


Figure S2. Plants received sgRNAs with targets *PDE334-1*, *PDE334-2* and *LEC1-3*. Translocation junctions detected using primers A and C from Figure S2-B. The chromatograms were aligned using Geneious Alignment. Each chromatogram is from a different primary transformant. Thus, each chromatogram represents independent events. The 10 alleles displayed in Figure 8 were generated with the chromatograms. This orientation could possibly cause increased *LEC1* expression.

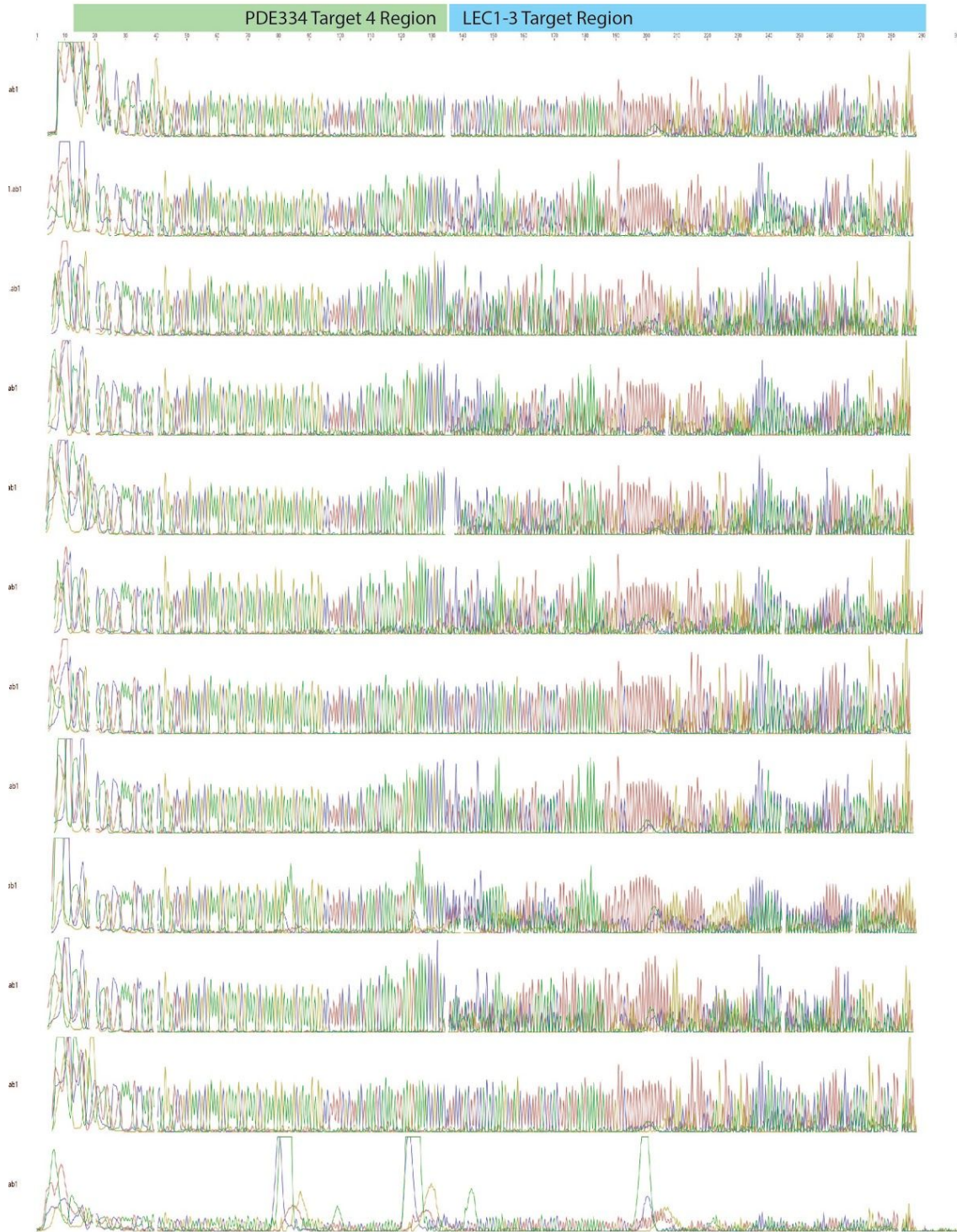


Figure S3. Plants received sgRNAs with targets *PDE334-1*, *PDE334-4* and *LEC1-3*. Translocation junctions detected using primers A and C from Figure S2-B. The chromatograms were aligned using Geneious Alignment. Each chromatogram is from a different primary transformant. Thus, each chromatogram represents an independent event. This orientation could possibly cause increased *LEC1* expression.

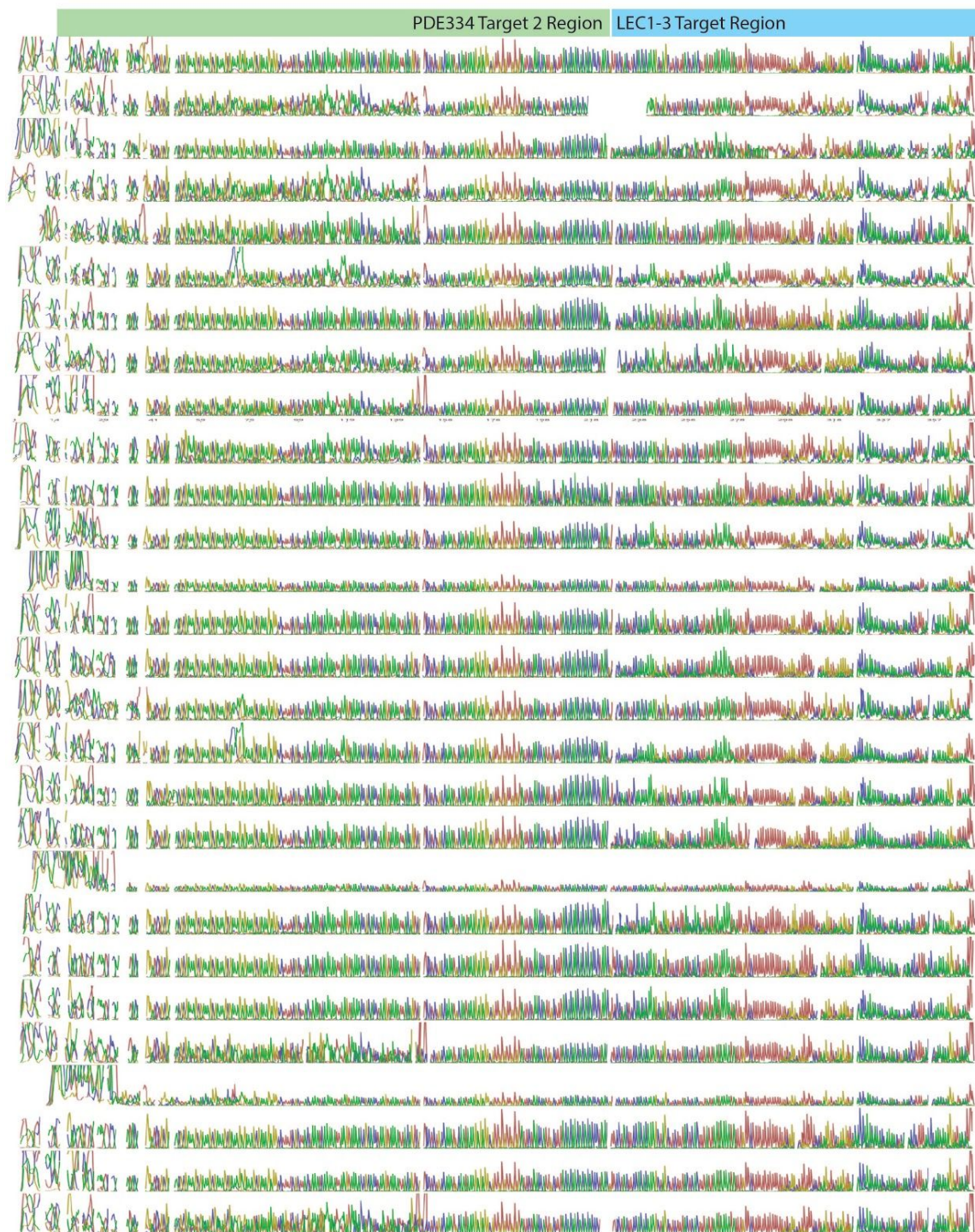


Figure S4. Plants received sgRNAs with targets *PDE334-3*, *PDE334-2* and *LEC1-3*. Translocation junctions detected using primers A and C from Figure S2-B. The chromatograms were aligned using Geneious Alignment. Each chromatogram is from a different primary transformant. Thus, each chromatogram represents an independent event. This orientation could possibly cause increased *LEC1* expression.

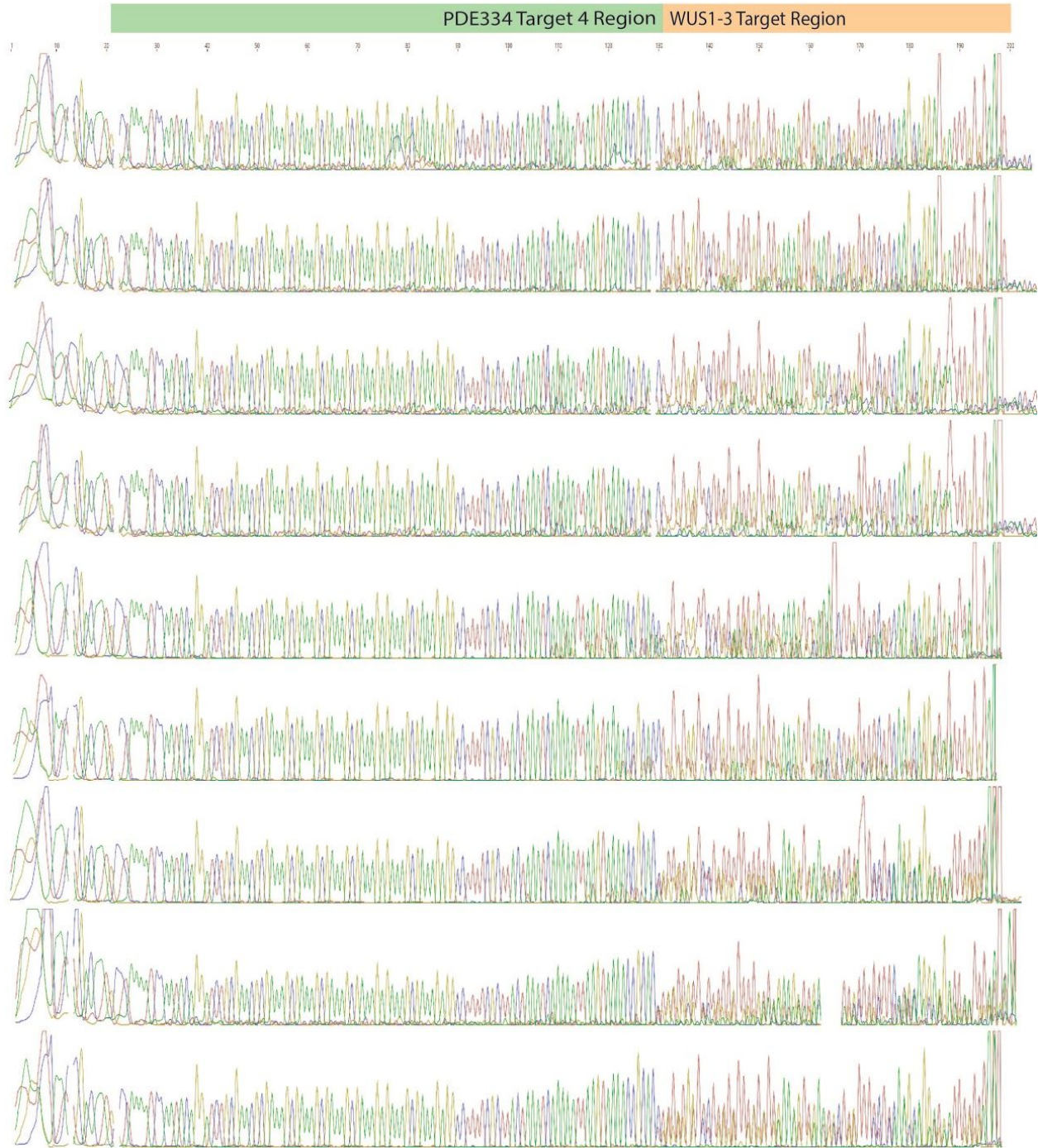


Figure S5. Plants received sgRNAs with targets *PDE334-1*, *PDE334-4* and *Wus1-1*. Translocation junctions detected using primers A and F from Figure S2-B. The chromatograms were aligned using Geneious Alignment. Each chromatogram is from a different primary transformant. Thus, each chromatogram represents an independent event. This orientation is not expected to increase *WUS1* expression.

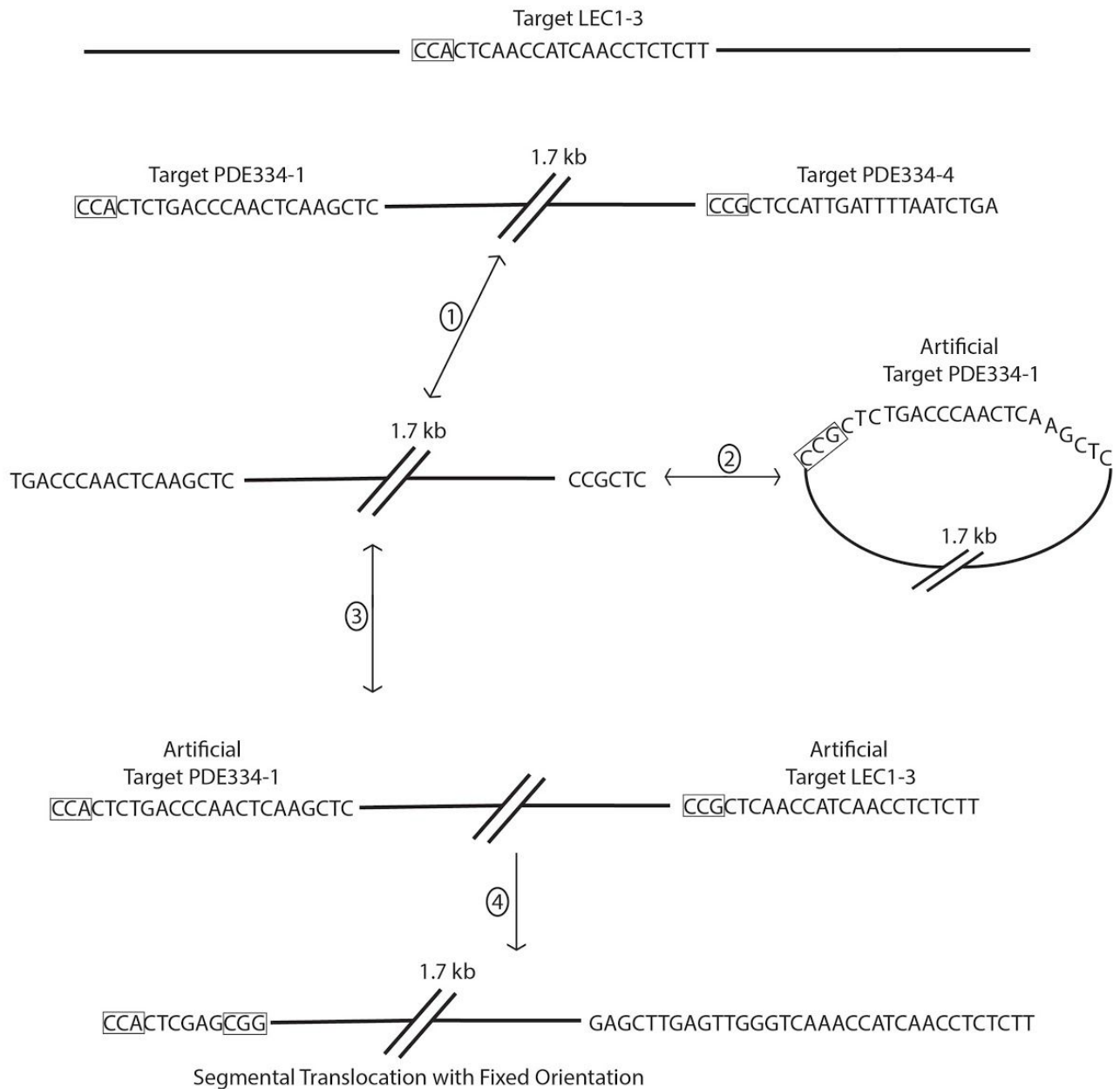


Figure S6. Example of recutting perfect junctions in attempt to control translocation orientation. Step #1 excises the segment. Step #2 is putative circularization of the segment, which reforms the PDE334-1 target sequence, which would be re-cut to form the linear fragment. Step #3 is a translocation of the linear segment into the original cut LEC1-3 locus. Step #4 forms an inversion of the inserted segment, and the resulting DNA contains no Cas9 targets. However, this approach made no detectable difference in translocation orientation or frequency.

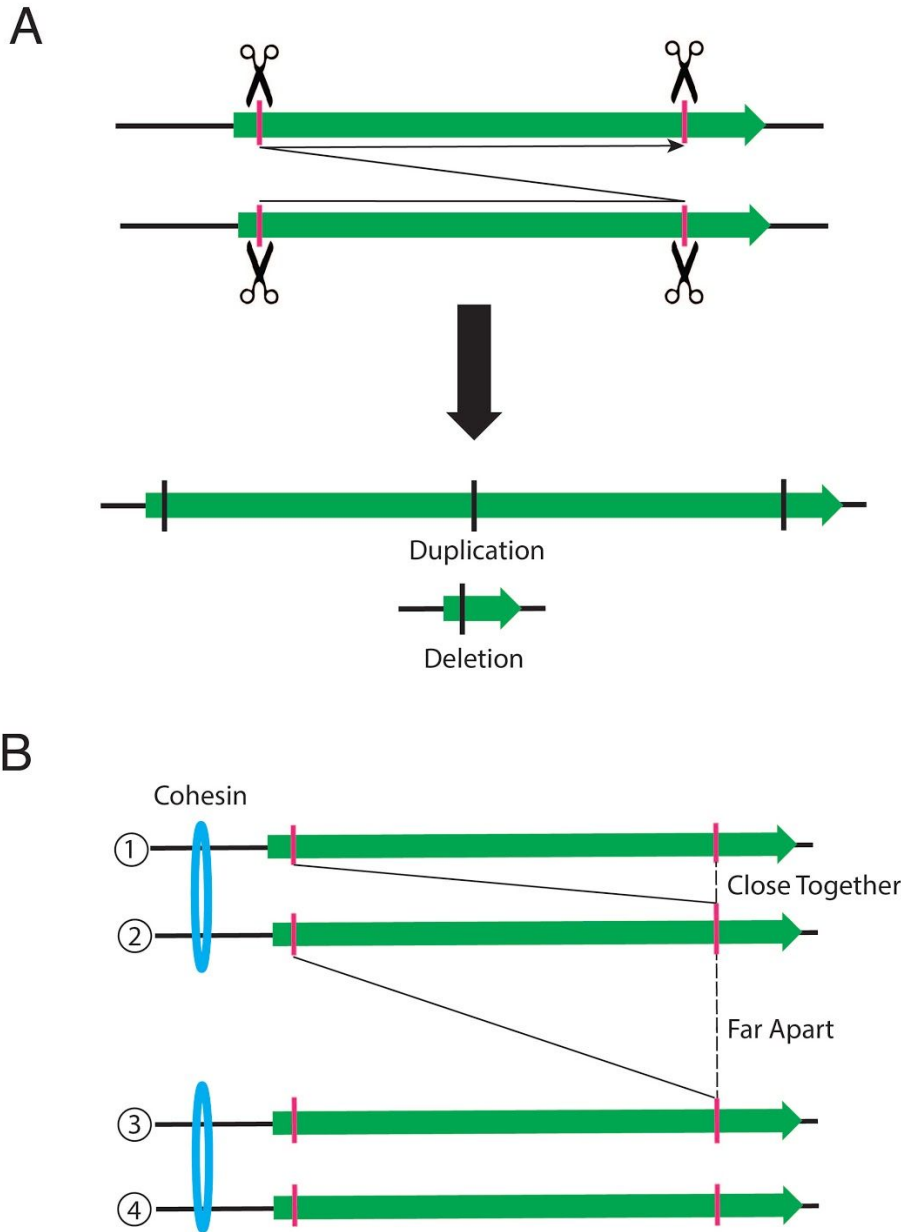


Figure S7. A. Simple translocations between homologs in G1 phase of the cell cycle result in a duplication and a deletion allele. There is no longer a Wild Type allele. B. The products of a simple translocation during G2 phase of the cell cycle depends on which chromatids are involved in the translocation. If a translocation occurs between chromatid #1 and chromatid #2 (sister chromatids), the duplication and deletion will move into separate cell lineages, as hypothesized to have occurred in Line #24 (Figure 4). If a translocation occurs between

chromatid #2 and a deletion in chromatid #3 (non-sister chromatids), the duplication and deletion alleles can possibly go to the same daughter cell. Translocations between sister chromatids are hypothesized to occur most frequently due to the close physical proximity between sister chromatids.

Table S1. CRISPR-Cas9 targets

Target	Target gene, chromosomal coordinates (TAIR10)	gRNA target + PAM	Construct Backbone	gRNA	Strategy
BRI1	At4g39400, Chr4:18326878	TTGGGTCATAACGATATCTCTGG	RC9 ¹	ProU6-26 + gRNA	Synthetic activation
NG1	Intergenic, Chr2:13228605	GGCGGCGTTCATGGCGTCACCGG	RC9 in pPLV02	(ProU6-26 + gRNA)*3	One vector
NG2	Intergenic, Chr3:16075544	AGGTCTTAAACGAAGATCTACGG			
NG3	Intergenic, Chr5:4065824	GATAAGATATAGAAGGTGTAGGG			
CH1-1	At1G44446, Chr1:16851141	TTTTCTTTTAGAGGGGAAGGGGG	RC9 in pEG302v2	(ProU6-26 + gRNA)*2	One vector
CH1-2	At1G44446, Chr1:16848808	ACGTGTTGGTGCGCGAGAGGTGG			
HB21	At2G18550, Chr2:8051138	CTAATGGAGTATAGACATTAGGG	RC9	ProU6-26 +(tRNA+gRNA)*3	Synthetic activation
HB40	At4G36740, Chr4:17315627	GTGGAAGAACAAAAGGCTCGAGG			
HB53	At5G66700, Chr5:26634496	ATCGTTAACCAGATCGATGGAGG			
PDE334-1	At4G32260, Chr4:15576459	GAGCTTGAGTTGGGTCAGAGTGG	RC9 in pEG302v2	(ProU6-26 + gRNA)*3	One vector, 3 sgRNA per vector
PDE334-2	At4G32260, Chr4:15574653	GTAATCTCTTTGGAGACGAATGG			
PDE334-3	At4G32260, Chr4:15576497	CATGGGAAGAGGATGAGGAAGGG			
PDE334-4	At4G32260, Chr4:15574747	TCAGATTAATCAATGGAGCGG			
RABE1B-1	At4G20360, Chr4:10990031	CCGAAATCGCCATGGGAAGATGG			
RABE1B-2	At4G20360, Chr4:10988071	AGATCAATGGCCGATGGAGAAGG			
RABE1B-3	At4G20360, Chr4:10988068	TCAATGGCCGATGGAGAAGGAGG			
WUS1-1	At2G17950, Chr2:7810727	AATAAGATCAAGACTTAGAGAGG			
LEC1-3	At1G21970, Chr1:7729658	AAGAGAGGTTGATGGTTGAGTGG			
LEC1-1	At1G21970, Chr1:7724042	TGACGCCGTTAGAAAGCGTACGG			
LEC1-2	At1G21970, Chr1:7732513	CCTCAAAGAGTAGTAGGAACAGG	RC9 in pEG302v2	(ProU6-26 + gRNA)*2	One vector

¹RC9 = proRPS5A-Cas9

Table S2. Summary of mutation efficiencies in T1 and T2 plants

Target	Gene target, Coordinates on TAIR10	Phenotype (T1) ¹		Mutation rate by sequencing						
				Leaf ² (T1)				Germline ³ (T2)		
		Biallelic knockout trait	Rate		Test		Control ³		% n	% n
%	n		%	n	%	n				
BRI1	<i>At4g39400</i> , Chr4:18326878	Dwarf	41.2	97	ND					
NG1	Intergenic, Chr2:13228605	None	N/A		99.1	18	0.6	1	90.9	11
NG2	Intergenic, Chr3:16075544				32.4	12	0.5	1	31.5	9
NG3	Intergenic, Chr5:4065824				96.8	18	1.1	1	88.9	10
CH1-I	<i>At1G44446</i> , Chr1:16851141	Yellow leaf	45.7	35	45.2	31	2.0	1	ND	
CH1-II	<i>At1G44446</i> , Chr1:16848808				55	31	1.6	1		
HB21	<i>At2G18550</i> , Chr2:8051138	Unknown	N/A		33	45	0	1	ND	
HB40	<i>At4G36740</i> , Chr4:17315627				56	43	0	1		
HB53	<i>At5G66700</i> , Chr5:26634496				60	46	0	1		

1: T1=primary transformant from flower dip, T2=selfed progeny of T1

2: Illumina Ampliseq from T1 plant

3: Sanger sequencing of amplicon from Cas9-negative T2

4: Wild-type leaf mutation rate in protospacer; i.e. noise.

n = Number of plants. All plants are primary transformants, except the control.

ND = no data

N/A = not applicable

Table S3. Number of germline mutations in T2 plants following targeting of Cas9 to three intergenic targets.

T1 ID#	21	22	26	27	36	38	41	42	43	47	57	Average (%)
# Cas9-Negative T2 Plants	3	1	1	4	1	1	3	3	1	1	3	2.3
NG1*	6/6	2/2	0/2	6/6	2/2	2/2	6/6	4/4	2/2	2/2	6/6	90.9
NG2*	N/A	0/2	0/2	3/6	0/2	N/A	0/4	4/6	2/2	0/2	4/6	31.5
NG3*	N/A	2/2	2/2	8/8	2/2	0/2	6/6	4/4	2/2	2/2	4/4	88.9

*Number of mutated alleles out of the total number of alleles tested.

Table S4. Phenotypic and PCR product classes of *CH1* CRISPR-Cas9 transgenic Arabidopsis

T2 family	Number of individuals	Phenotype	Duplication	Deletion
2	4	Wild type	-	-
3	3	Pale yellow	-	-
6	6	Pale yellow	-	+
6	1	Pale yellow	-	-
9	11	Pale yellow	-	+
10	1	Pale yellow	-	-
10	6	Wild type	-	-
11	1	Wild type	+	+
11	1	Wild type	-	+
11	5	Wild type	-	-
12	7	Wild type	-	-
13	1	Pale yellow	+	+
13	1	Pale yellow	+	-
13	1	Pale yellow	-	-
13	3	Wild type	-	-
16	2	Pale yellow	-	-
16	5	Wild type	-	-
19	2	Pale yellow	+	+
19	1	Pale yellow	-	+
19	1	Pale yellow	-	-
19	3	Wild type	-	-
20	4	Pale yellow	-	+
20	2	Wild type	+	+
20	1	Wild type	-	+
21	1	Pale yellow	-	+
21	4	Wild type	-	+
22	3	Pale yellow	+	+
22	1	Wild type	-	+
22	1	Wild type	-	-
23	2	Wild type	+	+
23	1	Wild type	+	-
23	4	Wild type	-	-
24	6	Pale yellow	+	-

24	6	Pale yellow	-	-
24	1	Wild type	+	-
25	1	Pale yellow	+	-
25	1	Pale yellow	-	-
25	1	Wild type	+	+
25	4	Wild type	-	-
26	4	Pale yellow	+	+
26	1	Pale yellow	-	+
26	1	Pale yellow	-	-
27	2	Pale yellow	-	-
27	5	Wild type	-	-
28	1	Wild type	+	+
28	6	Wild type	-	-
29	1	Pale yellow	-	-
29	2	Wild type	+	-
29	3	Wild type	-	-
30	1	Pale yellow	-	-
30	4	Wild type	+	+
30	1	Wild type	+	-
30	1	Wild type	-	+
31	1	Chimeric	+	+
31	1	Pale yellow	-	-
33	1	Wild type	+	+
33	1	Wild type	-	+
33	5	Wild type	-	-
34	2	Wild type	-	+
34	5	Wild type	-	-
35	3	Wild type	+	+
35	2	Wild type	+	-
35	2	Wild type	-	-

Table S5. PCR for translocations between NG1, NG2 and NG3

		PCR for Translocation Detection (+ or -)											
		Monocentric						Dicentric			No Centromere		
	Primers	2F,3R	3F,2R	2R,5R	3R,5R	2F,5F	3F,5F	2F,5R	2F,3F	3F,5R	5F,2R	5F,3R	2R,3R
	Expected Band Size (bp)	446	530	687	682	227	306	368	211	447	546	541	765
T1 plant #	PCR for cas9 (+ or -)												
1	+	-	-	-	-	-	-	-	-	-	-	-	-
2	+	-	-	-	-	-	-	-	-	-	-	-	-
9	+	-	-	-	-	-	-	-	-	-	-	-	-
10	+	-	-	-	-	-	-	-	-	-	-	-	-
12	+	-	-	-	-	-	-	-	-	-	-	-	-
13	+	-	-	-	-	-	-	-	-	-	-	-	-
14	+	-	-	-	-	-	-	-	-	-	-	-	-
17	+	-	-	-	-	-	-	-	-	-	-	-	-
20	+	-	-	-	-	-	-	-	-	-	-	-	-
21	+			-	-	-	-	-	-	-	-	-	-
22	+			-	-	-	-	-	-	-	-	-	-
23	+	-	-	-	-	-	-	-	-	-	-	-	-
26	+	-	-	-	-	-	-	-	-	-	-	-	-
27	+	-	-	+	-	+	-	-	-	-	-	-	-

29	+	-	-	-	-	-	-	-	-	-	-	-	-
31	+	-	-	-	-	-	-	-	-	-	-	-	-
32	+	-	-	-	-	-	-	-	-	-	-	-	-
34	+	-	-	-	-	-	-	-	-	-	-	-	-
35	+			-	-	-	-	-	-	-	-	-	-
36	+			-	-	-	-	-	-	-	-	-	-
38	+			+	-	-	-	+	-	-	-	-	-
41	+			-	-	-	-	-	-	-	-	-	-
42	+									-	-		
43	+									-	-		
47	+									-	-		
48	+									-	-		
55	+			-	-	-	-	-	-	-	-	-	-
56	+			-	-	-	-	-	-	-	-	-	-
57	+			-									
60	+			-	-	-	-	-	-	-	-	-	-
61	+			-	-	-	-	-	-	-	-	-	-
62	+			-	-	-	-	-	-	-	-	-	-
63	+			-	-	-	-	-	-	-	-	-	-
66	+			-	-	-	-	-	-	-	-	-	-
67	+			-	-	-	-	-	-	-	-	-	-
		0/16	0/16	2/31	0/3	1/30	0/30	1/3	0/30	0/34	0/34	0/30	0/3

Table S7. List of oligonucleotides

CH1-I_f	attgGGCGGCGTTCATGGCGTCAC	sgRNA building
CH1-I_r	aaacGTGACGCCATGAACGCCGCC	sgRNA building
CH1-II_f	attgAGGTCTTAAACGAAGATCTA	sgRNA building
CH1-II_r	aaacTAGATCTTCGTTTAAAGACCT	sgRNA building
NG1_f	attgGATAAGATATAGAAGGTGTA	sgRNA building
NG1_r	aaacTACACCTTCTATATCTTATC	sgRNA building
NG2_f	attgGTTTCTTTTAGAGGGGAAGG	sgRNA building
NG2_r	aaacCCTTCCCCTCTAAAAGAAAC	sgRNA building
NG3_f	attgGCGTGTTGGTGC GCGAGAGG	sgRNA building
NG3_r	aaacCCTCTCGCGCACCAACACGC	sgRNA building
LEC1-1_f	attgTGACGCCGTTAGAAAGCGTA	sgRNA building
LEC1-1_r	aaacTACGCTTTCTAACGGCGTCA	sgRNA building
LEC1-2_f	attgCCTCAAAGAGTAGTAGGAAC	sgRNA building
LEC1-2_r	aaacGTTCTACTACTCTTTGAGG	sgRNA building
LEC1-3_f	attgAAGAGAGGTTGATGGTTGAG	sgRNA building
LEC1-3_r	aaacCTCAACCATCAACCTCTCTT	sgRNA building
PDE334-1_f	attgGAGCTTGAGTTGGGTCAGAG	sgRNA building
PDE334-1_r	aaacCTCTGACCCA ACTCAAGCTC	sgRNA building
PDE334-2_f	attgGTAATCTCTTTGGAGACGAA	sgRNA building
PDE334-2_r	aaacTTCGTCTCCAAAGAGATTAC	sgRNA building
PDE334-3_f	attgCATGGGAAGAGGATGAGGAA	sgRNA building
PDE334-3_r	aaacTCCTCATCCTCTTCCCATG	sgRNA building
PDE334-4_f	attgTCAGATTAAAATCAATGGAG	sgRNA building
PDE334-4_r	aaacCTCCATTGATTTTAAATCTGA	sgRNA building
RABE1B-1_f	attgCCGAAATCGCCATGGGAAGA	sgRNA building
RABE1B-1_r	aaacTCTTCCCATGGCGATTTCCGG	sgRNA building
RABE1B-2_f	attgAGATCAATGGCCGATGGAGA	sgRNA building

RABE1B-2_r	aaacTCTCCATCGGCCATTGATCT	sgRNA building
RABE1B-3_f	attgTCAATGGCCGATGGAGAAGG	sgRNA building
RABE1B-3_r	aaacCCTTCTCCATCGGCCATTGA	sgRNA building
WUS1-1_f	attgAATAAGATCAAGACTTAGAG	sgRNA building
WUS1-1_r	aaacCTCTAAGTCTTGATCTTATT	sgRNA building
RPS5aPro_fwd	TCTAGTTGGAATGGGTCCATAATCGTGAGTAGATATATTACTC	for gibson making RPS5a-Cas9
RPS5aPro_rev	cttatccatGGCTGTGGTGAGAGAAAC	for gibson making RPS5a-Cas9
Cas9HP_fwd	ccacagccATGGATAAGAAGTACTCTATCGG	for gibson making RPS5a-Cas9
Cas9HP_rev	ccttatggagttgggttTCAAACCTTCCTCTTCTTC	for gibson making RPS5a-Cas9
2x_BsaI_f	attgGGGTCTTCGAGAAGACCT	BsaI palindrome for pEn_Comaira
2x_BsaI_r	aaacAGGTCTTCTCGAAGACCC	BsaI palindrome for pEn_Comaira
U6g_1/4_1f	ggggacaagtttgtacaaaaaagcaggcttaCTTTTTTCTTCTTCTTCGTTCA TACAG	U6:gRNA cassette with attB1/4 tags
U6g_1/4_1r	ggggacaactttgtatagaaaagttgggtgGGTCTAGAAAAAAGCACCGACTC	U6:gRNA cassette with attB1/4 tags
U6g_4r/3r_1f	ggggacaacttttctatacaaaagttgtaCTTTTTTCTTCTTCTTCGTTTCATAC AG	U6:gRNA cassette with attB4r/3r tags
U6g_4r/3r_1r	ggggacaactttattatacaaaagttgtaGGTCTAGAAAAAAGCACCGACTC	U6:gRNA cassette with attB4r/3r tags
U6g_3/2_1f	ggggacaactttgtataataaagttgtaCTTTTTTCTTCTTCTTCGTTTCATAC AG	U6:gRNA cassette with attB3/2 tags
U6g_3/2_1r	ggggaccactttgtacaagaagctgggtgGGTCTAGAAAAAAGCACCGACTC	U6:gRNA cassette with attB3/2 tags
RC9_3/2_f	ggggacaactttgtataataaagTTGTAAACGACGCCAGTGAATTG	RPS5a:Cas9 cassette with attB3/2 tags
RC9_3/2_r	ggggaccactttgtacaagaagctgggttGCCAAGCTCGGAATTAACCC	RPS5a:Cas9 cassette with attB3/2 tags
nos:kan_1/4_1f	ggggacaagtttgtacaaaaaagcaggcttaACGTATGACCCCCCGCAT	nos:kanR cassette with attB1/4 tags
nos.kan_1/4_1r	ggggacaactttgtatagaaaagttgggtgACCGTCACCGACTTGAGCCA	nos:kanR cassette with attB1/4 tags
Ampliseq NG2_f	acactctttccctacacgacgctcttGATCTTATGTCCGTGTAGATGCAAAC	

Ampliseq NG2_r	ctggagttcagacgtgtgctcttccgatctCATAGGCCAAGATCAATAGTCAAC	
Ampliseq NG3_f	acactctttccctacacgacgctcttCCGATCTCTATTTCCAACGCTTCTTCT AATCG	
Ampliseq NG3_r	ctggagttcagacgtgtgctcttccgatctCTAGGATGTAGCAAACCACACAT ATC	
Ampliseq NG5_f	acactctttccctacacgacgctcttCCGATCTTTCAAGGTAAGCAAATGTTG TGT	
Ampliseq NG5_r	ctggagttcagacgtgtgctcttccgatctAATGATATGAGTGTAATCGAGAA C	
NG Translocation Arm A	CTGCCTTCATATCAAAAGATTCCAC	See Figure 6B
NG Translocation Arm B	CCTAGTTCTTCTATCTAGATA	See Figure 6B
NG Translocation Arm C	CTATTTCCAACGCTTCTCTAATCG	See Figure 6B
NG Translocation Arm D	GAAACGTTCTCCTCATGTATGCTG	See Figure 6B
NG Translocation Arm E	TTTCAAGGTAAGCAAATGTTGTGT	See Figure 6B
NG Translocation Arm F	CAAGATTCAAGATTCAGTACTATTC	See Figure 6B
LEC1 A	ACAAAGTGACTAAGAACATAG	See Figure 4A
LEC1 B	TGGACAGTTCAGACGCTAATG	See Figure 4A
LEC1 C	TAGATTACCAGAGATGGATTTTGGTAGC	See Figure 7B
PDE334 A	GCATTGTGTATTTACGTTTGTAGTGG	See Figure 7B
Ampliseq CH1-1_f	acactctttccctacacgacgctcttccgatctAATACCGCCACGTGTTCAAT	
Ampliseq CH1-1_r	ctggagttcagacgtgtgctcttccgatctACTAAACCCATATCGCAAAAACGT	
Ampliseq CH1-2_f	acactctttccctacacgacgctcttccgatctACCGAATGTCACTAGACTTTG CT	
Ampliseq CH1-2_r	ctggagttcagacgtgtgctcttccgatctGCTGGTGAATAAGGAGCTGC	
CH1 A	ATACAATAAATACCATAGCAC	See Figure 3A

CH1 B	TGGACAACCAAACCAGGAAA	See Figure 3A
CH1 C	TCCTTTCACGCCCTGTAACC	See Figure 3A
CH1 D	AGAAGCTGAGCTACTAGAGTC	See Figure 3A
CH1 E	TTATGAGACTTCCTCTAACTC	See Figure 3A
CH1 F	AGAAGGCAAAGCACACAGC	See Figure 3A
CH1 G	ACCAGATTAATAACCTACACTTGTCTG	See Figure 3A
CH1 H	GCTGAAGAACTTTCGATTGG	See Figure 3A
PDE334 B	CGTGGTGGAACGTCATAGATTG	See Figure 7
PDE334 C	GCATTGTGTATTTACGTTGTAGTGG	See Figure 7
LEC1 E	CCACATAGCCAATGAGACAAGCC	See Figure 7
LEC1 F	TAGATTACCAGAGATGGATTTTGGTAGC	See Figure 7
WUS1 E	CCTAAAATCTCTTTACTACCAGCAAG	See Figure 7
WUS1 F	GATCTGATCGGCTGTTGGTGACC	See Figure 7
RABE1B B	ATTAAGCTAATTACTAGCAACGATTCG	See Figure 7
RABE1B Bb	AATTGAATCATAACTACAATCGTAGCC	See Figure 7; alternate to RABE1B B
RABE1B C	TTTCTTCTAAAGAGATAACATCAC	See Figure 7

Table S8. Knockout frequency for *BR11* in different types of crosses.

Egg \ Sperm	gRNA-5	gRNA-2	gRNA-7	gRNA-4	gRNA-8	gRNA-5	gRNA-3
Cas9-3	0/15	0/13	0/27	0/15	2/10	0/18	0/11
Cas9-4	12/15	13/24	0/39		7/10		8/9
Cas9-5	0/30	0/17	0/10		0/18		