

Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice

Huanbin Zhou¹, Bo Liu¹, Donald P. Weeks², Martin H. Spalding¹ and Bing Yang^{1,*}

¹Department of Genetics, Development and Cell Biology, Iowa State University, Ames, IA 50011, USA and

²Department of Biochemistry, University of Nebraska, Lincoln, NE 68588, USA

Received May 31, 2014; Revised August 21, 2014; Accepted August 22, 2014

ABSTRACT

The Cas9/sgRNA of the CRISPR/Cas system has emerged as a robust technology for targeted gene editing in various organisms, including plants, where Cas9/sgRNA-mediated small deletions/insertions at single cleavage sites have been reported in transient and stable transformations, although genetic transmission of edits has been reported only in *Ara-bidopsis* and rice. Large chromosomal excision between two remote nuclease-targeted loci has been reported only in a few non-plant species. Here we report in rice Cas9/sgRNA-induced large chromosomal segment deletions, the inheritance of genome edits in multiple generations and construction of a set of facile vectors for high-efficiency, multiplex gene targeting. Four sugar efflux transporter genes were modified in rice at high efficiency; the most efficient system yielding 87–100% editing in T0 transgenic plants, all with di-allelic edits. Furthermore, genetic crosses segregating Cas9/sgRNA transgenes away from edited genes yielded several genome-edited but transgene-free rice plants. We also demonstrated proof-of-efficiency of Cas9/sgRNAs in producing large chromosomal deletions (115–245 kb) involving three different clusters of genes in rice protoplasts and verification of deletions of two clusters in regenerated T0 generation plants. Together, these data demonstrate the power of our Cas9/sgRNA platform for targeted gene/genome editing in rice and other crops, enabling both basic research and agricultural applications.

INTRODUCTION

Precise genome modification with engineered, site-directed nucleases such as meganucleases, zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and Cas9/sgRNAs has emerged as a powerful genome-enabling tool for fundamental and applied biological

research. These technologies rely on DNA double strand breaks (DSBs) generated in a targeted fashion in the genomic region of interest by the potent nucleases (1–3). DSBs stimulate cellular DNA repair mechanisms and are repaired through error-prone nonhomologous end joining (NHEJ) and/or homologous recombination (HR) *in vivo*, leading to site-specific genetic alterations (4–7). NHEJ has been exploited to generate mutagenic insertions/deletions in transcription regulatory or protein-coding region of genes of interest that often result in gene inactivation. HR-associated molecules drawn to a specific DSB created by a site-specific engineered nuclease have been harnessed to allow incorporation of desired small or large genetic elements that are flanked by DNA sequences homologous to regions present on either side of the DSB. In most cases, NHEJ events are more predominant than HR in eukaryotic cells and thus are associated with the large majority of targeted genomic events (8–10). To generate a site-specific DSB in the genomic region, meganucleases, ZFNs and TALENs were developed successively (11–15). Both customizable hybrid nucleases are composed of programmable, sequence-specific DNA-binding modules fused with the non-specific cleavage domain of FokI nuclease. Such molecules, acting in pairs that recognize adjacent DNA sequences, have enabled precise genomic modifications in a broad range of species, including yeast, plants and animals [reviewed by (2,3)]. However, they often are effectively unavailable for routine use in many laboratories because of the time-consuming and labor-intensive process needed for engineering the nucleases genes; furthermore, their sensitivity to the DNA methylation state of the target sites limits the application of those nucleases in genome-wide analysis due to the widespread existence of cytosine methylation in complex genomes (16,17).

The type II CRISPR/Cas RNA-guided nucleases are the most recent addition to the tool kit of sequence-specific nucleases. Intense interest has been focused on the CRISPR/Cas9 system from *Streptococcus pyogenes* following initial reports of its successful use for gene editing (18–20). In this system, Cas9 nuclease coupled with tracrRNA (*trans*-activating crRNA) can be guided by an ~20-nt guide (or seed) sequence in crRNA (CRISPR RNA) to hybridize

*To whom correspondence should be addressed. Tel: +1 515 294 2968; Fax: +1 515 294 5256; Email: byang@iastate.edu

with a specific complementary DNA sequence (i.e. the target site) that is followed by a 5'-NGG or 5'-NAG PAM (protospacer adjacent motif) sequence to induce a precise cleavage of the target sequence 3–4 base pairs upstream of the PAM site (18). Alternatively, parts of crRNA and tracrRNA sequence can be fused in a synthetic gene to produce a single guide RNA (sgRNA) that is equally as effective as the separate crRNA and tracrRNA complex in targeting a specific DNA sequence for Cas9-directed cleavage (19,20). Unlike ZFNs and TALENs, the CRISPR/Cas9 system, referred to here as Cas9/sgRNA, is DNA methylation-insensitive. It is also more affordable, remarkably easier to use, and well-suited for multiplex gene targeting and high-throughput genome-wide gene editing at a similar or even higher efficiency than ZFNs and TALENs.

Recently, the Cas9/sgRNA system has been demonstrated to be effective for gene editing in several plants, including *Nicotiana benthamiana*, *Arabidopsis*, wheat, maize, sorghum and rice (21–33). Based primarily on results using marker genes that cause obvious phenotype changes when disrupted, investigators have suggested that it is feasible to modify most endogenous genes using the Cas9/sgRNA system. In a limited number of cases involving *Arabidopsis* and rice plants, it has been possible to demonstrate gene editing by Cas9/sgRNA in stable transformed T0 plants and only recently in the progeny of transgenic plants of the two species (22,27,33). In the most recent two studies using *Arabidopsis*, researchers found that Cas9/sgRNA-triggered gene modifications could be found in T1 generation plants mostly in somatic cells and that only 22–50% of T2 plants were homozygous for the modified gene (31,32). In both studies, the mutations were stably inherited in the T3 generation. Although the Cas9/sgRNA system has been shown to function in creating targeted gene disruption in rice (21,22,26,27,33), stable inheritance of the modified genes and their patterns of expression are only now being explored in depth (27).

In most aforementioned cases, genomic editing was shown to occur at a single locus leading to micro-deletions and/or insertions since the systems were not designed and constructed to express multiple guide RNAs for complex gene editing. However, Cas9/sgRNAs have potential to allow much more complex genome rearrangements that could allow studies, for example, into the function of large *cis*-regulatory domains in controlling plant gene expression, the roles of specific noncoding RNA genes in controlling plant development and the role of certainly closely clustered genes in plant chromosomes. Such studies require the ability to remove and/or replace entire large regions of chromosomes in a precisely targeted manner. The deletion of chromosomal segments between two widely spaced target sites using two pairs of ZFNs, two pairs of TALENs or two pairs of Cas9/sgRNAs to simultaneously cleave and remove a chromosome segment has been reported in cultured mammalian cells (including human cells), zebrafish and in *Arabidopsis* somatic cells when ZFN genes were chemically and transiently induced (34–38). However, the deletion of significant chromosome segments has not been reported in model plant systems or, more particularly, in a crop plant. Here we report proof-of-concept and proof-of-efficiency demon-

strations of just such chromosome segment deletions in the world's most widely grown crop, rice.

Here we also report a set of Gateway-based Cas9/sgRNA vectors that are highly efficient and easy to use for gene editing. The intermediate vector can be used to construct up to four sgRNA genes for single or multiplex gene editing in the rice genome. We demonstrated that most mutated T0 plants contain di-allelic (heterogeneous or homogeneous) modifications and that modified genes can be genetically segregated away from Cas9 and sgRNA transgenes and faithfully transmitted through ensuing generations in a Mendelian fashion. We further demonstrated that deletions of large chromosomal segments can be efficiently achieved in regenerated plants with our constructs. Together, these results suggested that the current set of Cas9/sgRNA constructs represents an important technical step forward that should be of significant benefit to the rice research community and agricultural industry in pursuing goals associated with customized genome editing and chromosome engineering.

MATERIALS AND METHODS

Plant materials

The *Japonica* rice cultivar Kitaake was used. Plants (wild type or transgenic) were grown in a growth chamber and greenhouse at a temperature of 28–30°C, in a relative humidity of 75% and with photoperiod of 12 h (h) in the growth chamber and with natural light plus supplemented light in the greenhouse. For protoplast isolation, rice seeds were sterilized and germinated in $\frac{1}{2}$ MS (Murashige Skoog) solid medium in a plastic container, grown at 30°C under a 12h/12h-light regime in the growth chamber for 8–10 days.

Construction of plasmids expressing Cas9 and guide RNAs

To construct a Cas9 expression cassette for the binary vector, the rice codon-optimized SpCas9 gene (21) was placed downstream of the maize ubiquitin 1 promoter (14) and upstream of the nopaline synthase (NOS) gene terminator by BamHI and SpeI restriction and DNA ligation. The Cas9 cassette and a Gateway recombination cassette that contained the attR1-ccdB-attR2 were cloned into pCAMBIA1300, which confers hygromycin B resistance (hptII) in transgenic plant (CAMBIA Health Solutions, Inc.), resulting in a destination vector, pUbi-Cas9. Constructs derived from pUbi-Cas9 were used for rice transgenics. For transient expression of Cas9 in rice mesophyll protoplasts, the pUC19 version of Cas9 was described earlier (21).

For construction of guide RNA genes, two different promoters of the rice small nuclear RNA U6 genes were commercially synthesized based on the genome sequence of the rice cultivar Nipponbare. For sgRNA constructs, the BtgZI or BsaI cutting site was placed downstream of the U6 promoter and used to fuse the seed sequences to the tracrRNA tail sequences of either 48 nucleotides [sgRNA(+48)] or 85 nucleotides [sgRNA(+85)], the guide RNA architectures based on the work as published (19). Those cassettes were synthesized by GenScript and cloned into pENTR4 (Life Technologies) by BamHI and EcoRI restriction and DNA

ligation, individually, resulting in a set of pENTR-sgRNA vectors in which the guide RNA cassettes could be mobilized to pUbi-Cas9 using the Gateway LR clonase (Life Technologies). For the dual tracrRNA and crRNA [dual guide RNA (dgRNA)] constructs, the DNA sequences for tracrRNA and the crRNA direct repeats were adapted from Jinek's work (18). The tracrRNA-coding DNA and the rice U6 promoter were synthesized from GenScript; the DNA for the crRNA repeat sequence and another rice U6 promoter was also synthesized from GenScript. The BsaI and BtgZI restriction sites were designed and located within the direct repeats of crRNA for cloning of seed sequences. Both crRNA and tracrRNA cassettes were cloned into pENTR4 in tandem by BamHI and EcoRI. The sequences of those synthetic guide RNAs (sgRNA or dgRNA) are provided in Supplementary data (Supplementary Figure S1).

The complementary oligos with appropriate 4-bp overhangs, after being annealed to each other, were phosphorylated, annealed and cloned into the respective pENTR-sgRNA or pENTR-dgRNA that was digested with appropriate restriction enzymes. Each sgRNA or crRNA was sequenced using primer (U6p-F1) for the accuracy of inserted seed sequences. The sgRNA or dgRNA cassettes were then used for transient expression with pUC19-Cas9 directly in rice protoplasts, or were recombined into pUbi-Cas9 for rice transformation. The sequences of those oligos are listed in Supplementary data (Supplementary Table S1).

Transient gene expression in rice protoplasts

Isolation and transfection of rice mesophyll protoplasts were carried out as previously described (21). A mixture of plasmid DNA containing pUC19-Cas9 and pENTR-sgRNA (10 μ g with a 1:1 molar ratio) was used.

Agrobacterium-mediated rice transformation

The binary vector-based plasmids were mobilized into the *Agrobacterium tumefaciens* strain EHA105 by electroporation. The immature embryo derived callus cells of Kitaake were used for tissue culture and transformation with a protocol modified from the method as described by Hiei *et al.* (39).

Detection of the targeted genomic editing events

Genomic DNA from the rice protoplasts or leaves of transgenic rice plants was extracted using the hexadecyltrimethylammonium bromide (CTAB) method (40) and further used for polymerase chain reaction (PCR) amplification with specific primers (see Supplementary Table S1 for sequence information). The PCR products were used for the T7 Endonuclease I (T7EI) assay. Briefly, the T7EI assay is as follows: PCR products (a mixture from mutant and wild type) were heated at 95°C for 5 min, then cooled from 95 to 85°C at 2°C/s, 85 to 25°C at 0.1°C/s and held at 25°C for 30 min. The denatured and reannealed PCR products were purified with phenol:chloroform extraction and ethanol precipitation, 100 ng of each purified product was digested with three units of T7EI for 2 h at 37°C and subjected to 2% agarose gel electrophoresis.

Detection of large fragment deletion

The transfected protoplast pools were used for genomic DNA extraction using the CTAB method. Similarly the transgenic calli and plant leaves were used for genomic DNA extraction. Genomic DNA was used for PCR amplification of the relevant region with primers, and the amplicons were subjected to Sanger sequencing. The amplicons were also cloned into the EcoRV-restricted pBluescriptSK, and clones with inserts were subjected to Sanger sequencing.

Genotyping of primary or progeny plants of transgenic rice

To screen the T0 transgenic rice plants for the presence of Cas9/sgRNA and for mutations it induced, the transgenes Cas9 and sgRNA (or dgRNA) were detected with PCR for each line, while PCR products from the targeted regions in each line were first analyzed in the T7EI assay and selectively subjected to Sanger sequencing.

For genotyping of T1 and T2 transgenic rice plants, a number of seeds from independent transgenic lines were grown in the greenhouse. Genomic DNA was extracted, and PCR amplification for Cas9, hptII and sgRNA/crRNA genes was performed on individual plants to detect the presence of transgenes. PCR products from the targeted region in individual plants were analyzed in the T7EI assay to determine the zygosity (homozygous or heterozygous) of progeny plants. A selection of PCR products from the targeted site in homozygous lines was subjected to Sanger sequencing.

RESULTS

Construction of CRISPR/Cas9 systems with different guide RNA architectures

As a follow-up to our earlier use of the TALEN system for editing rice genes (i.e. the SWEET genes) that are important susceptibility genes to plant disease (41), we recently demonstrated that the Cas9/sgRNA system was capable of inducing site-specific mutations in the promoter regions of both the *SWEET11* gene and *SWEET14* gene (21). In the previous work, the Cas9 gene from *S. pyogenes* (Sp-Cas9) and the SpCas9 gene codon-optimized for expression in rice each were co-expressed with sgRNAs derived from an architecture comprised of 48 nucleotides (nt) of the tracrRNA 3' 'tail' [referred to as sgRNA(+48)] and it induced site-specific mutagenesis efficiently. However, sgRNA with a longer tracrRNA tail (e.g. +85 nt) was reported recently to have higher efficiency of mutagenesis in cultured human cells (42), we also included sgRNA(+85) in our constructs for comparison of their relative efficiency in plant (Figure 1). Additionally, we constructed a dgRNA with separate crRNA and tracrRNA components similar to the native *S. pyogenes* CRISPR/Cas9 system, a system shown earlier to modify target sites with varying efficiencies (19,29). This platform was designed to offer the opportunity to produce multiple crRNAs under control of a single promoter to allow simultaneous targeting of multiple loci (Figure 1). We first tested the efficiency of bacterial Cas9 (SpCas9), rice codon optimized SpCas9, human codon optimized SpCas9 (19) and *Chlamydomonas* codon optimized SpCas9

SWEET11 in the other 14 independent lines by sequencing amplicons from two plants of each line. In contrast to these results, similar genotypic analyses with T7E1 and DNA sequencing using the same individual plants of the 16 lines (two plants each line) assayed above revealed no mutation at the *SWEET14* target site (Supplementary Table S2). The presence of Cas9 gene and guide RNA cassette in all 16 lines was confirmed by PCR analysis. Taken together, our data suggest that the dgRNA system has low efficiency in generating DSBs and subsequent mutations in the stably transformed rice plants.

sgRNA(+85), but not sgRNA(+48), causes high-efficiency mutations at the target sites in T0 transgenic lines

Next, we investigated the ability of two different sgRNA architectures [sgRNA(+85) and sgRNA(+48)] to mediate site-specific DSBs and NHEJ-based mutagenesis in stable transgenic rice lines. The first sgRNA construct contained one seed sequence in the +85 architecture and another in the +48 architecture, each targeting a different site in the *SWEET13* gene coding region (Supplementary Figure S2B). From eight independent T0 transgenic lines we obtained, all contained di-allelic mutations corresponding to the sgRNA(+85)/SWEET13-1 and none occurred at the sgRNA(+48)/SWEET13-2 targeted site, based on evaluation of five plants from each line (Figure 2). Similarly, a second sgRNA construct expressing a long sgRNA targeting the *SWEET1a* gene and a short sgRNA targeting the *SWEET1b* gene coding region (Supplementary Figure S2B), respectively, was used for generating transgenics. Ten independent transgenic lines were obtained and five plants from each line were evaluated for genomic editing events at the two target sites. As shown in Figure 3, seven lines contained di-allelic mutations at the sgRNA(+85)-targeted *SWEET1a* site. Two of the three lines that lacked the mutation did not contain either detectable Cas9 (line #9) or sgRNA (line #1) while only one non-mutated line, #6, contained both Cas9 and sgRNA (Figure 3A). Again, in marked contrast to results with the sgRNA(+85) construct, none of the same 50 plants showed a positive T7E1 assay for mutations at the sgRNA(+48)-targeted site in the *SWEET1b* gene. The lack of sgRNA activity was further confirmed by sequencing the relevant amplicons from two plants of each line. The sgRNA(+48) transgene could be detected in all of these lines. Finally, a construct expressing two long sgRNAs (one targeting a site in the *SWEET1b* gene and the other a site in the *SWEET13* gene) and two short sgRNAs (one targeting the *SWEET5* gene and the second a second site in the *SWEET13* gene) (Supplementary Figure S2B) was constructed and transferred into rice callus cells. Five transgenic lines were obtained and used for genotyping of the four target sites for possible mutations using both the T7E1 assay and DNA sequencing. As shown in Figure 4, three lines contained di-allelic mutations of *SWEET13* targeted by the sgRNA(+85)/SWEET13-1 and another line contained mutations of the *SWEET1b* gene targeted by another long sgRNA, sgRNA(+85)/SWEET1a. Mutations were found only in plants carrying a Cas9 gene and the respective sgRNA(+85) genes. On the other hand, neither of

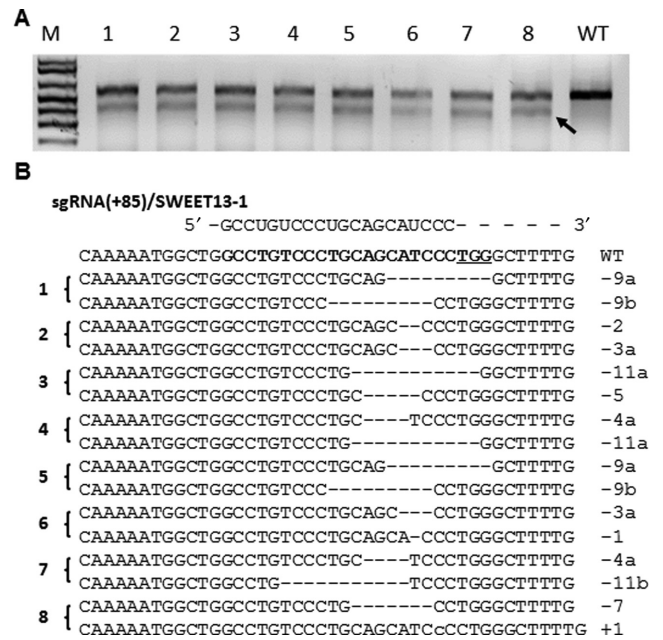


Figure 2. sgRNA(+85)/SWEET13-1 not sgRNA(+48)/SWEET13-2 in the cassette of two sgRNAs produces high-efficiency di-allelic mutations in rice *SWEET13* in T0 transgenic lines. (A) The gel picture of electrophoresis on the T7E1 treated rice *SWEET13* amplicons from the Cas9/sgRNA(+85)-SWEET13-1 T0 transgenic plants. The PCR products amplified from the sgRNA-targeted *SWEET13* locus were treated with T7E1 and subjected to 1% agarose gel electrophoresis. The arrow indicates the expected T7E1 cleavage product, suggesting possible mutations. M, 1-kb ladder DNA marker; WT, parental Kitaake control. (B) Sequences of *SWEET13* at the sgRNA-targeted site (bold and underlined in the wild-type sequence) in each T0 transgenic lines as in (A). The nucleotide changes (dashes for deletion and lower letter for insertion) are also indicated to the right side of each sequence suffixed with a letter, if needed, to distinguish different alleles.

Table 1. Efficiency and ratio of allelic mutations

Gene	Architecture	Efficiency ^a	Mono-allelic ^b mutation	Bi-allelic ^b mutation
<i>SWEET11</i>	crRNA	2/16 (12.5%)	2/2	0/2
<i>SWEET13</i>	sgRNA(+85)	8/8 (100%)	0/8	8/8
<i>SWEET1a</i>	sgRNA(+85)	7/10 (70%)	0/7	7/7
<i>SWEET13</i>	sgRNA(+85)	3/5 (60%)	0/3	3/3
	sgRNA(+85)	1/5 (20%)	0/1	1/1

^aThe number of transgenic lines with edited events versus the total number of T0 lines and percentage of gene editing in parenthesis.

^bThe number of mono- or di-allelic mutations versus the total number of editing events.

the two short sgRNA(+48) genes, sgRNA(+48)/SWEET5 and sgRNA(+48)/SWEET13-2, targeting *SWEET5* and *SWEET13* induced a site-specific mutation in the five transgenic lines. The gene editing efficiency in T0 plants carrying different guide RNA constructs is presented in Table 1.

We also investigated the potential off-target effect of sgRNA(+85)/SWEET13-1 that was 100% efficient in inducing site-specific di-allelic mutations in all eight independent rice T0 lines. Bioinformatic analyses found six sites that contained 16 or more nucleotides matching the 20 nu-

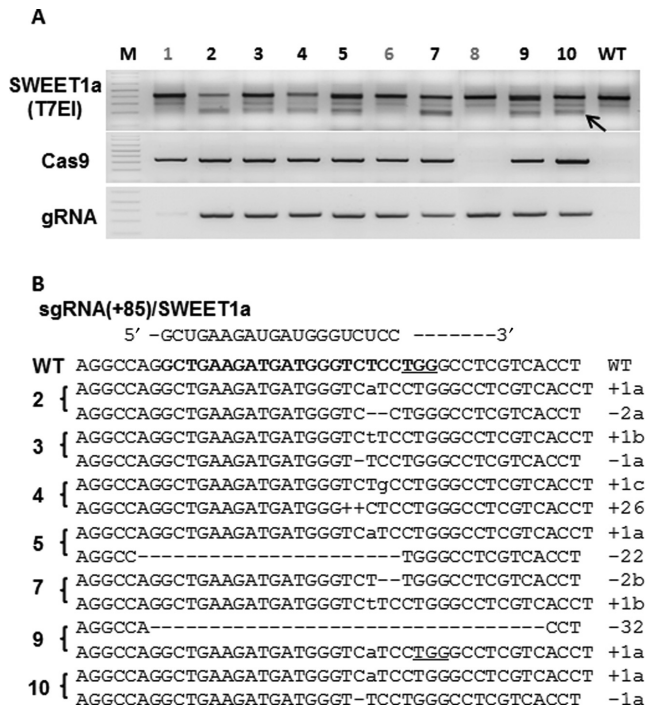


Figure 3. sgRNA(+85)/SWEET1a instead of sgRNA(+48)/SWEET16 in the cassette of two sgRNAs induces high-efficiency di-allelic mutations of *SWEET1a* in the primary (T0) transgenic rice plants. (A) Analyses of T0 transgenic lines for site-specific mutations of *SWEET1a* and for the presence of transgenes Cas9 and sgRNA. The first panel shows the agarose gel image of the T7 endonuclease I (T7EI) treated amplicon of *SWEET1a* with the expected ~200-bp cleavage products indicated by the arrow. M, 1-kb ladder DNA marker; lanes 1–10 each denote one representative plant from each independent line; WT, parental Kitaake. Lines 1, 6 and 8 (light black) are T7EI negative. The lower two panels show the PCR results of Cas9 and sgRNA transgenes in the 10 independent transgenic lines. (B) Sequencing confirmation of *SWEET1a* mutations in seven of the T7EI-positive T0 lines shown in (A). All seven lines contain di-allelic mutations (dash for deletion, lower letters and ‘++’ for insertions) at the sgRNA-targeted sites (bold and underlined). The number of altered nucleotides is shown at the right side of each sequence coupled with a lower-case letter, if needed, indicating the uniqueness of different alleles.

cleotides of the guide RNA sequence including two sites with 17 and 16 nucleotides, respectively, matching the guide RNA and also with a perfect PAM sequence (Supplementary Figure S4). Genomic DNAs of seven T0 lines were used for PCR amplification of the six relevant regions and replicons were sequenced. Analysis of the DNA sequencing data (including the sequencing traces from all amplicons representing 84 alleles) revealed no trace of nucleotide change at any of the six potential off-target sites (data not shown).

Inheritance and stability of targeted gene mutations in T1 and T2 generations

To investigate whether the genomic modification by Cas9/sgRNAs could be transmitted to progeny of the self-fertilized T0 plants, we genotyped the T1 progeny plants and, in some cases, the T2 progeny of a sampling of genotyped T1 plants. Detailed analysis was performed with the T1 plants of line #7, which harbored di-allelic (-4/-11) mutations in the *SWEET13* gene. Five of 10 T1 plants were

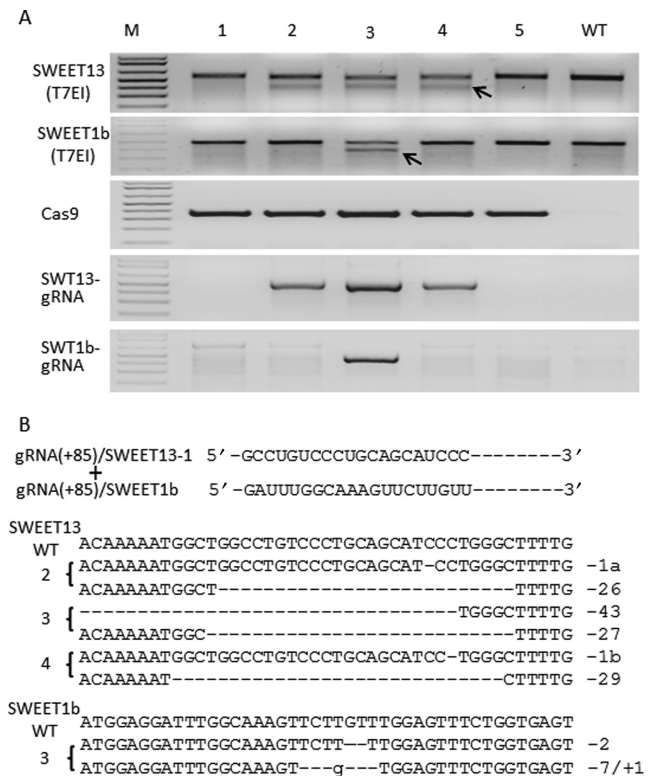


Figure 4. Two sgRNA(+85) versions produce mutations at two loci in T0 transgenic plants. (A) Detection of mutations and transgenes in five T0 transgenic lines. The upper two panels show the PCR products of *SWEET13* and *SWEET1b* at the relevant sites that were treated with T7EI and separated in 1% agarose gels. The arrow points to the DNA band of the T7EI cleavage products of similar size. The lower three panels show the PCR products of Cas9 and two sgRNAs in the five transgenic lines using specific respective primers. (B) Sequencing confirmation of the site-specific mutations (di-allelic) in the T0 transgenic lines. Dashes denote nucleotide deletions and lower-case letters denote nucleotide insertions. The number of altered nucleotides is shown at the right side of each sequence coupled with a letter, if needed, indicating the uniqueness of different alleles.

either homozygous for a 4-bp or 11-bp deletion, and the rest were heterozygous for di-allelic mutations based on the T7EI assay (data not shown); sequencing analysis revealed that all mutated alleles were either a 4-bp deletion or an 11-bp deletion, the same mutations found in the parental T0 line #7 (Figure 5A and Table 2). Furthermore, in genetic crosses, the two alleles segregated with a ratio of 3(-4):5(-11/-4):2(-11). A χ^2 test indicated there was no significant difference with the expected 1:2:1 ratio, suggesting Mendelian inheritance. Next, we assayed for the presence of T-DNA (transfer DNA, the specific segment of the *A. tumefaciens* Ti-plasmid that is transferred into the plant genome during *Agrobacterium*-mediated plant transformation) in the T1 population. PCR amplifications with primers targeting to the Cas9, sgRNA and hptII genes were conducted on DNA from progeny self-fertilized T0 plants. The results showed genetic segregation caused the loss of T-DNA in some T1 plants (Figure 5B). Subsequently, a sampling of T2 progeny plants derived from self-fertilized T1 plants #7-1 and #7-4 was randomly selected and genotyped for the presence of the *sweet13* gene (i.e. the mutated version of the *SWEET13* gene) and the presence of

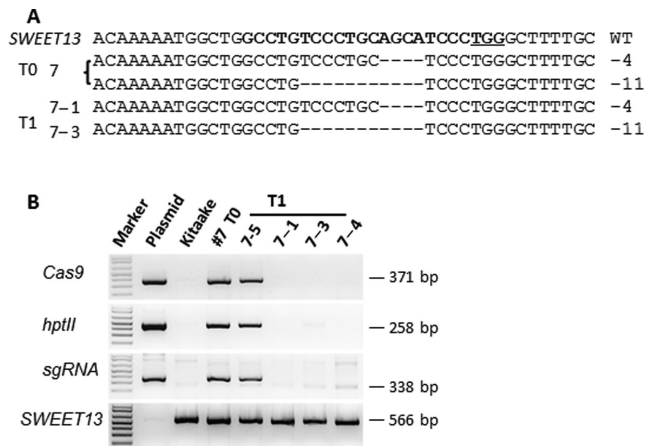


Figure 5. Inheritance of the Cas9/sgRNA modified *SWEET13* and removal of transgenes in T1 progeny. (A) Each of the di-allelic mutations (4- and 11-bp deletions) in one representative T0 line (#7) was transmitted to its progeny (T1 plant 7-1 and -3). The sgRNA-targeted site is denoted in the wild-type sequence as bold letters and the PAM sequence underlined. Nucleotide deletions are denoted as dashed lines. (B) PCR-detected presence and absence of individual transgenes (Cas9, *hptII* and sgRNA) from different sources as indicated above each lane of the agarose gel picture. *SWEET13* was used as a control for sample quality.

Cas9 and sgRNA transgenes in the modified genomes. PCR amplification and DNA sequencing analyses indicated that Cas9, sgRNA and *hptII* were not present in the genomes of the plants randomly selected for analysis and that the site-specific mutations (either 4- or 11-bp deletions) in the T2 plants were identical to those in progenitor T1 plants (data not shown). Detailed analysis of inheritance of the modified *SWEET13* gene in lines #3 and #5, as well as the modified *SWEET1a* gene in lines #2, #7 and #9, was similarly conducted. Similar results were obtained from both T1 and T2 progeny plants, as shown in Supplementary Figure S5 for the *SWEET1a* gene and summarized for all three genes in Table 2. The T-DNA segregation analysis of these lines revealed that line #5 carrying the *sweet13* gene mutation and lines #2 and #7 carrying the *sweet1a* gene mutation included some T1 plants lacking the T-DNA region (i.e. the region that carries the Cas9, sgRNA and *nptII* transgenes). However, line #3 carrying the *sweet13* mutant gene and line #9 carrying the *sweet1a* mutant gene did not exhibit T-DNA segregation in T1 plants, indicating the possibility of multiple T-DNA insertions or close linkage of the T-DNA insertion with the modified locus.

We next examined how the modified *SWEET11* gene was inherited in the T1 generation for line #3, whose mono-allelic 2-bp deletion in *SWEET11* mutation was induced by Cas9 and the tracrRNA/crRNA-SWEET11. T7EI assay and DNA sequencing results indicated a 4(-2):3(-2/WT):3(WT) segregation ratio and the presence of no new mutations in the T1 population. In this case, the T-DNA region was not segregated out for all of the T1 plants (Table 2).

We also investigated whether Cas9/sgRNA could induce site-specific mutations in the *SWEET1a* gene in T1 plants using the T0 line #6 that contained both Cas9 and sgRNA genes but no *SWEET1a* gene mutation (Figure 3A). No

Table 2. Segregation of edited events in T1 progeny

Gene	Lines	T0 Mutation	T1 segregation	
			Mutation ^a	T-DNA ^b
<i>SWEET1a</i>	7	Bi-allelic	3H	+/-
<i>SWEET1a</i>	3	Bi-allelic	8Ho/4He	+/-
<i>SWEET1a</i>	5	WT	7WT	+/-
<i>SWEET13</i>	2	Bi-allelic	2Ho	+/-
<i>SWEET13</i>	7	Bi-allelic	3Ho	+
<i>SWEET13</i>	6	Bi-allelic	5Ho/5He	+/-
<i>SWEET13</i>	9	Bi-allelic	3Ho/11He	+
<i>SWEET11</i>	3	Mono-allelic	4Ho/3He/3W	+

^a Ho stands for homozygous, He for heterozygous and WT for wild-type genotype of segregating plants.

^b '+' and '-' stand for presence and lack of T-DNA, respectively, in the progeny population.

plants tested positive for mutations at the targeted site using the T7EI assay even though Cas9 and sgRNA genes were detected in some of the T1 segregants (Supplementary Figure S5). Further check of Cas9 expression in T1 progenies of T0 plant #6 by RT-PCR indicated the presence of Cas9 transcript, indicating the lack of gene editing in this particular line was not due to silencing of the Cas9 gene.

Cas9/sgRNAs induce large chromosomal deletions in rice protoplasts

Since our Cas9/sgRNA(+85) system is highly efficient in induction of site-specific cleavage, we sought to test its ability to generating large segment (>100 kb) deletions between two genomic loci targeted by a combination of distinct sgRNAs. We postulated that the presence of widely spaced site-specific DSBs would facilitate the generation of large deletions between the two break sites. The segment deletion would require that the two independent DSBs be simultaneously generated at the two loci and that the DSBs rejoin through NHEJ repair mechanism in cells. In these experiments, we first replaced the short tracrRNA tail (+48) gene with the longer tail sequence (+85) in one sgRNA cassette, resulting in a vector containing two sgRNA(+85) cassettes that express two species of sgRNA(+85) under control of two different U6 promoters. This vector was also further modified such that the fragment containing two sgRNA cassettes (HindIII-HindIII) in one plasmid could be sub-cloned into another construct using HindIII and thus allow expression of four distinct sgRNAs in transgenic plant cells (Supplementary Figure S6).

The first multigene chromosome region to be targeted for deletion by Cas9/sgRNA activity contained a cluster of five biosynthetic genes that are involved in the production of labdane-related diterpenoids, a group of phytoalexins with antimicrobial activity, on chromosome 4 (43,44). The gene cluster spans about 170 kb (Figure 6A). One target site in both the *OsCPS4* and *CYP99A2* genes located at the ends

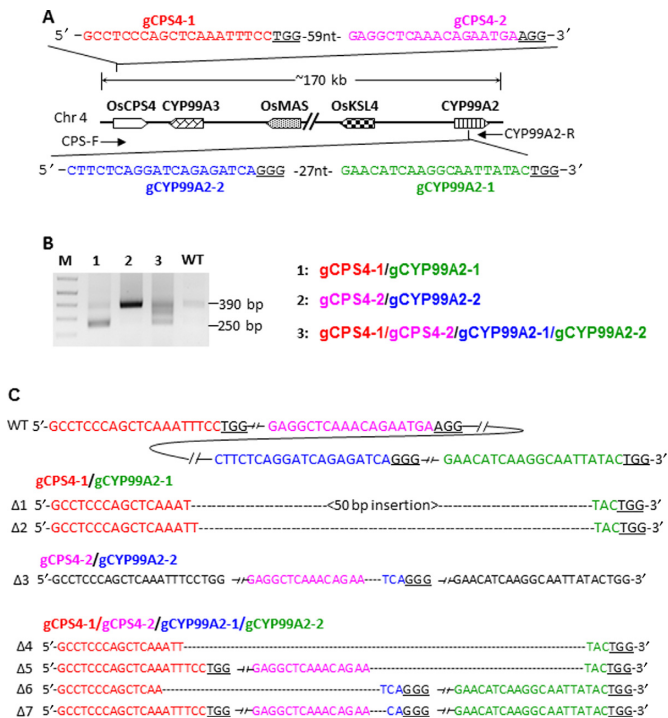


Figure 6. Large chromosomal deletion induced by Cas9/sgRNA in rice. (A) Schematic of a cluster of five diterpenoid genes within an ~170-kb region on rice chromosome 4, indicating sgRNA-targeted sites in colors and PAM sequences underlined. (B) Agarose gel electrophoresis image of the PCR products amplified from the genomic DNA of protoplast pools transfected with sgRNA constructs as indicated by a number corresponding to the sgRNA construct at the right side of picture. (C) Sequences of large DNA segment deletions induced by sgRNAs. The wild-type cluster is depicted as DNA sequence (including sgRNA-targeted sites in distinct colors) and lines (for the intervening sequences). The sequences of individual representative clones ($\Delta 1-7$) derived from PCR products shown in (B) are also depicted, with the deleted sequences indicated by dashes between the sgRNA-targeted sites.

of the cluster were chosen to design and construct sgRNAs gCPS4-1 and gCYP99A2-1, respectively (Figure 6A). Constructs containing the two sgRNAs and Cas9 genes were co-transferred into the rice mesophyll protoplasts. Genomic DNA was extracted 48 h post-transfection and used for PCR amplification with a pair of primers, one (CPS-F) specific to *OsCPS4* and the other (CYP99A2-R) to *CYP99A2*. The two primers were designed to amplify the genomic DNA derived from the protoplasts that underwent large chromosomal deletion between the two remote target sites, but not from the intact genomic region due to its large size (~170 kb). With the two primers, we amplified a product of the expected size (~250 bp) from protoplasts expressing the two sgRNA genes and the Cas9 gene, but not from those without Cas9/sgRNA (Figure 6B, lane 1 versus WT). Two clones containing these PCR amplicons were selected randomly for sequencing, and the results confirmed excision of the genomic segment between the gCPS4-1 and gCYP99A2-1 target sites (Figure 6C, $\Delta 1$ and $\Delta 2$; Supplementary Figure S7A).

We constructed a second pair of sgRNAs (gCPS4-2 and gCYP99A2-2) that also targeted *OsCPS4* and *CYP99A2*, but at different sites than used in the experiment described

above (Figure 6A). When co-expressed with Cas9 in rice protoplasts, these sgRNAs were active in inducing a large deletion, as detected by PCR and further confirmed with sequencing of the cloned PCR products (Figure 6B, lane 2; 6C, $\Delta 3$; Supplementary Figure S7B). We then subcloned and combined the first and the second pair of sgRNAs into a single construct (gCPS4-1/gCPS4-2/gCYP99A2-1/gCYP99A2-2), anticipating that the expression of four sgRNAs would lead to an increased success rate and diverse junctions of two cleavage sites when co-expressed with Cas9 in rice cells. Indeed, we detected a variety of large deletions that result from the combinatorial joining of two cleavage sites from both ends of the gene cluster (Figure 6B, lane 3; 6C, $\Delta 4-7$; Supplementary Figure S7C).

The sequencing data indicated that the two cleavage sites were consistent with the action of the two sgRNAs which guide Cas9 to cleave dsDNA about 3 nt upstream of the PAM sequences and that the junctions were derived from joining of the predicted breakage sites.

Cas9/sgRNAs induce deletions of two additional large gene clusters in rice protoplasts

Next, we extended our large chromosomal deletion experiments to target another cluster of 10 labdane-related diterpenoid synthetic genes spanning about 245 kb on rice chromosome 2 (45,46) (Figure 7A). Similarly, we designed one sgRNA targeting the cytochrome P450 gene *CYP76M5* (gCYP76M5-1) and another sgRNA targeting *CYP76M6* (gCYP76M6-1), respectively, at each end of the gene cluster (Figure 7A). Both sgRNA genes along with the Cas9 gene were co-transferred into rice protoplasts. Protoplasts were pooled 48 h post transfection for extraction of genomic DNA that was used for PCR amplification of the relevant region with the primers M5-F and M6-R. As expected, amplicons of ~260 bp were obtained from protoplasts producing the two sgRNAs and Cas9, but not from those without Cas9/sgRNA (Figure 7B, lane 1 versus WT). Directly sequencing the PCR product from one direction revealed two major products that were further confirmed with sequencing of the individual cloned PCR products (Figure 7C, $\Delta 1$ and $\Delta 2$; Supplementary Figure S7D). We designed a second pair of sgRNAs that also targeted *CYP76M5* (gCYP76M5-2) and *CYP76M6* (gCYP76M6-2), but at an adjacent location in each gene (Figure 7A). When co-expressed with Cas9 in rice protoplasts, sgRNA genes induced large deletions as detected by PCR and further confirmed with DNA sequencing of the cloned PCR products (Figure 7B, lane 2; 7C, $\Delta 3$; Supplementary Figure S7E). We then subcloned and combined the first and the second pair of sgRNAs into a single plasmid (gCYP76M5-1/gCYP76M5-2/gCYP76M6-1/gCYP76M6-2) and co-expressed them with Cas9 in rice cells. As expected from results of our previous experiment, we detected a variety of large deletions that resulted from the combinatorial joining of two cleavage sites from both ends of the gene cluster (Figure 7B, lane 3; 7C, $\Delta 4$ and $\Delta 5$; Supplementary Figure S7F).

We also sought to delete a third cluster of five diterpenoid synthetic genes within a 115-kb region on rice chromosome 6 using Cas9 and two sgRNAs (gOsKO1 and gOsKOL5) that independently targeted the border genes *OsKO1* and

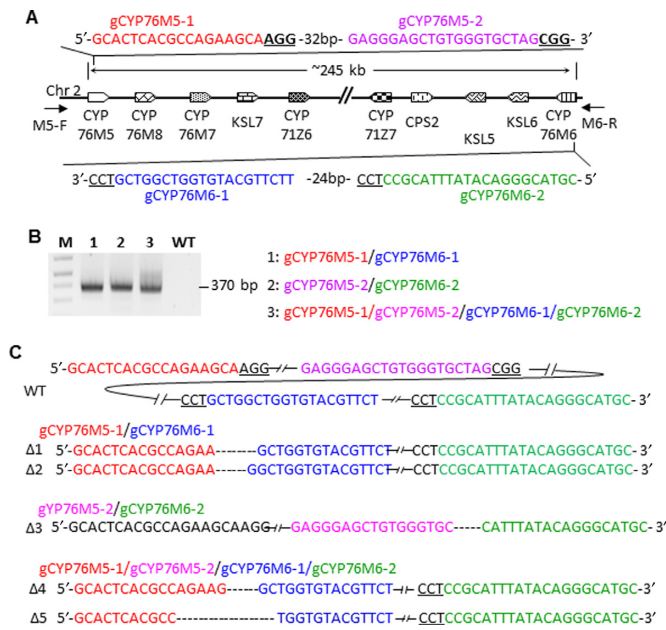


Figure 7. Large segment deletions induced by Cas9/sgRNAs in rice chromosome 2. (A) Schematic illustration of 10 diterpenoid genes spanning about 245 kb on rice chromosome 2, indicating the sgRNA-targeted sites in colors and PAM sequences underlined. (B) Agarose gel electrophoresis of the PCR products amplified from the genomic DNA of protoplasts transfected with sgRNA constructs as indicated by the number that corresponds to the sgRNA constructs at the right side of picture. (C) Sequences of large DNA segment deletions induced by sgRNAs. The wild-type cluster is depicted as DNA sequence (sgRNA-targeted sites in distinct colors) and lines (sequences not shown). The sequences of individual representative clones ($\Delta 1$ to $\Delta 7$) derived from PCR products shown in (B) are presented with the deleted sequences as dashed lines linking the junction sequences between the sgRNA-targeted sites.

OsKOL5 (Supplementary Figure S8A). The sgRNA gene construct was co-expressed with the Cas9 gene construct in rice protoplasts. Detection and confirmation of chromosomal deletions were performed similarly to that for the other two clusters described above. As shown in Supplementary Figure S8, these Cas9/sgRNAs also induced targeted chromosomal deletions (Supplementary Figure S8B and C).

Cas9/sgRNA activity produces rice plants containing large deletions of gene clusters

To determine if we could generate stable rice plants that contain the Cas9/sgRNA-induced deletions of large gene cluster involved in diterpenoid biosynthesis, we recombined the aforementioned cassette of gCYP76M5-1 and gCYP76M6-1 into the binary vector expressing Cas9 under the maize ubiquitin 1 promoter and introduced the resulting construct into rice callus cells through *Agrobacterium*-mediated gene transfer (Figure 8A). Twenty four independent hygromycin-resistant callus lines were obtained and screened for those with putative large segment deletions using PCR with the deletion specific primers M5-F and M6-R. Four positive lines were identified by the PCR approach and further confirmed by sequencing of the amplicons (Figure 8B). Three lines (#16, 17 and 21) contained identical sgRNA-mediated deletions and junction sequences at

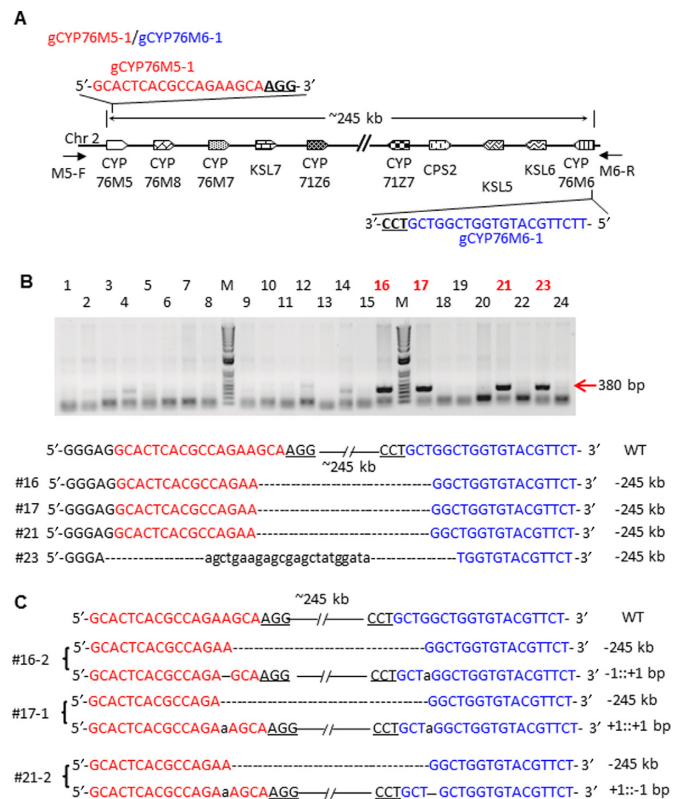


Figure 8. Deletion of 245 kb induced by Cas9/sgRNA activity in rice plants. (A) Schematic illustration of the 245-kb region with two sgRNA-targeted sites as denoted in colors and PAM sequences underlined and bold. (B) Callus lines containing the ~245-kb deletions identified with a PCR approach and confirmed with sequencing of amplicons. The agarose gel image indicates the four callus lines (number in red) are positive for large fragment deletions. The sequences beneath the image are the junction sequences with the deletions (dashed lines) or extra nucleotides (lower case letters). (C) DNA sequence changes in the representative plants generated from three callus lines (#16, 17 and 21) with the Cas9/sgRNA induced large deletions (dashed lines) in one chromosome and small nucleotide changes (dashed lines for deletions and lower case letters for insertions) in the homologous chromosome.

the two target sites, while the fourth line (#23) contained the large deletion and an extra sequence of 23 bp which matched a sequence 28 bp upstream of the PAM sequence for gCYP76M6-1 (Figure 8B). The sequencing trace of a single amplicon from each line indicated the homogeneity of the callus cells with the identical deletion. We detected no occurrence of potential segment inversions in any of the 24 callus lines using PCR with the inversion specific primer pair M5-F & M6-F or primer pair M5-R & M6-R. Out of the 24 callus lines, we were able to regenerate a total of 27 transgenic plants from 10 lines including #16, 17 and 21. Genotyping the individual plantlets using primers M5-F and M6-R and subsequent sequencing of amplicons confirmed the large segment deletions in seven plantlets derived from callus lines #16, 17 and 21. The 245-kb deletions were identical to each other and also identical to the ones detected in their initial callus lines (Figure 8C). To determine whether the large deletion is mono- or di-allelic, site-specific primers M5-F & M5-R and M6-F & M6-R were used to PCR-amplify regions of the two individual target sites,

respectively. Sequencing of the amplicons revealed single base changes (1-bp deletion or 1-bp insertion) at each of the predicted Cas9/sgRNA cleavage sites in the representative plants (Figure 8C).

Similarly we also assembled the four-sgRNA cassette (gCPS4-1/gCPS4-2/gCYP99A2-1/gCYP99A2-2) into the Cas9-expressing binary vector to target the 170-kb diterpenoid gene cluster on chromosome 4 for potential deletion in transgenic plants. The Cas9/sgRNA construct was transformed into callus cells and plants were generated from 16 hygromycin resistant callus lines. PCR analysis with the deletion specific primers CPS-F and CYP99A2-R identified four plants all from one callus line (#8) that were positive for deletions. Sequencing the amplicons confirmed the large deletion in each of the four T0 plants. Three plants contained the identical deletions through joining of the exactly predicted cleavage sites (3 bp upstream of PAM) for the guide RNA gCPS4-1 and gCYP99A2-2, respectively (Supplementary Figure S9, plant #8-2). The other plant contained a large deletion through joining of the cleavage site 10 bp upstream of PAM for gCPS4-1 and the predicted cleavage site of gCYP99A2-1 (Supplementary Figure S9, plant #8-3). There was a 1-bp insertion downstream of the deletion at the predicted gCYP99A2-1 cleavage site in all four plants. To determine if the deletions were mono- or diallelic, PCR amplification and sequencing of the relevant regions were performed and revealed four indels each at the predicted DSB sites on another chromosome in each of the four plants (Supplementary Figure S9). Taken together, our results demonstrate that our Cas9/sgRNA system is highly efficient in inducing large chromosomal deletions.

DISCUSSION

We have developed a simple and high-efficiency CRISPR/Cas system for targeted genomic editing in rice, a major staple crop and also a model species for cereal crops. The system consists of a common Cas9 expressing destination vector and an intermediate vector for cloning of up to four oligo-derived sgRNA genes. The major cloning work involves sequential insertion of two oligo-derived small dsDNAs in the intermediate vector (e.g. pENTR-sgRNA1 or pENTR-sgRNA2). The resulting sgRNA constructs can be easily combined with the master Cas9 binary vector using Gateway clonase into a single construct to target up to four distinct loci in transgenic cells, providing a system for high throughput plasmid construction and large scale genomic editing in rice. Furthermore, due to the compatibility of the rice-derived promoters driving the sgRNA and Cas9 genes with transcription machinery in other grass species, this system may be applicable to other grass-type crops used for food and biofuel production. We have successfully used this high-efficiency CRISPR/Cas system to create mutations in four rice SWEET genes—mutations that we have shown to be genetically transmitted to the second and third generations. Finally, we have demonstrated the feasibility of using this system to generate large chromosomal deletions by successfully deleting three separate clusters of genes.

A variety of different architectures of guide RNA scaffolds and many versions of Cas9 genes have been developed

by different groups for genome targeting in a wide range of organisms, including plant species (19,20,33,42,47). Based on our experimental results, the dual crRNA:tracrRNA guide RNA is significantly less efficient for inducing site-specific gene editing than the sgRNA system. Furthermore, the short sgRNA, such as sgRNA(+48), is much less efficient than sgRNA(+85). Both dual crRNA:tracrRNA and sgRNAs [sgRNA(+85) and sgRNA(+48)] are active to varying extents in rice protoplasts. Highest efficiency in generating targeted gene mutations was obtained with sgRNA(+85). Although the variation of efficiency obtained with different guide RNA architectures (crRNA:tracrRNA and sgRNAs with different lengths of the tracrRNA 3' tails) from our study in rice was established with different guide sequences, we tested their efficiency by using multiple seed sequences for each architecture and in a large number of transgenic plants. Our results with rice are consistent with results from prior studies in other systems using the same seed sequences in different guide architectures (19,31). In regard to Cas9, we demonstrated earlier that the bacterial Sp-Cas9 gene has weak activity in rice protoplasts (21). However, as shown in our present studies, use of the rice codon optimized SpCas9 gene increases the magnitude of mutation production remarkably. With stable transgenic plants, out of 15 plants, only one mono-allelic mutant plant was obtained using a human codon optimized Cas9 gene (data not shown), which was a significantly lower rate than the 87–100% efficiency achieved with the rice codon optimized Cas9. It should be noted that both the bacterial and rice codon optimized Cas9 protein may be somewhat toxic to *Agrobacterium* cells given that the transformed bacterial cells grew notably slower than cells with plasmids lacking the Cas9 gene—even though the Cas9 gene was driven by the intron-containing maize ubiquitin 1 gene promoter. In contrast, transgenic rice plants containing a Cas9 transgene grew normally, without any obvious phenotypic defects in growth and development.

Multiplexed gene targeting has been achieved in a number of species. Such targeting has been achieved using either one sgRNA targeting identical or closely matching sequence regions of homologous genes or with a number of sgRNA targeting multiple, different loci in the genome (42). Based on results in the present study, it appears prudent to use different RNA polymerase III dependent promoters that share no homology with one another, to avoid loss or rearrangement of sgRNA components during plasmid construction or rice transformation through potential DNA recombination between homologous regions in the sgRNA cassettes.

The reports using the CRISPR/Cas9 system in *Arabidopsis* shows that many of the mutations detected in T1 plants were created by Cas9/sgRNAs in somatic cells and that limited numbers of plants with heritable mutations could be obtained in T2 or later generations (31,32). However, our results with transgenic rice plants show that all the T0 plants carried di-allelic mutations and that T1 progeny could be readily obtained carrying heritable homozygous mutations. This means that targeted gene modification occurred in single embryonic rice cells that were regenerated into T0 plants that were capable of efficiently transmitting the mutated genes to subsequent generations. Of particular inter-

est are both the 1-bp insertion (+1) and the 2-bp deletion (−2) mutations of the *SWEET1a* gene found in the T0 line #7, which were separately and faithfully transmitted into T1 plants #4 and #7, suggesting that those minor changes in the PAM proximal target sequence impair Cas9 activity. More interest will be paid in future studies to the progeny of lines #3 and #10, in which another 1-bp deletion/insertion of the *SWEET1a* gene was found. We also will perform whole-genome re-sequencing of progeny from selected mutant lines to investigate the specificity (e.g. off-target effect) of the CRISPR/Cas9 system in rice.

In this study, we report a highly efficient Cas9/sgRNA(+85) platform for endogenous gene targeting in rice plants. Four different genomic loci were successfully mutated at a very high efficiency, including achievement of 100% di-allelic (homogeneous or heterogeneous) mutations. In addition, the mutations were transmitted stably into all T2 generation plants we examined. Our studies pave the way for genome editing by a broad spectrum of rice researchers. Since our Cas9/sgRNA(+85) platform achieves an extremely high efficiency in causing DSBs and subsequent on-target gene editing, including 100% di-allelic mutations, some of the T0 plants could be used directly for biological experimentation, greatly decreasing the time required for basic biology experiments. In addition, those constructs could be used for genome-wide knock-out screening in rice with sgRNA libraries (sgRNA constructs harboring a library of 20-bp oligo-derived seed sequences), resulting in mutant rice populations with much greater genetic diversity and precision than found in traditional insertion libraries made with T-DNA or transposons. Furthermore, with high efficiency gene targeting, endogenous gene tagging with fluorescent protein or small epitope tags likely can be achieved—a step that would greatly speed rice research. Importantly, all the materials created using Cas9/sgRNA(+85) editing could be released into the field without any concern about ecological risks as long as the transgenes were removed through classical genetic segregation.

Herein we provide the first evidence for Cas9/sgRNA-induced, large genomic deletions in a plant protoplast system and fertile, regenerated plants, which represents a first major step toward realizing the exciting potential of this type of targeted, major genomic modification for use in basic plant biology research and for crop bioengineering. The efficiency to generate the transgenic plants with large deletions in one chromosome and small site-specific nucleotide changes in another chromosome is relative high (three lines out 10 for chromosome 2 cluster; Figure 8). Technologies for efficient site-directed chromosomal segment deletion will greatly facilitate the functional analyses of clustered genes and noncoding genetic elements (genes or regulatory modules); such technologies also have practical implication in crop breeding, for example, by generating new agronomic traits in crops through deletion of unwanted loci that may be detrimental to fitness. We will be curious to see the transmission of chromosomal deletions in progeny of our multiple transgenic plants. Because the detected large deletions occurred in embryogenic cells, not in somatic cells, it is most likely they will be inherited as efficiently as the small indels we observed in the present

study. In addition, based on our large segment deletion results, our Cas9/sgRNA platform clearly is a valuable tool potential for chromosomal engineering, including production of investigator-designed chromosomal segment substitution lines, translocation lines, addition lines and deletion lines in rice.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

The authors thank CAMBIA Health Solutions, Inc. for pCAMBIA1300 and Pete J. Lelonek for maintaining greenhouse and growth chamber.

FUNDING

Iowa State University Plant Science Institute [to B.Y., M.H.S.]; National Institute of Food and Agriculture of the US Department of Agriculture [2014-67013-21720 to B.Y.]. Funding for open access charge: The Iowa State University Plant Science Institute Innovative Grant [701-6601-143000].

Conflict of interest statement. The authors declare no conflicts of interest.

REFERENCES

- Stoddard, B.L. (2011) Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure*, **19**, 7–15.
- Gaj, T., Gersbach, C.A. and Barbas, C.F. III (2013) ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.*, **31**, 397–405.
- Voytas, D.F. (2013) Plant genome engineering with sequence-specific nucleases. *Annu. Rev. Plant Biol.*, **64**, 327–350.
- Johnson, R.D. and Jasin, M. (2001) Double-strand-break-induced homologous recombination in mammalian cells. *Biochem. Soc. Trans.*, **29**, 196–201.
- Choulika, A., Perrin, A., Dujon, B. and Nicolas, J.F. (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.*, **15**, 1968–1973.
- Mengiste, T. and Paszkowski, J. (1999) Prospects for the precise engineering of plant genomes by homologous recombination. *Biol. Chem.*, **380**, 749–758.
- Fishman-Lobell, J., Rudin, N. and Haber, J.E. (1992) Two alternative pathways of double-strand break repair that are kinetically separable and independently modulated. *Mol. Cell. Biol.*, **12**, 1292–1303.
- Jasin, M. (1996) Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet.*, **12**, 224–228.
- Vasquez, K.M., Marburger, K., Intody, Z. and Wilson, J.H. (2001) Manipulating the mammalian genome by homologous recombination. *Proc. Natl Acad. Sci. U.S.A.*, **98**, 8403–8410.
- Bibikova, M., Beumer, K., Trautman, J.K. and Carroll, D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764.
- Kim, Y.G. and Chandrasegaran, S. (1994) Chimeric restriction endonuclease. *Proc. Natl Acad. Sci. U.S.A.*, **91**, 883–887.
- Kim, Y.G., Cha, J. and Chandrasegaran, S. (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proc. Natl Acad. Sci. U.S.A.*, **93**, 1156–1160.
- Niu, Y., Tenney, K., Li, H. and Gimble, F.S. (2008) Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. *J. Mol. Biol.*, **382**, 188–202.

14. Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., Bogdanove, A.J. and Voytas, D.F. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, **186**, 757–761.
15. Li, T., Huang, S., Jiang, W.Z., Wright, D., Spalding, M.H., Weeks, D.P. and Yang, B. (2011) TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res.*, **39**, 359–372.
16. Bultmann, S., Morbitzer, R., Schmidt, C.S., Thanisch, K., Spada, F., Elsaesser, J., Lahaye, T. and Leonhardt, H. (2012) Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. *Nucleic Acids Res.*, **40**, 5368–5377.
17. Kim, Y., Kweon, J., Kim, A., Chon, J.K., Yoo, J.Y., Kim, H.J., Kim, S., Lee, C., Jeong, E., Chung, E. *et al.* (2013) A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.*, **31**, 251–258.
18. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, **337**, 816–821.
19. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A. *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, **339**, 819–823.
20. Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. and Church, G.M. (2013) RNA-guided human genome engineering via Cas9. *Science*, **339**, 823–826.
21. Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B. and Weeks, D.P. (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Res.*, **41**, e188.
22. Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y. *et al.* (2013) Efficient genome editing in plants using a CRISPR/Cas system. *Cell Res.*, **23**, 1229–1232.
23. Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M. and Sheen, J. (2013) Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat. Biotechnol.*, **31**, 688–691.
24. Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D. and Kamoun, S. (2013) Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.*, **31**, 691–693.
25. Upadhyay, S.K., Kumar, J., Alok, A. and Tuli, R. (2013) RNA-guided genome editing for target gene mutations in wheat. *G3*, **3**, 2233–2238.
26. Xie, K. and Yang, Y. (2013) RNA-guided genome editing in plants using a CRISPR-Cas system. *Mol. Plant*, **6**, 1975–1983.
27. Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N. *et al.* (2014) The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.*, **12**, 797–807.
28. Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F. and Zhu, J.K. (2013) Application of the CRISPR-Cas system for efficient genome engineering in plants. *Mol. Plant*, **6**, 2008–2011.
29. Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H. and Qu, L.J. (2013) Targeted mutagenesis in rice using CRISPR-cas system. *Cell Res.*, **23**, 1233–1236.
30. Liang, Z., Zhang, K., Chen, K. and Gao, C. (2013) Targeted mutagenesis in *Zea mays* using TALENs and the CRISPR/Cas system. *J. Genet. Genomics*, **41**, 63–68.
31. Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L. *et al.* (2014) Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc. Natl Acad. Sci. U.S.A.*, **111**, 4632–4637.
32. Jiang, W., Yang, B. and Weeks, D.P. (2014) Efficient CRISPR/Cas9-mediated gene editing in Arabidopsis thaliana and inheritance of modified genes in T2 and T3 generations. *PLoS ONE*, **9**, e99225.
33. Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H. and Qu, L.J. (2013) Targeted mutagenesis in rice using CRISPR-Cas system. *Cell Res.*, **23**, 1233–1236.
34. Liu, P.Q., Chan, E.M., Cost, G.J., Zhang, L., Wang, J., Miller, J.C., Guschin, D.Y., Reik, A., Holmes, M.C., Mott, J.E. *et al.* (2010) Generation of a triple-gene knockout mammalian cell line using engineered zinc-finger nucleases. *Biotechnol. Bioeng.*, **106**, 97–105.
35. Gupta, A., Hall, V.L., Kok, F.O., Shin, M., McNulty, J.C., Lawson, N.D. and Wolfe, S.A. (2013) Targeted chromosomal deletions and inversions in zebrafish. *Genome Res.*, **23**, 1008–1017.
36. Xiao, A., Wang, Z., Hu, Y., Wu, Y., Luo, Z., Yang, Z., Zu, Y., Li, W., Huang, P., Tong, X. *et al.* (2013) Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.*, **41**, e141.
37. Choi, P.S. and Meyerson, M. (2014) Targeted genomic rearrangements using CRISPR/Cas technology. *Nat. Commun.*, **5**, 1–6.
38. Qi, Y., Li, X., Zhang, Y., Starker, C.G., Baltes, N.J., Zhang, F., Sander, J.D., Reyon, D., Joung, J.K. and Voytas, D.F. (2013) Targeted deletion and inversion of tandemly arrayed genes in *Arabidopsis thaliana* using zinc finger nucleases. *G3*, **3**, 1707–1715.
39. Hiei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (*Oryza sativa* L.) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *Plant J.*, **6**, 271–282.
40. Porebski, S., Bailey, L.G. and Baum, B. (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Rep.*, **15**, 8–15.
41. Li, T., Liu, B., Spalding, M.H., Weeks, D.P. and Yang, B. (2012) High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.*, **30**, 390–392.
42. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O. *et al.* (2013) DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.*, **31**, 827–832.
43. Wilderman, P.R., Xu, M., Jin, Y., Coates, R.M. and Peters, R.J. (2004) Identification of syn-pimara-7,15-diene synthase reveals functional clustering of terpene synthases involved in rice phytoalexin/allelochemical biosynthesis. *Plant Physiol.*, **135**, 2098–2105.
44. Shimura, K., Okada, A., Okada, K., Jikumaru, Y., Ko, K.W., Toyomasu, T., Sassa, T., Hasegawa, M., Kodama, O., Shibuya, N. *et al.* (2007) Identification of a biosynthetic gene cluster in rice for momilactones. *J. Biol. Chem.*, **282**, 34013–34018.
45. Swaminathan, S., Morrone, D., Wang, Q., Fulton, D.B. and Peters, R.J. (2009) CYP76M7 is an ent-cassadiene C11 α -hydroxylase defining a second multifunctional diterpenoid biosynthetic gene cluster in rice. *Plant Cell*, **21**, 3315–3325.
46. Wu, Y., Hillwig, M.L., Wang, Q. and Peters, R.J. (2011) Parsing a multifunctional biosynthetic gene cluster from rice: biochemical characterization of CYP71Z6 & 7. *FEBS Lett.*, **585**, 3446–3451.
47. Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Tsai, S.Q., Sander, J.D., Peterson, R.T., Yeh, J.R. and Joung, J.K. (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.*, **31**, 227–229.