Original research

Investigation of CRISPR/Cas9-induced *SD1* rice mutants highlights the importance of molecular characterization in plant molecular breeding

Sukumar Biswas [a], Jiaqi Tian [a], Rong Li [a], Xiaofei Chen [a], Zhijing Luo [a], Mingjiao Chen [a], Xiangxiang Zhao [b], Dabing Zhang [a, b], Staffan Persson [a, c], Zheng Yuan [a, b, *], Jianxin Shi [a, *]

[a] *Joint International Research Laboratory of Metabolic & Developmental Sciences, Shanghai Jiao Tong University-University of Adelaide Joint Centre for Agriculture and Health, State Key Laboratory of Hybrid Rice, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, 200240, China*

[b] *Key Laboratory of Crop Marker-Assisted Breeding of Huaian Municipality, Jiangsu Collaborative Innovation Center of Regional Modern Agriculture and Environmental Protection, Huaian, 223300, China*

[c] *School of Biosciences, University of Melbourne, Parkville VIC, 3010, Melbourne, Australia*

ARTICLE INFO

ABSTRACT

Although Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated 9 (Cas9) system has been widely used for basic research in model plants, its application for applied breeding in crops has faced strong regulatory obstacles, due mainly to a poor understanding of the authentic output of this system, particularly in higher generations. In this study, different from any previous studies, we investigated in detail the molecular characteristics and production performance of CRISPR/Cas9-generated *SD1* (*semi-dwarf 1*) mutants from T_2 to T_4 generations, of which the selection of T_1 and T_2 was done only by visual phenotyping for semidwarf plants. Our data revealed not only on- and off-target mutations with small or lager indels but also exogenous elements in T_2 plants. All indel mutants passed stably to T_3 or T_4 without additional modifications independent on the presence of Cas9, while some lines displayed unexpected hereditary patterns of Cas9 or some exogenous elements. In addition, effects of various *SD1* alleles on rice height and yield differed depending on genetic backgrounds. Taken together, our data showed that the CRISPR/Cas9 system is effective in producing homozygous mutants for functional analysis, but it may be not as precise as expected in rice, and that early and accurate molecular characterization and screening must be carried out for generations before transitioning of the CRISPR/Cas9 system from laboratory to field.

1. Introduction

Genetically modified (GM) crops via *Agrobacterium*-mediated T-DNA transfer have contributed greatly to agronomic, environmental, economic, health, and social benefits for farmers and consumers (ISAAA, 2019). However, disputes on food and environmental safety of GM crops have been steadily increasing (Huang, 2017). Therefore, many countries have developed regulatory systems for risk assessment and management of GM crops. While these regulatory systems are fully implemented, current GM regulatory frameworks increase not only the uncertainty but also the cost of GM crop development. Because homologous recombination via *Agrobacterium*-mediated T-DNA transfer in plants is typically not possible (Araki and Ishii, 2015), the identification of T-DNA inserts and other mutations is labor-intensive and time-consuming. New plant breeding techniques including genome editing offer opportunities for crop improvement without maintaining a transgene (Zhang et al., 2018b) and also bring about different challenges regarding the regulation and social acceptance of these new crop products (Araki and Ishii, 2015; Schaeffer and Nakata, 2015; Gao et al., 2018).

Genome editing systems use site-specific nucleases to introduce precisely targeted double-strand breaks (DSBs), while the desired modifications are subsequently obtained by endogenous DSB repair machinery. The site-specific nucleases include zinc finger nucleases (ZFNs), transcription activator–like effector nucleases (TALENs),

---

* Corresponding authors.
 E-mail addresses: zyuan@sjtu.edu.cn (Z. Yuan), jianxin.shi@sjtu.edu.cn (J. Shi).

and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated 9 (Cas9) (Lusser et al., 2012; Zhu et al., 2017). Among these, the CRISPR/Cas9 system has proven superior in specificity and precision and can modify single or multiple genes in plants (Feng et al., 2014, 2016; Xie et al., 2015; Ishizaki, 2016; Zong et al., 2017). In contrast with traditional GM techniques that rely on random recombination or integration, the CRISPR/Cas9 editing system is considered target specific and precise (Lusser et al., 2012). CRISPR/Cas9 nuclease cuts genomic target guided by a single guide RNA (sgRNA), the created DSBs are repaired by an error-apt nonhomologous end joining repair mechanism that can lead to indel (insertion-deletion) mutations in the absence of an exogenous DNA donor (Feng et al., 2014; Zhu et al., 2017). Therefore, the CRISPR/Cas9 system may facilitate crop improvement via possible escape from GM regulation because mutants generated by the CRISPR/Cas9 system do not necessarily contain any exogenous DNA insertions (Feng et al., 2014; Jones, 2015; Li et al., 2017; Wolter and Puchta, 2017; Zhu et al., 2017; Gao et al., 2018). On the other hand, if products obtained from mutagenesis intentionally and specifically alter the genetic materials of an organism in a way that does not occur naturally, they can, in principal, be subject to current GM regulation (Araki and Ishii, 2015; Court of Justice of the European Union, 2018). Hence, although CRISPR/Cas9 systems open up a new opportunity for quick and precise modification of crops, for example, to boost productivity (Jones, 2015; Li et al., 2017), to protect against pests and diseases (Wang et al., 2014), and to enhance nutrient content (Liang et al., 2014), the application of CRISPR/Cas9 in crop improvement has raised increasing concerns about the safety of its products. This highlights the needs for molecular characterization of CRISPR/Cas9 mutants.

Most molecular characterizations of CRISPR/Cas9-generated mutations in plants have focused mainly on the patterns and segregation of the targeted gene modifications either in transient systems or in early generations of stable transformants (Feng et al., 2014; Wang et al., 2014; Zhang et al., 2014; Zhou et al., 2014; Zhu et al., 2017). Much less is known about molecular characteristics in subsequent generations. In *Arabidopsis*, CRISPR/Cas9 is very effective in targeted gene mutagenesis (dominated by small indels), resulting homozygous lines are stable for the mutations in subsequent generations, and off-target mutations are rare (Feng et al., 2014). In rice, the stability and off-target sites of CRISPR/Cas9 editing are largely unknown, although there were several reports on the immediate generations after transformation (Zhang et al., 2014; Zhou et al., 2014; Ishizaki, 2016; Tang et al., 2018). In addition, none of these studies reported any nontarget DNA modifications in CRISPR/Cas9-generated rice mutants; such DNA modifications would be an important issue associated with GM regulation when breeding is considered (Araki and Ishii, 2015; Schaeffer and Nakata, 2015).

The short stature of the *semi-dwarf 1* (*SD1*) variety of rice IR8 results from a natural mutation in the *SD1* gene that encodes a GA20 oxidase 2 (GA20ox-1) associated with the biosynthesis of plant hormone gibberellin (Sasaki et al., 2002). The null allele of *SD1* was selected by nature during the rice green revolution, and a recent study reported that it also participates in the adaptation of rice to periodic flooding (Kuroha et al., 2018). Therefore, manipulation of *SD1* in elite rice cultivars has attracted remarkable interest (Li et al., 2018). In this study, we undertook a detailed molecular characterization of several generations of CRISPR/Cas9-generated *SD1* mutants in rice, with an aim to enhance elite rice performance using the CRISPR/Cas9 system. For this purpose, rice plants from several elite cultivars, including 9815B, JIAODA138, and HUAI-DAO1055, were transformed with a construct harboring the *Cas9* gene together with a single stranded guide RNA targeting *SD1*. We found that CRISPR/Cas9-induced mutation in rice was likely not as

precise as expected and that not every mutated *SD1* allele could be potentially used to enhance yield. Our results highlight the necessity of detailed early and accurate molecular characterization and performance evaluation of CRISPR/Cas9-generated mutants before the transition of CRISPR/Cas9-edited rice from laboratory to field.

2. Results

2.1. SD1-edited semidwarf T_2 plants harbor large chromosomal deletion and insertion

To explore the possibility to improve the production performance of some elite rice cultivars via manipulating the green revolution gene *SD1* by using the CRISPR/Cas9 system, rice plants from several elite cultivars, including 9815B, JIAODA138, and HUAIDAO1055 (Table 1), were transformed with a construct harboring the Cas9 (Fig. S1A) together with a single stranded guide RNA targeting the first exon of *SD1* (Fig. S1B). This construct contained also several exogenous elements including LacZ, HPT II, 35S, T-NOS, and F_1 fragment (Fig. S1A). Nontransformed callus or T_0 plants were excluded after antibiotic selection, and positive T_1 plants were visually selected based on the semidwarf phenotype without molecular characterization. Leaf samples from a total of 31 individual T_2 lines with the semidwarf phenotype were collected for molecular characterization. Based on genotyping data, selected T_2 seeds were collected and used for further analyses (Fig. 1A). This workflow, different from previous studies (Shan et al., 2014; Zhang et al., 2014; Zhou et al., 2014; Ishizaki, 2016), focused initially on phenotyping rather than genotyping, which allowed us to reexamine the outputs of CRISPR/Cas9 mutagenesis in rice and rethink the strategy for crop breeding using the CRISPR/Cas9 system.

Because all chosen T_2 lines were semidwarf, we expected to find on-target small indels or mutations closely associated with the *SD1* locus in accordance with the reported accuracy of similar systems (Zhang et al., 2014; Zhou et al., 2014). To our surprise, the *SD1* fragments in six of 31 T_2 lines could not be amplified (Fig. S2A) using the primer pair (SD1-F/SD1-R) flanking closely the expected sgRNA targeted site (Fig. S1C). Instead, we had to use another primer pair (SD1-F1/SD1-R1) that binds much further away from the expected editing site (Fig. S1C) to amplify them (Fig. S2B). This result indicated the occurrence of large chromosomal deletions or rearrangements at the expected sites in these lines. We sequenced the polymerase chain reaction (PCR) products, which confirmed several different large deletions in these lines (Figs. S3 and S4) where PCR amplification via SD1-F/SD1-R had failed. These data indicate that the CRISPR/Cas9 fidelity might not be as precise as previously suggested (Feng et al., 2014; Zhang et al., 2014; Zhou et al., 2014).

2.2. SD1-edited semidwarf T_2 plants display diverse genotypes

All *SD1* T_2 mutants could be classified into homozygous, heterozygous, biallelic, and chimeric as compared with nonedited plants based on Sanger sequencing results (Figs. 1A and S3), confirming high CRISPR/Cas9 efficiency (Zhang et al., 2014; Zhou et al., 2014). All 31 *SD1* T_2 mutant lines displayed 20 genotypes as compared with the genomic DNA sequence of the wild-type *SD1*, including four different insertions (10 lines), ten different deletions (14 lines), two bialleles (two lines), and four chimeras (five lines) (Fig. 1B). Nearly half of mutations (16 of 31) occurred at the expected 4th nucleotide position upstream of the protospacer adjacent motif (PAM), but the PAM was found to be absent in one-third of mutations (10 of 31) (Fig. 1B). Remarkably, one mutant (Q115) maintained an intact sgRNA target together with a large deletion (382 bp) occurred far downstream of the target (Fig. 1B). Notably,

Table 1
Signatures and segregations of CRISPR/Cas9-induced SD1 mutants in rice.

Line	Background	T2			T3		T4	
		Signature	Target genotype	Cas9	Mutation segregation	Cas9	Mutation segregation	Cas9
Q10	9815B	Homozygote	i1/i1	+	nt	nt	nt	nt
Q11	9815B	Homozygote	d7/d7	+	nt	nt	nt	nt
Q13	9815B	Homozygote	dL/dL	+	nt	nt	nt	nt
Q14	9815B	Homozygote	d7/d7	+	nt	nt	nt	nt
Q16	9815B	Heterozygote	d1/WT	+	nt	nt	nt	nt
Q18	9815B	Homozygote	i1/i1	+	10i1/i1	10+	10i1/i1	2+/8−
Q21	9815B	Heterozygote	i1/WT	+	3i1/i1:5h:2WT	8+/2−	10i1/i1	10−
Q23	9815B	Homozygote	d24/d24	+	10d24/d24	10+	10d24/d24	1+/9−
Q26	9815B	Homozygote	i1/i1	−	10i1/i1	10−	nt	nt
Q27	9815B	Homozygote	d257	+	10d257	10+	10d257	2+/8-
Q30	9815B	Homozygote	i1/i1	+	10i1/i1	9+/1−	10i1/i1	10−
Q31	9815B	Homozygote	d63/d63	+	10d63/d63	7+/3−	10d63/d63	10−
Q34	9815B	Homozygote	d4/d4	+	10d4/d4	10−	nt	nt
Q36	9815B	Homozygote	d7/d7	−	10d7/d7	10−	nt	nt
Q41	9815B	Chimera	d263,i194,r1	+	10d263/i194/r1	9+/1−	10d263/i194/r1	10−
Q46	JIAODA138	Homozygote	i1/i1	−	10i1i1	10−	nt	nt
Q48	JIAODA138	Homozygote	i1/i1	−	10i1/i1	10−	10i1/i1	10−
Q56	JIAODA138	Homozygote	i1/i1	+	10i1/i1	6+/4−	10i1/i1	10−
Q60	JIAODA138	Homozygote	i1/i1	+	10i1/i1	4+/6−	10i1/i1	10−
Q62	JIAODA138	Biallele	r1/i1	+	2r1/r1:5r1/i1:3i1/i1	6+/4−	10i1/i1	10−
Q71	HUAIDAO1055	Homozygote	i5/i5	+	10i5/i5	5+/5−	10i5/i5	10−
Q73	HUAIDAO1055	Homozygote	d7/d7	−	10d7/d7	10−	10d7/d7	10−
Q74	HUAIDAO1055	Biallele	r1/i1	+	3r1/r1:5r1/i1:2i1/i1	7+/3−	10i1/i1	10−
Q76	HUAIDAO1055	Homozygote	d1/d1	−	10d1/d1	10−	10d1/d1	10−
Q79	HUAIDAO1055	Homozygote	d19/d19	+	nt	nt	nt	nt
Q86	HUAIDAO1055	Chimera	d3,i1,r3	+	nt	nt	nt	nt
Q89	HUAIDAO1055	Homozygote	d63/d63	+	nt	nt	nt	nt
Q97	HUAIDAO1055	Homozygote	d2/d2	+	nt	nt	nt	nt
Q103	HUAIDAO1055	Chimera	d382,i1,r1	+	nt	nt	nt	nt
Q107	HUAIDAO1055	Chimera	d382,i1,r1	+	nt	nt	nt	nt
Q115	HUAIDAO1055	Chimera	WT,d382,i1	+	nt	nt	nt	nt

+, Cas9 detected; −, Cas9 not detected; nt, not tested. d#, deletion with # bp; i#, insertion with # bp; r#, replacement of # bp; h, heterozygous; WT, wild-type; #d, #i, #r, #+, #−, number of lines with identified deletion, insertion, replacement, presence of, and absence of Cas9, respectively.

one mutant (Q41) contained a large insertion (194 bp) in addition to a large deletion (257 bp); this large insertion was actually a rearrangement of a 194-bp fragment of SD1 upstream of the PAM (Fig. 1B). Therefore, although Q41 had the same length of deletion as that of Q27, the PCR products of Q41 were larger than those of Q27 (Fig. S2B). These data reflected again the complex outputs of CRISPR/Cas9 in rice.

In addition, mutants on different elite backgrounds differed in their genotype complexities. For example, CRISPR/Cas9-edited SD1 in 9815B and HUAIDAO1055 appeared relatively more complicated than those in JIAODA138 (Table 1). While it is difficult to explain, this result highlighted genotype-specific modification through the introduction of the CRISPR/Cas9 system.

2.3. Some SD1-edited semidwarf T2 plants display possible off-target mutation

The CRISPR/Cas9 system may also generate off-target mutation as it tolerates up to three mismatches between the sgRNA and the target (Zhu et al., 2017; Wu and Yin, 2019). To assess if such off-targets occurred in our SD1 lines, we screened possible off-target mutations for the five highest scoring targets from the CRISPR-P software (Liu et al., 2017) using genomic DNA extracted from the 31 T2 mutants. We found off-target mutations in three of 31 T2 mutants (Fig. 1C). Among them, Q115 had two off-targets: off-target 5 with a "C" instead of an "A" at the 7th position, and off-target 2 with a "G" instead of an "A" at the 6th position immediately upstream of the PAM, while Q103 and Q107 both had the same off-target 5 as that of Q115 (Fig. 1C). The two off-targets were found

in three independently transformed lines, while no such mutation was found in the corresponding wild type.

2.4. Most of SD1-edited semidwarf T2 plants still harbor exogenous elements

Because CRISPR/Cas9 systems were transferred to rice via conventional Agrobacterium-mediated transformation, the possible existence of exogenous T-DNA elements was further investigated. Of all 31 T2 lines examined (Table 2), most of them (25 of 31) contained simultaneously three exogenous T-DNA elements, namely, NOS, CaMV35S, and HPT. Among them, four lines in 9815B background contained additional exogenous T-DNA element LacZ. Other exogenous T-DNA elements examined, including NPTll, F1 fragment, and pBIN, were found to be absent in all T2 mutants. Notably, these 25 T2 lines that contained exogenous T-DNA elements also harbored Cas9 in their genomes. This result highlighted the necessity for early molecular screening to get rid of any exogenous elements (including Cas9) in CRISPR/Cas9-edited mutants in rice.

2.5. Genotypes of edited SD1 in semidwarf T2 plants pass over stably in a Mendel way

The inheritability of those CRISPR/Cas9-edited SD1 mutants was further investigated using identified 19 T2 lines, including 15 homozygous, two bialleles, one chimera, and one heterozygous (Fig. 1A and Table 1). All detected 180 T3 plants from 18 T2 homozygous lines including two bialleles and one chimera (10 plants each line) were homozygous, showing identical genotypes to their
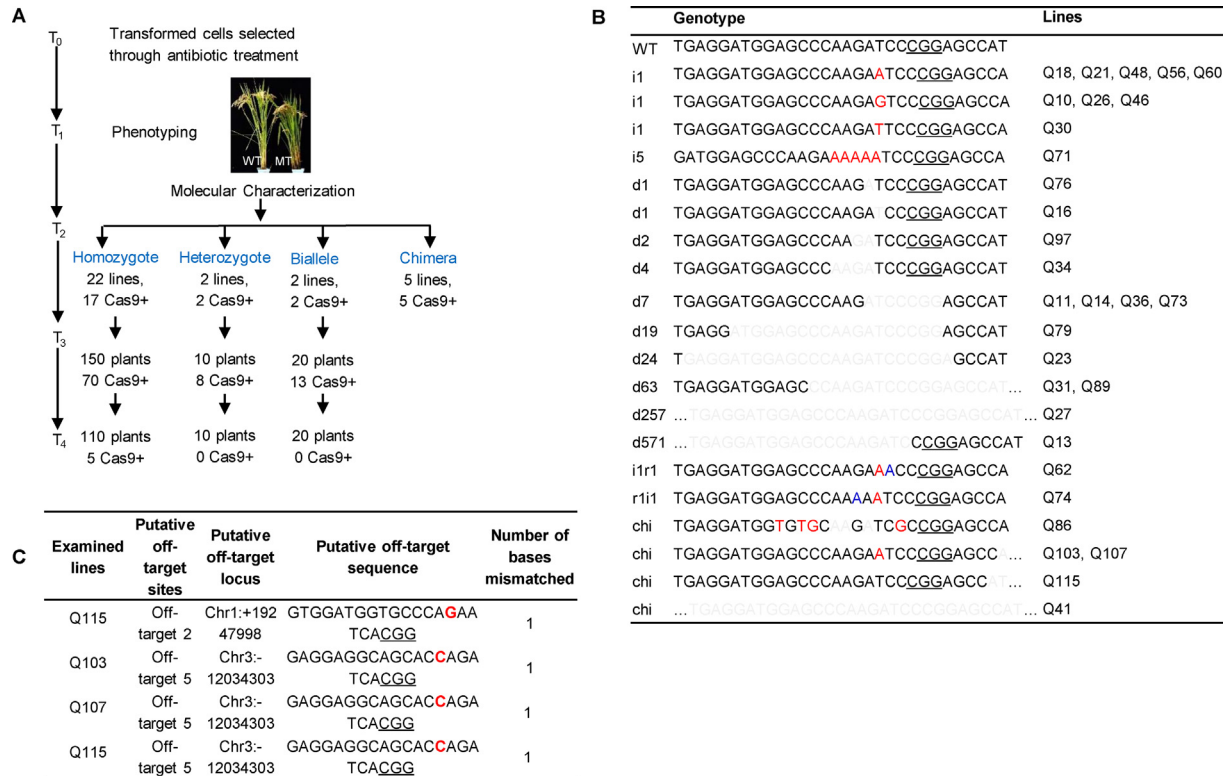
Fig. 1. The molecular characteristics of CRISPR/Cas9-induced *SD1* mutants in rice. **A**: Diagram summary of the experimental design and the final output of CRISPR/Cas9-induced *SD1* mutants. **B**: Sequencing results of the identified 20 genotypes in CRISPR/Cas9-induced *SD1* T_2 plants. WT, wild type; d#, deletion of # bp; i#, insertion of # bp; r#, replacement of # bp; chi, chimera. Ellipsis (…) indicates the occurrence of large chromosomal deletion. **C**: Summary of off-targets detected in *SD1* T_2 mutants.

corresponding T_2 plants, without any extra modifications, no matter if Cas9 was present or absent (Table 1). In addition, the T_2 heterozygous mutant (Q21) segregated in T_3 generation with a homozygous-to-heterozygous-to-wild type ratio of 3:5:2 (close to 1:2:1) without additional genotypes even in the presence of Cas9

(Fig. 1A; Table 1). The segregation patterns of 15 T_3 homozygous lines were further examined in T_4 generation, and the results showed that all of them pass stably from T_3 to T_4, without any novel modifications, regardless of the presence and absence of Cas9 (Fig. 1A; Table 1). As reported in *Arabidopsis* (Feng et al., 2014), our

Table 2
Signatures and segregations of detected exogenous elements in CRISPR/Cas9-induced *SD1* mutants in rice.

Line	T_2				T_3			
	T-DNA element			Vector backbone element	T-DNA element			Vector backbone element
	HPT	35S	NOS	LacZ	HPT	35S	NOS	LacZ
Q18	+	+	+	−	10+	10−	10+	10−
Q21	+	+	+	−	8+/2−	2+/8−	2+/8−	10−
Q23	+	+	+	−	10+	10−	10+	10−
Q26	−	−	−	−	10−	10−	10−	10−
Q27	+	+	+	−	10+	10−	10+	10−
Q30	+	+	+	−	9+/1−	10−	9+/1−	10−
Q31	+	+	+	+	7+/3−	10−	4+/6−	10−
Q34	+	+	+	+	10−	10−	10−	10−
Q36	−	−	−	−	10−	10−	10−	10−
Q41	+	+	+	+	9+/1−	6+/4−	9+/1−	10−
Q46	−	−	−	−	10−	10−	10−	10−
Q48	−	−	−	−	10−	10−	10−	10−
Q56	+	+	+	−	6+/4−	10−	6+/4−	10−
Q60	+	+	+	−	6+/4−	10−	6+/4−	10−
Q62	+	+	+	−	4+/6−	10−	3+/7−	10−
Q71	+	+	+	−	7+/3−	10−	7+/3−	10−
Q73	−	−	−	−	10−	10−	10−	10−
Q74	+	+	+	−	7+/3−	1+/9−	6+/4−	10−
Q76	−	−	−	−	10−	10−	10−	10−

+ and −, presence and absence of detected corresponding exogenous elements, respectively; nt, not tested; #+, #−, numbers of lines with detected exogenous elements; CRISPR, Clustered Regularly Interspaced Short Palindromic Repeats; Cas9, CRISPR-associated 9.

data indicated that CRISPR/Cas9-edited *SD1* mutation in rice is heritable in a Mendel way, whether it is heterozygous or homozygous, whether with or without Cas9.

2.6. Exogenous elements in semidwarf T_2 plants segregate irregularly

To follow the segregation of these detected T-DNA elements and Cas9, we continued our exogenous element identification in T_3 plants derived from 19 T_2 lines (including six lines without any exogenous elements, 13 lines carrying T-DNA and Cas9, 10 plants each) (Table 2). T_3 plants from these six lines without any exogenous elements were free of T-DNA and Cas9. However, although LacZ was totally absent in all 130 T_3 plants derived from those 13 T_2 lines carrying T-DNA elements, the presence of any one of NOS, CaMV35S, and HPT was detected in all of them. Notably, the Cas9 could be detected in T_4 plants of three lines. These results indicated possible multiple copies of T-DNA and Cas9 elements in these lines.

2.7. CRISPR/Cas9-edited SD1 alleles reduce plant height significantly but affect yield differently

The effects of various mutated *SD1* alleles on plant height and yield (grain weight per plant) were investigated in both T_3 and T_4 generations. As expected, all mutated *SD1* alleles significantly reduced plant height, and the resulting semidwarf traits passed stably from T_3 to T_4 generation. Generally, the effect of mutated *SD1* alleles on plant height was more evident in HUAIDAO1055 but less evident in JIAODA138 (Fig. 2A). Surprisingly, most of mutated *SD1* alleles significantly reduced the yield as well, which also passed stably from T_3 to T_4 generation. Generally, the effect of mutated *SD1* alleles on yield was more evident in 9815B and much less evident in JIAODA138 (Fig. 2B).

Incidentally, only one line (Q48 in JIAODA138 background) displayed both consistently decreased plant height and moderately elevated yield (Fig. 2). We noted that this line might be suitable as a future rice breeding line because Q48 was free of either Cas9 or any other tested exogenous elements. This specific allelic mutation resulted in a truncated and novel protein, which only shares the first 41 amino acids with wild-type SD1 (Fig. S5).

3. Discussion

It is well known that the predominant DSB repair pathway in plants is nonhomologous end joining, tending to generate short indels. In diploid plants, mutants induced by the CRISPR/Cas9 system generally can be either heterozygous (single allelic change), or homozygous (identical changes to both alleles), or biallelic (different changes at each allele) (Zhu et al., 2017). However, signatures occurring in crops still require further exploration because the outcome of the CRISPR/Cas9 system might vary with species, target sites, transformation methods, and CRISPR/Cas9 systems. The signatures of CRISPR/Cas9-induced gene mutations in later *Arabidopsis* generations (T_2 to T_3) have been intensively characterized (Feng et al., 2014; Jiang et al., 2014; Peterson et al., 2016; Wolt et al., 2016; Zhang et al., 2018b). However, data on rice, an important staple food crop, in later generations (T_1 to T_2) are preliminary and limited (Feng et al., 2014; Shan et al., 2014; Zhou et al., 2014; Xu et al., 2015; Ishizaki, 2016; Tang et al., 2018). In addition, little efforts were made on the molecular characteristics of exogenous elements in those CRISPR/Cas9-induced mutants, particularly in crops (Xu et al., 2015), an important issue that is highly associated with food safety (Convention on Biological Diversity, 2000).

3.1. Signatures

Among all 31 T_2 samples examined, the mutation pattern seemed to be different from those reported in early generations (Miao et al., 2013; Zhang et al., 2014; Zhou et al., 2014). All T_2 plants examined were mutants, largely due to the consequence of the selection for semidwarf phenotype, which were dominated by homozygotes including one biallele and five chimeras (Fig. 1A and B). This result confirmed the high efficiency of the CRISPR/Cas9 system in rice mutagenesis. However, CRISPR/Cas9-induced *SD1* signatures were quite different from those as revealed in early-generation *Arabidopsis* (Feng et al., 2014; Jiang et al., 2014; Zhang et al., 2018a), or rice (Mao et al., 2013; Miao et al., 2013; Zhang et al., 2014; Zhou et al., 2014). First, among these 31 semidwarf T_2 lines, in addition to small indels, namely small insertion (ranging from 1 to 5 bp, 10 of 31) and small deletion (ranging from 1 to 63 bp, 12 of 31), there was a relatively high frequency of large deletions (ranging from 257 to 571 bp, 6 of 31) (Figs. 1B and S4). Second, the edited sites revealed were not as precise as reported. Besides a high frequency of mutations (16 of 31) occurring at the 4th nucleotide position upstream of the PAM, additional one-third mutations (10 of 31) lost the PAM itself (Fig. 1B). Third, Q115 mutant displayed unaffected sgRNA target but a large deletion (382 bp) farther downstream of the sgRNA target. Fourth, in addition to a large deletion, Q41 also had a large rearranged insertion (194 bp). Therefore, our data indicated that the outcome of CRISPR/Cas9-generated mutants varies on species and loci, which needs careful molecular characterization on a case by case base.

Notably, all large deletion mutations detected in T_2 plants contained Cas9. Because we did not identify the genotype of T_1 plants, we did not know when large deletion occurred. Large deletion detected in T_2 could be the consequence of the function of persisted Cas9. However, further characterization of all T_3 progenies derived from 19 T_2 lines including large deletion lines with Cas9 (Q27 and Q41) did not find any new modifications in those lines, indicating that the large deletion could occur at earlier generations. Because Q27 and Q41 lacked the sgRNA target (Fig. 1B), we cannot exclude the possible silencing of the Cas9 and/or guide RNA transgenes in T_2 rice as reported previously (Zhang et al., 2014). Although large deletions were reported in T_0 (Zhou et al., 2014) or T_1 (Mao et al., 2013) rice plants using a vector containing two sgRNAs, reports on large deletion in CRISPR/Cas9-induced rice mutants with a single sgRNA are scarce. Data from this study, together with others (Miao et al., 2013), indicated that gene-specific factors affect the outcome of DSB repair and thus the CRISPR/Cas9 system (Zhu et al., 2017), which explained partially the mutation signatures observed in our study. The investigation on the reasons behind the signatures revealed in *SD1* mutants was ongoing, which included different CRISPR/Cas9 systems, different sgRNAs, and same sgRNA targeting different sites. Our preliminary results verified that large deletions including PAM occurred in two of ten T_1 mutants that were generated using the same CRISPR/Cas9 vector with a different sgRNA targeting the promoter region of the *SD1* (Fig. S6).

3.2. Off-targets

It is recently reported that the CRISPR-Cas9 system may lead to many unexpected mutations including high frequency of off-targets in several mammal cell lines (Kosicki et al., 2018), which raised serious safety concerns about the safety of the CRISPR/Cas9 system for clinical applications (Mattei, 2018). The off-target beyond the target gene has been carefully examined in plants including *Arabidopsis*, rice, cotton, and tobacco, and results indicated that off-target is rare (Feng et al., 2014; Zhang et al., 2014; Gao et al., 2015; Xu et al., 2015; Tang et al., 2018). Nonetheless, a recent
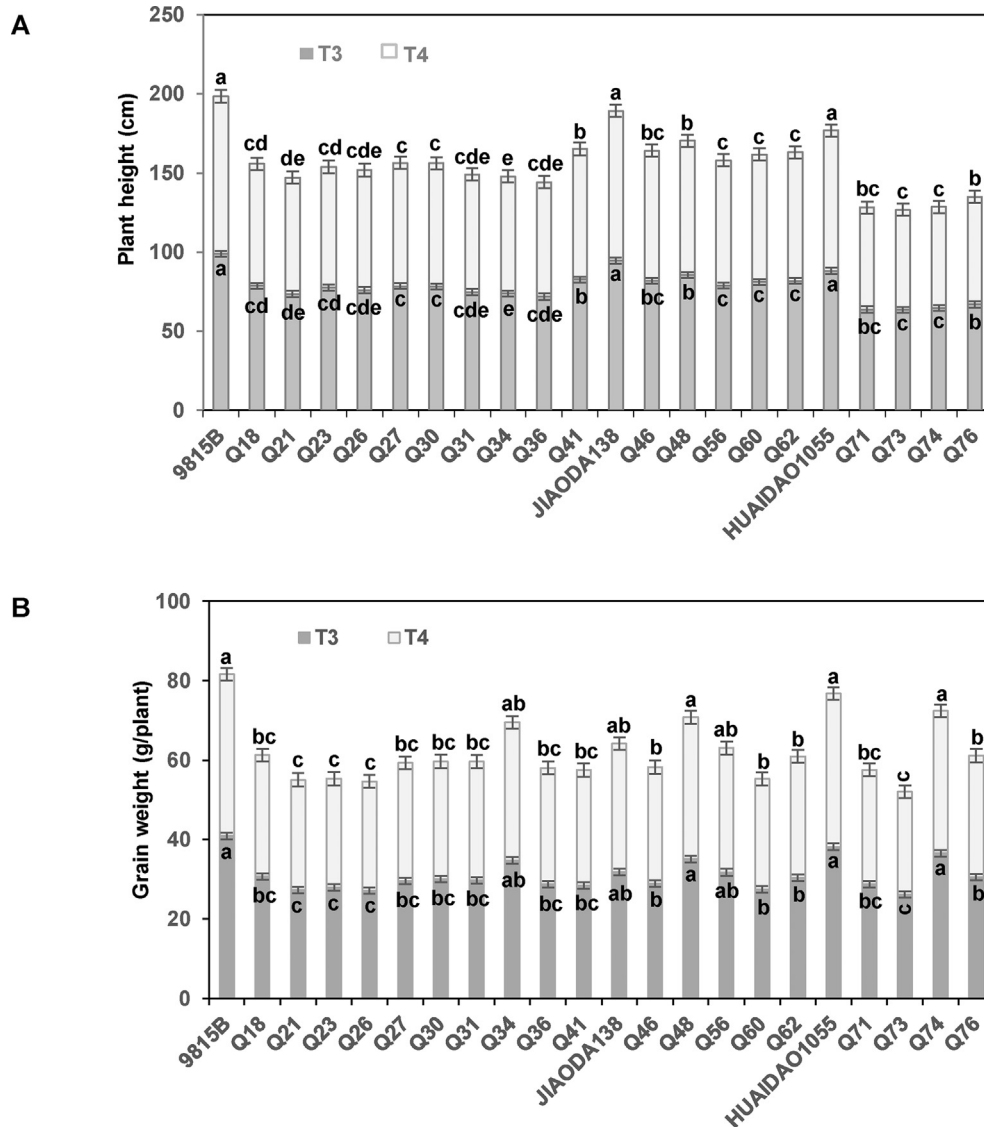
Fig. 2. Effects of mutated *SD1* alleles on plant height and yield in T_3 and T_4 generations. **A**: Height; **B**: Yield. 9815B, JIAODA138, or HUAIDAO1055 represents different genetic backgrounds. Values are mean ± SD (standard deviation of the mean, $n = 10$). Mean comparison was carried out using XLSTAT 2018 software, with different letters representing significant difference at $P < 0.01$.

study presented evidence that there is unexpected high frequency of off-target mutagenesis in CRISPR/Cas9-induced T_1 *Arabidopsis* mutants, which is further exacerbated in the T_2 progenies (Zhang et al., 2018b). In rice, off-target modifications are detectable, either rare (Xu et al., 2015) or frequent (Endo et al., 2015; Li et al., 2016), in positive T_1 plants. Therefore, the detected four off-target mutations in this study and the mutation that occurred within *SD1*, together with diverse and complex genotypes underlying the same semi-dwarf phenotype, pointed out that outputs of the CRISPR/Cas9 system need to be strictly monitored, and the edited sites must be characterized case by case to avoid unexpected modifications. Nevertheless, those four off-target mutations could result from the spontaneous mutations occurred during the tissue culture, which we did not investigate in this study. Whole-genome sequencing in the future will be useful to draw a conclusion.

3.3. Exogenous elements

The presence of exogenous elements (T-DNA elements and Cas9) in CRISPR/Cas9-induced T_1 mutants in rice can be completely segregated out in T_2 mutants (Zhou et al., 2014; Xu et al., 2015). Our results did not show the same trend, possibly due to the fact that we did not perform molecular characterization from the T_0 generation, the fact that there might be multiple copies of those exogenous elements in mutant genomes, or the fact that the PCR method used is not good enough to draw a conclusion. Nevertheless, our results highlighted the importance of the early and accurate molecular characterization and screening of these exogenous elements in CRISPR/Cas9-induced mutants. Because the presence of the transgene was found to be concurrent in mutants positive for Cas9 (Tables 1 and 2), the screening for the absence of Cas9 in T_1 could be a crucial measure to eliminate the exogenous elements in CRISPR/Cas9-induced mutants. It is worthy to note that the screening for Cas9 does not exclude the necessity to screen for other exogenous elements. In general, the current GM regulatory system requires a full risk analysis of organisms with exogenous elements; therefore, understanding of uncertainties and risks regarding genome editing is necessary and critical before a new global policy for the new biotechnology is established.

3.4. From laboratory to field

Because it is efficient, specific, and flexible, the CRISPR/Cas9 system has been widely used as the preferred genome-editing tool in plants for both basic and applied purpose (Araki and Ishii, 2015; Schaeffer and Nakata, 2015; Wolt, 2017). Our data corroborated the efficiency of the CRISPR/Cas9 system in rice mutagenesis for basic research but also pointed out the difficulty for its application in rice breeding. From a technical point of view, a potential applicable mutant must be inheritable, transgene-free, target-edited, and with desired traits, all affected by multiple factors, which have to be thoroughly characterized.

Although we identified a potential line Q48 with significantly reduced plant height and moderately elevated yield in this study, the mechanism underlying the lack of yield trade-off in Q48 remains uninvestigated. Indeed, all CRISPR/Cas9-induced SD1 mutants actually tended to significantly reduce the yield along with the reduction in plant height, and this effect apparently was influenced by genetic backgrounds in this case. Therefore, the application of the CRISPR/Cas9 system for successful rice breeding may be a long way than expected.

In this study, we carried out comprehensive molecular characterization on CRISPR/Cas9-induced mutants in later generations (T_2–T_4) in rice, in a journey to obtain edited alleles for potential enhancement of the production performance of current elite rice cultivars by manipulating the Green Revolution gene SD1. For this purpose, all mutants generated on different elite rice cultivar backgrounds at early generations before T_2 were selected by antibiotic (T_0) or phenotyping (T_0 to T_2) instead of molecular characterization, and the output of this study provided a useful basis for reoptimizing initial screening strategies for transgene-clean targeted genome editing in rice.

4. Materials and methods

4.1. Plasmid vector construction

The target sequence (GAGGATGGAGCCCAAGATCC) in the first exon of SD1 gene was amplified with primers SD1sgRNA-F and SD1sgRNA-R (Table S1) and cloned into pBIN-sgR-Cas9-OsU3 vector as described previously (Mao et al., 2013). The resulting construct was stably transformed to rice via Agrobacterium as reported (Hiei and Komari, 2008).

4.2. Plant materials and growth conditions

The rice (Oryza sativa) elite varieties 9815B, JIAODA138, and HUAIDAO1055 commonly cultured in Shanghai and Jiangsu province, China, were obtained from our seed library. SD1-targeted editing lines from T_0-T_4 generations, including 9815B^{SD1}, JIAODA138^{SD1}, and HUAIDAO1055^{SD1}, and their corresponding wild types were grown in the paddy field of Shanghai Jiao Tong University (30°N, 121°E), Shanghai, China, under natural rice-growing conditions.

4.3. Plant genomic DNA extraction

Genomic DNA from CRISPR/Cas9-generated rice mutants and wild-type tissues was extracted as previously described (Murray and Thompson, 1980). Rice tissues (mainly leaf tissues) were ground using mortar and pestle in the presence of liquid nitrogen, then incubated with 1.5 × lysis buffer cetyl trimethylammonium bromide (CTAB) and RNase for 60 min at 65 °C and centrifuged at 10,625 g (12,000 rpm) for 10 min. The upper phase (liquid) was collected and extracted again with phenol:chloroform and trichloromethane. The genomic DNA was precipitated by adding of isopropyl alcohol to the supernatant followed by centrifugation at 10,625 g (12,000 rpm) for 5 min. DNA pellets were washed twice with 70% ethanol and dissolved in ddH$_2$O. The qualities and quantities of extracted genomic DNA were measured and evaluated using both the NanoDrop 1000 UV/vis Spectrophotometer (Nano-Drop Technologies Inc., Wilmington, DE, USA) by OD260/OD280 and OD260/OD230 and the electrophoresis on 1% (w/v) agarose gel in 0.5 × Tris/Borate/EDTA (TBE) with GelRed staining. All purified genomic DNA was stored at −20 °C until used for analysis.

4.4. Genotype and exogenous elements analysis

The specific primers (Table S1) were used to amplify SD1 target, and each PCR reaction mixture (20 μL) contained 1× PCR buffer, 1× Q-solution (Qiagen, Germany), 0.2 μM dNTPs, 5 μM primer, 1 unit of HotStarTaq DNA Polymerase (Qiagen), and 60 ng of genomic DNA. The PCR program was initiated by heating at 95 °C for 15 min followed by 35 cycles of amplification at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 45 s, and a final step at 72 °C for 10 min. The amplified PCR fragments were sequenced directly or cloned into the pEASY-Blunt vector (Transgen Biotech, Shanghai, China) and then sequenced for genotype identification by Sanger method.

Additional specific primers (Table S1) were used to amplify exogenous and vector backbone elements in a 25-μL PCR reaction mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl$_2$, 0.2 mM dNTPs, 0.2 μM primer, 1.25 units of Taq DNA Polymerase (TaKaRa Biotechnology Co., Ltd., Japan), and 60 ng of genomic DNA. The PCR program was initiated by heating at 94 °C for 5 min, followed by 35 cycles of amplification at 94 °C for 30 s, 58 °C for 30 s, 72 °C for 30 s, and a final step at 72 °C for 7 min. The PCR products were investigated on 2% agarose gel.

4.5. Plant height and yield measurement

Plant heights (cm) of each mutant line and the corresponding wild type were measured from the highest panicle to the ground surface at mature stage (10 individual plants each genotype). Grain weights (g) of each mutant line and the corresponding wild type were weighed from all filled grains of each plant (10 individual plants each genotype). All numerical data presented here were expressed as the means ± standard deviation (SD) of the mean. Statistical analysis was carried out to compare the plant height (cm) and yields (g) with all individuals using Excel (2016). Mean comparison was carried out using XLSTAT 2018 software (a complete statistical add-in for Microsoft Excel), and the significant difference of two means was determined at $P < 0.01$.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jgg.2020.04.004.

References

Araki, M., Ishii, T., 2015. Towards social acceptance of plant breeding by genome editing. Trends Plant Sci. 20, 145−149.

Convention on Biological Diversity, 2000. Cartagena Protocol on Biosafety to the Convention on Biological Diversity: Text and Annexes. Secretariat of the Convention on Biological Diversity, Montreal, Canada.

Court of Justice of the European Union, 2018. PRESS RELEASE No 111/18: Organisms Obtained by Mutagenesis Are GMOs and Are, in Principle, Subject to the Obligations Laid Down by the GMO Directive. Judgm. Case C-528/16.

Endo, M., Mikami, M., Toki, S., 2015. Multigene knockout utilizing off-target mutations of the CRISPR/cas9 system in rice. Plant Cell Physiol. 56, 41−47.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L., Zeng, L., Liu, X., Zhu, J.K., 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.

Feng, C., Yuan, J., Wang, R., Liu, Y., Birchler, J.A., Han, F., 2016. Efficient targeted genome modification in maize using CRISPR/Cas9 system. J. Genet. Genomics 43, 37−43.

Gao, J., Wang, G., Ma, S., Xie, X., Wu, X., Zhang, X., Wu, Y., Zhao, P., Xia, Q., 2015. CRISPR/Cas9-mediated targeted mutagenesis in Nicotiana tabacum. Plant Mol. Biol. 87, 99−110.

Gao, W., Xu, W.T., Huang, K.L., Guo, M.Z., Luo, Y.B., 2018. Risk analysis for genome editing-derived food safety in China. Food Contr. 84, 128−137.

Hiei, Y., Komari, T., 2008. Agrobacterium-mediated transformation of rice using immature embryos or calli induced from mature seed. Nat. Protoc. 3, 824.

Huang, K., 2017. Safety Assessment of Genetically Modified Foods. Springer Singapore.

Ishizaki, T., 2016. CRISPR/Cas9 in rice can induce new mutations in later generations, leading to chimerism and unpredicted segregation of the targeted mutation. Mol. Breed. 36, 165.

ISAAA, 2019. Global Status of Commercialized Biotech/GM Crops in 2018: Biotech Crops Continue to Help Meet the Challenges of Increased Population and Climate Change. ISAAA Brief No. 54. ISAAA, Ithaca, NY.

Jiang, W.Z., Yang, B., Weeks, D.P., 2014. Efficient CRISPR/Cas9-mediated gene editing in Arabidopsis thaliana and inheritance of modified genes in the T2 and T3 generations. PLoS One 9, e0099225.

Jones, H.D., 2015. Regulatory uncertainty over genome editing. Native Plants 1, 1−3.

Kosicki, M., Tomberg, K., Bradley, A., 2018. Repair of double-strand breaks induced by CRISPR−Cas9 leads to large deletions and complex rearrangements. Nat. Biotechnol. 36, 765−771.

Kuroha, T., Nagai, K., Gamuyao, R., Wang, D.R., Furuta, T., Nakamori, M., Kitaoka, T., Adachi, K., Minami, A., Mori, Y., Mashiguchi, K., Seto, Y., Yamaguchi, S., Kojima, M., Sakakibara, H., Wu, J., Ebana, K., Mitsuda, N., Ohme-Takagi, M., Yanagisawa, S., Yamasaki, M., Yokoyama, R., Nishitani, K., Mochizuki, T., Tamiya, G., McCouch, S.R., Ashikari, M., 2018. Ethylene-gibberellin signaling underlies adaptation of rice to periodic flooding. Science 361, 181−186.

Li, M., Li, X., Zhou, Z., Wu, P., Fang, M., Pan, X., Lin, Q., Luo, W., Wu, G., Li, H., 2016. Reassessment of the four yield-related genes Gn1a, DEP1, GS3, and IPA1 in rice using a CRISPR/Cas9 system. Front. Plant Sci. 7, 377.

Li, S., Tian, Y., Wu, K., Ye, Y., Yu, J., Zhang, J., Liu, Q., Hu, M., Li, H., Tong, Y., Harberd, N.P., Fu, X., 2018. Modulating plant growth−metabolism coordination for sustainable agriculture. Nature 560, 595−600.

Li, X., Zhou, W., Ren, Y., Tian, X., Lv, T., Wang, Z., Fang, J., Chu, C., Yang, J., Bu, Q., 2017. High-efficiency breeding of early-maturing rice cultivars via CRISPR/Cas9-mediated genome editing. J. Genet. Genomics 44, 175−178.

Liang, Z., Zhang, K., Chen, K., Gao, C., 2014. Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. J. Genet. Genomics 41, 63−68.

Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K., Chen, L.L., 2017. CRISPR-P 2.0: an improved CRISPR/Cas9 tool for genome editing in plants. Mol. Plant 10, 530−532.

Lusser, M., Parisi, C., Plan, D., Rodríguez-Cerezo, E., 2012. Deployment of new biotechnologies in plant breeding. Nat. Biotechnol. 30, 231−239.

Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., Zhu, J.K., 2013. Application of the CRISPR-Cas system for efficient genome engineering in plants. Mol. Plant 6, 2008−2011.

Mattei, T.A., 2018. The CRISPR-cas9 genome editing system: not as precise as previously believed. World Neurosurg. 118, 377−378.

Miao, J., Guo, D., Zhang, J., Huang, Q., Qin, G., Zhang, X., Wan, J., Gu, H., Qu, L.J., 2013. Targeted mutagenesis in rice using CRISPR-Cas system. Cell Res. 23, 1233−1236.

Murray, M.G., Thompson, W.F., 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8, 4321−4326.

Peterson, B.A., Haak, D.C., Nishimura, M.T., Teixeira, P.J.P.L., James, S.R., Dangl, J.L., Nimchuk, Z.L., 2016. Genome-wide assessment of efficiency and specificity in crispr/cas9 mediated multiple site targeting in Arabidopsis. PLoS One 11, e0162169.

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G.S., Kitano, H., Matsuoka, M., 2002. A mutant gibberellin-synthesis gene in rice. Nature 416, 701−702.

Schaeffer, S.M., Nakata, P.A., 2015. CRISPR/Cas9-mediated genome editing and gene replacement in plants: transitioning from lab to field. Plant Sci. 240, 130−142.

Shan, Q., Wang, Y., Li, J., Gao, C., 2014. Genome editing in rice and wheat using the CRISPR/Cas system. Nat. Protoc. 9, 2395.

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X., Zhong, Z., Liu, B., Zheng, X., Zhang, D., Malzahn, A., Gong, Z., Qi, Y., Zhang, T., Zhang, Y., 2018. A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. Genome Biol. 19, 84.

Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., Qiu, J.L., 2014. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat. Biotechnol. 32, 947−951.

Wolt, J.D., 2017. Safety, security, and policy considerations for plant genome editing. In: Progress in Molecular Biology and Translational Science, vol. 149. Academic Press, pp. 215−241.

Wolt, J.D., Wang, K., Sashital, D., Lawrence-Dill, C.J., 2016. Achieving plant CRISPR targeting that limits off-target effects. Plant Genome 9, 1−8.

Wolter, F., Puchta, H., 2017. Knocking out consumer concerns and regulator's rules: efficient use of CRISPR/Cas ribonucleoprotein complexes for genome editing in cereals. Genome Biol. 18, 43.

Wu, J., Yin, H., 2019. Engineering guide RNA to reduce the off-target effects of CRISPR. J. Genet. Genomics 46, 523−529.

Xie, K., Minkenberg, B., Yang, Y., 2015. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc. Natl. Acad. Sci. U. S. A. 112, 3570−3575.

Xu, R.F., Li, H., Qin, R.Y., Li, J., Qiu, C.H., Yang, Y.C., Ma, H., Li, L., Wei, P.C., Yang, J.B., 2015. Generation of inheritable and "transgene clean" targeted genome-modified rice in later generations using the CRISPR/Cas9 system. Sci. Rep. 5, 11491.

Zhang, Y., Massel, K., Godwin, I.D., Gao, C., 2018a. Applications and potential of genome editing in crop improvement. Genome Biol. 19, 210.

Zhang, Q., Xing, H.L., Wang, Z.P., Zhang, H.Y., Yang, F., Wang, X.C., Chen, Q.J., 2018b. Potential high-frequency off-target mutagenesis induced by CRISPR/Cas9 in Arabidopsis and its prevention. Plant Mol. Biol. 96, 445−456.

Zhang, Hui, Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, Heng, Xu, N., Zhu, J.K., 2014. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. Plant Biotechnol. J. 12, 797−807.

Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H., Yang, B., 2014. Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. Nucleic Acids Res. 42, 10903−10914.

Zhu, C., Bortesi, L., Baysal, C., Twyman, R.M., Fischer, R., Capell, T., Schillberg, S., Christou, P., 2017. Characteristics of genome editing mutations in cereal crops. Trends Plant Sci. 22, 38−52.

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., Qiu, J.L., Wang, D., Gao, C., 2017. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. Nat. Biotechnol. 35, 438−440.