

Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing

Xu Tang^{1#}, Qiurong Ren^{1#}, Lijia Yang¹, Yu Bao^{2, 3}, Zhaohui Zhong¹, Binglin Liu¹, Yan Wang¹, Simon Sretenovic⁴, Yingxiao Zhang⁴, Xuelian Zheng¹, Tao Zhang^{2, 3*}, Yiping Qi^{4, 5*}, Yong Zhang^{1*}

¹Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 610054, China; ²Jiangsu Key Laboratory of Crop Genetics and Physiology, Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Agricultural College of Yangzhou University, Yangzhou 225009, China; ³Key Laboratory of Plant Functional Genomics of the Ministry of Education, Joint International Research Laboratory of Agriculture and Agri-Product Safety of the Ministry of Education, Yangzhou University, Yangzhou 225009, China; ⁴Department of Plant Science and Landscape Architecture, University of Maryland, College Park, Maryland 20742, USA; ⁵Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland 20850;

Running title: Improved single transcript unit (STU) CRISPR systems

[#]These authors contributed equally to this work

***Corresponding authors:**

Yong Zhang, Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Room 216, Main Building, No. 4, Section 2, North Jianshe Road, Chengdu, 610054, P.R. China; Email: zhangyong916@uestc.edu.cn

Yiping Qi, Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742, USA; Email: yiping@umd.edu

Tao Zhang, Jiangsu Key Laboratory of Crop Genetics and Physiology, Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding, Agricultural College of Yangzhou University, Yangzhou 225009, China; Email: zhangtao@yzu.edu.cn

Abstract

CRISPR-Cas9 and Cas12a are two powerful genome editing systems. Expression of CRISPR in plants is typically achieved with a mixed dual promoter system, in which Cas protein is expressed by a Pol II promoter and a guide RNA is expressed by a species-specific Pol III promoter such as U6 or U3. To achieve coordinated expression and compact vector packaging, it is desirable to express both CRISPR components under a single Pol II promoter. Previously, we demonstrated a first-generation single transcript unit (STU)-Cas9 system, STU-Cas9-RZ, which is based on hammerhead ribozyme for processing single guide RNAs (sgRNAs). In this study, we developed two new STU-Cas9 systems and one STU-Cas12a system for applications in plants, collectively called the STU CRISPR 2.0 systems. We demonstrated these systems for genome editing in rice with both transient expression and stable transgenesis. The two STU-Cas9 2.0 systems process the sgRNAs with Csy4 ribonuclease and endogenous tRNA processing system, respectively. Both STU-Cas9-Csy4 and STU-Cas9-tRNA systems showed more robust genome editing efficiencies than our first-generation STU-Cas9-RZ system and the conventional mixed dual promoter system. We further applied the STU-Cas9-tRNA system to compare two C to T base editing systems based on rAPOBEC1 and PmCDA1 cytidine deaminases. The results suggest STU-based PmCDA1 base editor system is highly efficient in rice. The STU-Cas12a system, based on Cas12a' self-processing of a CRISPR RNA (crRNA) array, was also developed and demonstrated for expression of a single crRNA and four crRNAs. Altogether, our STU CRISPR 2.0 systems further expanded the CRISPR toolbox for plant genome editing and other applications.

Key words: rice, single transcript unit (STU), CRISPR-Cas9, CRISPR-Cas12a, base editing

Introduction

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas (CRISPR associated) is a leading genome editing toolbox. Cas9 is an RNA guided sequence-specific nuclease (SSN) that mediates DNA targeting. The most popular *Streptococcus pyogenes* Cas9 (SpCas9) recognizes a target site containing an NGG protospacer adjacent motif (PAM) (Jinek et al., 2012). Since its demonstration in plants in late 2013 (Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013), CRISPR-Cas9 has been widely applied in many plant species (Malzahn et al., 2017; Yin et al., 2017). Most studies have relied on non-homologous end joining (NHEJ) DNA repair pathway to introduce insertion or deletion (InDel) mutations, achieving editing outcomes such as gene knockout (Malzahn et al., 2017), mutagenesis of microRNAs

(Zhou, 2017) or *cis*-regulatory sequences (Rodriguez-Leal et al., 2017; Zhang et al., 2018), as well as large chromosomal deletions (Zhou et al., 2014) and gene replacement (Li et al., 2016). In some cases, homology directed repair (HDR) was used for targeted gene replacement (Endo et al., 2016b; Gil-Humanes et al., 2017; Li et al., 2018a; Li et al., 2015; Miki et al., 2018; Schiml et al., 2014; Sun et al., 2016a; Svitashv et al., 2015). Cas9 nickase (Fauser et al., 2014; Ran et al., 2013) and base editors (Gaudelli et al., 2017; Komor et al., 2016; Li et al., 2017; Shimatani et al., 2017) have further expanded the applications of Cas9-based plant genome editing (Hua et al., 2018; Lu and Zhu, 2017; Ren et al., 2017; Ren et al., 2018; Yan et al., 2018; Zong et al., 2017).

CRISPR-Cas12a (formerly Cpf1), a class 2 type V-A CRISPR-Cas system, has also been applied for plant genome editing (Begemann et al., 2017; Endo et al., 2016a; Hu et al., 2017; Kim et al., 2017; Li et al., 2018b; Li et al., 2018c; Tang et al., 2017; Xu et al., 2016; Zhong et al., 2018). Unlike Cas9, Cas12a only requires CRISPR RNA (crRNA) without the need of trans-activating crRNA (tracrRNA) and it recognizes T-rich PAMs, resulting in staggered DNA double strand breaks (DSBs) (Fagerlund et al., 2015; Zetsche et al., 2015). In addition, Cas12a has ribonuclease activity that helps process the crRNA to maturity (Fonfara et al., 2016; Zetsche et al., 2017), which has been conveniently utilized for multiplexed plant genome editing (Wang et al., 2017b; Wang et al., 2018a). Despite many characteristic differences, Cas9 and Cas12a were both found to be very specific in mediating plant genome editing either by whole-genome sequencing (Tang et al., 2018) or by CIRCLE-seq (Lee et al., 2018; Tsai et al., 2017). In general, Cas9 and Cas12a have been demonstrated as highly efficient and specific SSNs in plants.

Most CRISPR-Cas9 studies used a mixed dual promoter system in which Cas9 is expressed by a Pol II promoter and the single guide RNA (sgRNA) is expressed by a Pol III promoter such as U6 or U3. While it is relatively easy to deploy CRISPR-Cas9 for multiplexed genome editing, stacking multiple sgRNA expression units quickly adds up to the length of an expression vector (Lowder et al., 2015; Ma et al., 2015; Xing et al., 2014; Zhang et al., 2016). It is challenging to package all components into a virus-based delivery vector (Ali et al., 2015; Baltes et al., 2014; Cody et al., 2017). In addition, repetitive use of multiple U6 or U3 promoters within one construct may cause variations on sgRNA expression levels and transgene silencing in plants (Ma et al., 2015). A lot of efforts have been put into development of a compact sgRNA expression system where multiple sgRNAs can be expressed from a single Pol III or Pol II promoter. For example, multiple sgRNAs can be expressed from a single Pol III promoter or a single Pol II promoter when spaced with tRNAs (Cermak et al., 2017; Xie et al., 2015). Alternatively, sgRNAs can be processed by hammer head (HH) and hepatitis delta virus (HDV) ribozymes (He et al., 2017) and Csy4 RNA ribonuclease (Cermak et al., 2017; Tsai et al., 2014). Among these CRISPR-Cas9 expression systems, Cas9 and sgRNAs are generally expressed in two separate expression units. Since Pol III promoters in many

organisms are not well characterized and such promoters are typically more suitable to express short transcripts, it is advantageous to use Pol II promoters to express multiple sgRNAs or crRNAs for multiplexed genome editing. Furthermore, higher genome editing efficiencies with Cas9 and Cas12a have been observed with selected constitutive Pol II promoters when compared to Pol III promoters (e.g. U6 or U3) in plants (Cermak et al., 2017; Mikami et al., 2017; Tang et al., 2016) and mammalian cells (Zhong et al., 2017). The use of Pol II promoters will also render guide RNAs under spatiotemporal control, enabling more sophisticated applications such as transcriptional regulation (Lowder et al., 2015; Lowder et al., 2018; Tang et al., 2017).

To achieve most simplified, compact and coordinated expression configuration, it is desirable to express the Cas gene and the guide RNAs from a single Pol II promoter. We previously demonstrated one such single transcript unit (STU) system, STU-Cas9-RZ, in which Cas9 and sgRNAs were linked by a poly A sequence and sgRNAs were processed by the HH ribozyme (Tang et al., 2016). While the HH ribozyme system seems self-sufficient for sgRNA processing, it has potential drawbacks of low *in vivo* processing activity (Mikami et al., 2017) and may not be suitable for expression in a virus-derived vector (Cody et al., 2017). In this study, we sought to develop improved STU systems that are highly efficient for plant genome editing while overcoming the potential drawbacks of the STU-Cas9-RZ system. Using rice as a test platform, two STU-Cas9 systems based Csy4 and tRNA were developed and closely compared with the STU-Cas9-RZ system for targeting one, two or three sites using protoplast transformation and stable transformation systems. The promising STU-Cas9-tRNA system was further demonstrated for high capacity multiplexed genome editing as well as targeted C to T base editing. Finally, we developed a STU-Cas12a system and demonstrated its effectiveness for genome editing in rice. We called these new systems collectively as the STU CRISPR 2.0 systems.

Results

Comparison of three STU-Cas9 systems in rice cells

Our previous STU-Cas9-RZ system utilized HH ribozyme for sgRNA processing (Tang et al., 2016). To develop second-generation STU-Cas9 2.0 systems, we decided to use the endoribonuclease Csy4 and tRNA for sgRNA processing for two reasons. First, these two systems rely on different mechanisms that are distinct from ribozyme: Csy4 is originated from a bacterial CRISPR transcript (pre-crRNA) processing system (Haurwitz et al., 2010) and the tRNA system relies on the plant endogenous tRNA-processing system (Xie et al., 2015). Second, efficient genome editing has been demonstrated in diverse eukaryotic organisms with Csy4-based sgRNA processing (Cermak et al., 2017; Ferreira et al., 2018; Qin et al., 2015;

Tsai et al., 2014) and tRNA-based sgRNA processing (Port and Bullock, 2016; Shiraki and Kawakami, 2018; Wu et al., 2018; Xie et al., 2015). Hence, there is a good chance that a STU-Cas9 2.0 system (STU-Csy4 or STU-tRNA) may outperform the STU-Cas9-RZ system. The three STU-Cas9 systems were expressed from the same Pol II promoter, maize ubiquitin promoter (pZmUbi), and compared to the conventional mixed dual promoter system (Figure 1a). For STU-Cas9-Csy4, a T2A ribosomal skipping peptide (Szymczak et al., 2004) was used to translate Cas4 and Cas9 from a single transcript (Figure 1a). With these four Cas9 systems, we targeted six sites in the rice genome. The resulting 24 constructs were used for transient transformation of rice protoplasts. NHEJ mutations were detected in all these samples by cleaved amplified polymorphic sequence (CAPS) analysis (Figure S1). Mutation frequencies, as sums of insertion and deletions at the target sites, were measured by deep sequencing of polymerase chain reaction (PCR) amplicons. Nucleotide substitutions were rare, and they were excluded for calculation of mutation frequencies since we could not distinguish them from sequencing errors. The three STU systems showed similar editing efficiencies at four target sites: *OsPDS*-sgRNA01, *OsPDS*-sgRNA02, *OsYSA*-sgRNA02 and *OsDEPI*-sgRNA01 (Figure 1b). However, at *OsYSA*-sgRNA01 and *OsEDPI*-sgRNA02 sites, higher editing efficiencies were observed for STU-Cas9-Csy4 and STU-Cas9-tRNA, when compared to STU-Cas9-RZ (Figure 1b). Across all six sites, STU-Cas9-Csy4 and STU-Cas9-tRNA had the same or higher editing efficiencies when compared to the conventional mixed dual promoter system (Figure 1b). These results suggest that STU-Cas9-Csy4 and STU-Cas9-tRNA are more robust in genome editing than the first-generation STU-Cas9-RZ system.

We analyzed the NHEJ mutations among all samples to further investigate the editing outcomes of different Cas9 expression systems. The results revealed that the deletion profiles varied greatly across the target sites (Figure 2). However, four different expression systems resulted in very similar deletion profiles, suggesting the NHEJ repair outcomes are largely dictated by the sequence composition of the target sites and but not influenced by the expression systems (Figure 2). Cas9 generates DNA DSBs with mostly blunt ends at 3 bp upstream from the NGG PAM. Interestingly, the most frequent deletion positions were at 4 or 5bp upstream of the PAM site (Figure 2). Also, deletion profiles at all six sites showed a rather asymmetric distribution where the deletion frequencies for positions between the DNA DSB and the PAM drastically dropped, and there were only few occasions that the PAM sites got deleted (Figure 2). The distribution of deletion size at these target sites further demonstrated that the NHEJ outcomes varied across target sites (Figure 3). While 1 bp deletions were the most predominant deletion type, deletions of more than 3 bp were also very common (Figure 3). These data suggest NHEJ repair outcomes are heavily dependent on target sequence composition, implying microhomology based alternative NHEJ (altNHEJ) is frequently used for repair of Cas9 induced DSBs.

Comparison of three STU-Cas9 systems in rice transgenic lines

We further compared the three STU-Cas9 systems with the mixed dual promoter system in stable transgenic rice plants. We focused on the *OsPDS*-sgRNA01 for testing editing efficiency of the four expression systems (Figure 4a). For each construct, fifty or more T0 transgenic plants were obtained for genotyping with Sanger sequencing. With the mixed dual promoter system, 35 out of 50 (70%) T0 plants were mutated and 26 plants (52%) carried biallelic mutations. For STU-Cas9-RZ system, 38 out of 51 (74.5%) were mutated and 31 plants (60.8%) carried biallelic mutations. For STU-Cas9-Csy4 system, 45 out of 51 (88.2%) were mutated and 36 plants (70.6%) carried biallelic mutations. For STU-Cas9-tRNA system, 51 out of 63 (81%) were mutated and 45 plants (71.4%) carried biallelic mutations. Because knockout of *OsPDS* results in albino phenotype, we indeed observed the albino phenotype among all biallelic mutants (Figure 4b). These results indicated the STU-Cas9-Csy4 and STU-Cas9-tRNA systems are more robust than the STU-Cas9-RZ and the mixed dual promoter system for targeted mutagenesis in rice. Hence, the results from stable transgenic lines are consistent with those from rice protoplasts.

Comparison of three STU-Cas9 systems for multiplexed editing of two and three sites

We next compared the three STU-Cas9 systems for multiplexed editing using the rice stable transformation system. First, we simultaneously targeted two sites: *OsPDS*-sgRNA01 and *OsDEP1*-sgRNA01 (Figure 5a). For each construct, over 30 independent T0 lines were genotyped by Sanger sequencing (Figure 5b). At the *OsPDS*-sgRNA01 site, mutation frequencies of T0 lines for STU-Cas9-RZ, STU-Cas9-Csy4 and STU-Cas9-tRNA were 79.5%, 90.6% and 86.7%, respectively. At the *OsDEP1*-sgRNA01 site, mutation frequencies of T0 lines for STU-Cas9-RZ, STU-Cas9-Csy4 and STU-Cas9-tRNA were 87.2%, 90.6% and 93.3%, respectively. Importantly, majority of the mutants carried biallelic mutations at both target sites (Figure 5b). Second, we simultaneously targeted three sites: *OsPDS*-sgRNA01, *OsPDS*-sgRNA02 and *OsDEP1*-sgRNA01 (Figure 6a). Again, more than 30 individual T0 plants were genotyped for each construct (Figure 6b). At the *OsPDS*-sgRNA01 site, mutation frequencies of T0 lines for STU-Cas9-RZ, STU-Cas9-Csy4 and STU-Cas9-tRNA were 83.3%, 93.9% and 89.2%, respectively. At the *OsPDS*-sgRNA02 site, mutation frequencies of T0 lines for STU-Cas9-RZ, STU-Cas9-Csy4 and STU-Cas9-tRNA were 58.3%, 60.6% and 67.6%, respectively. At *OsDEP1*-sgRNA01 site, mutation frequencies of T0 lines for STU-Cas9-RZ, STU-Cas9-Csy4 and STU-Cas9-tRNA were 86.1%, 84.8% and 89.2%, respectively. While *OsPDS*-sgRNA02 showed relatively low editing activity, majority of the mutants still carried biallelic mutations (Figure 6b). Impressively, the STU-Cas9-tRNA system resulted highest triple biallelic knockout frequency: 23 out of 25 triple mutants carried biallelic mutations at all three target sites (Figure 6b).

STU-Cas9-tRNA system for multiplexed editing of six target sites

Having identified that STU-Cas9-Csy4 and STU-Cas9-tRNA are robust in genome editing, we were inclined to further explore STU-Cas9-tRNA, which is simpler than STU-Cas9-Csy4, for high capacity multiplexing. To this end, we multiplexed all the six chosen sgRNAs into a single STU-Cas9-tRNA T-DNA construct (Figure 7a) and obtained 38 T0 transgenic lines for genotyping (Figure 7b). The mutation frequencies at *OsPDS*-sgRNA01, *OsPDS*-sgRNA02, *OsDEP1*-sgRNA01, *OsDEP1*-sgRNA02, *OsYSA*-sgRNA01 and *OsYSA*-sgRNA02 sites were 89.5%, 73.7%, 97.4%, 60.5%, 71.1% and 68.4%, respectively. The editing efficiencies at *OsPDS*-sgRNA01, *OsPDS*-sgRNA02 and *OsDEP1*-sgRNA01 sites were comparable to the results from the STU-Cas9-tRNA construct when only these three sgRNAs were multiplexed (Figure 6b), suggesting that high capacity multiplexing can be achieved without compromising editing efficiencies. Among 38 T0 lines, 18 lines (47.4%) contained mutations at all six target sites, and four lines carried biallelic mutations at each target site, indicating simultaneous mutagenesis at six sites are readily achievable with the STU-Cas9-tRNA system.

Application of the STU-Cas9 system for base editing

In recent years, both C to T and A to G base editing systems have been developed for inducing targeted base changes without DNA DSBs (Gaudelli et al., 2017; Komor et al., 2016). To test whether a STU-Cas9 system can be used for base editing, we focused on comparison of two C to T base editing systems which are based on two different cytidine deaminases: rAPOBEC1 (Komor et al., 2016) and PmCDA1 (Shimatani et al., 2017). We constructed the STU-Cas9-tRNA to express these two base editors and compared them by targeting four independent sites in the rice genome (Figure 8a). The resulting T-DNA constructs were used for transient transformation of rice protoplasts and the based editing efficiency, as sum of all targetable cytidines (Cs) within the targeting window for each target site, was calculated by deep sequencing. The results showed that PmCDA1 base editor significantly outperformed rAPOBEC1 base editor at three of the four target sites (Figure 8b). While the editing efficiencies by rAPOBEC1 were below 10%, the editing efficiencies by PmCDA1 were 30% and above for *OsCDC48*-sgRNA01, *OsROC5*-sgRNA04 and *OsROC5*-sgRNA05 (Figure 8b). A detailed analysis on individual targetable Cs within the base editing windows across these sites consistently pointed to much higher C to T conversion rates with the PmCDA1 base editor system (Figure 8c and Figure S2).

To obtain base-edited plants, we followed two T-DNA constructs with *OsCDC48*-sgRNA01 and *OsROC5*-sgRNA05 for rice stable transformation. Eighteen T0 lines for each construct were screened for editing at the target site. For *OsCDC48*-sgRNA01, seven out of 18 lines (38.9%) carried C to T base changes, while three lines contained deletions at the target site (Figure 8c and Figure S3). For *OsROC5*-sgRNA05,

11 out of 18 lines (68.8%) carried C to T base changes, while four lines contained deletions at the target site (Figure 8c and Figure S4). These results suggest the STU-Cas9-PmCDA base editing system is very efficient in generating plants with targeted C to T base changes.

A STU-Cas12a system for multiplexed genome editing

We also sought to demonstrate that CRISPR-Cas12a can be expressed as a single transcript unit. Self-processing of a crRNA array has been successfully used for multiplexed editing by Cas12a in human cells (Zetsche et al., 2017) and in plants (Wang et al., 2017b). It has been shown that the addition of an extra direct repeat (DR) at the end of the crRNA array helped process the last crRNA and hence boosted its activity (Zhong et al., 2017). Among all three Cas12a nucleases (AsCas12a, LbCas12a and FnCas12a) tested in plants, LbCas12a seems to have robust activity (Hu et al., 2017; Tang et al., 2017; Wang et al., 2017b; Xu et al., 2016; Zhong et al., 2018). We hence built an STU-Cas12a system in which LbCas12a and the crRNA array were driven by a single Pol II promoter (pZmUbi) and separated by poly A (Figure 9a). With this system, we targeted four independent sites in the rice genome. An average of ~20% mutation frequencies were achieved across these sites in protoplasts (Figure S5 and Figure 9b). We next used these four constructs to generate stable transgenic T0 lines and genotyped them by CAPS analysis (Figure S6) and Sanger sequencing. At *OsDEPI*-crRNA01, *OsDEPI*-crRNA02, *OsROC5*-crRNA01 and *OsROC5*-crRNA02, mutation frequencies of T0 lines were 61.1%, 82.4%, 53.8% and 54.5%, respectively; biallelic mutation frequencies were 38.9%, 64.7%, 15.4% and 22.7%, respectively (Figure 9c). We next generated a multiplexed STU-Cas12a construct targeting four crRNAs simultaneously (Figure 9d). Using a protoplast assay (Figure S7), we found an average of ~20% mutation frequencies at all four target sites (Figure 9e). This construct was further used to generate T0 transgenic lines followed by CAPS analysis (Figure S8) and Sanger sequencing. Analysis of 24 T0 lines revealed mutation frequencies of 29.2%, 50%, 33.3% and 50% at *OsDEPI*-crRNA01, *OsDEPI*-crRNA02, *OsROC5*-crRNA01 and *OsROC5*-crRNA02, respectively (Figure 9f). Biallelic mutations were identified for each crRNA at variant efficiencies and one T0 line contained biallelic mutations at all four target sites (Figure 9f).

Discussion

In a previous review, we proposed three promising STU-Cas9 systems to be tested and deployed for plant genome editing, which are based on ribozyme, Csy4 and tRNA, respectively (Lowder et al., 2016). We demonstrated STU-Cas9-RZ system in plants earlier (Tang et al., 2016). In this study, we developed the STU-Cas9-Csy4 and STU-Cas9-tRNA systems and closely compared them with STU-Cas9-RZ and the conventional mixed dual promoter system. Our data suggest the STU-Cas9-RZ system had similar editing

efficiency with the conventional mixed dual promoter system (Figure 1), which is consistent with our previous report (Tang et al., 2016). However, we found the new STU-Cas9-Csy4 and STU-Cas9-tRNA systems showed more robust genome editing as compared to STU-Cas9-RZ (Figure 1, 5 and 6). Hence, these two STU-Cas9 2.0 systems represent improved STU-Cas9 expression systems for plant applications. Our results were consistent with a previous study, which used a dual Pol II promoter system and showed that Cas9-Csy4 and Cas9-tRNA systems had higher editing efficiencies than the Cas9-RZ system (Cermak et al., 2017). We however want to note that in these RZ systems the sgRNA is flanked by two HH ribozyme cleavage sites with only one full HH ribozyme sequence at the 3' end, which thus requires trans-cleavage activity of the HH ribozyme. By contrast, the HH ribozyme was positioned at the 5' end of a guide RNA for self-cleavage in the dual ribozyme system that we used for the development of a highly efficient Cas12a system (Tang et al., 2017; Zhong et al., 2018).

The use of deep sequencing allowed us to thoroughly compare NHEJ repair outcomes of the four Cas9 expression systems across six different target sites. As anticipated, our results suggest Cas9 expression systems themselves did not affect the editing outcomes as the mutation types and frequencies across four expression systems were strikingly similar (Figure 2 and 3). Hence, the four different Cas9 expression systems rather served as additional biological replicates for studying NHEJ repair outcomes. We found the repair outcomes are heavily dependent of target site sequences. For example, the mutation profiles at the six target sites are all different from each other, in terms of deletion positions and sizes (Figure 2 and 3). This observation indicates the frequent involvement of the altNHEJ pathway which depends on sequence microhomology. Previously, we showed microhomology based altNHEJ in repair of zinc finger nuclease (ZFN) induced DNA DSBs in *Arabidopsis* and the use of altNHEJ was drastically boosted when the Ku70/Ku80 based canonical NHEJ pathway was blocked (Qi et al., 2013). Our study here, however, suggests that microhomology-based altNHEJ seems to be more frequent than previously thought in repair of Cas9-generated DNA DSBs in wild-type rice cells. Microhomology-based altNHEJ has been cleverly used for precise genome editing in human cells, given the editing outcomes were partly predictable (Bae et al., 2014). Since HDR frequency is low in plants, it is appealing to use altNHEJ to introduce desired deletions at target sites in plants.

While we were preparing this manuscript, three groups have published additional STU-Cas9 systems in rice (Ding et al., 2018; Mikami et al., 2017; Wang et al., 2018a). Mikami et al. made a striking finding that sgRNAs, when directly fused to the Cas9 coding sequence, can be processed by unknown ribonucleases in rice to result in functional CRISPR-Cas9 ribonuclear protein complex (Mikami et al., 2017). They showed multiplexed genome editing by expressing two sgRNAs. Inspired by this research, the second study by Wang et al. developed a simplified STU-Cas9 system in which the sgRNAs were separated

by 6-bp linkers (Wang et al., 2018a). The third study by Ding et al. cleverly used the intron of the Cas9 cassette to express one or a few sgRNAs (Ding et al., 2018), an idea similar to the one previously demonstrated in *Chlamydomonas* (Jiang, 2017). It is hard for us to make direct comparison between our STU-Cas9 2.0 systems and the STU-Cas9 system developed by Mikami et al., since they only genotyped rice calli, not independent T0 plants (Mikami et al., 2017). However, the first two studies used a similar STU strategy that relies on *in planta* processing of sgRNAs (Mikami et al., 2017; Wang et al., 2018a). We compared the work by Wang et al. with ours as both studies generated T0 transgenic lines for analysis. Wang et al. targeted three sites for multiplexed editing with their simple STU-Cas9 system. While the mutation frequencies at the three sites in T0 lines are relatively comparable to our study, none of their analyzed T0 lines contained biallelic mutations at all three target sites at once. Also, chimeric lines were found in some cases (Wang et al., 2018a). Similarly, Ding et al. achieved simultaneous biallelic mutation rates of 9.7% (3/31) with PTG-6 and 0% (0/8) with PTG-7 when two sgRNAs and four sgRNAs were multiplexed in their intron-based STU vectors even though the tRNA processing system was also used (Ding et al., 2018). By contrast, our STU-Cas9 2.0 systems resulted in biallelic mutations at high frequencies at all targeted sites among the T0 lines. For example, when two sgRNAs were multiplexed with STU-Cas9-Csy4 and STU-Cas9-tRNA, we observed simultaneous biallelic mutation frequencies of 75% (24/32) and 73.3% (22/30), respectively (Figure 5). When three sgRNAs were multiplexed with STU-Cas9-Csy4 and STU-Cas9-tRNA, we found simultaneous biallelic mutation frequencies of 39.4% (13/33) and 62.2% (23/37) (Figure 6). While none of the three published studies have expressed more than four sgRNAs with their STU-Cas9 systems, we multiplexed six target sites with STU-Cas9-tRNA and could readily identified T0 lines (4/38; 10.5%) in which all six target sites contain biallelic mutations (Figure 7). Hence, our new STU-Cas9 2.0 systems showed highest editing efficiency especially in terms of achieving biallelic knockouts. Further, Csy4 and tRNA-based processing systems have been demonstrated with high editing activities in plants and mammalian systems (Cermak et al., 2017; Port and Bullock, 2016; Qin et al., 2015; Shiraki and Kawakami, 2018; Tsai et al., 2014; Xie et al., 2015). It is conceivable that our STU-Cas9-Csy4 and STU-Cas9-tRNA are likely to work well in other organisms. In the future, it will be interesting to compare these STU-Cas9 systems with other STU-Cas9 systems (Ding et al., 2018; Mikami et al., 2017; Wang et al., 2018a) in many other plant species like dicots.

CRISPR-Cas9 can be converted to nickases for promoting targeting specificity (Fauser et al., 2014; Ran et al., 2013), stimulating HDR (Cermak et al., 2015; Miki et al., 2018; Sun et al., 2016b; Wang et al., 2017a), and repurposing for base editing (Gaudelli et al., 2017; Komor et al., 2016; Shimatani et al., 2017). Further, deactivated Cas9 (dCas9) has been used for engineering synthetic transcriptional regulators (Lowder et al., 2015; Lowder et al., 2018; Piatek et al., 2014; Tang et al., 2017) and DNA labeling in plants (Dreissig et al., 2017). It is conceivable that our STU-Cas9 2.0 systems can be applied in all these

applications. As a proof-of-concept, we applied the STU-Cas9-tRNA system to express two C to T base editing systems in rice. Our data suggest the STU-Cas9 PmCDA1 base editing system is highly efficient, resulting in up to 50% base editing frequencies in its target window when analyzed in rice protoplasts (Figure 8). While the rAPOBEC1 base editor is frequently used in plants (Lu and Zhu, 2017; Zong et al., 2017), PmCDA1 base editor resulted in much higher base editing efficiencies in our study. However, we are cautioned to claim that PmCDA1 base editor is inherently more efficient than rAPOBEC1 base editor in plants, because the low activity of rAPOBEC1 base editor could be due to suboptimal codon optimization of the rAPOBEC1 sequence. A recent study in human cells suggest codon optimization can significantly improve base editing efficiency (Koblan et al., 2018). Nevertheless, we have developed a highly efficient STU-Cas9-tRNA based PmCDA1 base editor system for base editing in plants. In the future, it will be interesting to compare PmCDA1 with human APOBEC3A as the latter was recently shown to have high C to T base editing activities in plants (Zong et al., 2018) and can efficiently target methylated regions with minimized bystander and off-target activities in mammalian cells (Gehrke et al., 2018; Wang et al., 2018b). We also observed indels in T0 plants transformed with the STU-Cas9-PmCDA1 base editor, and the relatively high frequency of indel mutations (~20% on average) is consistent with the previous reports of the PmCDA1 base editor (Shimatani et al., 2017) as well as the rAPOBEC1 base editor (Li et al., 2017; Lu and Zhu, 2017). Such deletions were likely resulted from NHEJ repair of chromosomal nicks in plant cells. Future efforts should also be directed to minimize such unintended NHEJ outcomes without compromising base editing efficiency.

In this study, we also developed a STU-Cas12a system and compared it for expressing one and four crRNAs in rice for genome editing. We found editing frequencies were slightly dropped when the same crRNA was moved into a crRNA array for multiplexed editing. The two recent reports also explored STU-Cas12a systems (Ding et al., 2018; Wang et al., 2018a). Ding et al. expressed a crRNA array with an intron to multiplex two crRNAs for targeted gene deletion (Ding et al., 2018). However, the study was done in rice protoplasts and no stable transformation data were available for comparison with ours. Wang et al. reported a STU-Cas12a system in which they flanked the crRNA array with two tRNA sequences (Wang et al., 2018a). They found their STU-Cas12a system resulted in similar editing frequencies when multiplexing nine crRNAs, and the editing frequencies ranged from 4.2% to 70.8% in rice T0 lines (Wang et al., 2018a). Given LbCas12a can self-process a crRNA array, we didn't flank our crRNA array with tRNA sequences. Otherwise, the two STU-Cas9 strategies are similar as both link the LbCas12a and the crRNA array together under a single Pol II promoter. We achieved editing efficiencies from 29.2% to 50% at the four target sites, within a similar range as their work (Wang et al., 2018a). With a dual Pol II promoter and a double ribozyme system, we previously showed very high biallelic mutation rates in rice with LbCas12a (Tang et al., 2017) and FnCas12a (Zhong et al., 2018). Given all the STU-Cas12a systems

reported to date are all based on the self-processing of crRNA arrays by Cas12a, in the future, it may be worthwhile to test other multiplexing strategies to further improve Cas12a mediated genome editing in plants.

Conclusion

In conclusion, we developed two STU-Cas9 2.0 systems that process sgRNAs with Csy4 and tRNA. The STU-Cas9 2.0 systems displayed more robust genome editing activities when compared to our first-generation STU-Cas9-RZ system and the conventional mixed dual promoter system. We achieved high-frequency base editing using the new STU-Cas9-tRNA system with the PmCDA1 base editor. We also developed a STU-Cas12a system for multiplexed plant genome editing. While the study was done in rice, we anticipate the systems that we developed here will have wide applications for expressing CRISPR systems in diverse plant species.

Experimental procedures

Construction of the vectors

The Cas9 vectors of this study were constructed based on the STU Cas9 vector pTX172 in our previous study (Tang et al., 2016). To construct the STU-Cas9-RZ system, the Hsp terminator was amplified from pZHY988B and then cloned into the *SacI* restriction site of pTX172. To construct the STU-Cas9-Csy4 system, the Csy4-P2A fragment was amplified from pMOD_A0501 (Cermak et al., 2017) and the Cas9-polyA fragment and the Csy4 site-ccdB-gRNA-Csy4 site fragment were amplified from pTX172. The above three fragments were cloned into *SbfI*+*HindIII* linearized STU-Cas9-RZ plasmids by Gibson assembly. To generate the STU-Cas9-tRNA system, the Cas9-polyA fragment was amplified from pTX172 and the tRNA-ccdB-gRNA-tRNA fragment was amplified from pTC315. These two fragments were cloned into *SbfI*+*HindIII* linearized STU-Cas9-RZ plasmids by Gibson assembly. For single gene editing, sgRNAs were synthesized as duplexed oligonucleotides (Table S1). Oligos were annealed and cloned into *BsaI* linearized Cas9 vectors. For multiplex gene editing, sgRNA arrays were generated by PCR and cloned into the Cas9 vectors by Golden Gate cloning (Table S1).

The Cas9 base editing constructs were constructed based on the STU-Cas9-tRNA system. To construct rAPOBEC1 base editing system, rAPOBEC1 fragment, nCas9 fragment and UGI fragment were amplified separately, and then cloned into the STU-Cas9-tRNA system by Gibson Assembly. To construct PmCDA1 base editing system, nCas9 fragment and PmCDA1-UGI fragment were amplified and cloned into the STU-Cas9-tRNA system by Gibson Assembly.

To construct the STU Cas12a system, the LbCas12a fragment was amplified from pYPQ230 (Tang et al., 2017) and the polyA-ccdB-DR fragment was synthesized by Genscript Nanjing Ltd. These two fragments were cloned into *SbfI*+*HindIII* linearized STU-Cas9-RZ plasmids by Gibson assembly. For single gene editing, crRNAs were synthesized as duplexed oligonucleotides (Table S1). Oligos were annealed and cloned into *BsaI* linearized STU Cas12a vectors. For multiplex gene editing, crRNA arrays were generated by overlap extension PCR and cloned into the STU Cas12a vectors by Golden Gate cloning (Table S1).

The plasmid maps of all the T-DNA vectors used in this study were provided in Supplemental files.

Rice protoplast transformation and stable transformation

The *Oryza sativa* Geng/Japonica cultivar Nipponbare was used in this study. Rice protoplast transformation was performed as described previously (Tang et al., 2016; Zhang et al., 2013; Zhong et al., 2018). After

transformation, rice protoplasts were incubated at 28 °C for 2 days before DNA extraction. The T-DNA constructs were introduced into *Agrobacterium* EHA105 by the freeze-thaw method. Rice stable transformation was carried out as previously published protocol (Tang et al., 2016; Zheng et al., 2016; Zhou et al., 2018).

Detection of targeted gene mutations

Genomic DNA was extracted from protoplasts or transgenic plants by the CTAB method (Murray and Thompson, 1980). Genomic regions of targeted sites were amplified with specific primers (Table S1). Then the PCR products were digested by corresponding restriction enzymes overnight for CAPS analysis (Zhang et al., 2013). The digested products were analyzed on 1% agarose gels. T0 mutant lines were further genotyped by Sanger sequencing.

High-throughput sequencing analysis

High-throughput sequencing analysis was carried out as published previously (Tang et al., 2017; Zhong et al., 2018). Genome regions of targeted sites were PCR-amplified using barcoded primers (Table S1). Purified DNA samples were quantified by Qubit 2.0 Fluorometer (Life Technologies) and were sequenced using Illumina HiSeq 2,500 platform. Data processing was carried out using CRISPRMatch (You et al., 2018).

Data availability

The raw data of deep sequencing have been deposited to the Sequence Read Archive in National Center for Biotechnology Information (NCBI) under the accession number PRJNA494622.

Acknowledgements

This work was supported by the National Science Foundation of China (31771486), the Sichuan Youth Science and Technology Foundation (2017JQ0005), the Program for International Science and Technology Cooperation and Exchanges of Sichuan Province (2018HH0112) and the Science Strength Promotion Program of UESTC to Yong Z., the National Transgenic Major Project (SQ2018ZD08019-003) to Yong Z. and T.Z., the Jiangsu Specially-Appointed Professor and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) to T.Z., and the National Science Foundation Plant Genome Research Program (IOS-1758745), USDA-NIFA Biotechnology Risk Assessment Research Program (2018-33522-28789), Foundation for Food and Agriculture Research (593603), and a Syngenta Biotechnology Research Grant to Y.Q.

Competing interests

The authors declare no competing interests.

References

- Ali, Z., Abul-faraj, A., Li, L., Ghosh, N., Piatek, M., Mahjoub, A., Aouida, M., Piatek, A., Baltes, N.J., Voytas, D.F., Dinesh-Kumar, S. and Mahfouz, M.M. (2015) Efficient Virus-Mediated Genome Editing in Plants Using the CRISPR/Cas9 System. *Mol Plant* **8**, 1288-1291.
- Bae, S., Kweon, J., Kim, H.S. and Kim, J.S. (2014) Microhomology-based choice of Cas9 nuclease target sites. *Nature methods* **11**, 705-706.
- Baltes, N.J., Gil-Humanes, J., Cermak, T., Atkins, P.A. and Voytas, D.F. (2014) DNA replicons for plant genome engineering. *Plant Cell* **26**, 151-163.
- Begemann, M.B., Gray, B.N., January, E., Gordon, G.C., He, Y., Liu, H., Wu, X., Brutnell, T.P., Mockler, T.C. and Oufattole, M. (2017) Precise insertion and guided editing of higher plant genomes using Cpf1 CRISPR nucleases. *Sci Rep* **7**, 11606.
- Cermak, T., Baltes, N.J., Cegan, R., Zhang, Y. and Voytas, D.F. (2015) High-frequency, precise modification of the tomato genome. *Genome Biol* **16**, 232.
- Cermak, T., Curtin, S.J., Gil-Humanes, J., Cegan, R., Kono, T.J.Y., Konecna, E., Belanto, J.J., Starker, C.G., Mathre, J.W., Greenstein, R.L. and Voytas, D.F. (2017) A Multipurpose Toolkit to Enable Advanced Genome Engineering in Plants. *Plant Cell* **29**, 1196-1217.
- Cody, W.B., Scholthof, H.B. and Mirkov, T.E. (2017) Multiplexed Gene Editing and Protein Overexpression Using a Tobacco mosaic virus Viral Vector. *Plant Physiol* **175**, 23-35.
- Ding, D., Chen, K., Chen, Y., Li, H. and Xie, K. (2018) Engineering Introns to Express RNA Guides for Cas9- and Cpf1-Mediated Multiplex Genome Editing. *Mol Plant* **11**, 542-552.
- Dreissig, S., Schiml, S., Schindele, P., Weiss, O., Rutten, T., Schubert, V., Gladilin, E., Mette, M.F., Puchta, H. and Houben, A. (2017) Live-cell CRISPR imaging in plants reveals dynamic telomere movements. *Plant J* **91**, 565-573.
- Endo, A., Masafumi, M., Kaya, H. and Toki, S. (2016a) Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci Rep* **6**, 38169.
- Endo, M., Mikami, M. and Toki, S. (2016b) Biallelic Gene Targeting in Rice. *Plant Physiol* **170**, 667-677.
- Fagerlund, R.D., Staals, R.H. and Fineran, P.C. (2015) The Cpf1 CRISPR-Cas protein expands genome-editing tools. *Genome Biol* **16**, 251.
- Fausser, F., Schiml, S. and Puchta, H. (2014) Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J* **79**, 348-359.
- Ferreira, R., Skrekas, C., Nielsen, J. and David, F. (2018) Multiplexed CRISPR/Cas9 Genome Editing and Gene Regulation Using Csy4 in *Saccharomyces cerevisiae*. *ACS synthetic biology* **7**, 10-15.
- Fonfara, I., Richter, H., Bratovic, M., Le Rhun, A. and Charpentier, E. (2016) The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature* **532**, 517-521.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I. and Liu, D.R. (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* **551**, 464-471.
- Gehrke, J.M., Cervantes, O., Clement, M.K., Wu, Y., Zeng, J., Bauer, D.E., Pinello, L. and Joung, J.K. (2018) An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat Biotechnol* **36**, 977-982.

- Gil-Humanes, J., Wang, Y., Liang, Z., Shan, Q., Ozuna, C.V., Sanchez-Leon, S., Baltes, N.J., Starker, C., Barro, F., Gao, C. and Voytas, D.F. (2017) High-efficiency gene targeting in hexaploid wheat using DNA replicons and CRISPR/Cas9. *Plant J* **89**, 1251-1262.
- Haurwitz, R.E., Jinek, M., Wiedenheft, B., Zhou, K. and Doudna, J.A. (2010) Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355-1358.
- He, Y., Zhang, T., Yang, N., Xu, M., Yan, L., Wang, L., Wang, R. and Zhao, Y. (2017) Self-cleaving ribozymes enable the production of guide RNAs from unlimited choices of promoters for CRISPR/Cas9 mediated genome editing. *J Genet Genomics* **44**, 469-472.
- Hu, X., Wang, C., Liu, Q., Fu, Y. and Wang, K. (2017) Targeted mutagenesis in rice using CRISPR-Cpf1 system. *J Genet Genomics* **44**, 71-73.
- Hua, K., Tao, X., Yuan, F., Wang, D. and Zhu, J.K. (2018) Precise A.T to G.C Base Editing in the Rice Genome. *Mol Plant* **11**, 627-630.
- Jiang, W.W., D.P. (2017) A gene-within-a-gene Cas9/sgRNA hybrid construct enables gene editing and gene replacement strategies in *Chlamydomonas reinhardtii*. *Algal Research* **26**, 474-480.
- 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.
- Kim, H., Kim, S.T., Ryu, J., Kang, B.C., Kim, J.S. and Kim, S.G. (2017) CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat Commun* **8**, 14406.
- Koblan, L.W., Doman, J.L., Wilson, C., Levy, J.M., Tay, T., Newby, G.A., Maianti, J.P., Raguram, A. and Liu, D.R. (2018) Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol* **36**, 843-846.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. and Liu, D.R. (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424.
- Lee, K., Zhang, Y., Kleinstiver, B.P., Guo, J.A., Aryee, M.J., Miller, J., Malzahn, A., Zarecor, S., Lawrence-Dill, C.J., Joung, J.K., Qi, Y. and Wang, C. (2018) Activities and specificities of CRISPR-Cas9 and Cas12a nucleases for targeted mutagenesis in maize. *Plant Biotechnol J*.
- Li, J., Meng, X., Zong, Y., Chen, K., Zhang, H., Liu, J., Li, J. and Gao, C. (2016) Gene replacements and insertions in rice by intron targeting using CRISPR-Cas9. *Nat Plants* **2**, 16139.
- Li, J., Sun, Y., Du, J., Zhao, Y. and Xia, L. (2017) Generation of Targeted Point Mutations in Rice by a Modified CRISPR/Cas9 System. *Mol Plant* **10**, 526-529.
- Li, J., Zhang, X., Sun, Y., Zhang, J., Du, W., Guo, X., Li, S., Zhao, Y. and Xia, L. (2018a) Efficient allelic replacement in rice by gene editing: A case study of the NRT1.1B gene. *J Integr Plant Biol* **60**, 536-540.
- 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.
- Li, S., Li, J., Zhang, J., Du, W., Fu, J., Sutar, S., Zhao, Y. and Xia, L. (2018b) Synthesis-dependent repair of Cpf1-induced double strand DNA breaks enables targeted gene replacement in rice. *J Exp Bot* **69**, 4715-4721.
- Li, S., Zhang, X., Wang, W., Guo, X., Wu, Z., Du, W., Zhao, Y. and Xia, L. (2018c) Expanding the Scope of CRISPR/Cpf1-Mediated Genome Editing in Rice. *Mol Plant* **11**, 995-998.
- Li, Z., Liu, Z.B., Xing, A., Moon, B.P., Koellhoffer, J.P., Huang, L., Ward, R.T., Clifton, E., Falco, S.C. and Cigan, A.M. (2015) Cas9-Guide RNA Directed Genome Editing in Soybean. *Plant Physiol* **169**, 960-970.
- Lowder, L., Malzahn, A. and Qi, Y. (2016) Rapid Evolution of Manifold CRISPR Systems for Plant Genome Editing. *Front Plant Sci* **7**, 1683.

- Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W., Tang, X., Zheng, X., Voytas, D.F., Hsieh, T.F., Zhang, Y. and Qi, Y. (2015) A CRISPR/Cas9 toolbox for multiplexed plant genome editing and transcriptional regulation. *Plant Physiol* **169**, 971-985.
- Lowder, L.G., Zhou, J., Zhang, Y., Malzahn, A., Zhong, Z., Hsieh, T.F., Voytas, D.F., Zhang, Y. and Qi, Y. (2018) Robust Transcriptional Activation in Plants Using Multiplexed CRISPR-Act2.0 and mTALE-Act Systems. *Mol Plant* **11**, 245-256.
- Lu, Y. and Zhu, J.K. (2017) Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Mol Plant* **10**, 523-525.
- Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., Wang, B., Yang, Z., Li, H., Lin, Y., Xie, Y., Shen, R., Chen, S., Wang, Z., Chen, Y., Guo, J., Chen, L., Zhao, X., Dong, Z. and Liu, Y.G. (2015) A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant* **8**, 1274-1284.
- Malzahn, A., Lowder, L. and Qi, Y. (2017) Plant genome editing with TALEN and CRISPR. *Cell Biosci* **7**, 21.
- Mikami, M., Toki, S. and Endo, M. (2017) In Planta Processing of the SpCas9-gRNA Complex. *Plant Cell Physiol* **58**, 1857-1867.
- Miki, D., Zhang, W., Zeng, W., Feng, Z. and Zhu, J.K. (2018) CRISPR/Cas9-mediated gene targeting in Arabidopsis using sequential transformation. *Nat Commun* **9**, 1967.
- Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res* **8**, 4321-4325.
- 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.
- Piatek, A., Ali, Z., Baazim, H., Li, L., Abulfaraj, A., Al-Shareef, S., Aouida, M. and Mahfouz, M.M. (2014) RNA-guided transcriptional regulation in planta via synthetic dCas9-based transcription factors. *Plant Biotechnol J* **13**, 578-589.
- Port, F. and Bullock, S.L. (2016) Augmenting CRISPR applications in *Drosophila* with tRNA-flanked sgRNAs. *Nature methods* **13**, 852-854.
- Qi, Y., Zhang, Y., Zhang, F., Baller, J.A., Cleland, S.C., Ryu, Y., Starker, C.G. and Voytas, D.F. (2013) Increasing frequencies of site-specific mutagenesis and gene targeting in Arabidopsis by manipulating DNA repair pathways. *Genome Res* **23**, 547-554.
- Qin, W., Liang, F., Feng, Y., Bai, H., Yan, R., Li, S. and Lin, S. (2015) Expansion of CRISPR/Cas9 genome targeting sites in zebrafish by Csy4-based RNA processing. *Cell Res* **25**, 1074-1077.
- Ran, F.A., Hsu, P.D., Lin, C.Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y. and Zhang, F. (2013) Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380-1389.
- Ren, B., Yan, F., Kuang, Y., Li, N., Zhang, D., Lin, H. and Zhou, H. (2017) A CRISPR/Cas9 toolkit for efficient targeted base editing to induce genetic variations in rice. *Sci China Life Sci* **60**, 516-519.
- Ren, B., Yan, F., Kuang, Y., Li, N., Zhang, D., Zhou, X., Lin, H. and Zhou, H. (2018) Improved base editor for efficiently inducing genetic variations in rice with CRISPR/Cas9-guided hyperactive hAID mutant. *Mol Plant* **11**, 623-626.
- Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering Quantitative Trait Variation for Crop Improvement by Genome Editing. *Cell* **171**, 470-480 e478.
- Schimpl, S., Fauser, F. and Puchta, H. (2014) The CRISPR/Cas system can be used as nuclease for in planta gene targeting and as paired nickases for directed mutagenesis in Arabidopsis resulting in heritable progeny. *Plant J* **80**, 1139-1150.
- Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L. and Gao, C. (2013) Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **31**, 686-688.

- Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., Teramura, H., Yamamoto, T., Komatsu, H., Miura, K., Ezura, H., Nishida, K., Ariizumi, T. and Kondo, A. (2017) Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol* **35**, 441-443.
- Shiraki, T. and Kawakami, K. (2018) A tRNA-based multiplex sgRNA expression system in zebrafish and its application to generation of transgenic albino fish. *Sci Rep* **8**, 13366.
- Sun, Y., Li, J. and Xia, L. (2016a) Precise Genome Modification via Sequence-Specific Nucleases-Mediated Gene Targeting for Crop Improvement. *Front Plant Sci* **7**, 1928.
- Sun, Y., Zhang, X., Wu, C., He, Y., Ma, Y., Hou, H., Guo, X., Du, W., Zhao, Y. and Xia, L. (2016b) Engineering Herbicide Resistant Rice Plants through CRISPR/Cas9-mediated Homologous Recombination of the Acetolactate Synthase. *Mol Plant* **9**, 628-631.
- Svitashev, S., Young, J.K., Schwartz, C., Gao, H., Falco, S.C. and Cigan, A.M. (2015) Targeted Mutagenesis, Precise Gene Editing, and Site-Specific Gene Insertion in Maize Using Cas9 and Guide RNA. *Plant Physiol* **169**, 931-945.
- Szymczak, A.L., Workman, C.J., Wang, Y., Vignali, K.M., Dilioglou, S., Vanin, E.F. and Vignali, D.A. (2004) Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nat Biotechnol* **22**, 589-594.
- 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. and 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.
- Tang, X., Lowder, L.G., Zhang, T., Malzahn, A.A., Zheng, X., Voytas, D.F., Zhong, Z., Chen, Y., Ren, Q., Li, Q., Kirkland, E.R., Zhang, Y. and Qi, Y. (2017) A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants. *Nat Plants* **3**, 17018.
- Tang, X., Zheng, X., Qi, Y., Zhang, D., Cheng, Y., Tang, A., Voytas, D.F. and Zhang, Y. (2016) A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Mol Plant* **9**, 1088-1091.
- Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J. and Joung, J.K. (2017) CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR-Cas9 nuclease off-targets. *Nature methods* **14**, 607-614.
- Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J. and Joung, J.K. (2014) Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol* **32**, 569-576.
- Wang, M., Lu, Y., Botella, J., Mao, Y., Hua, K. and Zhu, J.K. (2017a) Gene Targeting by Homology-directed Repair in Rice using a Geminivirus-based CRISPR/Cas9 System. *Mol Plant* **10**, 1007-1010.
- Wang, M., Mao, Y., Lu, Y., Tao, X. and Zhu, J.K. (2017b) Multiplex Gene Editing in Rice using the CRISPR-Cpf1 System. *Mol Plant* **10**, 1011-1013.
- Wang, M., Mao, Y., Lu, Y., Wang, Z., Tao, X. and Zhu, J.K. (2018a) Multiplex gene editing in rice with simplified CRISPR-Cpf1 and CRISPR-Cas9 systems. *J Integr Plant Biol* **60**, 626-631.
- Wang, X., Li, J., Wang, Y., Yang, B., Wei, J., Wu, J., Wang, R., Huang, X., Chen, J. and Yang, L. (2018b) Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat Biotechnol* **36**, 946-949.
- Wu, H., Liu, Q., Shi, H., Xie, J., Zhang, Q., Ouyang, Z., Li, N., Yang, Y., Liu, Z., Zhao, Y., Lai, C., Ruan, D., Peng, J., Ge, W., Chen, F., Fan, N., Jin, Q., Liang, Y., Lan, T., Yang, X., Wang, X., Lei, Z., Doevendans, P.A., Sluijter, J.P.G., Wang, K., Li, X. and Lai, L. (2018) Engineering CRISPR/Cpf1 with tRNA promotes genome editing capability in mammalian systems. *Cell Mol Life Sci* **75**, 3593-3607.

- Xie, K., Minkenberg, B. and Yang, Y. (2015) Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 3570-3575.
- Xing, H.L., Dong, L., Wang, Z.P., Zhang, H.Y., Han, C.Y., Liu, B., Wang, X.C. and Chen, Q.J. (2014) A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC Plant Biol* **14**, 327.
- Xu, R., Qin, R., Li, H., Li, D., Li, L., Wei, P. and Yang, J. (2016) Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol J* **15**, 713-717.
- Yan, F., Kuang, Y., Ren, B., Wang, J., Zhang, D., Lin, H., Yang, B., Zhou, X. and Zhou, H. (2018) Highly Efficient A.T to G.C Base Editing by Cas9n-Guided tRNA Adenosine Deaminase in Rice. *Mol Plant* **11**, 631-634.
- Yin, K., Gao, C. and Qiu, J.L. (2017) Progress and prospects in plant genome editing. *Nat Plants* **3**, 17107.
- You, Q., Zhong, Z., Ren, Q., Hassan, F., Zhang, Y. and Zhang, T. (2018) CRISPRMatch: An Automatic Calculation and Visualization Tool for High-throughput CRISPR Genome-editing Data Analysis. *Int J Biol Sci* **14**, 858-862.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V. and Zhang, F. (2015) Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759-771.
- Zetsche, B., Heidenreich, M., Mohanraju, P., Fedorova, I., Kneppers, J., DeGennaro, E.M., Winblad, N., Choudhury, S.R., Abudayyeh, O.O., Gootenberg, J.S., Wu, W.Y., Scott, D.A., Severinov, K., van der Oost, J. and Zhang, F. (2017) Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol* **35**, 31-34.
- Zhang, H., Si, X., Ji, X., Fan, R., Liu, J., Chen, K., Wang, D. and Gao, C. (2018) Genome editing of upstream open reading frames enables translational control in plants. *Nat Biotechnol* **36**, 894-898.
- Zhang, Y., Zhang, F., Li, X., Baller, J.A., Qi, Y., Starker, C.G., Bogdanove, A.J. and Voytas, D.F. (2013) Transcription activator-like effector nucleases enable efficient plant genome engineering. *Plant Physiol* **161**, 20-27.
- Zhang, Z., Mao, Y., Ha, S., Liu, W., Botella, J.R. and Zhu, J.K. (2016) A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. *Plant Cell Rep* **35**, 1519-1533.
- Zheng, X., Yang, S., Zhang, D., Zhong, Z., Tang, X., Deng, K., Zhou, J., Qi, Y. and Zhang, Y. (2016) Effective screen of CRISPR/Cas9-induced mutants in rice by single-strand conformation polymorphism. *Plant Cell Rep* **35**, 1545-1554.
- Zhong, G., Wang, H., Li, Y., Tran, M.H. and Farzan, M. (2017) Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. *Nature chemical biology* **13**, 839-841.
- Zhong, Z., Zhang, Y., You, Q., Tang, X., Ren, Q., Liu, S., Yang, L., Wang, Y., Liu, X., Liu, B., Zhang, T., Zheng, X., Le, Y., Zhang, Y. and Qi, Y. (2018) Plant genome editing using FnCpf1 and LbCpf1 nucleases at redefined and altered PAM sites. *Mol Plant* **11**, 999-1002.
- Zhou, H., Liu, B., Weeks, D.P., Spalding, M.H. and Yang, B. (2014) Large chromosomal deletions and heritable small genetic changes induced by CRISPR/Cas9 in rice. *Nucleic Acids Res* **42**, 10903-10914.
- Zhou, J., Xin, X., He, Y., Chen, H., Li, Q., Tang, X., Zhong, Z., Deng, K., Zheng, X., Akher, S.A., Cai, G., Qi, Y. and Zhang, Y. (2018) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep.*
- Zhou, J.D., K.; Cheng, Y.; Zhong, Z.; Tian, L.; Tang, X.; Tang, A.; Zeng, X.; Zhang, T.; Qi, Y.; Zhang, Y. (2017) CRISPR-Cas9 based genome editing reveals new insights into microRNA function and regulation in rice. *Front Plant Sci* **8**, 1598.
- Zong, Y., Song, Q., Li, C., Jin, S., Zhang, D., Wang, Y., Qiu, J.L. and Gao, C. (2018) Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat Biotechnol* **36**, 950-953.

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

Figure legends

Figure 1. Comparison of three STU-Cas9 systems in rice cells. (a) Schematics of four CRISPR-Cas9 expression systems: conventional dual promoter system, STU-Cas9-RZ system, STU-Cas9-Csy4 system, and STU-Cas9-tRNA system. (b-g) Mutation frequencies at six target sites by four Cas9 systems. The experiments were carried out in rice protoplasts and the frequencies were measured by amplicon based deep sequencing. Error bars represent standard deviations of two biological replicates.

Figure 2. Comparison of positional deletion frequencies at six target sites. The experiments were carried out in rice protoplasts and the frequencies were measured by amplicon based deep sequencing. The PAM sites are highlighted in red. Each line represents the same target site, while each column represents the same Cas9 expression strategy. Error bars represent standard deviations of two biological replicates.

Figure 3. Frequencies of deletions of different sizes at six target sites. The experiments were carried out in rice protoplasts and the frequencies were measured by amplicon based deep sequencing. Each line represents the same target site, while each column represents the same Cas9 expression strategy. Error bars represent standard deviations of two biological replicates.

Figure 4. Evaluation of three STU-Cas9 systems with OsPDS-sgRNA01 in rice transgenic lines. (a) Mutation rates in stable transgenic T0 lines. Each line was genotyped by Sanger sequencing of PCR amplicons. (b) Phenotype of wild type control and example lines that contain biallelic mutations at the *OsPDS* target site.

Figure 5. Evaluation of three STU-Cas9 systems with multiplexed two sgRNAs in rice transgenic lines. (a) Schematics of the STU-Cas9 expression constructs. (b) Upper panel: a summary table of mutation frequencies at two target sites by three STU-Cas9 systems; lower panel: a schematic presentation of the genotyping results for all T0 lines categorized as biallelic mutation, monoallelic mutation and wild type.

Figure 6. Evaluation of three STU-Cas9 systems with multiplexed three sgRNAs in rice transgenic lines. (a) Schematics of the STU-Cas9 expression constructs. (b) Upper panel: a summary table of mutation frequencies at three target sites by three STU-Cas9 systems; lower panel: a schematic presentation of the genotyping results for all T0 lines categorized as biallelic mutation, monoallelic mutation and wild type.

Figure 7. STU-Cas9-tRNA system with multiplexed six sgRNAs in rice transgenic lines. (a) Schematics of the STU-Cas9-tRNA expression construct. (b) Upper panel: a summary table of mutation frequencies at three target sites by the STU-Cas9-tRNA system; lower panel: a schematic presentation of the genotyping results for all T0 lines categorized as biallelic mutation, monoallelic mutation and wild type.

Figure 8. Comparison of two STU-Cas9 base editing systems in rice. (a) Schematics of the two STU-Cas9 base editing vectors for expressing one sgRNA each. (b) C to T base editing frequencies at four target sites. The editing frequencies were calculated by combining all the base changes within the target windows by analyzing the deep sequencing data. Error bars represent standard deviations of two biological replicates. (c) Pie chart presentation of the frequencies of base changes at all the targetable cytosine sites by both STU-Cas9 base editors at four target sites. Note the average frequencies of two biological replicates are shown in red. (d) Summary of genome editing in T0 lines at two target sites, *OsCDC48*-sgRNA01 and *OsROC5*-sgRNA05.

Figure 9. A STU-Cas12a system for plant genome editing. (a) Schematics of the STU-Cas12a system expressing one crRNA, which is flanked by direct repeats (DRs) for processing by Cas12a. (b) Mutation frequencies at four target sites by the STU-Cas12a system. The experiments were carried out in rice protoplasts and the frequencies were measured by amplicon based deep sequencing. Error bars represent standard deviations of two biological replicates. (c) Mutation frequencies at four target sites by the STU-Cas12a system. (d) Schematics of the STU-Cas12a system expressing four crRNAs, which are flanked by DRs for processing by Cas12a. (e) Mutation frequencies at four target sites by the multiplexed STU-Cas12a system. The experiments were carried out in rice protoplasts and the frequencies were measured by amplicon based deep sequencing. Error bars represent standard deviations of two biological replicates. (f) Upper panel: a summary table for mutation frequencies at four target sites; lower panel: a schematic presentation of the genotyping results for all T0 lines categorized as biallelic mutation, monoallelic mutation and wild type.

Supporting information

Table S1. Oligos used in this study

Figure S1. CAPS analysis of STU-Cas9 systems. Transiently transformed rice protoplasts were assayed by CAPS at six target sites by four Cas9 systems with two biological replicates. WT: wild type; DU: Dual promoter system; RZ: STU-Cas9-RZ system; Csy4: STU-Cas9-Csy4 system; tRNA: STU-Cas9-tRNA system. DL2000 DNA marker consists of 6 fragments: 2000, 1000, 750, 500, 250 and 100 bp.

Figure S2. Base editing efficiency at four target sites by rAPOBEC1 and PmCDA1 base editors.

C to T base editing frequencies at four target sites with rAPOBEC1 and PmCDA1 base editors. Note only results of one representative biological replicate are shown here. The 20 nt protospacer and the PAM (in blue) are shown. The cytosines (Cs) within the base editing window are highlighted in red.

Figure S3. Detection of base editing in T0 rice lines by STU-Cas9-PmCDA with *OsCDC48*-sgRNA01.

(a) Base editing and indel events were first identified by single-strand conformation polymorphism (SSCP) analysis. (b) Sanger sequencing results for six example T0 lines. The target sequence is shown in blue. Base editing events are highlighted in red in lower case. The PAM is highlighted in red in upper case.

Figure S4. Detection of base editing in T0 rice lines by STU-Cas9-PmCDA with *OsROC5*-sgRNA05.

(a) Base editing and indel events were first identified by single-strand conformation polymorphism (SSCP) analysis. (b) Sanger sequencing results for six example T0 lines. The target sequence is shown in blue. Base editing events are highlighted in red in lower case. The PAM is highlighted in red in upper case.

Figure S5. CAPS analysis of the STU-Cas12a system. Transiently transformed rice protoplasts were assayed by CAPS at four target sites by the STU-Cas12a system with two biological replicates. DL2000 DNA marker consists of 6 fragments: 2000, 1000, 750, 500, 250 and 100 bp.

Figure S6. CAPS analysis of T0 lines generated by the STU-Cas12a system. STU-Cas12a targeting with *OsDEP1*-crRNA01, *OsDEP1*-crRNA02, *OsROC5*-crRNA01 and *OsROC5*-crRNA were analyzed in 18, 17, 13, and 22 T0 transgenic lines, respectively. DL2000 DNA marker consists of 6 fragments: 2000, 1000, 750, 500, 250 and 100 bp.

Figure S7. Analysis of transiently transformed rice protoplasts for multiplexing four crRNAs using the STU-Cas12a system. (a) CAPS analysis of transiently transformed rice protoplasts for multiplexed editing with two biological replicates. (b) Detection of targeted deletions by two crRNAs targeting the same genes (*OsDEP1* and *OsROC5*) in transiently transformed rice protoplasts with two biological replicates.

The arrows point to deletion products. DL2000 DNA marker consists of 6 fragments: 2000, 1000, 750, 500, 250 and 100 bp.

Figure S8. CAPS analysis of T0 lines generated by the multiplexed STU-Cas12a system. A total of 24 T0 transgenic lines were assayed by CAPS for mutations at the targeted sites of OsDEP1-crRNA01, OsDEP1-crRNA02, OsROC5-crRNA01 and OsROC5-crRNA02. DL2000 DNA marker consists of 6 fragments: 2000, 1000, 750, 500, 250 and 100 bp.

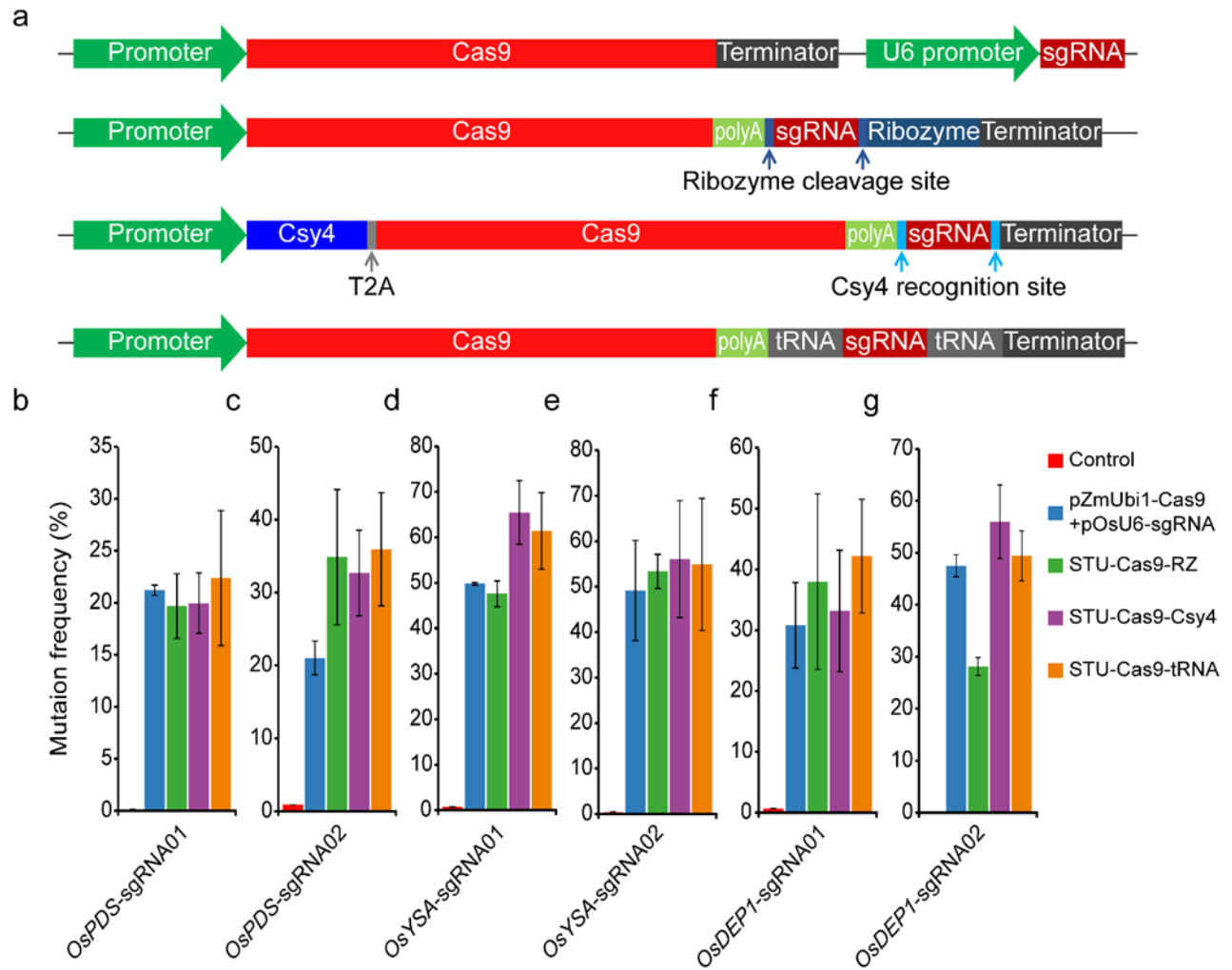


Figure 1. Comparison of three STU-Cas9 systems in rice cells

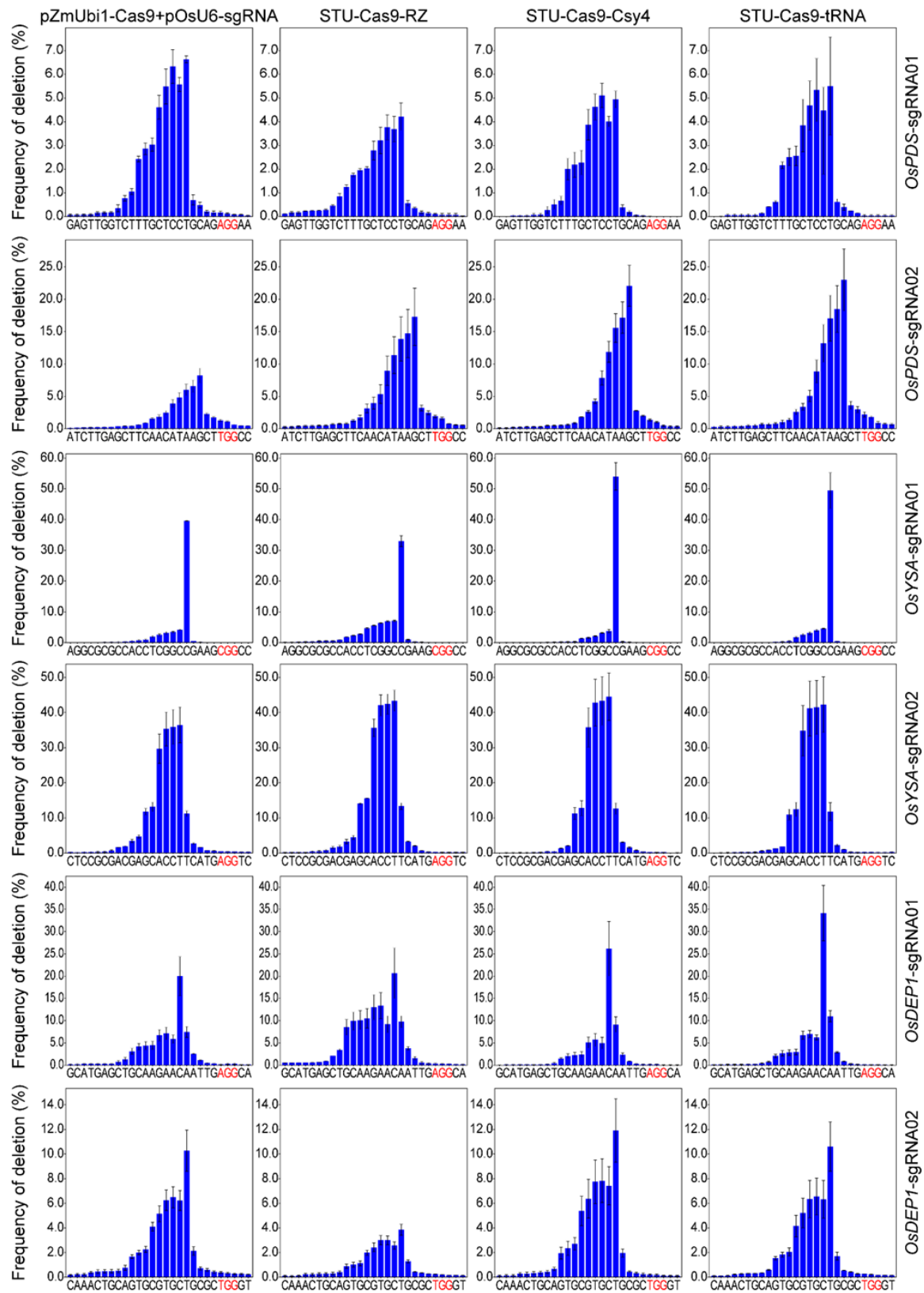


Figure 2. Comparison of positional deletion frequencies at six target sites

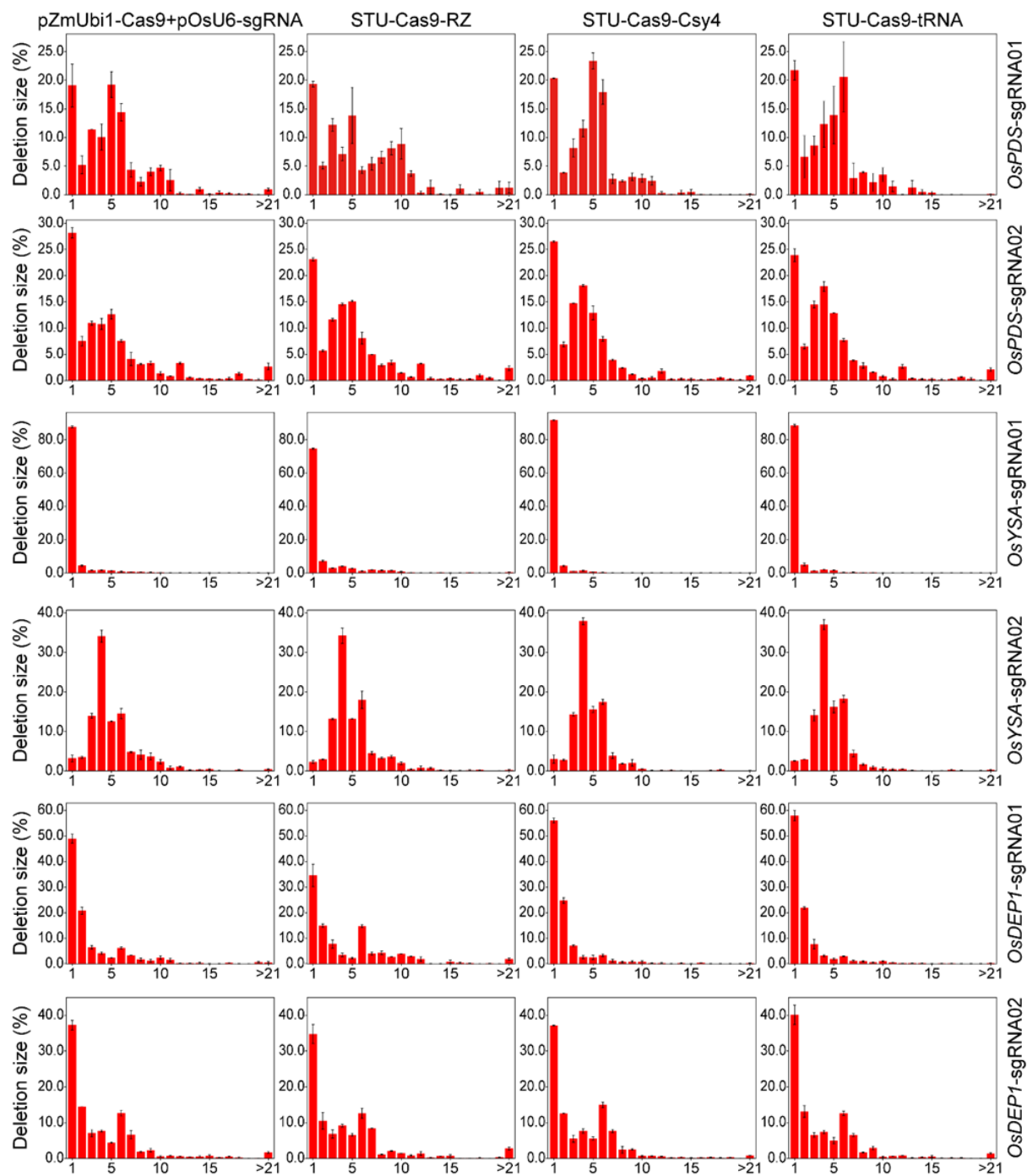


Figure 3. Frequencies of deletions of different sizes at six target sites

a

Construct	Tested T0 lines	Mutated T0 lines (number; ratio)	Biallelic mutation lines (number; ratio)
pZmUbi1-Cas9+pOsU6-sgRNA	50	35; 70.0%	26; 52.0%
STU-Cas9-RZ	51	38; 74.5%	31; 60.8%
STU-Cas9-Csy4	51	45; 88.2%	36; 70.6%
STU-Cas9-tRNA	63	51; 81.0%	45; 71.4%

b



Figure 4. Evaluation of three STU-Cas9 systems with OsPDS-sgRNA01 in rice stable transgenic lines

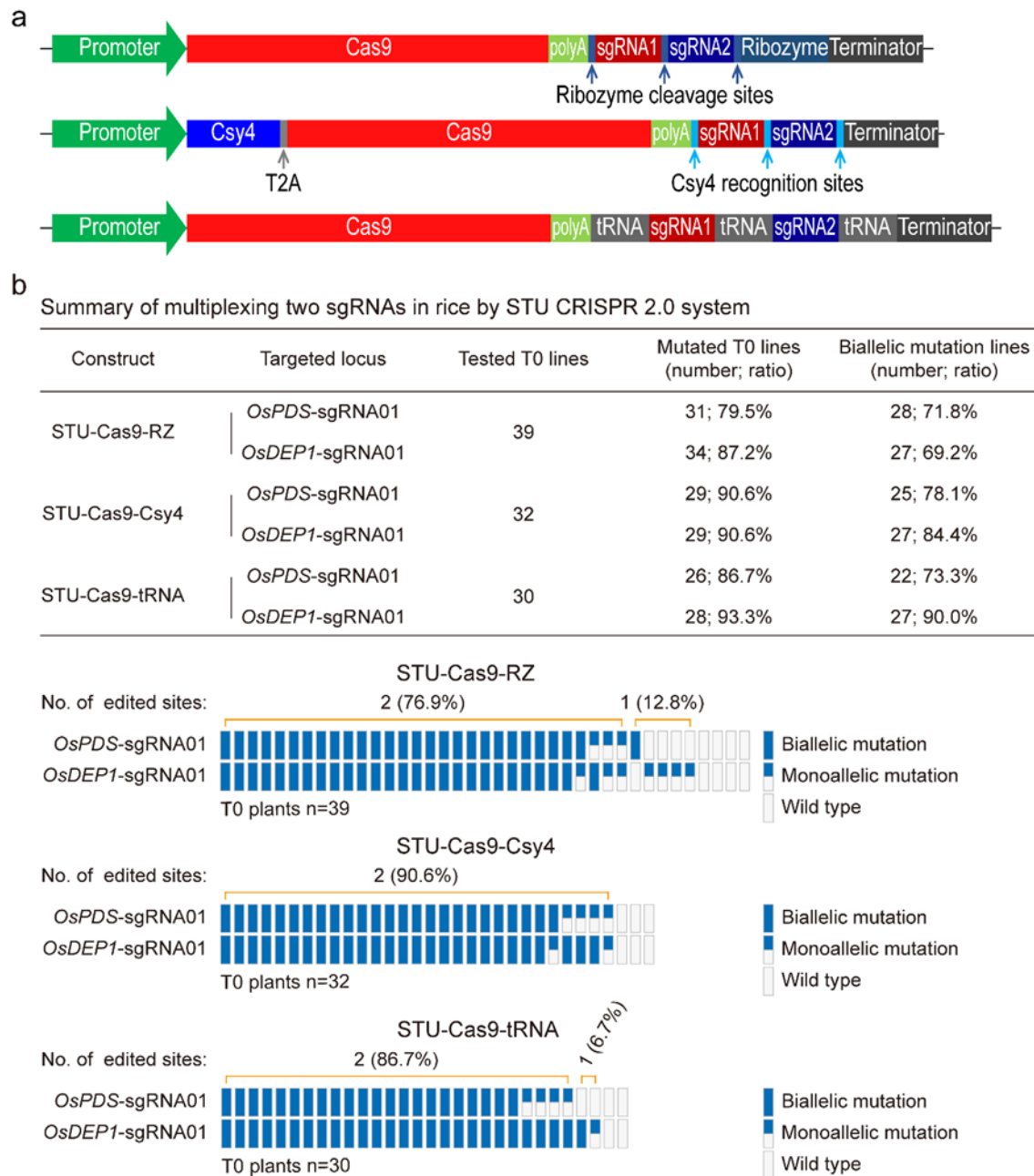


Figure 5. Evaluation of three STU-Cas9 systems with multiplexing two sgRNAs in rice transgenic lines



b Summary of multiplexing three sgRNAs in rice by STU CRISPR 2.0 system

Construct	Targeted locus	Tested T0 lines	Mutated T0 lines (number; ratio)	Biallelic mutation lines (number; ratio)
STU-Cas9-RZ	<i>OsPDS</i> -sgRNA01	36	30; 83.3%	25; 69.4%
	<i>OsPDS</i> -sgRNA02		21; 58.3%	12; 33.3%
	<i>OsDEP1</i> -sgRNA01		31; 86.1%	25; 69.4%
STU-Cas9-Csy4	<i>OsPDS</i> -sgRNA01	33	31; 93.9%	27; 81.8%
	<i>OsPDS</i> -sgRNA02		20; 60.6%	13; 39.4%
	<i>OsDEP1</i> -sgRNA01		28; 84.8%	21; 63.6%
STU-Cas9-tRNA	<i>OsPDS</i> -sgRNA01	37	33; 89.2%	31; 83.8%
	<i>OsPDS</i> -sgRNA02		25; 67.6%	23; 62.2%
	<i>OsDEP1</i> -sgRNA01		33; 89.2%	30; 81.1%

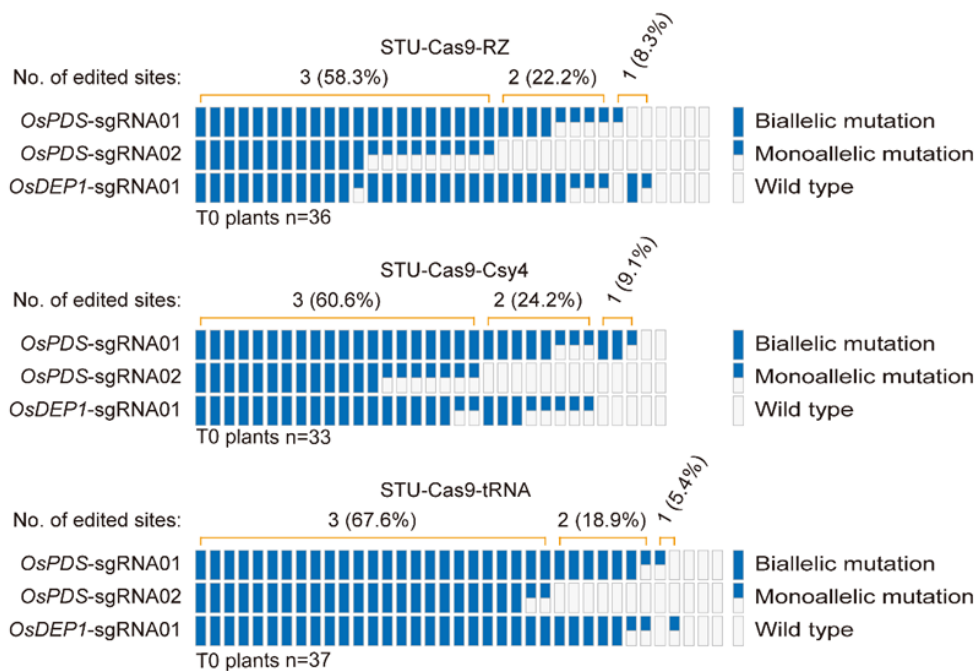


Figure 6. Evaluation of three STU-Cas9 systems with multiplexing three sgRNAs in rice transgenic lines

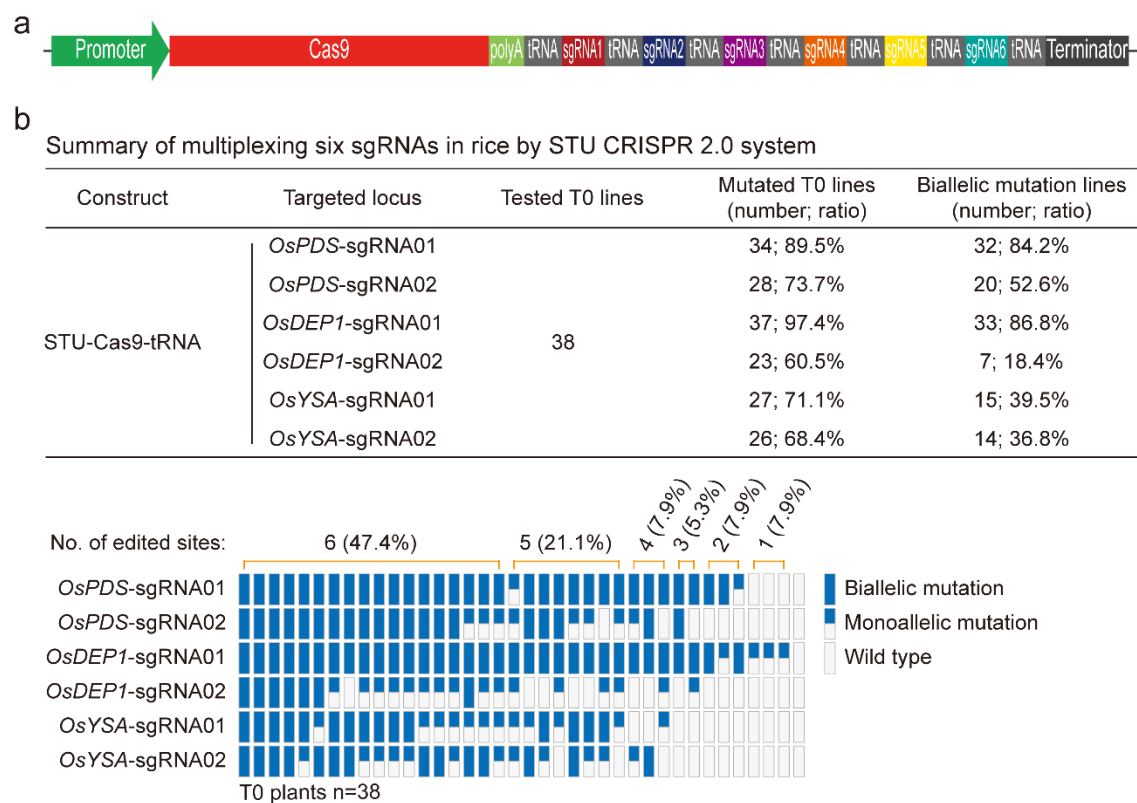


Figure 7. Evaluation of the STU Cas9-tRNA system with multiplexing six sgRNAs by stable transformation