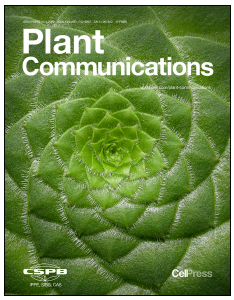
Journal Pre-proof



Boosting genome editing in plants with single transcript unit surrogate reporter systems

Xu Tang, Qiurong Ren, Xiaodan Yan, Rui Zhang, Li Liu, Qinqin Han, Xuelian Zheng, Yiping Qi, Hongyuan Song, Yong Zhang

PII: S2590-3462(24)00191-3

DOI: https://doi.org/10.1016/j.xplc.2024.100921

Reference: XPLC 100921

To appear in: *PLANT COMMUNICATIONS*

Received Date: 2 January 2024

Revised Date: 20 March 2024

Accepted Date: 11 April 2024

Please cite this article as: Tang, X., Ren, Q., Yan, X., Zhang, R., Liu, L., Han, Q., Zheng, X., Qi, Y., Song, H., Zhang, Y., Boosting genome editing in plants with single transcript unit surrogate reporter systems, *PLANT COMMUNICATIONS* (2024), doi: https://doi.org/10.1016/j.xplc.2024.100921.

Boosting genome editing in plants with single transcript unit surrogate reporter systems

Xu Tang[1, 2, 3, 4#], Qiurong Ren[3, 5#], Xiaodan Yan[2, 3], Rui Zhang[4], Li Liu[4], Qinqin Han[4], Xuelian Zheng[1,4], Yiping Qi[6, 7*], Hongyuan Song[2, 3*], Yong Zhang[1, 2, 3, 4*]

[1]Chongqing Key Laboratory of Plant Resource Conservation and Germplasm Innovation, Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, School of Life Sciences, Southwest University, Chongqing 400715, China; [2]Key Laboratory of Agricultural Biosafety and Green Production of Upper Yangtze River, Ministry of Education, Chongqing 400715, China; [3]College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400715, China; [4]Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 610054, China; [5]School of Synbiology, School of Life Science, Shanxi University, Taiyuan 030006, China; [6]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, Maryland 20742, USA; [7]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland 20850, USA

[#]These authors contributed equally to this work

Running title: Improved surrogate reporter systems in plants

Short summary: Our study introduces multiple STU-SR systems to improve selecting genome-edited plants. By using identical sgRNAs for reporter and endogenous genes, a direct correlation between their editing activities is established. STU-SR systems employ strategies like single strand annealing and base editing to restore functional reporter genes post-genome editing in rice and *Brassica oleracea*, demonstrating their versatility and effectiveness in enhancing genome-edited plants.

***Corresponding authors:**

Yong Zhang, Chongqing Key Laboratory of Plant Resource Conservation and Germplasm Innovation, Integrative Science Center of Germplasm Creation in Western China (Chongqing) Science City, School of Life Sciences, Southwest University, Chongqing 400715, China; Email: zhangyong916@swu.edu.cn

Hongyuan Song, Key Laboratory of Agricultural Biosafety and Green Production of Upper Yangtze River, Ministry of Education, College of Horticulture and Landscape Architecture, Southwest University, Chongqing 400715, China; Email: yuahs@swu.edu.cn

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

Abstract: CRISPR-Cas-based genome editing holds immense promise for advancing plant genomics and crop enhancement. However, the challenge of low editing activity complicates the identification of edited events. In this study, we introduce multiple Single Transcript Unit Surrogate Reporter (STU-SR) systems to enhance the selection of genome-edited plants. These systems utilize the same sgRNAs designed for endogenous genes to edit reporter genes, establishing a direct link between reporter gene editing activity and that of endogenous genes. Various strategies are employed to restore functional reporter genes post-genome editing, including efficient single strand annealing (SSA) for homologous recombination in STU-SR-SSA systems. STU-SR-BE systems leverage base editing to reinstate the start codon, enriching C-to-T and A-to-G base editing events. Our results showcase the effectiveness of these STU-SR systems in enhancing genome editing events in monocot rice, encompassing Cas9 nuclease-based targeted mutagenesis, cytosine base editing, and adenine base editing. The systems exhibit compatibility with Cas9 variants, such as the PAM-less SpRY, and are demonstrated to boost genome editing in *Brassica oleracea*, a dicot vegetable crop. In summary, we have developed highly efficient and versatile STU-SR systems for enrichment of genome-edited plants.

Key words: CRISPR-Cas9, single transcript unit (STU), surrogate reporter (SR), single strand annealing (SSA), cytosine base editing, adenine base editing

Introduction

Genome editing has become a powerful tool for modifying plant genomes, offering significant potential for plant genomics research and crop genetic improvement (Zhang et al., 2019). Among these technologies, the CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR Associated) genome editing system has gained considerable attention due to its efficiency, precision, and ease to make. The CRISPR-Cas systems such as Cas9 or Cas12a (Tang and Zhang, 2023), guided by guide RNAs (gRNAs), can target specific positions in the genome, enabling precise editing of targeted genes. Over the years, genome editing techniques have continued to evolve and expand. Two revolutionary editing technologies have emerged in the field of genome editing: base editing and prime editing (Anzalone et al., 2019; Gaudelli et al., 2017; Komor et al., 2016). Base editing allows direct conversion of one base to another, while prime editing enables base transition, transversion and precise short indels without double-strand breaks (DSB) and donor repair template (Anzalone et al., 2019). The introduction of these technologies further broadens the application scope of genome editing in plants, such as fine-tuning of plant gene expression (Tang and Zhang, 2023; Zhou et al., 2023).

Despite the significant breakthroughs achieved with CRISPR-Cas technologies, there are still challenges that limit their widespread applications in plants. One such challenge is that editing efficiency may vary significantly with different gRNAs under different conditions. This instability of editing efficiency complicates the screening of editing events and increases the risk of experimental failure. High genome editing efficiency based on CRISPR-Cas requires optimal delivery and expression of editing reagents, typically via methods such as Agrobacterium-mediated transformation. In these experiments, it often involves the use of selectable marker genes (typically antibiotic or herbicide resistance markers such as kanamycin, hygromycin, Basta, etc.) to select for transgenic events. These transgenic marker genes can help researchers identify which cells or tissues have successfully received the constructs, but they do not directly reflect the editing status of the target gene. Therefore, researchers often need to further screen the transformation events to determine which events have undergone the desired gene edits. As a result, if the CRISPR-Cas expression levels are low or the editing efficiency at the target site is low in the transgenic events, it becomes challenging to obtain editing events within the transgenic population. Unfortunately, this happens to be the case when the conventional CRISPR-Cas9 system is constructed and used, where the selection marker, Cas9, and single guide RNA (sgRNA) cassettes are driven by separate promoters (Figure 1A) (Hassan et al., 2021).

Previously, surrogate reporters were demonstrated to enrich cells with targeted mutations by nucleases such as zinc-finger nucleases (ZFNs) and TAL-effector nucleases (TALENs) (Kim et al., 2011). Later, this strategy has been applied to enrich CRISPR-Cas9 based genome editing events in mammalian cells (Liao et al., 2015; Ramakrishna et al., 2014). More recently, similar surrogate reporter systems were established in plants to enrich genome editing events by base editing (Xu et al., 2020b), prime editing (Li et al., 2020; Xu et al., 2020a), or Cas9 mediated mutagenesis via a viral vector system (Tian et al., 2022). Among these surrogate reporter approaches, dedicated sgRNAs were employed to target the broken surrogate reporter genes while different sgRNAs were used to edit endogenous genes of interest. Compared to the conventional CRISPR-Cas9 expression system (Figure 1A), these surrogate reporter systems allow for the selection of Cas9 expression because only successful editing of the broken marker genes will result in regenerated plants (Figure 1B). However, because different sgRNAs are used to edit surrogate reporter genes and endogenous genes, these reporter systems only select for events with high CRISPR-Cas expression levels but do not necessarily exhibit high genome editing activity at the endogenous target sites (Figure 1B). Hence, such surrogate reporter systems fall short in directly reporting genome editing activity for the genes of interest, highlighting a need for more integrated and efficient systems.

Addressing this gap, we introduce the Single Transcript Unit Surrogate Reporter (STU-SR) systems, a leap forward in CRISPR-Cas9 technology (Tang et al., 2019; Tang et al., 2016). Unlike preceding methods,

the STU-SR systems utilize identical sgRNAs for both the reporter and target genes, directly linking reporter gene editing with the modification of endogenous genes. This design not only simplifies the selection process by ensuring that successful plant regeneration is intrinsically tied to effective genome editing but also enhances overall efficiency by coordinated expression of Cas9 and sgRNAs under a single Pol II promoter (Figure 1C). Tailored for versatility across various plant species (Figure 1C), the STU-SR systems' streamlined vector construction markedly improves user accessibility for diverse applications in plant biotechnology. Demonstrated to significantly boost genome editing outcomes in both monocot (rice) and dicot (*Brassica oleracea*) plants, these systems represent a significant advancement in our capacity to precisely edit plant genomes. Thus, these innovative STU-SR systems improve editing efficiency and reduce the efforts in screening genome edited plants. This approach opens new possibilities for improving precise genome editing that can go beyond plants.

Results

Establishment of an STU-SR-SSA system for enriching editing events by CRISPR-Cas9

To effectively enrich CRISPR-Cas9 editing events, we tested an STU-SR strategy using the sgRNAs designed to target endogenous genes of interest to edit the surrogate reporter genes. Specifically, we adopted a broken reporter configuration that requires single strand annealing (SSA)-based homologous recombination (HR) to reconstitute the functional reporter, such as a Hygromycin phosphotransferase (*HYG*) reporter gene (Figure 2A). SSA was chosen because it is a very efficient HR repair mechanism in both dividing and non-dividing cells, likely owning to its independence of exogenous repair template (Puchta, 2005; Roth et al., 2012; Zhang et al., 2013). Initially, we tested this concept in rice by inserting the endogenous target site with the protospacer adjacent motif (PAM), called surrogate site, within the *HYG* gene, flanked by 90 bp homologous sequences, to construct the *HY::YG* surrogate reporter gene. When the *HY::YG* surrogate reporter gene is cleaved by sgRNAs, it can be restored to a complete Hygromycin phosphotransferase coding sequence via SSA, leading to the expression of functional Hygromycin phosphotransferase. Simultaneously, the sgRNA targeting the endogenous gene may induce mutations at the target gene of interest. This allows for the enrichment of editing events, achieved by selecting hygromycin-resistant plants (Figure 2A). For the construction of the STU-SR-SSA vector in rice, we employed a tail-to-tail design of STU Cas9 expression unit and *HY::YG* reporter expression unit (Figure 2B). This design allows simultaneous cloning of sgRNAs and their target sites into the STU Cas9 expression unit and the *HY::YG* reporter expression unit, respectively. These corresponding DNA fragments can be amplified and ligated into the STU-SR backbone vector after enzymatic digestion with BsaI and SpeI, resulting in the STU-SR expression vectors (Figure 2B).

To test the effectiveness of the STU-SR-SSA system, we compared the editing efficiencies of the STU-SR-SSA system with the STU CRISPR-Cas9 control system at three endogenous rice gene target sites in stable transgenic rice plants (Figure 2C). The results showed that the editing efficiencies for the STU system at the *OsPDS*-sgRNA01, *OsPDS*-sgRNA02, and *OsDEP1*-sgRNA01 loci were 56.3%, 77.3%, and 77.8%, respectively. In contrast, the STU-SR-SSA system achieved 100% editing efficiency at all three sites, representing an enhancement of editing efficiency by 28.5% to 77.6% compared to the STU system. The biallelic editing efficiencies at these three loci increased from 43.8%-72.2% to 80%-100%, with *OsPDS*-sgRNA02 and *OsDEP1*-sgRNA01 both reaching 100% biallelic editing efficiency (Figure 2C, Supplemental Figure 1). Sanger sequencing results confirmed successful editing at both the *HY::YG* reporter gene and the endogenous gene loci (Figure 2D-2E). *OsPDS* and *OsDEP1* biallelic mutant plants exhibited photobleaching and dwarf phenotype, respectively (Figure 2F-2G), consistent with our previous report on the knockout phenotypes of both genes (Lowder et al., 2015).

Previously, we demonstrated that SpRY could achieve PAM-less genome editing in plants (Ren et al., 2021b). However, the overall editing efficiency of SpRY is lower than the wildtype SpCas9, likely due to its PAM-less nature and self-editing when delivery in DNA constructs (Ren et al., 2021b; Walton et al., 2020). It is very appealing to further improve the SpRY expression system for more robust editing outcomes in plants. We investigated whether the STU-SR-SSA system could enhance the editing efficiency of SpRY. The *OsPDS*-sgRNA04 locus was used for testing with stable rice transformation, and the results showed that the STU-SR-SSA system increased editing efficiency by 33.3% compared to the STU system (100% vs. 75%) and enhanced biallelic editing efficiency by 53.8% (100% vs. 65%) (Figure 2C, Supplemental Figure 2). Similarly, Sanger sequencing results confirmed successful editing at both the *HY::YG* reporter gene and the endogenous target gene, *OsPDS* (Figure 2D-E). These results collectively indicate that the STU-SR-SSA system works effectively with Cas9 and its variants for improved genome editing in rice.

Enrichment of genome editing events independent of T-DNA copy number

In the process of Agrobacterium-mediated stable transformation, the integration of one or more T-DNA copies can occur (De Buck et al., 2009; Jupe et al., 2019; Lee and Gelvin, 2008). When a single T-DNA copy is inserted, the *HY::YG* reporter gene should be corrected, resulting in a single band in the PCR detection (Figure 3A). However, when two or more T-DNA copies are integrated, the *HY::YG* reporter genes may experience the correction of one or multiple copies or even all copies, corresponding to one or two bands in the PCR detection (Figure 3A). We conducted PCR detection on 8 randomly selected T0 plants per STU-SR-SSA construct. At the *OsPDS*-sgRNA04 site edited by SpRY, two plants (lines 1 and 5) showed a single *HYG* band, exhibiting complete correction of the *HY::YG* reporter gene(s) (Figure 3B). Four plants (lines 2, 3, 6 and 7) showed both *HYG* and *HY::YG* bands with nearly equivalent intensity,

suggesting these lines carry two copy of *HY::YG* reporters and only one was corrected (Figure 3B). For lines 4 and 8, the corrected *HYG* band was faint and the *HY::YG* was much brighter, indicating that multiple copy of the *HY::YG* reporter were present in these two lines and most of them remained unedited or not corrected (Figure 3B). The presence of edited and/or unedited alleles of *HY::YG* and editing at the endogenous site were confirmed by Sanger sequencing (Figure 3C). For the *OsPDS*-sgRNA01 and *OsPDS*-sgRNA02 plants, the *HY::YG* reporter gene(s) was fully restored in all cases (Supplemental Figure 3). Regarding the 8 *OsDEP1*-sgRNA01 T0 plants, five plants exhibited complete correction of the *HY::YG* reporter gene(s), while the other three plants showed partial copy correction (Supplemental Figure 3). Hence, regardless the copy number of the *HY::YG* reporter in the regenerated plants, correction of at least one broken reporter copy seems sufficient to report robust genome editing at the endogenous loci. These data indicate that the STU-SR system is capable of enriching gene editing events regardless of the number of T-DNA copies, providing flexibility in its application.

Application of STU-SR-BE systems for enriching base editing events

Base editing, either via cytosine base editors (CBEs) or adenine base editors (ABEs), allows for the transition of one base to another without introducing DNA double-strand breaks. Base editing provides greater precision than targeted mutagenesis by Cas nucleases (Molla et al., 2021; Ren et al., 2021a; Ren et al., 2021b; Wu et al., 2022a). Therefore, we aimed to investigate whether the STU-SR system could also effectively enrich base editing events. Unlike the SSA enrichment strategy used for enriching CRISPR-Cas9 editing events, we developed STU-SR-CBE and STU-SR-ABE for enriching base editing events. In both cases, the surrogate reporter gene was created by strategically integrating the endogenous target site with the PAM (also known as the surrogate site) to the beginning of the hygromycin resistance (*HYG*) coding sequence, with removing the original ATG start codon (hyg^{-ATG}) (Figure 4A). For the CBE system, when the 5' end of the sgRNA contains "ACG," we can directly add the surrogate site in front of the hyg^{-ATG} coding sequence. If "CAC" is present at the 5' end of the sgRNA, we will add the reverse complementary sequence of the surrogate site before hyg^{-ATG}. In both cases, "ACG" and "CAC" must be in the same reading frame as hyg^{-ATG} (Figure 4A). If needed, appropriate bases should be added between the surrogate site and hyg^{-ATG} to ensure they are in the same reading frame. When C-to-T editing occurs in the base C of "ACG" or the base C close to PAM in "CAC", a new ATG start codon is generated, initiating the translation of the *HYG* gene. For the ABE system, if the sgRNA's 5' end contains "ATA," we can directly add the surrogate site in front of hyg^{-ATG}. When A-to-G editing occurs in the base A near PAM in "ATA", a new ATG start codon will be generated, initiating the translation of the *HYG* gene (Figure 4A). If the 5' end of the sgRNA lacks the corresponding sequence, such as "ACG" or "CAC" for cytosine base editing, or "ATA" for adenine base editing, one will need to adjust the protospacer sequence at the PAM distal end, allowing the presence

of a 1-2 bp mismatch to form the corresponding sequence, as base editors tolerate most of the single or double mismatches at the distal end (Kim et al., 2019; Talas et al., 2021). Upon successful expression of the active hygromycin phosphotransferase and potential base editing at the surrogate site, hygromycin-resistant plants would be obtained, which presumably should also carry high frequency base editing at the endogenous target site because the same sgRNA is used for editing both sites (Figure 4A).

In our demonstration, we used a PmCDA1-based CBE (Nishida et al., 2016; Tang et al., 2019) because it is a highly efficient CBE in plants with undetectable genome-wide off-target effects (Randall et al., 2021; Ren et al., 2021a). For A-to-G editing, ABE8e was chose due to its high editing activities (Lapinaite et al., 2020; Richter et al., 2020). Also, ABE8e's genome-wide off-target effects have been comprehensively studies in plants (Sretenovic et al., 2023; Wu et al., 2022b). Similar to the STU-SR-SSA system, we designed the CBE/ABE expression cassette and hyg^{-ATG} expression cassette tail-to-tail for ease of cloning sgRNAs and surrogate sites into the STU-SR-BE backbone. Using forward and reverse oligos as primers, we amplified fragments from the STU-SR-BE backbones. After digestion with BsaI and AvrII (KpnI), these PCR fragments were ligated into the STU-SR-BE backbone vectors cut with the same enzymes, resulting in the STU-SR-CBE and STU-SR-ABE expression vectors (Figure 4B). Our initial experiments evaluated the enrichment efficiency of STU-SR-CBE at four rice endogenous target sites. The results demonstrated that T0 transgenic plants generated by STU-SR-CBE consistently exhibited the correction of the surrogate reporter gene (Figure 4C). At the *OsPDS*-sgRNA02 site, the STU-SR-CBE system showed a 2.09-fold enhancement in editing efficiency compared to the CBE system (11.1% vs 5.3%). Similarly, at the *OsPDS*-sgRNA07 site, a 2.30-fold improvement was observed with STU-SR-CBE over STU-CBE (88.9% vs 38.6%). At the *OsCDC48*-sgRNA01 site, 1.53-fold enhancement in editing efficiency was observed with STU-SR-CBE over STU-CBE (80% vs 52.4%). At *OsDEP1*-sgRNA03 site, the STU-CBE system already achieved 95.5% editing efficiency, and STU-SR-CBE further improved it to 100% editing efficiency (Figure 4C). Sanger sequencing confirmed successful C-to-T editing at both the reporter gene and endogenous gene loci (Figure 4D, Supplemental Figure 4).

We also expanded the editing range by replacing Cas9 in the STU-SR-CBE system with SpRY, generating STU-SR-CBE_SpRY for PAM-less C-to-T base editing. Testing at the *OsPDS*-sgRNA04 site, the STU-SR-CBE_SpRY system showed a 1.98-fold increase in editing efficiency compared to STU-CBE_SpRY (9.5% vs 4.8%) (Figure 4C, 4E, Supplemental Figure 4). Similarly, we also generated STU-SR-SpRY-ABE for enriching PAM-less A-to-G base editing. At the *OsMPK6*-sgRNA01 site, we observed a 3.02-fold enhancement in editing efficiency with STU-SR-ABE_SpRY (13.6% vs 4.5%) (Figure 4C, 4F, Supplemental Figure 4). Together, the STU-SR-BE systems effectively enriched base editing events in every case tested and the systems can be combined with SpRY to expand the target range in base editing.

Enhanced genome editing in *Brassica oleracea* with the STU-SR-SSA system

To further validate the versatility of the STU-SR system for enriching genome editing events in other plant species, we chose *Brassica oleracea*, a dicotyledonous plant. In *Brassica oleracea* transformation, the commonly used selection marker is the Bialaphos resistance (*BAR*) gene. Therefore, we integrated endogenous gene editing sites into the *BAR* gene, flanked by homologous sequences, to construct the *BA::AR* surrogate reporter gene. The STU Cas9 expression unit and the *BA::AR* reporter expression unit were designed to be tail to tail arranged, facilitating the construction of STU-SR-SSA expression vectors that utilize the *BAR* selection marker gene (Figure 5A). When the *BA::AR* surrogate reporter gene is cleaved by a targeting sgRNA, it can be restored to a functional *BAR* gene through SSA homologous recombination. This reconstitution enables the expression of the active phosphinothricin acetyltransferase (PAT) enzyme. Simultaneously, the same sgRNA will target the endogenous locus to render editing. By selecting for Basta-resistant plants, we could achieve the enrichment of events where mutations occurred at the endogenous target gene (Figure 5A).

We selected three endogenous target loci, *BoPDS*-sgRNA01, *BoPDS*-sgRNA02, and *BoBIK*-sgRNA01, and compared the editing efficiencies in stable transgenic plants. Like what we did in rice, we compared the STU-SR-SSA system to the STU system. Transgenic T0 *Brassica oleracea* plants were obtained with Agrobacterium-mediated transformation. Our results revealed that at the *BoPDS*-sgRNA01 site, the STU system did not generate any mutations among 60 T0 lines, whereas the STU-SR-SSA system achieved editing in 35% of the 80 T0 plants. At the *BoPDS*-sgRNA02 site, the STU-SR-SSA system exhibited a 3.34-fold improvement in editing efficiency compared to the STU system (5% vs 16.7%). Moreover, at the *BoBIK*-sgRNA01 site, the STU-SR-SSA system demonstrated a remarkable 48.4% increase in editing efficiency, reaching nearly 80% (Figure 5B). Hence, the STU-SR-SSA system effectively enriched editing events at all three endogenous target sites (Figure 5B). Based on phenotypic analysis of the regenerated T0 plants, the STU-SR-SSA system generated biallelic and homozygous mutations much more frequently than the STU system, either for *BoPDS* (Figure 5C) and *BoBIK* (Figure 5D), where the loss of function phenotypes as photobleaching and dwarfism respectively were previously reported (Ma et al., 2019). Further validation through Sanger sequencing confirmed the restoration of the *BA::AR* surrogate reporter gene (Figure 5E) and demonstrated successful editing at the endogenous gene loci (Figure 5F). In summary, our experiments in *Brassica oleracea* demonstrate the effectiveness of the STU-SR-SSA system in enriching gene editing events in dicotyledonous plants.

Discussion

CRISPR-Cas based genome editing systems have been rapidly adopted by plant researchers to investigate gene function and develop improved crops. As DNA targeting and editing is programmed by gRNAs, it is inevitable that editing efficiency at different target sites may vary significantly. Not to mention, chromatin status and epigenetic features also play a role in influencing editing outcomes (Liu et al., 2019; Weiss et al., 2022). Currently, we have not fully understood how each CRISPR-Cas system functions optimally in plants and the rules for designing the most efficient gRNAs for each application. Furthermore, the target sites may be pre-defined in base editing or prime editing experiments and there is lack of flexibility of altering the protospacer sequences in order to boost the genome editing activity. Considering these circumstances, it is immensely significant if we can enrich gene editing events. It will help accelerate the pace of scientific investigations and plant breeding initiatives, enabling researchers to obtain edited plants more swiftly and efficiently, thereby saving valuable time and resources. In the context of complex plant genomes, where specific editing events can be exceedingly rare and challenging to obtain, enriching these events substantially increases the likelihood of achieving the desired genetic modifications, particularly for intricate editing tasks.

Previously, surrogate reporter systems have been developed to enrich genome editing outcomes for targeted mutagenesis by CRISPR-Cas9 (Tian et al., 2022), base editing (Xu et al., 2020b), and prime editing (Li et al., 2020; Xu et al., 2020a). However, these surrogate reporter systems each rely on a pre-defined sgRNA to revert the broken reporter to a functional one. These systems thus can enrich transgenic lines with high levels of Cas expression but cannot necessarily enrich lines with high sgRNA expression and activity for the endogenous target sites (Figure 1B). We reasoned that an ideal and efficient surrogate reporter system would allow for direct selection of edited lines based on the sgRNAs targeting the genes of interest (Figure 1C). However, at a first glance, it is very challenging to develop such surrogate report systems. After all, the sequences of the reporter genes are pre-defined, but the target sequences would vary from experiment to experiment, depending on the genes of interest. We employed different strategies to solve this problem. To develop surrogate reporter systems for CRISPR-Cas9, we used single strand annealing (SSA) that relies on the tandem repeat sequences flanking the DNA DSB, as SSA is an efficient DSB repair mechanism conserved in a variety of organisms, including yeast (Paques and Haber, 1999), humans (Bhargava et al., 2016), and plants (Puchta, 2005; Zhang et al., 2013). Indeed, our STU-SR-SSA reporter systems, based on different selection markers, worked well and significantly enriched editing events in rice (Figure 2) and *Brassica oleracea* (Figure 5). Furthermore, we found that the STU-SR-SSA reporter systems worked robustly, regardless of the copy number of T-DNA (Figure 3). There data suggest that the SSA-based surrogate systems are widely applicable in plants. It should be compatible to biolistic

transformation methods which often generate transgenic plants with multiple copies of the transgene. Although we demonstrated the STU-SR-SSA systems with CRISPR-Cas9, the same SSA principle can be used for developing surrogate reporter systems for other CRISPR systems such as CRISPR-Cas12a (Tang et al., 2017; Zhong et al., 2018), CRISPR-Cas12b (Ming et al., 2020) and compact CRISPR-Cas12j2 (Liu et al., 2022).

Compared to CRISPR-Cas mediated targeted mutagenesis, base editing and prime editing are more precise genome editing tools for installing base-precision changes in the genomes (Anzalone et al., 2020). However, these genome editing tools often suffer from low editing efficiency, making their use in plant applications more laborious and riskier (Molla et al., 2021; Tang et al., 2020). In this study, we developed STU-SR-CBE and STU-SR-ABE systems to enrich C-to-T and A-to-G base editing events, respectively. In our strategy, the surrogate site (encompassing the protospacer and PAM) was added to the 5' end of the reporter gene with its original start codon deleted. The start codon-less reporter gene is revived when a start codon is created via base editing (Figure 4A). Our data showed such STU-SR-BE systems significantly enriched base editing events, and the systems are compatible with Cas9 variants such as SpRY (Figure 4C). It is of note that the target 'C' for the CBE reporter and 'A' for the ABE reporter resides at 2nd and 3rd position of the protospacer respectively (Figure 4A), which are not necessary within the optimal base editing windows for all base editors. However, the fact base editing at these positions can happen in the base editors tested in this study suggests the STU-SR-BE strategies could work for many, if not all base editors.

Prime editing has been under constant improvement in plants. For example, prime editing becomes a relatively reliable tool for precise genome editing in monocot plants such as rice, thanks to innovative strategies employed for these improvements (Gupta et al., 2023; Jiang et al., 2022; Lin et al., 2021; Qiao et al., 2023; Tang et al., 2020; Xu et al., 2022). However, little success has been reported for prime editing in dicot plants (Lu et al., 2021). Due to low editing efficiency, prime editing could greatly benefit from surrogate reporter-based enrichment systems, as demonstrated recently (Li et al., 2020; Li et al., 2022; Xu et al., 2020a). Since the targeted DNA changes in prime editing vary from site to site and the editing is more complex than base editing, a different strategy is needed to develop authentic surrogate reporters that install the surrogate site into the reporter gene. We envision that the surrogate site can be added to the start codon-less reporter gene, as the case for base editing reporter systems. Specific prime editing guide RNA (pegRNA) needs to be designed to create a start codon, likely around or downstream the Cas9 cleavage site (or 3 bp upstream of the PAM), as that region is most amenable to prime editing (Sretenovic and Qi, 2022). This strategy in principle should work and is worth testing. Such surrogate reporter-based prime editing systems may enable reliable prime editing in dicot plants in the future.

In summary, we developed multiple Single Transcript Unit-Surrogate Reporter (STU-SR) systems for CRISPR-Cas mediated targeted mutagenesis, cytosine base editing, and adenine base editing in plants. The successful application of the STU-SR system in rice (a monocot) and *Brassica oleracea* (a dicot) demonstrates its versatility and effectiveness, indicating the potential to be applied across a broad spectrum of plant species. Continued improvement and expansion of this technology are expected to enable a wide range of genome editing applications in plants. Our successful STU-SR strategies reported here may serve as a valuable reference and inspiration for the development of other genome editing technologies to augment plant research and applications.

Experimental procedures

Vector construction

The STU-SR systems used in this study were generated based on these plasmids: pTX172 (Addgene #89259), pGEL031 (Addgene #137900), pGEL035 (Addgene #137903), pYPQ166-SpRY (Addgene #161520), and pYPQ262B-ABE8e (Addgene #161524). The rice STU-SR-SSA system was created by assembling a PCR-amplified backbone fragment into pTX172 (Zhou et al., 2022) via Gibson Assembly, removing the original *HYG* expression cassette. Subsequently, the *HY::YG* surrogate reporter gene was obtained through overlap PCR. Together with the 35S promoter and 35S terminator, it was assembled via Gibson Assembly into the vector behind the STU-Cas9 expression cassette, resulting in the construction of the STU-SR-SSA backbone vector pGEL901 (Addgene #218546). Similarly, the STU-SR-SSA_SpRY backbone vector pGEL902 (Addgene #218547) was constructed by replacing the Cas9 and *HYG* expression units on pGEL031 with the PCR-amplified SpRY from pYPQ166-SpRY and the *HY::YG* surrogate reporter gene expression unit. For the STU-SR-CBE system, the *HYG* expression cassette in pGEL035 was removed. Subsequently, the hyg^{-ATG} surrogate reporter unit was cloned at the SacI site of pGEL035 to produce the STU-SR-CBE backbone pGEL903 (Addgene #218548). The PCR-amplified SpRY fragment from pYPQ166-SpRY was then used to replace the Cas9 segment, resulting in the generation of the STU-SR-CBE_SpRY backbone vector pGEL904 (Addgene #218549). Similarly, the STU-SR-ABE_SpRY backbone vector pGEL905 (Addgene #218550) was constructed by replacing the CBE and *HYG* expression units on pGEL035 with the PCR-amplified SpRY-ABE8e from pYPQ262B-ABE8e and the hyg^{-ATG} surrogate reporter unit. The ZmUbi1 promoter and *HY::YG* surrogate reporter gene on the rice STU-SR-SSA system were replaced with the 35S promoter and *BA:AR-2A-MYB* surrogate reporter gene, resulting in the construction of the Brassica oleracea STU-SR-SSA backbone vector pGEL906 (Addgene #218551). The construction of the STU-SR system expression vector, as shown in Figures 2B and 4B, begins with the specific cleavage of the vector backbone using restriction endonucleases to generate sticky ends.

Subsequently, forward and reverse oligonucleotides are designed to incorporate flanking restriction enzyme sites, sgRNA sequences, and surrogate sites. These oligonucleotides are amplified via PCR to obtain insertion fragments containing the required sgRNA, the terminator of the Cas9 expression unit, and part of the surrogate reporter unit (including corresponding surrogate sites). The PCR products are then digested with the corresponding restriction endonucleases to create ends that are complementary to the sticky ends of the vector backbone. Ligation of these digested fragments to the vector backbone via T4 DNA ligase facilitates the assembly of the expression vector by the pairing of compatible ends. The primer sequences are listed in Supplemental Table 1. All vectors used in this study are available from Addgene.

Rice stable transformation

The Oryza sativa Japonica cultivar *Nipponbare* was used in this study. Rice stable transformation was carried out following previously published protocols (Zheng et al., 2023; Zhou et al., 2017). Binary vectors were transformed into Agrobacterium strain EHA105 using the freeze-thaw method. Rice seeds were sterilized and cultured on N6-D solid medium in the light at 32°C for 2-3 weeks. Rice calli were immersed in the Agrobacterium suspension, dried on filter paper, and co-cultured on solid medium at 25°C in the dark for 3 days. Infected calli were washed and transferred to N6D-S screening medium at 32°C for 2 weeks. Actively growing calli were moved to RE-III regenerative medium at 28°C with a 16-hour light/8-hour dark cycle for 3-4 weeks. Regenerated seedlings were transferred to rooting medium for 2-3 weeks. Transgenic rice plants were grown in a growth chamber at 28°C under a 16-hour light/8-hour dark cycle.

***Brassica oleracea* stable transformation**

The *Brassica oleracea* line “159”, previously developed in our laboratory, served as the plant material. Agrobacterium mediated transformation was carried out as previous described (Ma et al., 2019). Briefly, hypocotyls from 7- to 10-day-old seedlings were chosen as the target explants and were pre-cultured on callus initiation medium for 2 days. Subsequently, the preincubated hypocotyls were immersed in an Agrobacterium-infection buffer and cocultivated in the dark at 25°C for 48 hours. After cocultivation, the explants were transferred to medium optimized for callus and shoot induction. Once the regenerating shoots reached a height of 1-2 cm, they were carefully moved to rooting medium to obtain transgenic T0 plants. Transgenic plants were transplanted into soil.

Mutagenesis analysis

Genomic DNA was extracted from transgenic plants using the cetyltrimethylammonium bromide (CTAB) method as previous reported (Zheng et al., 2023; Zhou et al., 2022), and subsequent PCR amplification of the target gene was performed with specific primers as listed in Supplemental Table 1. To detect mutations,

both Single-Strand Conformation Polymorphism (SSCP) analysis (Zheng et al., 2016) and Sanger sequencing were employed (Zhou et al., 2019). The genotypes of T0 mutant lines were analyzed using the CRISPR-GE DSDecodeM software (Xie et al., 2017).

Author Contributions

Y.Z. proposed the project and designed the experiments. X. T., Q. R. and R.Z. constructed all the plasmids. X.T., Q.R., L.L., Q.H. and X.Z. performed the rice stable transformation. X.T., Q.R. and R.Z. analyzed the rice T0 plants. X.Y. and H.S. performed the *Brassica oleracea* stable transformation and analyzed the *Brassica oleracea* T0 plants. Y.Z., Y.Q., X.T. and H.S. analyzed the data and wrote the manuscript with input from other authors. All authors read and approved the final version of the manuscript.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (award no. 2023YFD1202900), the National Science Foundation of China (award no. 32270433 and 32101205), the Natural Science Foundation of Sichuan Province (award no. 2022NSFSC0143) to Y.Z. and X.T., Joint Science and Technology Project between Sichuan Province and Chongqing Municipality (award no. CSTC2021JSCX-CYLHX0001) to H.S. and X.T., Modern Seed Industry Project of Chongqing Municipal Science and Technology Bureau (award no. CSTB2023TIAD-KPX0025) to H.S., the National Science Foundation of China (award no. 32301248) to Q.R., the National Science Foundation of China (award no. 32072045) to X.Z. It also supported by the NSF Plant Genome Research Program (award no. IOS-2029889 and IOS-2132693) to Y.Q.

Declaration of Interests

The authors declare no competing interests.

References

- Anzalone, A.V., Koblan, L.W. and Liu, D.R. (2020) Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol* **38**, 824-844.
- Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A. and Liu, D.R. (2019) Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149-157.
- Bhargava, R., Onyango, D.O. and Stark, J.M. (2016) Regulation of Single-Strand Annealing and its Role in Genome Maintenance. *Trends Genet* **32**, 566-575.
- De Buck, S., Podevin, N., Nolf, J., Jacobs, A. and Depicker, A. (2009) The T-DNA integration pattern in Arabidopsis transformants is highly determined by the transformed target cell. *Plant J* **60**, 134-145.
- 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.
- Gupta, A., Liu, B., Chen, Q.J. and Yang, B. (2023) High-efficiency prime editing enables new strategies for broad-spectrum resistance to bacterial blight of rice. *Plant Biotechnol J* **21**, 1454-1464.
- Hassan, M.M., Zhang, Y., Yuan, G., De, K., Chen, J.G., Muchero, W., Tuskan, G.A., Qi, Y. and Yang, X. (2021) Construct design for CRISPR/Cas-based genome editing in plants. *Trends Plant Sci* **26**, 1133-1152.
- Jiang, Y., Chai, Y., Qiao, D., Wang, J., Xin, C., Sun, W., Cao, Z., Zhang, Y., Zhou, Y., Wang, X.C. and Chen, Q.J. (2022) Optimized prime editing efficiently generates glyphosate-resistant rice plants carrying homozygous TAP-IVS mutation in EPSPS. *Mol Plant* **15**, 1646-1649.
- Jupe, F., Rivkin, A.C., Michael, T.P., Zander, M., Motley, S.T., Sandoval, J.P., Slotkin, R.K., Chen, H., Castanon, R., Nery, J.R. and Ecker, J.R. (2019) The complex architecture and epigenomic impact of plant T-DNA insertions. *PLoS Genet* **15**, e1007819.
- Kim, D., Kim, D.E., Lee, G., Cho, S.I. and Kim, J.S. (2019) Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat Biotechnol* **37**, 430-435.
- Kim, H., Um, E., Cho, S.R., Jung, C., Kim, H. and Kim, J.S. (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. *Nat. Methods* **8**, 941-943.
- 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.
- Lapinaite, A., Knott, G.J., Palumbo, C.M., Lin-Shiao, E., Richter, M.F., Zhao, K.T., Beal, P.A., Liu, D.R. and Doudna, J.A. (2020) DNA capture by a CRISPR-Cas9-guided adenine base editor. *Science* **369**, 566-571.
- Lee, L.-Y. and Gelvin, S.B. (2008) T-DNA Binary Vectors and Systems. *Plant Physiol.* **146**, 325-332.
- Li, H., Li, J., Chen, J., Yan, L. and Xia, L. (2020) Precise Modifications of Both Exogenous and Endogenous Genes in Rice by Prime Editing. *Mol Plant* **13**, 671-674.
- Li, H., Zhu, Z., Li, S., Li, J., Yan, L., Zhang, C., Ma, Y. and Xia, L. (2022) Multiplex precision gene editing by a surrogate prime editor in rice. *Mol Plant* **15**, 1077-1080.
- Liao, S., Tammaro, M. and Yan, H. (2015) Enriching CRISPR-Cas9 targeted cells by co-targeting the HPRT gene. *Nucleic Acids Res* **43**, e134.
- Lin, Q., Jin, S., Zong, Y., Yu, H., Zhu, Z., Liu, G., Kou, L., Wang, Y., Qiu, J.L., Li, J. and Gao, C. (2021) High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat Biotechnol* **39**, 923-927.

- Liu, G., Yin, K., Zhang, Q., Gao, C. and Qiu, J.L. (2019) Modulating chromatin accessibility by transactivation and targeting proximal dsgRNAs enhances Cas9 editing efficiency in vivo. *Genome Biol* **20**, 145.
- Liu, S., Sretenovic, S., Fan, T., Cheng, Y., Li, G., Qi, A., Tang, X., Xu, Y., Guo, W., Zhong, Z., He, Y., Liang, Y., Han, Q., Zheng, X., Gu, X., Qi, Y. and Zhang, Y. (2022) Hypercompact CRISPR-Cas12j2 (CasPhi) enables genome editing, gene activation, and epigenome editing in plants. *Plant Commun* **3**, 100453.
- Lowder, L.G., Zhang, D., Baltes, N.J., Paul, J.W., 3rd, 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.
- Lu, Y., Tian, Y., Shen, R., Yao, Q., Zhong, D., Zhang, X. and Zhu, J.K. (2021) Precise genome modification in tomato using an improved prime editing system. *Plant Biotechnol J* **19**, 415-417.
- Ma, C., Zhu, C., Zheng, M., Liu, M., Zhang, D., Liu, B., Li, Q., Si, J., Ren, X. and Song, H. (2019) CRISPR/Cas9-mediated multiple gene editing in Brassica oleracea var. capitata using the endogenous tRNA-processing system. *Horticulture Research* **6**.
- Ming, M., Ren, Q., Pan, C., He, Y., Zhang, Y., Liu, S., Zhong, Z., Wang, J., Malzahn, A.A., Wu, J., Zheng, X., Zhang, Y. and Qi, Y. (2020) CRISPR-Cas12b enables efficient plant genome engineering. *Nat Plants* **6**, 202-208.
- Molla, K.A., Sretenovic, S., Bansal, K.C. and Qi, Y. (2021) Precise plant genome editing using base editors and prime editors. *Nat Plants* **7**, 1166-1187.
- Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., Shimatani, Z. and Kondo, A. (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* **353**.
- Paques, F. and Haber, J.E. (1999) Multiple pathways of recombination induced by double-strand breaks in Saccharomyces cerevisiae. *Microbiol Mol Biol Rev* **63**, 349-404.
- Puchta, H. (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot* **56**, 1-14.
- Qiao, D., Wang, J., Lu, M.H., Xin, C., Chai, Y., Jiang, Y., Sun, W., Cao, Z., Guo, S., Wang, X.C. and Chen, Q.J. (2023) Optimized prime editing efficiently generates heritable mutations in maize. *J Integr Plant Biol* **65**, 900-906.
- Ramakrishna, S., Cho, S.W., Kim, S., Song, M., Gopalappa, R., Kim, J.S. and Kim, H. (2014) Surrogate reporter-based enrichment of cells containing RNA-guided Cas9 nuclease-induced mutations. *Nat Commun* **5**, 3378.
- Randall, L.B., Sretenovic, S., Wu, Y., Yin, D., Zhang, T., Eck, J.V. and Qi, Y. (2021) Genome- and transcriptome-wide off-target analyses of an improved cytosine base editor. *Plant Physiol* **187**, 73-87.
- Ren, Q., Sretenovic, S., Liu, G., Zhong, Z., Wang, J., Huang, L., Tang, X., Guo, Y., Liu, L., Wu, Y., Zhou, J., Zhao, Y., Yang, H., He, Y., Liu, S., Yin, D., Mayorga, R., Zheng, X., Zhang, T., Qi, Y. and Zhang, Y. (2021a) Improved plant cytosine base editors with high editing activity, purity, and specificity. *Plant Biotechnol J* **19**, 2052-2068.
- Ren, Q., Sretenovic, S., Liu, S., Tang, X., Huang, L., He, Y., Liu, L., Guo, Y., Zhong, Z., Liu, G., Cheng, Y., Zheng, X., Pan, C., Yin, D., Zhang, Y., Li, W., Qi, L., Li, C., Qi, Y. and Zhang, Y. (2021b) PAM-less plant genome editing using a CRISPR-SpRY toolbox. *Nat Plants* **7**, 25-33.

- Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., Doudna, J.A. and Liu, D.R. (2020) Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol* **38**, 883-891.
- Roth, N., Klimesch, J., Dukowic-Schulze, S., Pacher, M., Mannuss, A. and Puchta, H. (2012) The requirement for recombination factors differs considerably between different pathways of homologous double-strand break repair in somatic plant cells. *Plant J* **72**, 781-790.
- Sretenovic, S., Green, Y., Wu, Y., Cheng, Y., Zhang, T., Van Eck, J. and Qi, Y. (2023) Genome- and transcriptome-wide off-target analyses of a high-efficiency adenine base editor in tomato. *Plant Physiol* **193**, 291-303.
- Sretenovic, S. and Qi, Y. (2022) Plant prime editing goes prime. *Nat Plants* **8**, 20-22.
- Talas, A., Simon, D.A., Kulcsar, P.I., Varga, E., Krausz, S.L. and Welker, E. (2021) BEAR reveals that increased fidelity variants can successfully reduce the mismatch tolerance of adenine but not cytosine base editors. *Nat. Commun.* **12**, 6353.
- 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., Ren, Q., Yang, L., Bao, Y., Zhong, Z., He, Y., Liu, S., Qi, C., Liu, B., Wang, Y., Sretenovic, S., Zhang, Y., Zheng, X., Zhang, T., Qi, Y. and Zhang, Y. (2019) Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. *Plant Biotechnol J* **17**, 1431-1445.
- Tang, X., Sretenovic, S., Ren, Q., Jia, X., Li, M., Fan, T., Yin, D., Xiang, S., Guo, Y., Liu, L., Zheng, X., Qi, Y. and Zhang, Y. (2020) Plant prime editors enable precise gene editing in rice cells. *Mol Plant* **13**, 667-670.
- Tang, X. and Zhang, Y. (2023) Beyond knockouts: fine-tuning regulation of gene expression in plants with CRISPR-Cas-based promoter editing. *New Phytol* **239**, 868-874.
- 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.
- Tian, Y., Zhong, D., Li, X., Shen, R., Han, H., Dai, Y., Yao, Q., Zhang, X., Deng, Q., Cao, X., Zhu, J.K. and Lu, Y. (2022) High-throughput genome editing in rice with a virus-based surrogate system. *J. Integr. Plant Biol.*
- Walton, R.T., Christie, K.A., Whittaker, M.N. and Kleinstiver, B.P. (2020) Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* **368**, 290-296.
- Weiss, T., Crisp, P.A., Rai, K.M., Song, M., Springer, N.M. and Zhang, F. (2022) Epigenetic features drastically impact CRISPR-Cas9 efficacy in plants. *Plant Physiol* **190**, 1153-1164.
- Wu, Y., He, Y., Sretenovic, S., Liu, S., Cheng, Y., Han, Y., Liu, G., Bao, Y., Fang, Q., Zheng, X., Zhou, J., Qi, Y., Zhang, Y. and Zhang, T. (2022a) CRISPR-BETS: a base-editing design tool for generating stop codons. *Plant Biotechnol J* **20**, 499-510.
- Wu, Y., Ren, Q., Zhong, Z., Liu, G., Han, Y., Bao, Y., Liu, L., Xiang, S., Liu, S., Tang, X., Zhou, J., Zheng, X., Sretenovic, S., Zhang, T., Qi, Y. and Zhang, Y. (2022b) Genome-wide analyses of PAM-relaxed Cas9 genome editors reveal substantial off-target effects by ABE8e in rice. *Plant Biotechnol J* **20**, 1670-1682.
- Xie, X., Ma, X., Zhu, Q., Zeng, D., Li, G. and Liu, Y.-G. (2017) CRISPR-GE: A Convenient Software Toolkit for CRISPR-Based Genome Editing. *Molecular Plant* **10**, 1246-1249.

- Xu, R., Li, J., Liu, X., Shan, T., Qin, R. and Wei, P. (2020a) Development of Plant Prime-Editing Systems for Precise Genome Editing. *Plant Commun* **1**, 100043.
- Xu, W., Yang, Y., Liu, Y., Kang, G., Wang, F., Li, L., Lv, X., Zhao, S., Yuan, S., Song, J., Wu, Y., Feng, F., He, X., Zhang, C., Song, W., Zhao, J. and Yang, J. (2020b) Discriminated sgRNAs-Based SurroGate System Greatly Enhances the Screening Efficiency of Plant Base-Edited Cells. *Mol Plant* **13**, 169-180.
- Xu, W., Yang, Y., Yang, B., Krueger, C.J., Xiao, Q., Zhao, S., Zhang, L., Kang, G., Wang, F., Yi, H., Ren, W., Li, L., He, X., Zhang, C., Zhang, B., Zhao, J. and Yang, J. (2022) A design optimized prime editor with expanded scope and capability in plants. *Nat. Plants* **8**, 45-52.
- Zhang, Y., Malzahn, A.A., Sretenovic, S. and Qi, Y. (2019) The emerging and uncultivated potential of CRISPR technology in plant science. *Nat Plants* **5**, 778-794.
- 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.
- 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.
- Zheng, X., Zhang, S., Liang, Y., Zhang, R., Liu, L., Qin, P., Zhang, Z., Wang, Y., Zhou, J., Tang, X. and Zhang, Y. (2023) Loss-function mutants of OsCKX gene family based on CRISPR-Cas systems revealed their diversified roles in rice. *Plant Genome* **16**, e20283.
- 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, J., Deng, K., Cheng, Y., Zhong, Z., Tian, L., Tang, X., Tang, A., Zheng, X., Zhang, T., Qi, Y. and Zhang, Y. (2017) CRISPR-Cas9 based genome editing reveals new insights into MicroRNA function and regulation in rice. *Front Plant Sci* **8**, 1598.
- Zhou, J., Liu, G., Zhao, Y., Zhang, R., Tang, X., Li, L., Jia, X., Guo, Y., Wu, Y., Han, Y., Bao, Y., He, Y., Han, Q., Yang, H., Zheng, X., Qi, Y., Zhang, T. and Zhang, Y. (2023) An efficient CRISPR-Cas12a promoter editing system for crop improvement. *Nat Plants* **9**, 588-604.
- 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. (2019) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. *Plant Cell Rep* **38**, 475-485.
- Zhou, J., Zhang, R., Jia, X., Tang, X., Guo, Y., Yang, H., Zheng, X., Qian, Q., Qi, Y. and Zhang, Y. (2022) CRISPR-Cas9 mediated OsMIR168a knockout reveals its pleiotropy in rice. *Plant Biotechnol J* **20**, 310-322.

Figure legends

Figure 1. Comparison of three CRISPR-Cas9 editing systems. (a) In a conventional CRISPR-Cas9 expression system, the plant selection marker, Cas9 and sgRNAs are driven by separate promoters. There is little correlation between edited and regenerated events. (b) In a conventional surrogate reporter system, CRISPR-Cas9 mediated genome editing events can be enriched because Cas9 expression and activity is selected during plant regeneration. However, there is only indirect correlation between edited and regenerated events. (c) In the single transcript unit-surrogate reporter (STU-SR) system, both Cas9 and sgRNA are expressed under a single Pol II promoter. Editing at the reporter gene and at the endogenous target gene is done by the same sgRNA. There is direct correlation between edited and regenerated events. Hence, STU-SR systems are stringent and powerful in enriching genome editing events.

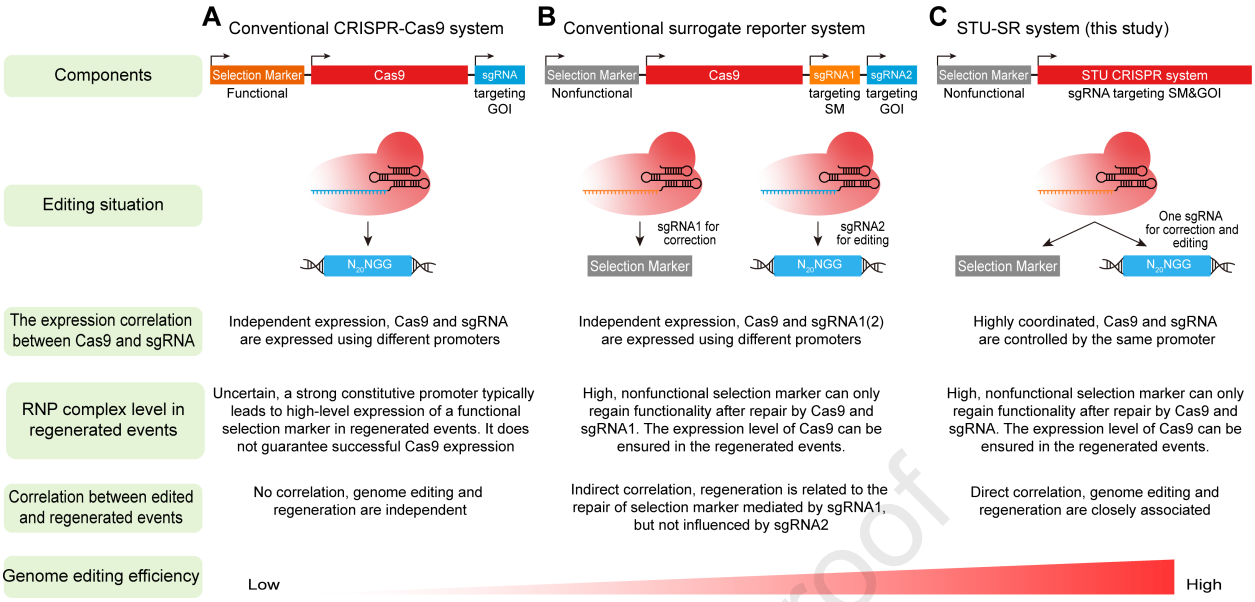
Figure 2. Efficient gene editing in rice by STU-SR-SSA system. (a) Enrichment diagram depicting the implementation of the STU-SR-SSA system for targeted gene editing in rice. (b) Schematic illustration of the stepwise cloning strategy employed for constructing the rice STU-SR-SSA vector for targeted mutagenesis. (c) Summary of genome editing by STU and the STU-SR-SSA systems in rice T0 lines. Representative Sanger sequencing chromatograms of editing at the *HY::YG* reporter (d) and the endogenous gene loci (e) in obtained plants using the STU-SR-SSA system. (f) The phenotypes of the wild type (WT; left) and a *OsPDS* mutant (right) generated by STU-SR-SSA system. (g) The phenotypes of the WT (left) and a *OsDEP1* mutant (right) generated by STU-SR-SSA system.

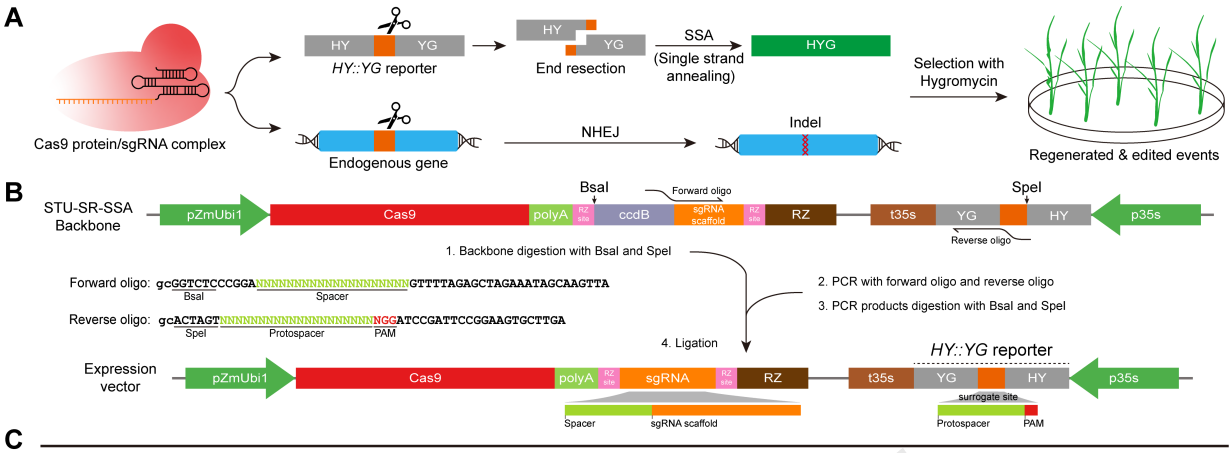
Figure 3. Enrichment of gene editing events achieved regardless of T-DNA copy number. (a) Schematic illustration depicting the editing outcomes of the *HY::YG* reporter(s) in single or multiple T-DNA-integrated plant events. (b) PCR detection results of the *HY::YG* reporter in the 8 *OsPDS*-sgRNA04 T0 plants. (c) Representative Sanger sequencing chromatograms and genotypes of the *OsPDS*-sgRNA04 T0 lines with editing at the *HY::YG* reporter and the endogenous gene loci by STU-SR-SSA system.

Figure 4. Efficient base editing in rice by STU-SR system. (a) Enrichment diagram depicting the implementation of the STU-SR-BE system for base editing in rice. (b) Schematic representation of the stepwise cloning strategy employed for constructing the STU-SR-BE vectors for base editing in rice. When designing the vectors, ensure that the newly generated ATG start codon resulting from base editing aligns with the hyg^{-ATG} CDS within the same open reading frame. Otherwise, adjust the sequence between the surrogate site and the hyg^{-ATG} CDS to ensure they are in the same ORF. If the 5' end of the sgRNA lacks the corresponding sequence, such as 'ACG' or 'CAC' in CBE or 'ATA' in ABE, this can be addressed by adjusting the protospacer sequence at the PAM distal end, allowing the presence of a 1-2 bp mismatch to form the corresponding sequence. (c) Summary of base editing by STU and the STU-SR-BE systems in

rice T0 lines. (d) Examples of base edited rice T0 lines by the STU-SR-BE systems at the hygromycin site and representative endogenous gene loci. (e) Examples of base edited rice T0 lines by the STU-SR-BE systems working with SpRY variant at the hygromycin site and representative endogenous gene loci.

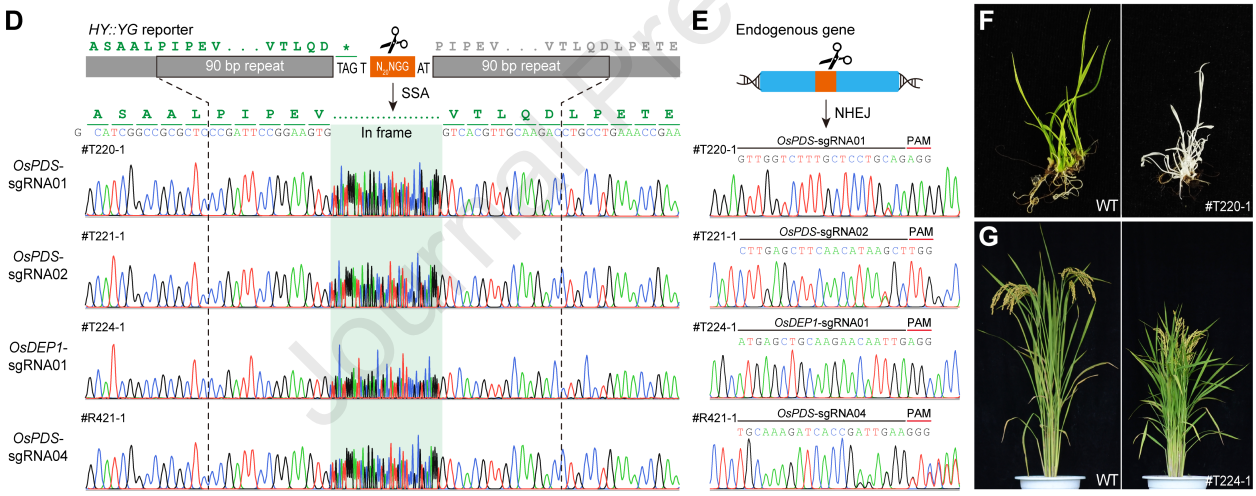
Figure 5. Efficient gene editing in *Brassica oleracea* by STU-SR system. (a) Schematic representation of the STU-SR-SSA vector and its implementation for achieving enrichment of gene editing events in *Brassica oleracea*. (b) Comparison of editing efficiencies at three endogenous gene loci using STU and STU-SR-SSA systems. (c) Representative images of *BoPDS* gene editing using STU and STU-SR-SSA systems. (d) Representative images of *BoBIK* gene editing using STU and STU-SR-SSA systems. (e) Representative Sanger sequencing chromatograms of T0 lines with editing at the BA::AR reporter locus. (f) Genotypes of T0 lines with gene editing by the STU-SR-SSA system in *Brassica oleracea*.s

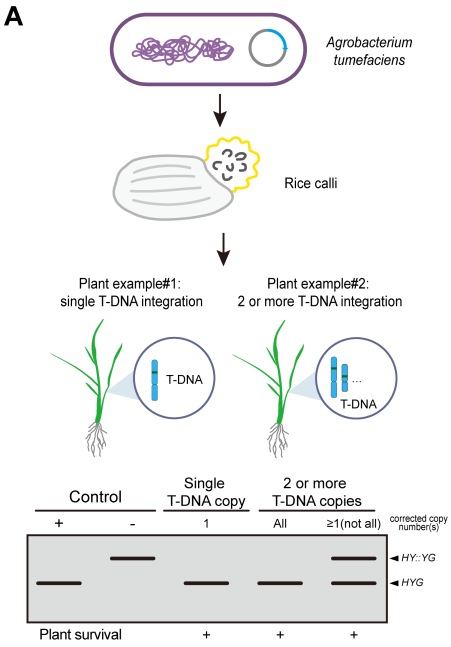


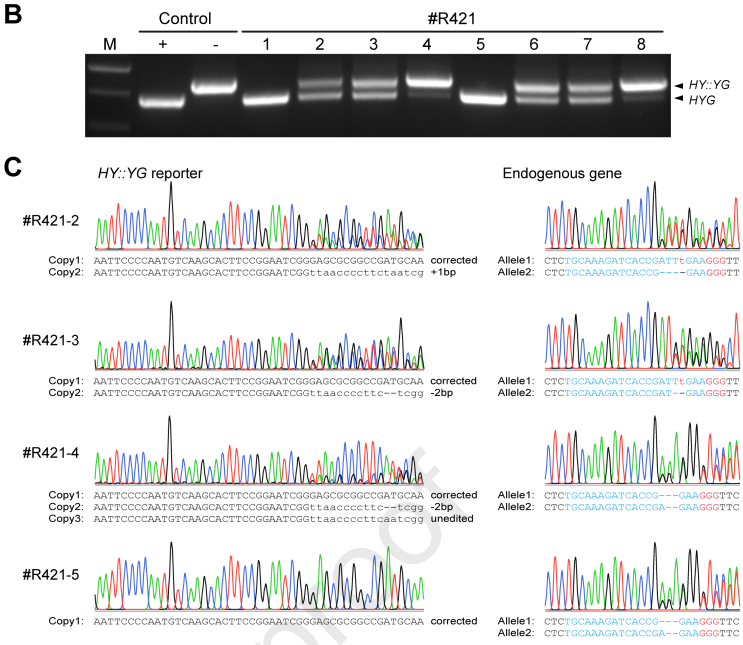


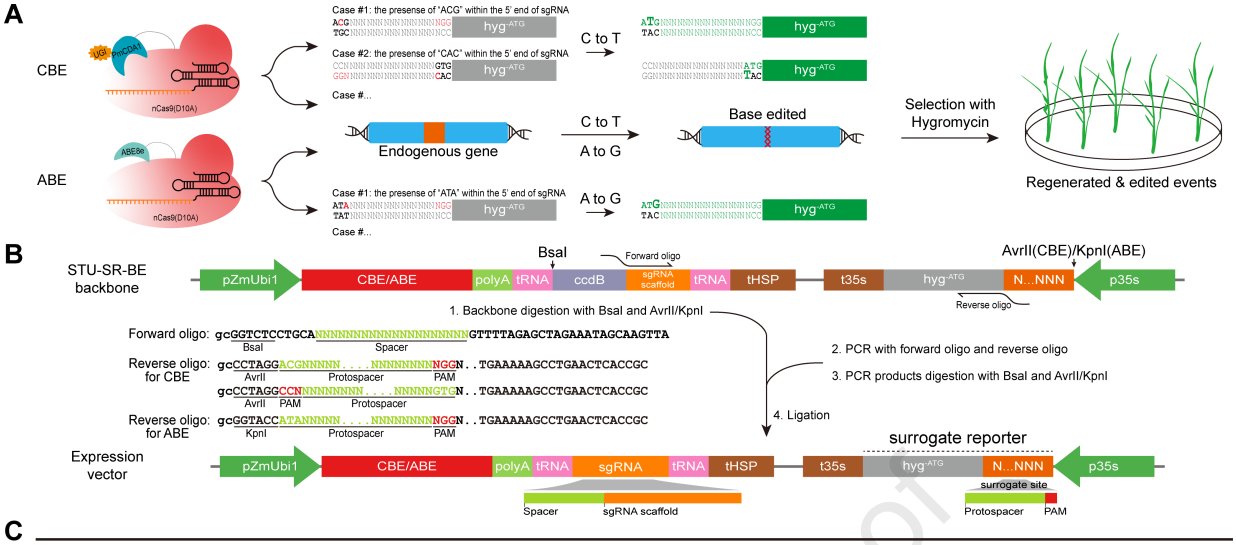
C

CRISPR-Cas9	Targeted gene	Construct	sgRNA	sgRNA+PAM	Tested T0 lines	*HYG* corrected T0 lines	Edit number	Edit ratio	Biallelic edit number	Biallelic edit ratio
STU	*OsPDS*	#T198	*OsPDS*-sgRNA01	GTTGGTCTTTGCTCCTGCAGAGG	16	-	9	56.3%	7	43.8%
STU-SR-SSA		#T220			20	20	20	100%	16	80%
STU	*OsPDS*	#T204	*OsPDS*-sgRNA02	CTTGAGCTTCAACATAAGCTTGG	22	-	12	77.3%	14	63.6%
STU-SR-SSA		#T221			6	6	6	100%	6	100%
STU	*OsDEP1*	#T202	*OsDEP1*-sgRNA01	ATGAGCTGCAAGAACAATTGAGG	18	-	14	77.8%	13	72.2%
STU-SR-SSA		#T224			18	18	18	100%	18	100%
STU_SpRY	*OsPDS*	#R420	*OsPDS*-sgRNA04	TGCAAAGATCACCGATTGAAGGG	20	-	15	75%	13	65%
STU-SR-SSA_SpRY		#R421			16	16	16	100%	16	100%









C

BE systems	Targeted gene	Construct	sgRNA	sgRNA+PAM	Tested T0 lines	*HYG* corrected T0 lines	Base edited number	Base edited ratio
STU-CBE	*OsPDS*	#R204	*OsPDS*-sgRNA02	CTTGAGCTTCAACATAAGCTTGG	19	-	1	5.3%
STU-SR-CBE		#R198			18	18	2	11.1%
STU-CBE	*OsPDS*	#R205	*OsPDS*-sgRNA07	ACTCCAGCATTGAACGGTTTGG	19	-	7	38.6%
STU-SR-CBE		#R199			18	18	16	88.9 %
STU-CBE	*OsCDC48*	#R206	*OsCDC48*-sgRNA01	ACCAGCCAGCGTCTGGCGCCGG	21	-	11	52.4%
STU-SR-CBE		#R200			20	20	16	80%
STU-CBE	*OsDEP1*	#R208	*OsDEP1*-sgRNA03	TACGAGGCGGATCGAGCTAGCGG	22	-	21	95.5%
STU-SR-CBE		#R202			23	23	23	100%
STU-CBE_SpRY	*OsPDS*	#R433	*OsPDS*-sgRNA04	TGCAAAGATCACCGATTGAAGGG	21	-	1	4.8%
STU-SR-CBE_SpRY		#R434			21	21	2	9.5%
STU-ABE_SpRY	*OsMPK6*	#R427	*OsMPK6*-sgRNA01	ACATACTCAGTCATAAAATCGGT	22	-	1	4.5%
STU-SR-ABE_SpRY		#R428			22	22	3	13.6%