SCIENCE CHINA
Life Sciences

•COVER ARTICLE•

November 2020 Vol.63 No.11: 1619–1630
https://doi.org/10.1007/s11427-020-1722-0

Shortening the sgRNA-DNA interface enables SpCas9 and eSpCas9(1.1) to nick the target DNA strand

Rong Fan[1,2,3†], Zhuangzhuang Chai[2,3†], Sinian Xing[2], Kunling Chen[2], Fengti Qiu[2,4], Tuanyao Chai[4], Jin-Long Qiu[5], Zhengbin Zhang[1*], Huawei Zhang[2*#] & Caixia Gao[2,3*]

[1]*Key Laboratory of Agricultural Water Resources, Hebei Laboratory of Agricultural Water-Saving, Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang 050022, China;*
[2]*State Key Laboratory of Plant Cell and Chromosome Engineering, Center for Genome Editing, Institute of Genetics and Developmental Biology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing 100101, China;*
[3]*College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;*
[4]*College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China;*
[5]*State Key Laboratory of Plant Genomics, Institute of Microbiology, Innovation Academy for Seed Design, Chinese Academy of Sciences, Beijing 100101, China*

Received April 7, 2020; accepted May 1, 2020; published online June 24, 2020

The length of the sgRNA-DNA complementary sequence is a key factor influencing the cleavage activity of *Streptococcus pyogenes* Cas9 (SpCas9) and its variants. The detailed mechanism remains unknown. Here, based on *in vitro* cleavage assays and base editing analysis, we demonstrate that reducing the length of this complementary region can confer nickase activity on SpCas9 and eSpCas9(1.1). We also show that these nicks are made on the target DNA strand. These properties encouraged us to develop a dual-functional system that simultaneously carries out double-strand DNA cleavage and C-to-T base conversions at separate targets. This system provides a novel tool for achieving trait stacking in plants.

SpCas9, eSpCas9(1.1), truncated spacer, DSB, nickase, co-editing

Citation: Fan, R., Chai, Z., Xing, S., Chen, K., Qiu, F., Chai, T., Qiu, J.L., Zhang, Z., Zhang, H., and Gao, C. (2020). Shortening the sgRNA-DNA interface enables SpCas9 and eSpCas9(1.1) to nick the target DNA strand. Sci China Life Sci 63, 1619–1630. https://doi.org/10.1007/s11427-020-1722-0

INTRODUCTION

Streptococcus pyogenes Cas9 (SpCas9) is a sequence specific nuclease that interacts with specific DNA targets via sgRNAs (Jiang and Doudna, 2017). The sequence of nucleotides at the 5′ end of the sgRNA (the spacer sequence) base-pairs with the target sequence upstream of the PAM (the protospacer adjacent motif), therefore determining the specificity of SpCas9 (Jiang and Doudna, 2017; Puchta, 2018). The spacer sequence in SpCas9 sgRNAs usually consists of 20 nucleotides (nt) ("full size sgRNAs") (Doudna and Charpentier, 2014). Changing the length of this segment is an effective way to reduce the off-target mutations (Cho et al., 2014; Fu et al., 2014). Compared with full-length sgRNAs, sgRNAs bearing spacers of 17 or 18 nucleotides (nt) significantly reduce off-target activity while maintaining comparable on-target activity (Fu et al., 2014). However, sgRNAs, of which spacer is less than 17 nt, can rarely enable Cas9 to induce insertion and deletion mutations (indels) (Fu et al., 2014). To enhance the specificity of SpCas9, a series of variants have been developed, including eSpCas9, SpCas9-HF, evoCas9, HypaCas9, and Sniper-Cas9, etc. (Casini et al.,

†Contributed equally to this work
*Corresponding authors (Zhengbin Zhang, email: zzb@sjziam.ac.cn; Huawei Zhang, email: hwzhang@genetics.ac.cn; Caixia Gao, email: cxgao@genetics.ac.cn)
#Present address: Institute of Advanced Agricultural Science, Peking University, Weifang 261325, China (email: huawei.zhang@pku-iaas.edu.cn)

© Science China Press and Springer-Verlag GmbH Germany, part of Springer Nature 2020

2018; Chen et al., 2017; Kleinstiver et al., 2016; Lee et al., 2018; Slaymaker et al., 2016). Both increasing and reducing the length of the sgRNA-DNA interface dramatically reduced the editing efficiency of four out of five Cas9 variants, the exception being Sniper-Cas9 (Lee et al., 2018). However, the basis of the effect is not understood. Recently, Fu et al. observed that sgRNAs with extensive mismatches with their targets were able to guide SpCas9 to nick double-strand DNA (Fu et al., 2019). Similarly, Szczelkun et al. described a truncated sgRNA (Δ7 nt in the complementary region) in conjunction with *Streptococcus thermophiles* Cas9 (StCas9) led to the accumulation of nicked molecules (Szczelkun et al., 2014). These observations suggested that truncated/elongated spacer-derived segments impose variable extents of influence on the HNH and RuvC cleavage domains of the nucleases, such that under some conditions they nick the target DNA, rather than cutting it. Here, we attempted to test this hypothesis.

RESULTS

Truncating the sgRNA-DNA interface induces eSpCas9(1.1) to nick the target DNA *in vitro*

We examined the influence of sgRNA-DNA interface length on the nuclease activities of SpCas9 and its variant eSpCas9(1.1). We initially investigated their activities *in vitro*. To generate sgRNAs with different lengths of complementary sequence, we employed a pre-sgRNA molecule, HH-sgRNA-HDV, in which a hammerhead ribozyme (HH) and hepatitis delta virus ribozyme (HDV) are fused to the 5′ and 3′ ends of the sgRNA, respectively (Gao and Zhao, 2014). The HH and HDV can be removed accurately by self-catalyzed cleavage. We placed the HH-sgRNA-HDV fusion under the control of the T7 promotor, and transcribed it in a cell free system using T7 RNA polymerase. In this way we generated a series of sgRNAs consisting of 17–24 complementary nucleotides (hereafter referred to as $sgRNA_{17}$–$sgRNA_{24}$) to target a well-studied site on the BFP (blue fluorescent protein) reporter gene (Zong et al., 2017). We expressed Cas9 protein and its variants in *Escherichia coli* and purified it (Figure S1 in Supporting Information) (Liang et al., 2017). We then tested whether the length of the sgRNA-DNA complementary sequence influenced the efficiency of double-strand breaks using plasmid pUbi-BFPm linearized by *Eco*R V as substrate. We found that the linearized plasmid was totally cleaved by SpCas9 when it was combined with $sgRNA_{20}$ (Figure 1A). Using a longer sgRNA-DNA did not affect the double-strand DNA cleavage, but shorter sgRNAs reduced cleavage (Figure 1A). With $sgRNA_{20}$–$sgRNA_{23}$, the cleavage activity of eSpCas9(1.1) was only a little lower than that of SpCas9 (Figure 1A and B). However, with $sgRNA_{24}$, more than half of the substrate molecules remained uncleaved (Figure 1B). When the sgRNA-DNA interface was <20 nt, double strand DNA cleavage by eSpCas9(1.1) decreased dramatically (Figure 1B) and with $sgRNA_{17}$ and $sgRNA_{18}$, it was almost undetectable (Figure 1B). Clearly, eSpCas9(1.1) is more sensitive to reduction of the length of the sgRNA-DNA interface than SpCas9.

We performed additional digestion assays *in vitro* using the supercoiled plasmid pUbi-BFPm as substrate, which yielded similar results. Due to conformational changes, the nicked circular plasmids move more slowly than the linearized plasmids in DNA-agarose gel, while supercoiled plasmids move faster. In this assay, eSpCas9(1.1), but not SpCas9, produced nicked plasmid (Figure 1C and D). Very weak nicked bands were detected when eSpCas9(1.1) was coupled with $sgRNA_{20}$–$sgRNA_{24}$ (Figure 1D), but they increased when it was coupled with truncated sgRNAs (Figure 1D). With $sgRNA_{17}$, most of the supercoiled plasmids were nicked (Figure 1D). The data in Figure 1 indicate that eSpCas9(1.1) cuts the two strands of DNA with unequal efficiency when combined with truncated sgRNAs, so that in many cases only one strand is nicked while the other remains intact.

eSpCas9(1.1) nicks the target DNA strand *in vivo* when combined with truncated sgRNA-DNA interface

We have showed that reducing the length of the sgRNA-DNA interface leads eSpCas9(1.1) to nick the target DNA. However, both the linearized and supercoiled plasmid digestion assays were carried out *in vitro*. To examine the situation *in vivo*, a novel strategy was developed. Recently, a highly efficient base editor has been developed based on the observation that catalytically inactivated Cas9 (dCas9), which lacks all cleavage activity, is not able to trigger base editing efficiently (Zong et al., 2017). Deamination causes C-to-U conversions on the non-target strand, and nicking of the target strand by the SpCas9 (D10A) nickase stimulates the synthesis of a new DNA strand that base-pairs with the edited strand to replace the original one, thus facilitating the desired C:G-to-T:A base conversions (Komor et al., 2016). Hence if eSpCas9(1.1) nicked the target DNA strand, it should behave as a nickase like nCas9(D10A) (not like dCas9) and promote efficient base editing when fused with a cytosine deaminase and uracil glycosylase inhibitor (UGI). We therefore made a series of constructs, including APOBEC3A-eSpCas9(1.1)-UGI, APOBEC3A-SpCas9-UGI, and APOBEC3A-dSpCas9-UGI (Figure 2A) and chose the well-studied site in BFP-to-GFP alteration as the reporter system for cytosine deamination (Figure S2 in Supporting Information) (Zong et al., 2018; Zong et al., 2017). It has been shown that $sgRNA_{20}$ can guide the base editors PBE (APOBEC1-nCas9(D10A)-UGI) and A3A-PBE (APOBE-

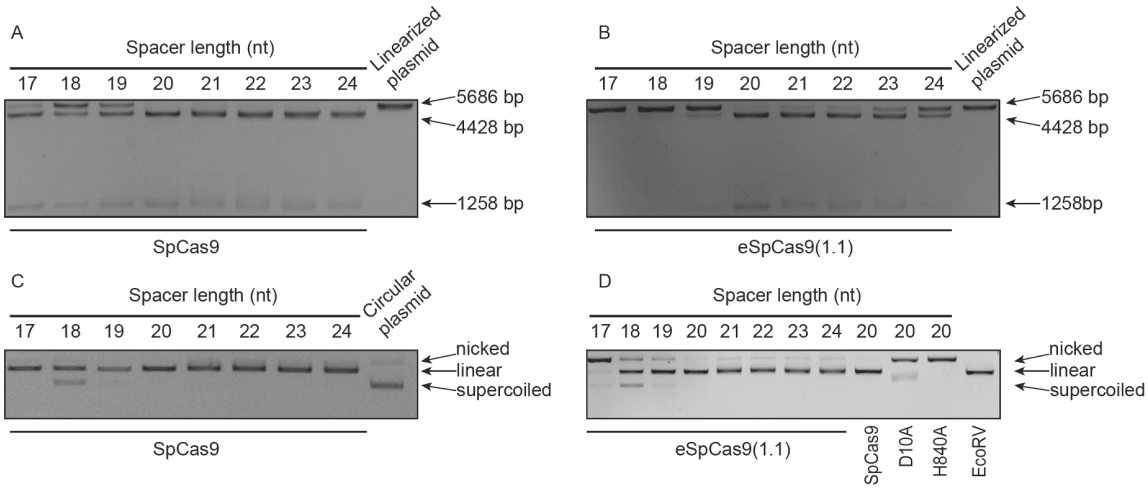


Figure 1 Truncated sgRNA-DNA interfaces lead eSpCas9(1.1) to nick the double-strand DNA *in vitro*. A, Assays of cleavage of linearized pUbi-BFPm plasmid by SpCas9 in combination with sgRNAs of different lengths. B, Assays of cleavage of linearized pUbi-BFPm plasmid by eSpCas9(1.1) in combination with sgRNAs of different lengths. A and B, The size of linearized pUbi-BFPm plasmid is 5,686 bp. If cleaved by SpCas9 or eSpCas9(1.1) in combination with sgRNAs of different lengths at the target site, it would be divided into a 4,428 bp fragment and a 1,258 bp fragment. C, Assays of cleavage of circular pUbi-BFPm plasmid by SpCas9 in combination with sgRNAs of different lengths. D, Assays of cleavage circular pUbi-BFPm plasmid by eSpCas9 (1.1) in combination with sgRNAs of different lengths.

C3A-nCas9(D10A)-UGI) to convert C4 to T4 thus achieving BFP-to-GFP conversion (Zong et al., 2018; Zong et al., 2017). To generate $sgRNA_{17}$–$sgRNA_{20}$ *in vivo*, we used the OsU3 promoter to drive the expression of a tRNA-sgRNA cassette in which the tRNA can be accurately removed from the tRNA-sgRNA precursor by endogenous RNaseZ (Xie et al., 2015). The three base editors in Figure 2 (APOBEC3A-eSpCas9(1.1)-UGI, APOBEC3A-SpCas9-UGI, APOBEC3A-dSpCas9-UGI) were separately transfected into rice protoplasts with $sgRNA_{17}$–$sgRNA_{20}$ and pUbi-BFPm. PBE and A3A-PBE were separately transfected into rice protoplasts with $sgRNA_{20}$ and pUbi-BFPm. Flow cytometry showed that APOBEC3A-dSpCas9-UGI and APOBEC3A-SpCas9-UGI yielded virtually no GFP-expressing cells, whereas APOBEC3A-eSpCas9(1.1)-UGI yielded 0.1%, 0.9%, 15.8%, and 0.3% of GFP-expressing cells, respectively, with $sgRNA_{17}$–$sgRNA_{20}$ (Figure 2B). These results indicate that eSpCas9(1.1) may nick the target-strand of the BFP target when truncated sgRNAs are used, and that it does so relatively efficiently when combined with $sgRNA_{19}$. We note that the fact that it does not induce a high level of BFP-to-GFP conversions with $sgRNA_{17}$–$sgRNA_{18}$ is not consistent with the results of the cleavage assay *in vitro*. It is possible that the truncated sgRNAs might narrow the deamination window of the target so that C4 is no longer located in the most effective position for editing.

Altering the length of the sgRNA-DNA interface leads APOBEC3A-eSpCas9(1.1)-UGI to act as a dual-functional system

The above observations show that when conjugated with $sgRNA_{19}$, eSpCas9(1.1) may nick the target DNA strand like nCas9(D10A), so that APOBEC3A-eSpCas9(1.1)-UGI can perform base editing in this situation. Since eSpCas9(1.1) yields indels as efficiently as wild-type SpCas9 (Zhang et al., 2017), APOBEC3A-eSpCas9(1.1)-UGI should also be able to produce indels efficiently and so act as a dual-functional system achieving double strand DNA cleavage and base editing on separate targets in a single procedure (Figure 3).

To examine this possibility, we chose targets in *OsCDC48* and *OsNRT1.1B*. We first tested whether truncated sgRNAs reduced the number of DSBs caused by eSpCas9(1.1). To this end, we cloned 663 bp and 545 bp genomic fragments harboring these targets into pEASY-blunt vector, and prepared sgRNAs with 15–20 nt complementary nucleotides to these targets ($sgRNA_{15}$–$sgRNA_{20}$). We then linearized pEASY-blunt-*OsCDC48* and pEASY-blunt-*OsNRT1.1B* with *Bsa* I. To test their activities on these linearized plasmids, we incubated these plasmids with either SpCas9 or eSpCas9(1.1) along with $sgRNA_{15}$–$sgRNA_{20}$. We found that the double-strand DNA cleavage activity of SpCas9 on both targets decreased in parallel with truncation of the sgRNAs (Figure 4A and B); with $sgRNA_{15}$ and $sgRNA_{16}$, they failed altogether to induce any double-strand DNA breaks on either targets, while with $sgRNA_{17}$, they produced very few breaks in the *OsNRT1.1B* target (Figure 4B). In contrast, eSpCas9(1.1) did not create any double-strand breaks on either targets with any of the truncated $sgRNA_{15–19}$ (Figure 4A and B). This finding is consistent with our observation that eSpCas9(1.1) was less tolerant of sgRNAs truncation than SpCas9 in terms of making double-strand breaks in *BFP* (Figure 1).

We showed earlier that with $sgRNA_{17}$ and $sgRNA_{19}$, SpCas9 was only capable of partially cleaving linearized

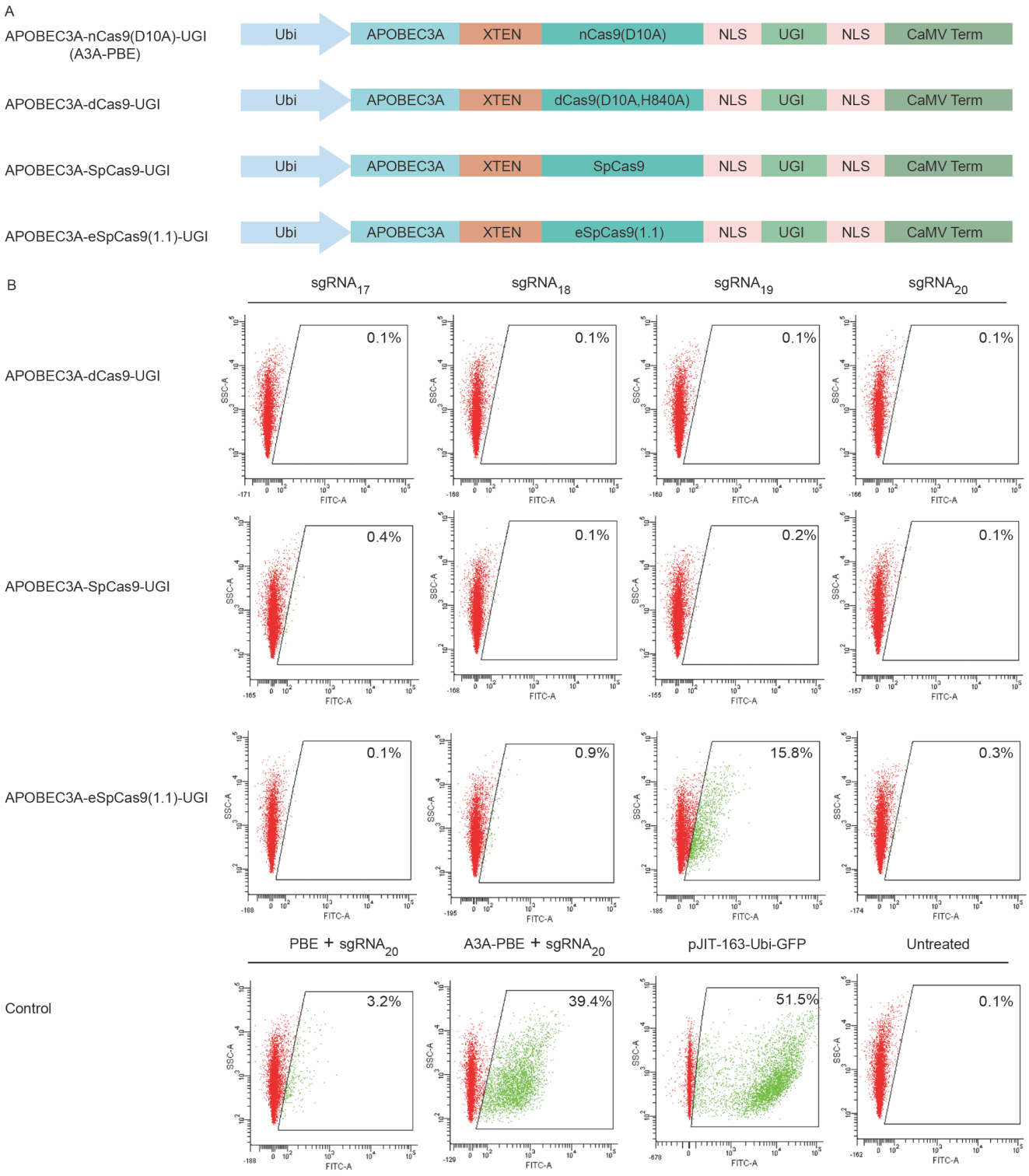


Figure 2 Truncated sgRNA-DNA interface leads eSpCas9(1.1) to nick the target DNA strand *in vivo*. A, Schematic of the constructs used in these experiments. B, Frequencies of C-to-T conversions in *BFP* target were examined in rice protoplasts by flow cytometry. C-to-T conversions were introduced by APOBEC3A-dCas9-UGI, APOBEC3A-SpCas9-UGI, and APOBEC3A-eSpCas9(1.1)-UGI in combination with sgRNAs of different lengths. PBE, A3A-PBE, SpCas9 along with $sgRNA_{20}$, and an untreated protoplast sample served as controls.

pUbi-BFPm *in vitro*, whereas it completely cleaved super-coiled pUbi-BFPm (Figure 1A and C); on the other hand, with $sgRNA_{18}$ and $sgRNA_{19}$, eSpCas9(1.1) cleaved super-coiled pUbi-BFPm efficiently but not the linearized plasmid (Figure 1B and D). Evidently, DNA topology dramatically influences the nuclease activities of SpCas9 and

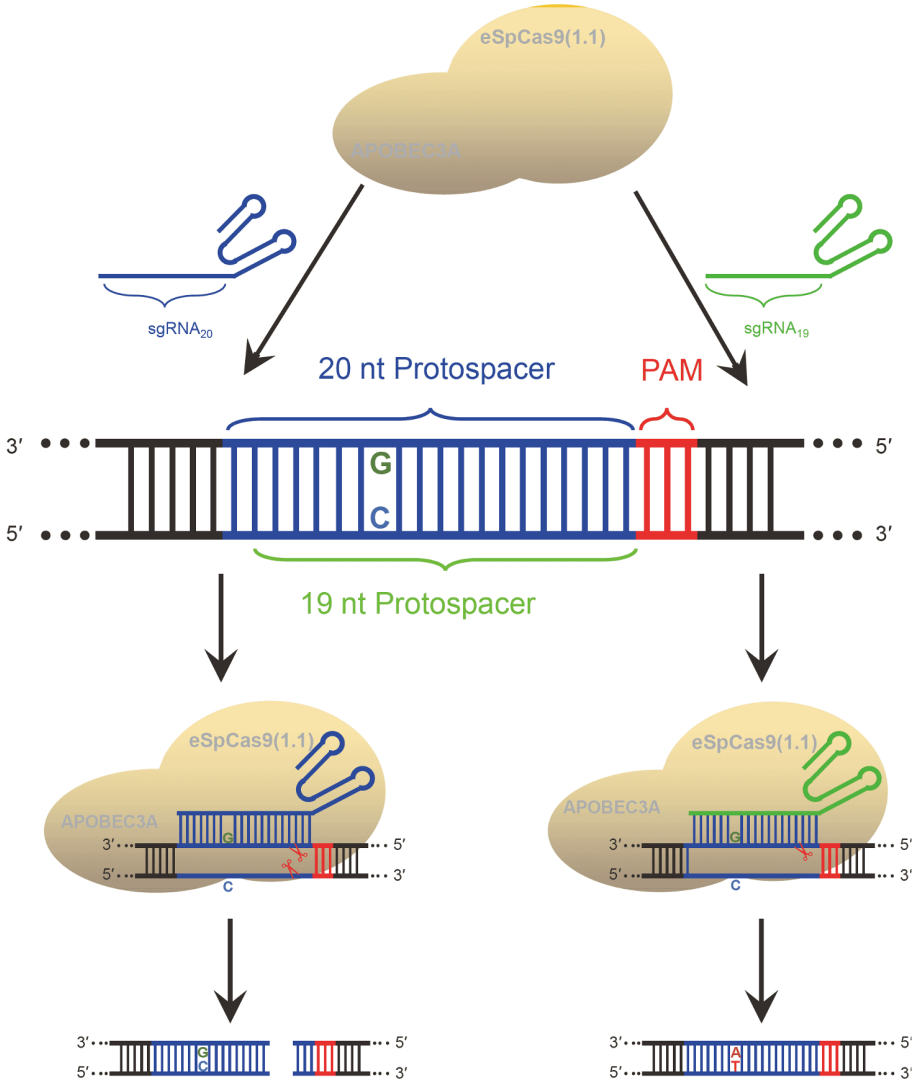


Figure 3 Mode of the APOBEC3A-eSpCas9(1.1)-UGI dual-functional system. The system can simultaneously perform DNA base editing with an $sgRNA_{19}$ and simultaneously produce indels by double-strand DNA cleavage with an $sgRNA_{20}$.

eSpCas9(1.1). Since the structure of genomic DNA in chromosomes is much more complex than that of plasmid DNA, it was unclear whether SpCas9 and eSpCas9(1.1) could nick chromosomal DNA and also whether in combination with APOBEC3A they could perform base editing on chromosomal DNA *in vivo*. We therefore used APOBEC3A-eSpCas9(1.1)-UGI and APOBEC3A-SpCas9-UGI along with $sgRNA_{15}$–$sgRNA_{20}$ to target *OsCDC48* and *OsNRT1.1B* in rice protoplasts with SpCas9, eSpCas9(1.1), PBE, and A3A-PBE were used as controls. Amplicon deep sequencing was used to identify indels and base editing changes in the endogenous targets. Both APOBEC3A-eSpCas9(1.1)-UGI and APOBEC3A-SpCas9-UGI were found to induce C-to-T conversions (Figure 4C and D). APOBEC3A-eSpCas9(1.1)-UGI was most efficient with $sgRNA_{19}$, while APOBEC3A-SpCas9-UGI was most efficient with $sgRNA_{17}$ (Figure 4C and D). Evidently, both eSpCas9(1.1) and SpCas9 could nick the target strand, but SpCas9 needed shorter sgRNAs than eSpCas9(1.1). APOBEC3A-eSpCas9(1.1)-UGI/$sgRNA_{20}$ produced as many indels as SpCas9/$sgRNA_{20}$ on *OsNRT1.1B*, but many fewer indels on *OsCDC48* (Figure 4E and F).

To extend the generality of the APOBEC3A-eSpCas9(1.1)-UGI system, we used APOBEC3A-eSpCas9(1.1)-UGI and APOBEC3A-SpCas9-UGI with $sgRNA_{18–20}$ to target two additional rice gene targets, *OsACC* and *OsSPL14*. Consistent with the above assays, APOBEC3A-eSpCas9(1.1)-UGI gave rise to quite efficient base editing with the truncated sgRNAs (Figure 5A and B) and produced fewer indels than SpCas9/$sgRNA_{20}$ with $sgRNA_{20}$ (Figure 5A and B). We tested the efficiency of base editing and indel production on five additional endogenous targets further (Figure S3 in Supporting Information). With $sgRNA_{20}$, APOBEC3A-eSpCas9(1.1)-UGI introduced fewer indels than SpCas9 at all five targets. Also its base editing efficiency with $sgRNA_{19}$ was lower than that of A3A-PBE/$sgRNA_{20}$.

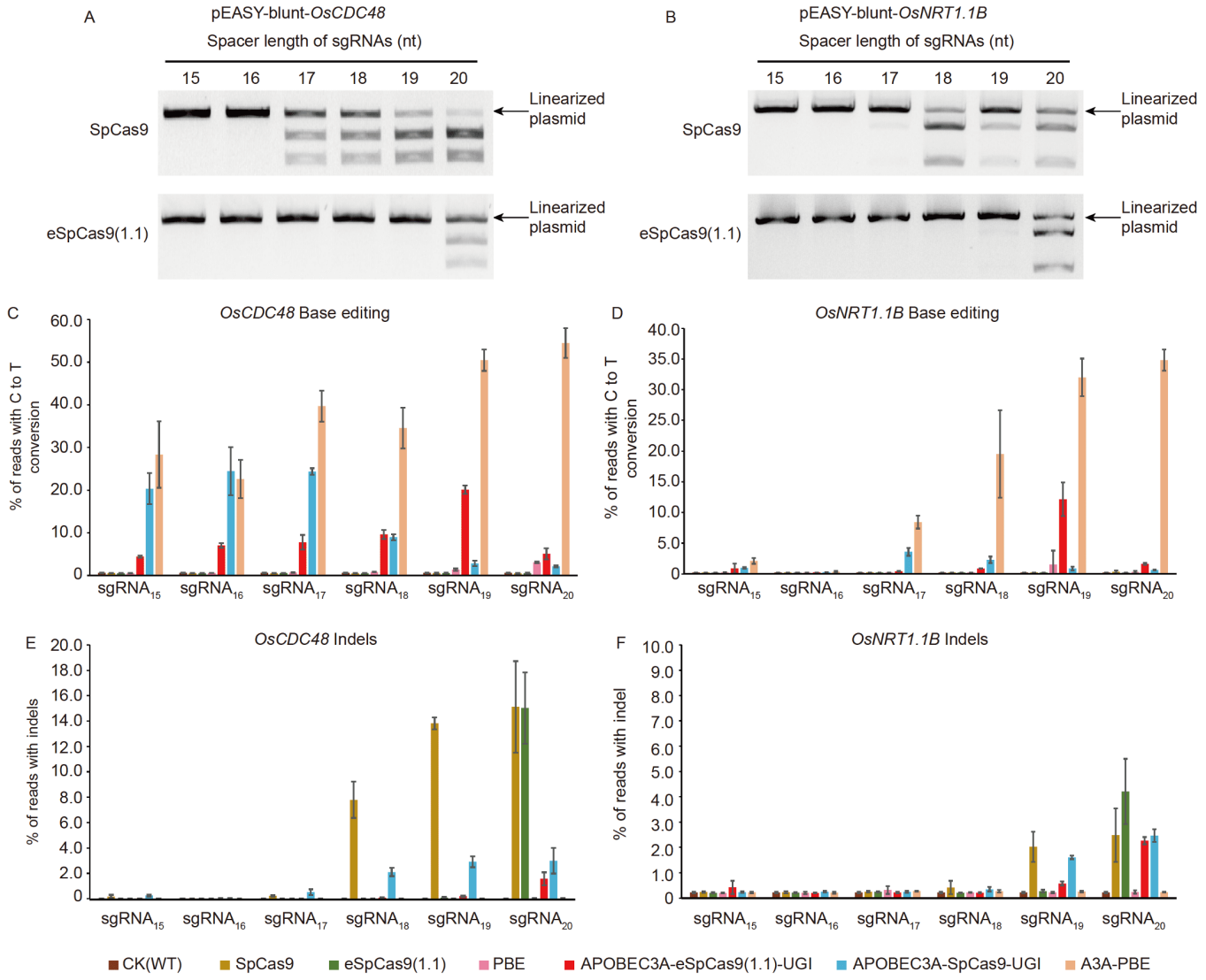


Figure 4 APOBEC3A-eSpCas9(1.1)-UGI generates base edits/indels in combination with different lengths of sgRNAs. A, Assays of cleavage of *Bsa* I-linearized pEASY-blunt-*OsCDC48* by SpCas9 and eSpCas9(1.1) in combination with different lengths of sgRNAs. B, Assays of cleavage of *Bsa* I-linearized pEASY-blunt-*OsNRT1.1B* by SpCas9 and eSpCas9(1.1) in combination with different lengths of sgRNAs. The gels show double-strand DNA cleavage assay by SpCas9 and eSpCas9(1.1) in combination with different lengths of sgRNAs. C, Frequencies of C-to-T conversions in reads of the *OsCDC48* target region. D, Frequencies of C-to-T conversions in reads of the *OsNRT1.1B* target region. E, Frequencies of indels in reads of the *OsCDC48* target region. F, Frequencies of indels in reads of the *OsNRT1.1B* target region. C–F, SpCas9, eSpCas9(1.1), PBE, pAPOBEC3A-eSpCas9(1.1)-UGI, pAPOBEC3A-SpCas9-UGI, and A3A-PBE were transformed into rice protoplasts along with tRNA-sgRNA$_{15–20}$ to target *OsCDC48* and *OsNRT1.1B*. An untreated protoplasts sample served as control (n=3 biologically independent experiments). All values are means±SD.

This result suggested that the fusion of eSpCas9(1.1) with APOBEC3A and UGI might have reduced both double-strand DNA cleavage activity and C-to-T base editing efficiency of eSpCas9(1.1) at some targets. However, C-to-T base editing efficiency was more efficient than that of PBE/sgRNA$_{20}$ at eight out of the ten targets used above, indicating that APOBEC3A-eSpCas9(1.1)-UGI/sgRNA$_{19}$ was an efficient base editor.

We also examined base editing windows and found that APOBEC3A-SpCas9-UGI could perform base editing with sgRNA$_{18}$ and even shorter sgRNAs (Figure 6), suggesting that shortening the sgRNA-DNA interface might also enable wild-type SpCas9 to nick genomic DNA on the target DNA strand. On most targets, the editing windows of APOBEC3A-eSpCas9(1.1)-UGI/sgRNA$_{19}$ were similar to or narrower than that of A3A-PBE/sgRNA$_{20}$, even though the editing was less efficient (Figure 6; Figure S4 in Supporting Information). We found that on some targets, APOBEC3A-eSpCas9(1.1)-UGI/sgRNA$_{19}$ could perform base editing and create indels, as could APOBEC3A-eSpCas9(1.1)-UGI/sgRNA$_{20}$. This is a drawback of this dual-functional system. Nevertheless, simultaneous base editing and indel production should be very useful for trait stacking in plants.

To explore whether APOBEC3A-eSpCas9(1.1)-UGI could

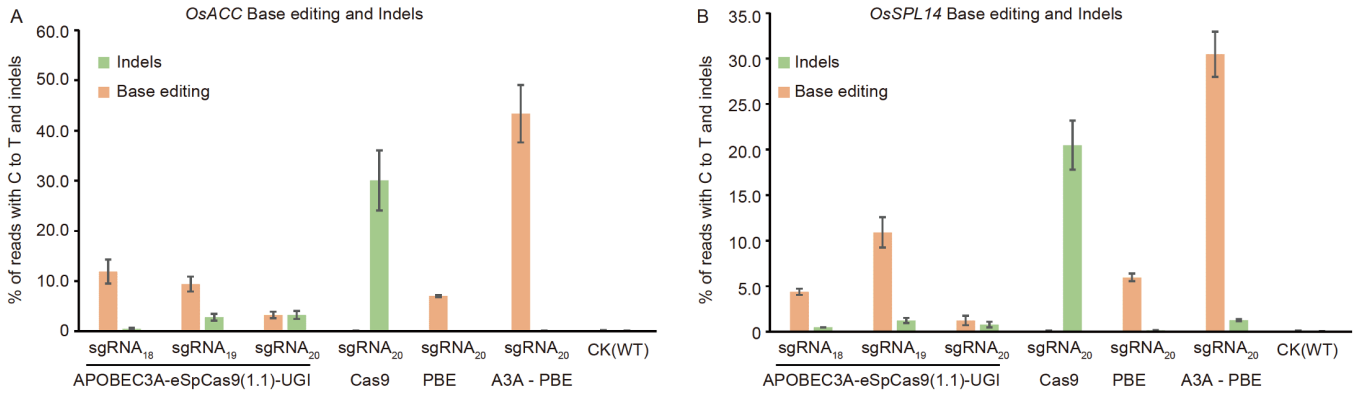


Figure 5 APOBEC3A-eSpCas9(1.1)-UGI generates base edits and indels in combination with different lengths of sgRNA. A and B, Frequencies of C-to-T conversions and indels in reads of the *OsACC* and *OsSPL14* target regions introduced by APOBEC3A-eSpCas9(1.1)-UGI along with tRNA-sgRNA$_{18–20}$, SpCas9, PBE and A3A-PBE in combination with tRNA-sgRNA$_{20}$. An untreated protoplast sample served as control (n=3 biologically independent experiments). All values are means±SD.

generate deletions, we designed a pair of sgRNA$_{20}$ to target *OsEPSPS* in rice protoplasts. APOBEC3A-eSpCas9(1.1)-UGI indeed generated the expected 327 bp deletion although at a slightly lower frequency than SpCas9 (Figure 7).

We wished to see whether APOBEC3A-eSpCas9(1.1)-UGI could perform trait stacking, by simultaneously carrying out several different types of editing on separate genes. Base editing of wheat ALS-P174 by C-to-T base editor can confer tolerance to the herbicide, nicosulfuron (Zhang et al., 2019). Recently, we described a wheat ALS-P174-based co-editing strategy using nicosulfuron to select mutants with desired base substitutions (Zhang et al., 2019). We attempted to use APOBEC3A-eSpCas9(1.1)-UGI to create a similar system for screening indel mutants. We constructed a vector expressing both an sgRNA$_{19}$ targeting TaALS-P174 and an sgRNA$_{20}$ targeting *TaGW2*, a key gene controlling wheat grain weight (Zhang et al., 2018). The vector and APOBEC3A-eSpCas9(1.1)-UGI were co-transferred into immature wheat embryos by particle bombardment and the embryos were incubated on non-selective medium. Emerging seedlings were transferred to a selective medium containing 0.254 mg L^{-1} nicosulfuron, and eight seedlings grew on the selective medium (Figure 8A and B). We found that in each case, base conversions had occurred in *TaALS* alongside an indel mutation in *TaGW2* (Figure 8C and D). Transiently-expressed SpCas9/sgRNA is also to generate such mutants, but isolating them without a selection procedure is time-consuming and labor-intensive (Ma et al., 2016). We believe that the APOBEC3A-eSpCas9(1.1)-UGI-based nicosulfuron-resistance selection system could make it easier to isolate transgene-free indel mutants.

DISCUSSION

In-depth understanding of the influence of DNA-sgRNA interactions on SpCas9 nuclease activity is critical for improving genome editing efficiency and reducing off-target effects. Various efforts have been made to improve the specificity of SpCas9, such as changing the length of the DNA-sgRNA interface, using high-fidelity SpCas9 variants, etc. (Casini et al., 2018; Chen et al., 2017; Cho et al., 2014; Fu et al., 2014; Kleinstiver et al., 2016; Lee et al., 2018; Slaymaker et al., 2016). Guide-seq and whole-genome sequencing (WGS) are often used to analyze the off-target effect for genome editing (Feng et al., 2014; Tang et al., 2018; Tsai et al., 2015). Indel rates due to double-strand DNA cleavage were used to quantify off-target activity in nearly all of these studies. However, this quantification strategy is imperfect, because single nicks in DNA can sometimes also give rise to indels, and this method would not distinguish between dCas9 and Cas9 nickase activity.

We showed that truncated sgRNA-DNA interface sequences can confer nickase activity on both SpCas9 and eSpCas9(1.1) (Figure 1C and D) and that eSpCas9(1.1) is more sensitive than SpCas9 to the length of sgRNA-DNA interface sequences (Figure 1). For the first time we demonstrated that the nicking occurs on the target DNA strand. This means that truncated sgRNAs affect both the RuvC domain and the HNH domain of SpCas9, but the RuvC domain is more affected than the HNH domain, with the result that it mainly acts as a nickase, though still performing a little double-strand DNA cleavage. With full size sgRNAs (20 nt spacer), SpCas9 and eSpCas9(1.1) both carry out mainly double-strand DNA cleavage. So when either SpCas9 or eSpCas9(1.1) is used, nicking and double-strand DNA cleavage can take place concurrently. This finding inspired us to develop the dual-functional editing system, APOBEC3A-eSpCas9(1.1)-UGI/sgRNA$_{19}$-sgRNA$_{20}$, which can perform base editing and generate indels at the same time. However, the editing efficiency of the system was sub-optimal at some targets and needs optimizing. Meanwhile,

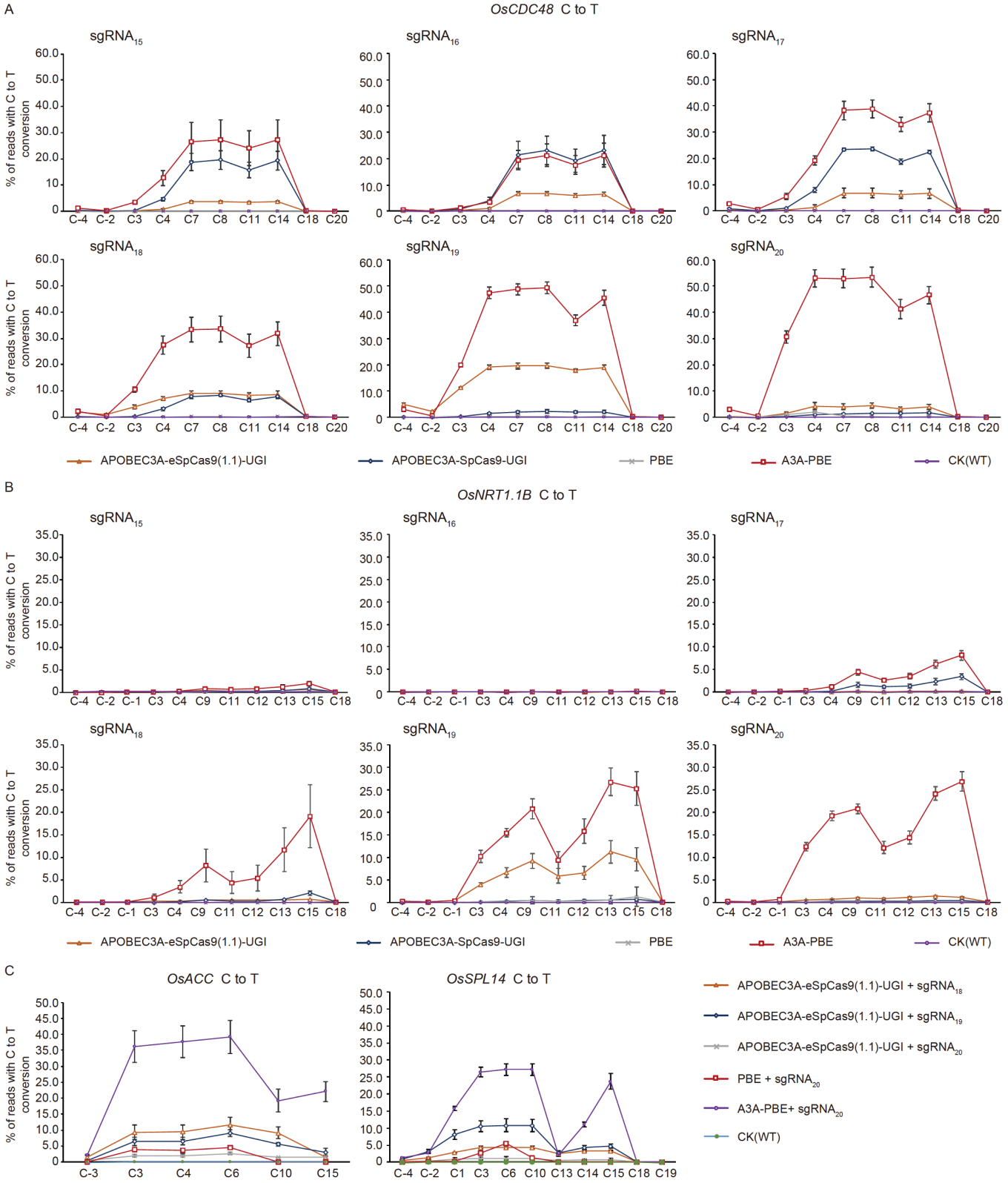


Figure 6 Comparison of single C-to-T conversion rates. A and B, Frequencies of single C-to-T conversions in reads of the *OsCDC48* and Os*NRT1.1B* target regions. APOBEC3A-eSpCas9(1.1)-UGI, APOBEC3A-SpCas9-UGI, PBE, and A3A-PBE were transformed into rice protoplasts along with tRNA-sgRNA$_{15–20}$ to target *OsCDC48* and Os*NRT1.1B*, respectively. C, Frequencies of single C-to-T conversions in reads of the *OsACC* and *OsSPL14* target regions. APOBEC3A-eSpCas9(1.1)-UGI along with tRNA-sgRNA$_{18–20}$ and PBE, A3A-PBE along with tRNA-sgRNA$_{20}$ were transformed into rice protoplasts to target *OsACC* and *OsSPL14*. An untreated protoplast sample served as control (n=3 biologically independent experiments). All values are means±SD.

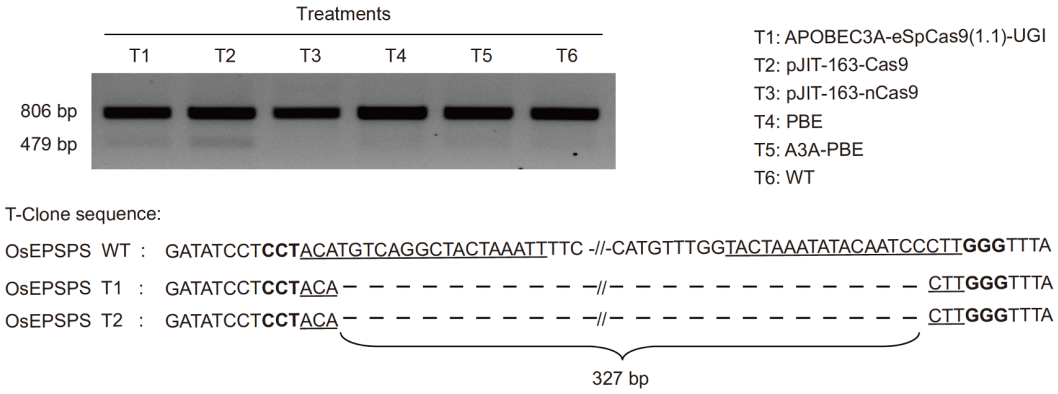


Figure 7 DNA agarose gel electrophoresis and T-clone sequencing analyses of DNA deletions introduced by APOBEC3A-eSpCas9(1.1)-UGI combined with tRNA-dual-sgRNA$_{20}$. SpCas9, nCas9(D10A), PBE, A3A-PBE and an untreated protoplast sample served as controls.

some efforts have been made to improve the efficiency in plants (Hao et al., 2020; Wang et al., 2019; Zhang et al., 2020). We imagine that the complex of APOBEC3A-eSpCas9(1.1)-UGI is defective, such that the two catalytic domains interfere with each other. In the present work, we have mainly focused on SpCas9 and eSpCas9(1.1). We also wonder whether the truncated sgRNAs enable the other SpCas9 variants to lead to the same results and to create more efficient dual-functional editing system. It will be important to investigate other variants and homologs of SpCas9.

MATERIALS AND METHODS

Plasmid constructs

pJIT163-Ubi-2xNLS-Cas9, pJIT163-Ubi-3xNLS-nCas9, pAPOBEC1-nCas9(D10A)-UGI (PBE), pAPOBEC3A-nCas9(D10A)-UGI (A3A-PBE), OsU3/TaU6-tRNA-sgRNA-PUC57, and pUbi-BFPm were described previously (Shan et al., 2013; Zhang et al., 2017; Zong et al., 2018; Zong et al., 2017). The primers used in this work are listed in Tables S1 and S2 in Supporting Information. All of the primers were synthesized by BGI (Beijing Genomics Institute, Beijing, China).

To construct OsU3/TaU6-dual-tRNA-sgRNAs-PUC57, the tRNA-sgRNA fragment in OsU3/TaU6-tRNA-sgRNA-PUC57 was amplified and cloned into OsU3/TaU6-tRNA-sgRNA-PUC57 using *Bam*H I and *Kpn* I (Thermo Scientific, USA)/*Kpn* I and *Mlu* I (Thermo Scientific, USA). It was then transformed into *E. coli* trans1-T1 competent cells (TransGen Biotech, Beijing, China).

Point mutations were introduced into the coding sequence of nCas9(D10A) using a Fast Mutagenesis System (TransGen Biotech, Beijing, China), to generate expression cassettes for SpCas9, nCas9(H840A), dSpCas9(D10A, H840A), and eSpCas9(1.1) (K848A, K1003A, R1060A).

To construct APOBEC3A-SpCas9-UGI, APOBEC3A-dSpCas9-UGI, and APOBEC3A-eSpCas9(1.1)-UGI plasmids, SpCas9, dSpCas9(D10A, H840A), and eSpCas9(1.1) (K848A, K1003A, R1060A) were cloned into A3A-PBE using *Bam*H I and *Mlu* I (Thermo Scientific, USA) to replace nCas9(D10A), using a ClonExpressII One Step Cloning kit (Vazyme, Nanjing, China).

To construct pJIT-163-Ubi-HH-sgRNA-HDV for cleavage assays *in vitro*, the HH-gRNA scaffold-HDV backbone was synthesized commercially (Genescript, Nanjing, China) and target sequences were cloned into HH-HDV-pUC57 with *Bpi* I (Thermo Scientific, USA). Then HH-sgRNA-HDV was cloned into pJIT-163-Ubi-BFPm using *Bam*H I and *Eco*R I (Thermo Scientific, USA) to replace the BFPm coding sequence.

Protoplasts transfection

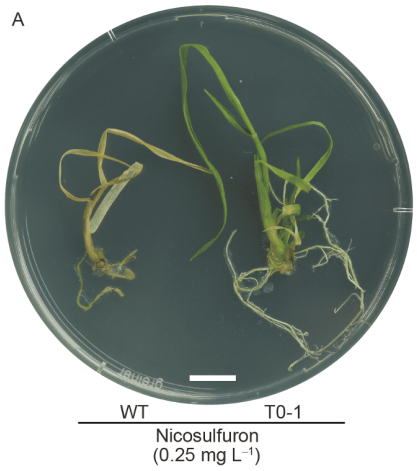
We used the *Japonica* rice variety Zhonghua11 to prepare protoplasts. Protoplasts were isolated and transformed as described (Shan et al., 2013). Plasmid DNA (10 μg per construct) was introduced by PEG-mediated transfection, with a mean transformation efficiency of 30%–50% by flow cytometry. The transfected protoplasts were incubated at 23°C. At 48 h post-transfection, they were collected to extract genomic DNA for deep amplicon sequencing and PCR restriction enzyme digestion assays (PCR-RE assays).

Flow cytometry analysis

Flow cytometry analysis was carried out using a FACSAria III (BD Biosciences, USA). Rice protoplasts were transfected with sgRNA expression plasmids, fluorophore expression plasmids, and editor expression plasmids. All samples were sorted for GFP-positive cells.

DNA extraction

Genomic DNA was extracted with a DNA Quick Plant System (Tiangen Biotech, Beijing, China). Targeted se-



B

Plant ID	*TaALS-A* subgenome	*TaGW2-A* subgenome	*TaGW2-B* subgenome	*TaGW2-D* subgenome
T0-1	Hetero	Hetero	Biallelic	Biallelic
T0-2	Hetero	Biallelic	Biallelic	Mosaic
T0-25	C to T / Indel	Biallelic	Homo	Biallelic
T0-33	C to T / Indel	Biallelic	Biallelic	Biallelic
T0-34	C to T / Indel	Biallelic	Mosaic	Mosaic
T0-35	Hetero	Homo	Biallelic	Mosaic
T0-36	C to T / Indel	Biallelic	Biallelic	Mosaic
T0-38	Hetero	Mosaic	Biallelic	Biallelic

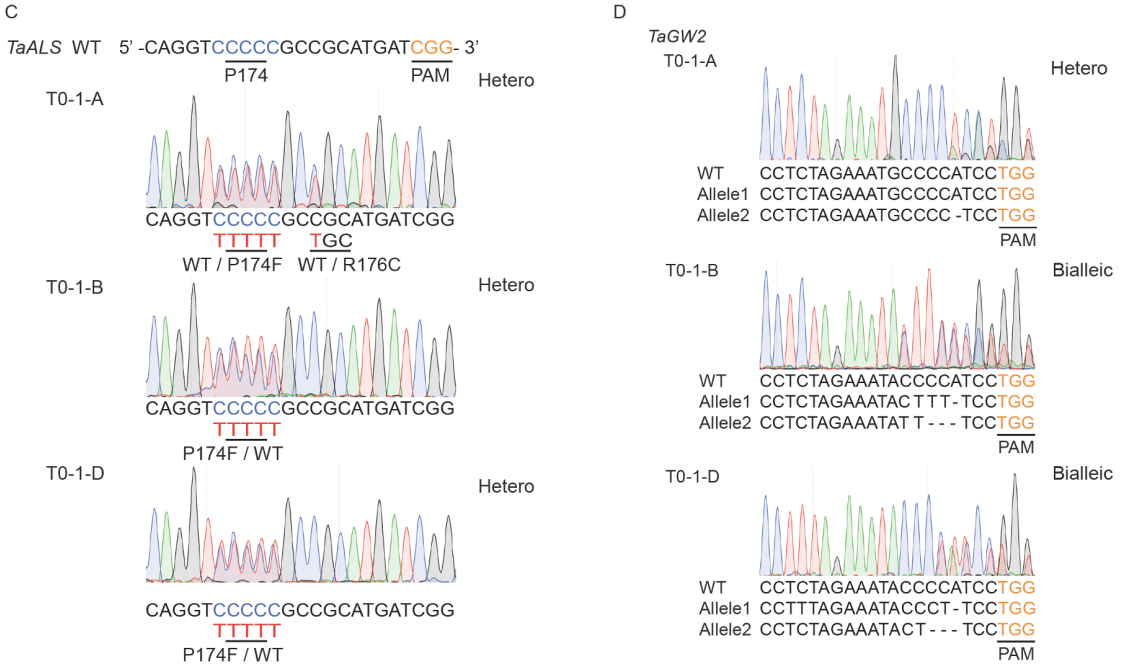


Figure 8 A TaALS-P174-based co-editing strategy for introducing indels with the dual-functional system. A, The T0-1 mutant with missense edits at position TaALS-P174 (right) that survived in a regeneration medium with 0.254 mg L^{-1} nicosulfuron. Wild-type (WT) plants were killed (left). Scale bar, 1 cm. B, Genotypes of all the T0 mutants. C, Sanger sequencing of *TaALS* in mutant T0-1. Target sequences, PAM sequences, DNA modifications, and editable cytosines are indicated in bold, brown, red, and blue, respectively. D, Sanger sequencing of *TaGW2* in mutant T0-1. Target sequences and PAM sequences are indicated in bold and brown, respectively.

quences were amplified with specific primers, and the amplicons were purified with an EasyPure PCR Purification kit (TransGen Biotech, Beijing, China) and quantified with a NanoDrop 2000 spectrophotometer (Thermo Scientific, USA).

Amplicon deep sequencing and data analysis

Genomic DNA was extracted from protoplast samples at 48 h post-transfection and used as template. In the first-round PCR, the target region was amplified with Fastpfu DNA polymerase (TransGen Biotech, Beijing, China) using site-specific primers. In the second round, both forward and reverse barcodes were added to the ends of the PCR products for library construction. Equal amounts of the PCR products were pooled, and the samples were sequenced commercially (Sangon, Shanghai, China) using the Illumina NovaSeq platform. The sgRNA target sites in the sequenced reads were examined for C-to-T substitutions and indels. Amplicon sequencing was repeated three times for each target site, using genomic DNA extracted from three independent protoplast samples. Analyses of base-editing and indel processivity were performed as previously described (Komor et al., 2016).

Biolistic delivery of DNA constructs into wheat immature embryo cells

DNA of plasmids pAPOBEC3A-eSpCas9(1.1)-UGI and pTaU6-tRNA-dual-sgRNAs was delivered simultaneously into immature embryos of Kenong199 via particle bombardment, as previously described (Liang et al., 2017; Ma et al., 2016). After bombardment, the embryos were cultured for plantlet regeneration on a medium with 0.254 mg L^{-1} nicosulfuron and incubated in a growth chamber (23°C, 16 h light/8 h dark) (Zhang et al., 2019).

Mutant identification by PCR-RE and Sanger sequencing

PCR-RE and Sanger sequencing were used to identify wheat mutants with indels in target regions, as described previously (Shan et al., 2013; Wang et al., 2014). For wheat, we detected mutations by PCR-RE or Sanger sequencing with homolog-specific primers.

Expression and purification of SpCas9, eSpCas9(1.1), nCas9(D10A), and nCas9(H840A) proteins

To express SpCas9, eSpCas9(1.1), nCas9(D10A), and nCas9 (H840A), we cloned the coding sequences of SpCas9, eSpCas9(1.1), nCas9(D10A), and nCas9(H840A) into pET28a (+) to generate pET28a-Cas9-His, pET28a-eSpCas9(1.1)-His, pET28a-nCas9(D10A)-His, and pET28a-nCas9 (H840A)-His, respectively (Liang et al., 2017). Each plasmid was then transformed into BL21 Star *E. coli* cells (TransGen Biotech, Beijing, China). Protein expression was induced with 0.5 mmol L^{-1} IPTG (Inalco, USA) at 18°C for 14–16 h. After induction, cells were harvested and resuspended in lysis buffer (25 mmol L^{-1} Tris-HCl (pH 8.0), 500 mmol L^{-1} NaCl, 1 mmol L^{-1} DTT, 10 mmol L^{-1} imidazole, 0.1% Triton X-100, 1 mmol L^{-1} PMSF), broken with a sonicator (Fisher Scientific, USA) and centrifuged at 25,000×*g* for 30 min. Supernatants were rotated with Ni-NTA beads at 4°C for 1 h, and proteins were purified according to the manufacturer's manual (GE Healthcare, USA). Purified proteins were concentrated in an Amicon Ultra-15 with a 30-kD cut-off (Millipore, USA), flash-frozen in liquid nitrogen and stored at –80°C.

In vitro transcription of HH-sgRNA-HDV sgRNA

The templates for transcription were amplified from pJIT-163-Ubi-HH-sgRNA-HDV with introduction of a T7 promoter at the 5′ end of the HH-sgRNA-HDV coding sequence using the relevant primers. Transcription was accomplished with an HiScribe T7 High Yield RNA Synthesis kit (New England Biolabs, USA) according to the manufacturer's instructions.

In vitro cleavage of SpCas9 and eSpCas9(1.1)

As previously described (Liang et al., 2017), DNA fragments containing the target site were amplified, purified, and eluted with RNase-free water. Cas9 protein (50 ng) and sgRNA (50 ng) were mixed with the purified target DNA (100–150 ng) in NEB buffer™ 3.1 (New England Biolabs, USA) (100 mmol L^{-1} NaCl, 50 mmol L^{-1} Tris-HCl, 10 mmol L^{-1} $MgCl_2$, 100 μg μL^{-1} BSA, pH 7.9) in a total volume of 20 μL, followed by digestion at 37 °C for 20 min. 1 μL Proteinase K (Thermo Scientific, USA) was added to the digested products at 37 °C for 0.5 h, and the digestion products were immediately separated on a 1.2% agarose gel.

Statistical analysis

All numerical values are presented as means±SD.

Compliance and ethics *The authors have filed a patent application based on the results reported in this paper. The patent does not restrict the research use of the methods in this article.*

Acknowledgements *This work was supported by grants from the Strategic Priority Research Program of the Chinese Academy of Sciences (Precision Seed Design and Breeding, XDA24020102), the National Transgenic Science and Technology Program (2018ZX0801002B), the National Natural Science Foundation of China (31788103 and 31971370), the Chinese Academy of Sciences (QYZDY-SSW-SMC030), and the National Key R&D Program of China (2018YFA0900600, 2016YFD0100102-11, and 2016YFD0100605).*

References

Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demichelis, F., et al. (2018). A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. Nat Biotechnol 36, 265–271.

Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. Nature 550, 407–410.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res 24, 132–141.

Doudna, J.A., and Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. Science 346, 1258096.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L., et al. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. Proc Natl Acad Sci USA 111, 4632–4637.

Fu, B.X.H., Smith, J.D., Fuchs, R.T., Mabuchi, M., Curcuru, J., Robb, G.B., and Fire, A.Z. (2019). Target-dependent nickase activities of the CRISPR-Cas nucleases Cpf1 and Cas9. Nat Microbiol 4, 888–897.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat Biotechnol 32, 279–284.

Gao, Y., and Zhao, Y. (2014). Self-processing of ribozyme-flanked RNAs into guide RNAs *in vitro* and *in vivo* for CRISPR-mediated genome editing. J Integr Plant Biol 56, 343–349.

Hao, Y., Zong, W., Zeng, D., Han, J., Chen, S., Tang, J., Zhao, Z., Li, X., Ma, K., Xie, X., et al. (2020). Shortened snRNA promoters for efficient

CRISPR/Cas-based multiplex genome editing in monocot plants. Sci China Life Sci https://doi.org/10.1007/s11427-019-1612-6.
Jiang, F., and Doudna, J.A. (2017). CRISPR-Cas9 structures and mechanisms. Annu Rev Biophys 46, 505–529.
Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495.
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, J.K., Jeong, E., Lee, J., Jung, M., Shin, E., Kim, Y., Lee, K., Jung, I., Kim, D., Kim, S., et al. (2018). Directed evolution of CRISPR-Cas9 to increase its specificity. Nat Commun 9, 3048.
Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., Liu, J., Zhang, H., Liu, C., Ran, Y., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. Nat Commun 8, 14261.
Ma, H., Tu, L.C., Naseri, A., Huisman, M., Zhang, S., Grunwald, D., and Pederson, T. (2016). Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nat Biotechnol 34, 528–530.
Puchta, H. (2018). Broadening the applicability of CRISPR/Cas9 in plants. Sci China Life Sci 61, 126–127.
Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.L., et al. (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. Nat Biotechnol 31, 686–688.
Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. Science 351, 84–88.
Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. Proc Natl Acad Sci USA 111, 9798–9803.
Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X., Zhong, Z., Liu, B., Zheng, X., et al. (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.
Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat Biotechnol 33, 187–197.
Wang, X., Lu, J., Lao, K., Wang, S., Mo, X., Xu, X., Chen, X., and Mo, B. (2019). Increasing the efficiency of CRISPR/Cas9-based gene editing by suppressing RNAi in plants. Sci China Life Sci 62, 982–984.
Wang, Y., Cheng, X., Shan, Q., Zhang, Y., Liu, J., Gao, C., and Qiu, J.L. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. Nat Biotechnol 32, 947–951.
Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proc Natl Acad Sci USA 112, 3570–3575.
Zhang, D., Zhang, H., Li, T., Chen, K., Qiu, J.L., and Gao, C. (2017). Perfectly matched 20-nucleotide guide RNA sequences enable robust genome editing using high-fidelity SpCas9 nucleases. Genome Biol 18, 191.
Zhang, Q., Yin, K., Liu, G., Li, S., Li, M., and Qiu, J.L. (2020). Fusing T5 exonuclease with Cas9 and Cas12a increases the frequency and size of deletion at target sites. Sci China Life Sci https://doi.org/10.1007/s11427-020-1671-6.
Zhang, R., Liu, J., Chai, Z., Chen, S., Bai, Y., Zong, Y., Chen, K., Li, J., Jiang, L., and Gao, C. (2019). Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. Nat Plants 5, 480–485.
Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L., Liu, J., Chen, K., Zhang, H., Gao, C., et al. (2018). Analysis of the functions of *TaGW2* homoeologs in wheat grain weight and protein content traits. Plant J 94, 857–866.
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.

SUPPORTING INFORMATION

The supporting information is available online at https://doi.org/10.1007/s11427-020-1722-0. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.