Genome-wide analyses of PAM-relaxed Cas9 genome editors reveal substantial off-target effects by ABE8e in rice

Yuechao Wu1,2,3,†, Qiurong Ren4,†, Zhaohui Zhong4,†, Guanqing Liu1,2,3,†, Yangshuo Han1,2,3,†, Yu Bao1,2,3,†, Li Liu4,†, Shuyue Xiang5,†, Shuo Liu1,2,3,†, Xu Tang4, Jianping Zhou4, Xuelian Zheng4, Simon Sretenovic5, Tao Zhang1,2,3,*, Yiping Qi5,6,*,† and Yong Zhang1,4,*,†

1Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding/Jiangsu Key Laboratory of Crop Genetics and Physiology, Agricultural College of Yangzhou University, Yangzhou, China
2Key Laboratory of Plant Functional Genomics of the Ministry of Education/Joint International Research Laboratory of Agriculture and Agri-Product Safety, The Ministry of Education of China, Yangzhou University, Yangzhou, China
3Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops, Yangzhou University, Yangzhou, China
4Department of Biotechnology, School of Life Sciences and Technology, Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu, China
5Institute of Plant Science and Landscape Architecture, University of Maryland, College Park, Maryland, USA
6Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, Maryland, USA

Received 12 February 2022; accepted 28 April 2022.

*Correspondence (Tel +86-28-87996817; email zhangyong916@uestc.edu.cn (Y.Z.); Tel +86-514-87977229; fax +86-514-87996817; email zhangtao@yzu.edu.cn (T.Z.); Tel +1-301-405-8683; email Yiping@umd.edu (Y.Q.))
†These authors contributed equally to this work.

Keywords: PAM-relaxed Cas9 nucleases, cytosine base editor, adenine base editor, off-target effect, whole-genome sequencing, genome editing, rice.

Summary
PAM-relaxed Cas9 nucleases, cytosine base editors and adenine base editors are promising tools for precise genome editing in plants. However, their genome-wide off-target effects are largely unexplored. Here, we conduct whole-genome sequencing (WGS) analyses of transgenic plants edited by xCas9, Cas9-NGv1, Cas9-NG, SpRY, nCas9-NG-PmAQR1, nSpRY-PmAQR1 and nSpRY-ABE8e in rice. Our results reveal that Cas9 nuclease and base editors, when coupled with the same guide RNA (gRNA), prefer distinct gRNA-dependent off-target sites. De novo generated gRNAs by SpRY editors lead to additional, but insubstantial, off-target mutations. Strikingly, ABE8e results in ~500 genome-wide A-to-G off-target mutations at TA motif sites per transgenic plant. ABE8e’s preference for the TA motif is also observed at the target sites. Finally, we investigate the timeline and mechanism of somaclonal variation due to tissue culture, which chiefly contributes to the background mutations. This study provides a comprehensive understanding on the scale and mechanisms of off-target and background mutations occurring during PAM-relaxed genome editing in plants.

Introduction
CRISPR-Cas9 genome editing tools have greatly revolutionized plant genetics and breeding. Streptococcus pyogenes Cas9 (SpCas9) is the predominant Cas9 widely used, partly due to its high genome editing efficacy and simple NGG protospacer adjacent motif (PAM) requirement (Cong et al., 2013; Jainek et al., 2012; Mali et al., 2013). To broaden the targeting scope, many PAM-relaxed SpCas9 variants have been engineered, including xCas9 (recognizing NG, GAA and GAT PAMs; Hu et al., 2018), SpCas9-NGv1 and SpCas9-NG (recognizing NG PAM; Nishimatsu et al., 2018), and PAM-less SpRY (Walton et al., 2020). These nucleases have been widely adopted for genome editing in plants (Hassan et al., 2021; Zhang et al., 2019). However, relaxed PAM requirements could make these nucleases prone to guide RNA (gRNA)-dependent off-targeting, which awaits a comprehensive investigation in plants.

The development of cytosine base editors (CBEs) and adenine base editors (ABEs) further expanded the genome editing toolbox (Anzalone et al., 2020), enabling precise base changes in plants (Molla et al., 2021). Cytidine deaminases and adenosine deaminases used in CBEs and ABEs could potentially catalyse deamination reactions nonspecifically in the genomes, causing gRNA-independent off-target effects. For example, whole-genome sequencing (WGS) revealed off-target effects caused by rAPOBEC1-based CBEs in rice (Jin et al., 2019; Ren et al., 2021b) and mouse (Zuo et al., 2019). CBEs engineered with different cytidine deaminases showed less off-target effects in human cells (Doman et al., 2020; Yu et al., 2020) and in rice (Jin et al., 2020; Ren et al., 2021b). ABE8e, a highly processive ABE (Lapinaite et al., 2020), catalyses highly efficient A-to-G base transitions in human cells (Richter et al., 2020) and in plants (Li et al., 2021; Ren et al., 2021c; Wang et al., 2021; Wei et al., 2021; Xu et al., 2021). Although elevated A-to-I conversions were reported in the transcriptomes of ABE8e-treated human cells (Richter et al., 2020), it is unknown whether, or to what extent, gRNA-independent off-target mutations in plants would be generated by ABE8e.

Merging PAM-relaxed Cas9 variants and highly efficient cytidine/adenosine deaminases opens the door for highly flexible base editing in plants (Molla et al., 2021). CBEs based on xCas9 were reported in rice to edit NGN PAM sites, albeit with very low efficiency (Hua et al., 2019; Li et al., 2019; Zeng
et al., 2020a; Zhong et al., 2019). SpCas9-NGv1- and SpCas9-
NG-based CBEs were tested in different plant species (Endo et al.,
2019; Hua et al., 2019; Zeng et al., 2020a, 2020b; Zhong et al.,
2019), generally outperforming xCas9-based CBEs at relaxed PAM
sites (Molla et al., 2021). SpRY CBEs were demonstrated to edit
NRN PAMs better than NYN PAMs in rice (Li et al., 2021; Ren et al.,
2021c; Xu et al., 2021; Zhang et al., 2021). Similarly, ABEs were
demonstrated in plants with SpCas9-NGv1 (Negishi et al., 2019) or
SpCas9-NG (Hua et al., 2019; Wang et al., 2019b; Zeng et al.,
2020a) and SpRY (Li et al., 2021; Ren et al., 2021a, 2021c; Xu et al.,
2021; Zhang et al., 2021). Despite the wide demonstration of these PAM-
relaxed CBEs and ABES in plants, their potential genome-wide
off-target effects have not been reported. To fill this critical
knowledge gap, we comprehensively assessed gRNA-dependent
and -independent off-target effects of these PAM-relaxed
nucleases and base editors using WGS in rice. We also
investigated the generation of somaclonal variation in the
context of genome editing.

Results
Observing off-target effects of PAM-relaxed genome
editing in rice through whole-genome sequencing
Our previous study revealed that xCas9 largely retained the NGG
PAM requirement of SpCas9 with improved editing specificity
(Zhong et al., 2019). To simply validate this observation, we
included an xCas9 construct for editing an NGG PAM site with
OsDEP1-gR02-GGG. Although SpCas9-NGv1 and SpCas9-NG
both recognize NGN PAMs (Endo et al., 2019; Negishi et al.,
2019; Nishimatsu et al., 2018), SpCas9-NG has higher editing
efficiency than SpCas9-NGv1 (Nishimatsu et al., 2018; Zhong
et al., 2019). It is however unknown for the off-target effects of
SpCas9-NGv1 and SpCas9-NG variants. Thus, we targeted two
independent sites OsDEP1-gR01-GGT and OsDEP1-gR02-CGC
with both variants. Genome-integrated T-DNAs are prone to self-
editing by SpRY and its derived base editors, leading to de novo
generated gRNAs (Ren et al., 2021c), we wanted to investigate the
scale of off-target mutagenesis due to such de novo
generated gRNAs by SpRY at four different target sites
(OsDEP1-gR01-CGC, OsDEP1-gR04-CGC, OsPDS-gR01-TCA and
OsPDS-gR03-TAA). For off-target analysis of PAM-relaxed CBEs,
we focused on SpCas9-NG and SpRY with the highly efficient and
specific PmCDA1 cytidine deaminase (Ren et al., 2021b). nSpCas9-
NG-PmCDA1 and nSPRY-PmCDA1 each edited two target
sites (OsDEP1-gR01-TGT and OsDEP1-gR02-CGC for
nSpCas9-NG-PmCDA1; OsALS-gR21-GCA and OsALS-gR22-
AGC for nSPRY-PmCDA1). By contrast, off-target effects of the
highly efficient adenosine deaminase, ABE8e, are largely
unknown. Using nSPRY-ABE8e to edit two independent sites
(OsPDS-gR01-TTG and OsPDS-gR04-TAA) and including the
nSPRY-ABE8e control without a gRNA, we hoped to reveal both
gRNA-dependent and -independent off-target effects by this
highly efficient PAM-less ABE.

These constructs, along with corresponding controls that didn’t
contain targeting gRNAs (Table S1), were used to generate
transformed rice plants through Agrobacterium-mediated trans-
formation. Genome editing frequencies were calculated for most
constructs including PAM-relaxed Cas9 nucleases (SpCas9-NGv1,
SpCas9-NG and SpRY; Figure 1a), and CBEs (nSpCas9-NG-
PmCDA1 and nSPRY-PmCDA1) (Figure 1b), and nSPRY-ABE8e
(Figure 1c). As expected, SpCas9-NG showed higher editing
efficiency than SpCas9-NGv1 (Figure 1a). Different numbers (one to
four) of the edited T0 lines and the corresponding controls
without targeting gRNAs were chosen for WGS (Figure 1d and
Table S1). The resulting sequencing data showed >95% sequenc-
ing depth, >99% mapping ratio, and >97% genome coverage for
all 58 samples (Table S2), which were processed according to a
rigid bioinformatics pipeline to call out single nucleotide variations
(SNVs) and insertions or deletions (INDELs) for further compar-
sions and analyses (Figure 1e; Ren et al., 2021b; Tang et al.,
2018). We analysed the three T0 lines edited by xCas9 at OsDEP1-
gR02-GGG site and did not find gRNA-dependent off-target
mutations (Table S3), which is consistent with its high targeting
specificity reported in human cells (Hu et al., 2018) and in rice
(Zhong et al., 2019).

Comparison of SpCas9-NGv1, SpCas9-NG and nSpCas9-
NG-PmCDA1 reveals differential gRNA-dependent off-
target effects dictated by nuclease activity and editor
types
According to our previous research results (Randall et al., 2021;
Ren et al., 2021b; Tang et al., 2018), gRNA-dependent off-
target mutations may occur at sites with less than 5-bp
mismatches compared with the protospacers. To assess gRNA-
dependent off-target effects of SpCas9-NG-based editors, we
compared SpCas9-NGv1, SpCas9-NG and nSpCas9-NG-
PmCDA1 at editing NGN PAM sites. At the OsDEP1-gR02-
CGC site, WGS discovered six off-target sites that were edited by
SpCas9-NGv1, five out of six being shared among two T0
lines (Figure 2a). These six off-target sites contain NGN PAMs
and no more than 1 mismatch mutation in the 3–20 nt region of
the protospacers, suggesting high likelihood of off-target
editing. The resulting off-target mutations are small deletions
and -bp insertions around the Cas9 cleavage site located 3 bp
upstream of the PAM (Figure 2a). These indels are common of
Cas9 editing outcomes. A total of 11 off-target sites with NGN
PAMs were discovered among the two T0 lines edited by
SpCas9-NG, including the four identified with SpCas9-NGv1
(Figure 2b). Only one off-target mutation was shared by the two
T0 lines (Figure 2b). Many of the newly discovered off-
target sites with SpCas9-NG contain two or more mismatches
to the protospacer (Figure 2b), suggesting that SpCas9-NG has
higher nuclease activity than SpCas9-NGv1, which is consistent
with previous studies (Nishimatsu et al., 2018; Zhong et al.,
2019). Six off-target sites were identified in the two T0 lines
tedited by nSpCas9-NG-PmCDA1, with three different off-target
sites in each line (Figure 2c). Unlike SpCas9-NGv1 and SpCas9-
NG that shared four off-target sites, the six off-target sites
identified with nSpCas9-NG-PmCDA1 are all different from
those identified with the nucleases (Figure 2d), suggesting
gRNA-dependent off-target mutations by Cas9 nucleases and
base editors follow different mechanisms. Of note, four of the
six off-target sites carry deletions spreading across the proto-
spacer (Figure 2c), supporting the off-target mutations were
curred by cytidine deaminase activity and base excision repair.
Interestingly, none of the T0 lines analysed here showed
evidence of T-DNA self-editing. This could be explained by the
fact that the GTT PAM in the gRNA scaffold is not an optimal
PAM for SpCas9-NGv1, SpCas9-NG and nSpCas9-NG-PmCDA1,
although self-editing by SpCas9-NG was previously reported in
rice (Qin et al., 2020).
Figure 1: Assessment of PAM-less genome editing in rice by whole-genome sequencing. (a-c) Genome editing frequencies in T0 lines by PAM-relaxed Cas9-NGv1, Cas9-NG and SpRY (a) by PAM-relaxed cytosine base editors based on nCas9-NG and nSpRY (b), and by PAM-less nSpRY-ABE8e adenine base editor (c). (d) Summary of plants used for whole-genome sequencing. (e) The bioinformatic pipeline for analysis of whole-genome sequencing (WGS) data. NA, editing frequency in T0 lines was not scored for the constructs xCas9-OsDEP1-gr02-GGG and nSpRY-PmCDA1-OsALS-gr21-GCA. Different editing systems targeting the same target site are indicated by using the same sgRNA name (e.g., OsDEP1-gr01-TGG for SpCas9-NG, SpCas9-NGv1 and nSpCas9-NG-PmCDA1).
Comparison of SpRY and nSpRY-ABE8e reveals gRNA-dependent off-target mutations by de novo generated gRNAs

To investigate gRNA-dependent off-target effects of SpRY editors, we first investigated the gRNA-dependent off-target effects by SpRY-derived base editors. The results showed that no gRNA-dependent off-targeting was found in the edited T₀ lines by nSpRY-PmCDA1 (Table S3). However, 18 and 5 potential off-target sites with up to 5 mismatches were edited by SpRY and SpRY-ABE8e, respectively (Table S3). Among these edited off-target sites, 21 out of 23 contain no more than 3 mismatch mutations in the 3–20 nt region of the protospacers (Fig. S1A–B and Fig. S2). Thus, the off-target effect of SpRY could be minimized by improving the specificity of protospacers.

**Figure 2** Different sequence preference of gRNA-dependent potential off-target editing by Cas9-NG nucleases and cytosine base editors. (a–c) gRNA-dependent off-target mutations in edited T₀ lines at the OsDEP1-gR02-CGC site by SpCas9-NGv1 (a), SpCas9-NG (b) and nSpCas9-PmCDA1 (c). Off-target sites that were shared between SpCas9-NGv1 and SpCas9-NG are marked in red. Top panel, sequence comparison of target gRNA and potential off-target sites. Middle panel, the genotype of the off-target sites. Bottom panel, the number of potential off-target sites in two T₀ plants. (d) Venn diagram depicting many shared off-target sites induced by the OsDEP1-gR02-CGC gRNA in SpCas9-NGv1 and SpCas9-NG, while not in nCas9-NG-PmCDA1.

© 2022 The Authors. *Plant Biotechnology Journal* published by Society for Experimental Biology and The Association of Applied Biologists and John Wiley & Sons Ltd., 20, 1670–1682
We next focused our analysis on de novo generated gRNAs due to T-DNA self-editing, a common phenomenon caused by the PAM-less nature of SpRY (Ren et al., 2021c). Ten lines were analysed at four target sites (Figure 1a and b). New gRNAs were generated at all four target sites among eight T0 lines (Figure 3a and Figure S3). Based on these new protospacers, we identified potential off-target sites with 0–5 nucleotide mismatches using Cas-OFFinder (Bae et al., 2014). However, only two new gRNAs resulted in off-target mutations at these predicted off-target sites (Figure 3a). At the OsDEP1-gR01-CGC site, one new gRNA appeared to cause one SNV mutation at a target site with multiple nucleotide mismatches (Figure 3b). Similarly, at the OsDEP1-gR04-CGC site, one new gRNA seemed to generate either SNV or INDEL mutations at five off-target sites (Figure 3c). These off-target sites showed significant difference to the protospacer of the original target gRNA (Figure 3c), suggesting that the mutations at these sites were unlikely to be caused by the original gRNA, rather more likely to be created by the new gRNA. Given that detected mutations at these off-target sites are located upstream relative to the Cas9 cleavage site (Figure 3b and c), it is possible that some of these mutations were caused by tissue culture, not by gRNA-dependent SpRY editing.

We also investigated self-editing related off-target effects of SpRY-based CBE and ABE. For nSpRY-PmCDA1, T-DNA self-editing of the OsALS-gR21-GCA construct and the OsALS-gR22-AGC construct was detected in one out of two T0 lines each (Fig. S4), generating one and two new gRNAs, respectively (Figure 3d). For all three new gRNAs, WGS did not detect off-target mutations at the off-target sites predicted by Cas-OFFinder (Figure 3d). For nSpRY-ABE8e, T-DNA self-editing was detected in most T0 lines for the OsPDS-gR01-TTG and the OsPDS-gR04-TAA constructs (Figures 3d and S5). Interestingly, in both cases, no off-target mutations were detected at Cas-OFFinder-predicted off-target sites with three or fewer nucleotide mismatches (Figure 3d). However, for nSpRY-ABE-OsPDS-gR01-TTG, mutations were detected in line 2 at two predicted off-target sites with four and five nucleotide mismatches to the protospacer of the new gRNA and with six nucleotide mismatches to the protospacer of the original target gRNA.
of the original target gRNA (Figure 3e). Similarly, one off-target mutation was detected for nSpRY-ABE-OsPDS-gR01-TAA in line 4, where the off-target site showed two fewer mismatches (five vs. seven) to the protospacer of the new gRNA than the original target gRNA (figure S7). All three off-target events are A-to-G conversions at target sites with NRN PAMs (Figure 3e and f), consistent with high purity base conversion by ABE8e (Richter et al., 2020) and SpRY PAM preference of NRN PAMs over NYN PAMs (Walton et al., 2020). Together, these data suggest that very few gRNA-dependent off-target mutations were induced by PAM-relaxed SpRY base editors.

Comparison of PAM-relaxed nucleases and base editors reveals gRNA-independent genome-wide off-target A-to-G mutations by ABE8e

We next pursued our analyses to reveal any off-target effects of these PAM-relaxed editors that are independent of gRNAs. For xCas9, SpCas9-NGv1, SpCas9-NG, SpRY and nSpRY-PmCDA1 constructs, both genome-edited plants and control plants shared similar numbers of SNVs (ranging from 86 to 322, on average 187), INDELs (ranging from 48 to 108, on average 75; Figures 4a and S6) and frequencies of deletions for different sizes (Figure S7). These mutations appeared to be present in all genomic regions across the genome (Figure 4b and Figure S8). Importantly, the numbers of SNVs and INDELs observed are in the same range as those observed in previous studies (Jin et al., 2019; 2020; Ren et al., 2021b; Tang et al., 2018), supporting these mutations were somaclonal variation due to tissue culture. Strikingly, both genome-edited plants and control plants expressing nSpRY-ABE8e showed many more SNVs, averaging 700 per plant (Figure 4a) and being present in all genomic regions (Figure 4b). By contrast, nSpRY-ABE8e expressing plants showed similar numbers of INDELs (on average 77) to other plant groups (Figure 4d). A close analysis showed the excessive amount of SNVs in nSpRY-ABE8e expressing plants are A-to-G mutations, and the high enrichment of A-to-G mutations and decreased fractions of other nucleotide substitutions were only observed with plants expressing nSpRY-ABE8e (Figure 4c). These A-to-G mutations were randomly spread across all 12 chromosomes of rice genome (Figure 4d). About 95% of these A-to-G mutations belong to the category of 25%-75% allele frequencies (Fig. S9), suggesting these are largely germline transmittable mutations. Our results hence demonstrated genome-wide gRNA independent A-to-G off-target mutagenesis in rice by the highly processive ABE8e.

ABE8e favours TA motif sites for both off-target and on-target editing

To further study the off-target effects by ABE8e, we analysed all the A-to-G off-target editing sites in 10 T0 lines. The results showed unambiguously that ABE8e favours conversion of A to G in TA motifs on either Watson strand (Figure 5a) or Crick strand (Fig. S10). We reasoned that such a preference of editing TA motifs by ABE8e could also be reflected at on-target sites. To this end, we tested nCas9-ABE8e at editing an NPG PAM site in rice protoplasts and the data showed A-to-G conversions at both A4 and A12 (Figure 5b), with both positions being at the edge of the editing window known for ABE8e (Richter et al., 2020). The editing frequency at A12 proceeded by a ‘T’ is significantly higher than A4 proceeded by a ‘G’ (Figure 5b), supporting that ABE8e also favours TA motifs for on-target editing. We then analysed all 11 edited alleles in T0 lines by nSpRY-ABE8e-OsPDS-gR01-TTG (Figure 5c) and found A4 proceeded by a ‘T’ was edited at much higher frequency than A7 proceeded by an ‘A’ (Figure 5d), although both A4 and A7 are within the ABE8e editing window. Furthermore, we analysed the gRNA-dependent off-target editing outcomes discovered at four off-target sites by the same construct (Figure 5e). A-to-G conversions were only found at TA sites, not at AA, CA and GA sites (Figure 5f). Taken together, these analyses indicate that ABE8e has a strong preference of the TA motif for both off-target and on-target editing.

Investigation of the somaclonal variation production timeline in rice tissue culture

Since most SNVs (except those from ABE8e-expressing plants) and INDELs are derived from tissue culture, it would be helpful to understand the geneit mechanism and timeline for somaclonal variation. Like many other plants, rice genome editing involves the generation of embryogenic callus, followed by Agrobacterium-mediated transformation and regeneration (Nishimura et al., 2006). We reasoned that somaclonal variation mutations would be collectively generated before (termed as ‘Phase I somaclonal variation’) and after Agrobacterium-mediated transformation (termed as ‘Phase II somaclonal variation’; Figure 6a). Based on the WGS data, we mapped all the T-DNA insertion sites to the rice genome among all the T0 lines. Although most plants contained only one T-DNA insertion, 16 plant pairs shared the same T-DNA insertion for each pair (Figure 6b), suggesting each pair of these plants were derived from the same T-DNA transformation event. We hypothesize that shared mutations among such plant pairs would largely represent Phase I somaclonal variations. Our analysis largely confirmed this as the T0 plants that share the same T-DNA insertion sites showed high proportion of shared mutations (Figure 6c and Fig. S11). Although the numbers of shared mutations for the T0 lines with the same T-DNA insertions vary greatly (from 23 to 168), the average number (98) is significantly higher than the average number of shared mutations (7.4) among T0 lines with diverse T-DNA insertion sites (Figure 6d).

We next sought to understand the timeline of genome editing in the context of Phase II somaclonal variation production (Figure 6a). We took advantage of the genome-wide off-target editing by ABE8e and identified three T0 plant pairs that were derived from the same transgenic events, based on the shared T-DNA insertion sites (Figure 6b). In all three cases, the sum of whole-genome SNVs are more than 1300, with about 70% being A-to-G mutations (Figure 6e), consistent with the genome-wide A-to-G off-target mutations by ABE8e (Figure 4). If the ABE8e-based off-target editing were to occur before the transformed callus being developed into two T0 lines, the shared mutations between the two T0 lines would contain a high percentage of A-to-G mutations. This is indeed the case for the two T0 lines edited by nSpRY-ABE8e at OsPDS-gR01-TTG site, where over 70%
shared mutations were A-to-G mutations (Figure 6e). For the two remaining cases, about 20% total shared mutations among the two single-event T0 lines were A-to-G mutations (Figure 6e), indicating most of the A-to-G off-targeted mutations in these lines were largely independently induced by the same ABE8e transgenic event. These data suggest variable timelines for genome editing to occur in the developmental stage that generates Phase II somaclonal variation. The collective analyses here elucidate the details and timelines of genome editing and somaclonal variation in rice tissue culture: About 100 mutations are Phase I somaclonal variation mutations and about 253 (ranging from 62 to 854) mutations are Phase II somaclonal variation mutations. Genome editing can occur at different timepoints during the Phase II tissue culture stage.

Figure 5 ABE8e favours A-to-G conversion at TA motifs at both off-target and on-target sites. (a) Preference of a TA motif by ABE8e at gRNA-independent off-target A-to-G base editing in Watson strand, 0 indicates the A-to-G SNV position. (b) Base editing frequencies at different protospacer positions by ABE8e at a target site in rice protoplasts, n represents biological replicates. Data reanalysed from ref (Ren et al., 2021c). Error bars represent s.e.m. P-value was calculated by the one-sided paired Student’s t-Test, *P < 0.05, **P < 0.01. (c) The genotype of mutation alleles in T0 stable transformation plants. (d) Base editing frequencies at different protospacer positions by ABE8e at a target site in rice T0 lines. (e) Presence of TA motifs at the target site appears to increase gRNA-dependent off-target A-to-G editing. (f) The frequency of A-to-G SNV with different di-nucleic acids in T0 stable transformation plants.
PAM-relaxed Cas9 variants such as SpCas9-NG and SpRY greatly increase the targeting scope in plant genome editing (Endo et al., 2019; Hua et al., 2019; Li et al., 2021; Negishi et al., 2019; Ren et al., 2021a, 2021c; Wang et al., 2019b; Xu et al., 2021; Zeng et al., 2020a, 2020b; Zhang et al., 2021; Zhong et al., 2019). However, off-target risks also increase with their relaxed PAM restriction and tendency for T-DNA self-editing (Qin et al., 2020; Ren et al., 2021c). Based on WGS analyses in rice, we have found very few off-target mutations induced by SpCas9-NG, SpRY and their derived CBEs based on PmCDA1, a highly specific cytidine deaminase (Ren et al., 2021b). Our WGS analyses also revealed that SpRY and its derived base editors had higher tendency than SpCas9-NG editors to self-edit their T-DNA (Qin et al., 2020; Ren et al., 2021c). Yet, very limited numbers of off-target mutations

Figure 6  Investigation of somaclonal variation production in rice tissue culture. (a) A model that divides the generation of somaclonal variation into two phases, which points to potential of minimizing Phase II somaclonal variation with the use of organic factors to accelerate plant regeneration. (b) Genome-wide mapping of T-DNA integration sites for all T0 lines. Constructs that contain more than one T-DNA integration site are highlighted in red. The two T0 lines that carry the same T-DNA integration site were grouped by a solid line on the right, indicating they are from the same transgenic event. (c) Four examples for the analysis of T0 lines for shared mutations revealed by WGS. The T0 lines resulting from the same transgenic event (highlighted in red) share a significant portion of mutations (termed Phase I somaclonal variation). (d) T0 lines with the same T-DNA integration sites share an average of 98 mutations, while T0 lines with different T-DNA integration sites barely share any mutations. (e) The frequency of A-to-G SNVs in shared SNVs and whole-genome SNVs from the nSpRY-ABE8e T0 lines with the same transgenic events, the number above of each bar represents A-to-G SNVs versus all SNVs in a pair of T0 lines. *P-value was calculated by the Wilcoxon rank sum test, **P < 0.05, ***P < 0.01, NS represents not significant.

Discussion

PAM-relaxed Cas9 variants such as SpCas9-NG and SpRY greatly increase the targeting scope in plant genome editing (Endo et al., 2019; Hua et al., 2019; Li et al., 2021; Negishi et al., 2019; Ren et al., 2021a, 2021c; Wang et al., 2019b; Xu et al., 2021; Zeng et al., 2020a, 2020b; Zhang et al., 2021; Zhong et al., 2019). However, off-target risks also increase with their relaxed PAM restriction and tendency for T-DNA self-editing (Qin et al., 2020; Ren et al., 2021c). Based on WGS analyses in rice, we have found very few off-target mutations induced by SpCas9-NG, SpRY and their derived CBEs based on PmCDA1, a highly specific cytidine deaminase (Ren et al., 2021b). Our WGS analyses also revealed that SpRY and its derived base editors had higher tendency than SpCas9-NG editors to self-edit their T-DNA (Qin et al., 2020; Ren et al., 2021c). Yet, very limited numbers of off-target mutations
were detected in the edited plants by the de novo generated new gRNAs. Hence, our results benchmark these genome editing tools for broadened editing scope without significant off-target effects in plants.

The development of the highly processive ABE8e (Lapinaite et al., 2020; Richter et al., 2020) has greatly boosted precise adenine base editing in plants, with up to 100% editing efficiency and extremely low occurrence of INDEL by-products, which collectively contributed to high frequency of homozygous editing in plants within a single generation (Li et al., 2021; Ren et al., 2021c; Wang et al., 2021; Wei et al., 2021; Xu et al., 2021). Recently, transcriptome-wide analysis in human cells revealed off-target A-to-I conversions caused by ABE8e at the RNA level (Richter et al., 2020), a phenomenon that was previously reported for ABE7.10 (Zhou et al., 2019a). However, significant genome-wide off-target effects have not been previously reported for ABE8e in any organism. Remarkably, we discovered substantial genome-wide off-target effects induced by ABE8e in rice, ~500 A-to-G off-target mutations generated per plant (Figure 4a and d). These off-target mutations greatly outweigh the somaclonal variation mutations, presenting a significant implication for the use of ABE8e in plant research. Unlike RNA mutations which are transient and non-inheritable, the resulting A-to-G mutations at the DNA level are largely inheritable (Fig. S9). Such off-target effects of ABE8e must be addressed before its safe use in plant genetics and crop breeding. Encouragingly, engineered point mutations in the adenosine deaminase have been shown to reduce transcriptome off-target effects by ABE7.10 (Zhou et al., 2019a), ABE8e (Richter et al., 2020) and other ABE8 variants (Gaudelli et al., 2020). It awaits further testing whether genome-wide off-target A-to-G conversions could be largely mitigated by adopting a highly specific ABE8e variant that carries a promising mutation such as V106W (Gaudelli et al., 2020; Richter et al., 2020).

Interestingly, we found that ABE8e favours editing of TA motifs on DNA, which is consistent with the previous observation that ABE7.10 prefers TA motifs for off-target editing on RNA (Zhou et al., 2019a). Importantly, we found that such a TA motif preference by ABE8e also applies to the target sequence. Hence, this exciting discovery can be applied to improve on-target editing by ABE8e or its further engineered variants by intentionally targeting ‘A’ in a TA motif to achieve high editing efficiency. A CBE was previously used to fine-tune gene expression in strawberry to increase the sugar content (Xing et al., 2020).

Given the high abundance of TA motifs in the cis-regulatory elements (e.g., the TATA box) of many plant genes, ABE8e would be a promising tool for engineering quantitative trait variation by editing cis-regulatory elements, an innovative genome editing application that has been conventionally achieved with the Cas9 nuclease(s) (Molla et al., 2021; Rodriguez-Leal et al., 2017; Zhang et al., 2019).

Our WGS analyses, along with the previous studies (Fossi et al., 2019; Jin et al., 2019, 2020; Ren et al., 2021b; Tang et al., 2018, 2019), uncovered the scale of somaclonal variation derived from the tissue culture process, which by itself is a bottleneck for genome editing in plants (Altpeter et al., 2016). Since somaclonal variation is present in all genome-edited plants that are generated by tissue culture, effective strategies are needed to reduce such background mutations, of which many are germline-transmittable (Tang et al., 2018). Here, we took a unique approach to investigate the generation of somaclonal variation before and after Agrobacterium-mediated transformation, which should be applicable to other plants. For the Phase I somaclonal variation mutations, existing before plant transformation (Figure 6a), we may have limited means of reducing them. However, there are often more Phase II somaclonal variation mutations generated, which occur after Agrobacterium-mediated transformation. We hypothesize that Phase II somaclonal variation may be reduced by accelerating plant regeneration with the expression of morphogenic or growth factors, as recently demonstrated in different plant species (Debernardi et al., 2020; Lowe et al., 2016; Maher et al., 2020). It will be promising to test this idea.

In summary, the comprehensive WGS analyses of PAM-relaxed Cas9 nucleases and their derived base editors revealed highly specific genome editing in rice. However, ABE8e, despite its promise for highly efficient and high-purity base editing, showed substantial genome-wide off-target A-to-G conversions that are independent of gRNAs. This study also points to promising approaches of enhancing on-target and reducing off-target A-to-G editing by ABE8e or its variants, as well as potentially reducing Phase II somaclonal variation in genome-edited plants.

## Experimental procedures

### Plant material and growth condition

The Nipponbare rice cultivar (Oryza sativa L. ssp. Japonica cv. Nipponbare) was used in this study as the WT control and transformation host. All plants for the WGS assay were grown in growth chambers under a controlled environmental condition of 60% relative humidity with a 16/8 h and 32/28 °C regime for under the light/dark cycle.

### Construction of T-DNA vectors

The PAM-relaxed CRISPR-Cas9 plant genome editing systems used in this study were reported in our previous studies (Ren et al., 2021c; Zhong et al., 2019). Target sites were inserted by Golden Gate reaction using Bsal HF V02 and T4 DNA Ligase (New England Biolabs) per our previous description (Zhou et al., 2017, 2021, 2022). Briefly, the synthesized pair oligos (10 μM) were annealed and cool down to room temperature (23 °C). The annealed mixture was diluted to 50 nM for a total 15 cycles in the Golden Gate reaction (Zhou et al., 2021, 2022). The reaction mixture was transformed to Escherichia coli DH5α competent cells followed by mini prep and Sanger sequencing.

### Rice transient and stable transformation

Rice protoplast isolation, transformation and editing activity evaluation were performed as described previously (Tang et al., 2017; You et al., 2018; Zhang et al., 2013). The Agrobacterium-mediated rice stable transformation was based on previously published protocols with minor modifications (Hei et al., 1994; Wang et al., 2019a; Zhou et al., 2019b). Briefly, the rice callus was induced and the binary T-DNA vectors were transformed into Agrobacterium tumefaciens EHA105 strain. The transformed EHA105 strain was cultured in the flask until the OD600 = 0.1 at 28 °C and collected by centrifuge. The collected Agrobacterium was resuspended with AAM-AS medium for calli transformation. After 3 days of co-incubation, the transformed calli were washed by sterile water and transferred to N6-S solid medium for 14 days under continuous light at 32 °C. The grown calli were collected and incubated at REIII solid medium. After a 14-day regeneration, the newly grown individual plants were transferred to HF solid medium for root induction. Then, the generated plants were moved into pods and grown in soil at growth chamber under...
18 h light at 32 °C and 6 h dark at 28 °C. After 4 weeks’ growth, the leaf was collected both for targeted mutagenesis assay and whole-genome sequencing.

**Mutation detection and analysis**

The genomic DNA was extracted using the CTAB method (Stewart and Via, 1993). About 100 ng genomic DNA and a 50 μL PCR reaction was used to amplify the transgene and target sequence for detection of transgenic plants and genome editing events. The oligos used in this study were shown in Table S4. PCR was done with 2x Rapid Taq Mix (Vazyme) and examined using SSCP strategy (Zheng et al., 2016). The genotype at the target sites of each plant was confirmed by Sanger sequencing.

**Whole-genome sequencing and data analysis**

One gram of fresh leaves was obtained from each edited rice plant for WGS. Genomic DNA was extracted using Plant Genome DNA Kit (Tiangen). All plant samples were sequenced by the Illumina NovaSeq platform (Novogene, Beijing, China). The average sequencing clean data generated for each sample was 20 Gb, with the average depth being ~50× to ~70×. For data processing, adapters and low-quality reads were first trimmed and filtered using SKEWER (v. 0.2.2) (Jiang et al., 2014). Cleaned reads were then mapped to rice reference sequence TIGR7 (MSU7) with BWA mem (v. 0.7.17) software (Li and Durbin, 2010). Picard (https://broadinstitute.github.io/picard/) software (v. 2.22.4) and Samtools (v. 1.9) (Li et al., 2009) were employed to mark duplicate reads and generate sort BAM files, respectively. The Genome Analysis Toolkit (GATK v. 3.8) (McKenna et al., 2010) was applied to realign the reads near INDELS and recalibrate base quality scores against known SNPs and INDELS databases (http://snp-seek.irri.org/). After the raw BAM files were processed by GATK, analysis-ready BAM files were generated. To identify genome-wide somatic mutations with high confidence, we applied three software each to identify SNVs and INDELs, respectively. Whole-genome SNVs were detected by LoFreq (v. 2.1.2) (Wilms et al., 2012), MuTect2 (Cibulskis et al., 2013) and VarScan2 (v. 2.4.3) (Koboldt et al., 2012). Whole-genome INDELS were detected by MuTect2 (Cibulskis et al., 2013), VarScan2 (v. 2.4.3) (Koboldt et al., 2012) and Pindel (v. 0.2) (Kim et al., 2018). The Bedtools (v. 2.27.1) (Li, 2011) was used to overlap all SNVs/INDELS among replicates or different software. SNVs and INDELS identified by all three corresponding software were retained for further analysis. Cas-OffFinder in silico (v. 2.4) (Ba et al., 2014) was used to predict putative off-target sites in the rice genome. The PAM type of SpRY, SpCas9-NG and xCas9 were set to NNNN, NGN and NGN, respectively, allowing up to 5-nt mismatches in the protospacer. IGV (v. 2.8.4) software (Thorvaldsdottir et al., 2013) was applied to visualize discovered mutations with the generated BAM and VCF files. To identify the insertion locations of T-DNA in each line, the cleaned reads were first aligned to the rice reference genome and vector sequences simultaneously. Then, the BAM files were visualized using the IGV software and ‘Group Alignments by’ mode was set to ‘chromosome of mate’ in IGV. Lastly, each T-DNA insertion site was confirmed by manual checking of paired reads aligned to both vector sequences and specific chromosomes. The genome-wide distribution of mutations was drawn by Circos (v.0.69) (Krzywinski et al., 2009). The adjacent 3-bp sequences of the A-to-G SNVs were extracted from the reference genome sequence, and then submitted to WebLogo3 (http://weblogo.threeplusone.com/; Crooks et al., 2004) to plot motif weblogo. Data processing, analyses and figure plotting were completed by using R and Python.

**Acknowledgements**

We thank Aimee Malzahn for proofreading this manuscript. This research was supported by the Sichuan Science and Technology Program (award no. 2021JRC0032, 2021YH0084 and 2021YFYZ20016) to J.Z. and Y.Z., the National Natural Science Foundation of China (award no. 32101205, 32072045 and 31960423) to X.T. and X.Z., the Technology Innovation and Application Development Program of Chongqing (award no. CSTC2021JSCX-CYLX0001) to X.T. and Y.Z., and the Open Foundation of Jiangsu Key Laboratory of Crop Genetics and Physiology (award no. YCSL202009) to I.Z. and Y.Z. T.Z. is also supported by the National Science Foundation Plant Genome Research Program (award no. IOS-2029889) and the U.S. Department of Agriculture Biotechnology Risk Assessment Grant Program competitive grant (award no. 2018-33522-28789) to Y.Q. S.S. is a Foundation for Food and Agriculture Research Fellow.

**Conflicts of interest**

The authors declare no competing interests.

**Authors’ contributions**

Y.Z. proposed the project. Y.Z., T.Z. and Y.Q. conceived and designed the experiments. Q.R., Z.Z., X.T. and S.S. made the vectors for rice transformation. Q.R. and Z.Z. conducted rice protoplast transformation and data analysis. Q.R., Z.Z., L.L., S.X. and X.Z. did the rice stable transformation and mutagenesis assays. Q.R., Z.Z., L.L., S.X. and J.Z. prepared rice seedling samples for WGS. Y.W. and G.L. performed WGS data analysis and X.Z. did the rice stable transformation and mutagenesis assays. Q.R., Z.Z., L.L., S.X. and J.Z. prepared rice seedling samples for WGS. Y.W. and G.L. performed WGS data analysis and generated the figures. Y.H., Y.B. and S.L. assisted with data analysis. Y.Z., T.Z. and Y.Q. supervised the research and wrote the manuscript. All authors participated in the discussion and revision of the manuscript.

**Data availability statement**

The WGS data have been deposited in the Sequence Read Archive in National Center for Biotechnology Information (NCBI) under the accession number PRJNA792795 and Beijing Institute of Genomics Data Center (http://bigd.big.ac.cn) under BioProject PRJCA007564.

**References**

Altpeter, F., Springer, N.M., Bartley, L.E., Blechl, A.E., Brutnell, T.P., Citovsky, V., Conrad, L.J. et al. (2018) Advancing crop breeding in the era of genome editing. Plant Cell, 28, 1510–1520.

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.

Bae, S., Park, I. and Kim, I.S. (2014) Cas-OffFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics, 30, 1473–1475.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S. et al. (2013) Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. 31, 213–219.
Off-target effects of PAM-relaxed Cas9 in rice

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D. et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. Science, 339, 819–823.

Crooks, G.E., Hon, G., Chandonia, J.M. and Brenner, S.E. (2004) WebLogo: A sequence logo generator. Genome Res. 14, 1188–1190.

Debernardi, J.M., Tricoli, D.M., Ercoli, M.F., Hayta, S., Ronald, P., Palatnik, J.F. and Dubovskiy, J. (2020) A GFP-GFP chimeric protein improves the regeneration efficiency of transgenic plants. Nat. Biotechnol. 38, 1274–1279.

Doman, J.L., Raguram, A., Newby, G.A. and Liu, D.R. (2020) Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. Nat. Biotechnol. 38, 620–628.

Endo, M., Miikami, M., Endo, A., Kaya, H., Itoh, T., Nishimatsu, H., Nureki, O. et al. (2019) Genome editing in plants by engineered CRISPR-Cas9 recognizing NGM. Nat. Plants, 5, 14–17.

Fossi, M., Amundson, K., Kuppu, S., Brit, A. and Comai, L. (2019) Regeneration of Solanum tuberosum plants from protoplasts induces widespread genome instability. Plant Physiol. 180, 78–86.

Gaudelli, N.M., Lam, D.K., Rees, H.A., Sola-Esteves, N.M., Barrera, L.A., Born, D.A., Edwards, A. et al. (2020) Directed evolution of adenine base editors with increased activity and therapeutic application. Nat. Biotechnol. 38, 892–900.

Hanusa, M.M., Zhang, Y., Yuan, G., De, K., Chen, J.G., Muchero, W., Tuskan, G.A. et al. (2021) Construct design for CRISPR/Cas-based genome editing in plants. Trends Plant Sci. 26(11), 1133–1152.

Hei, Y., Ohta, S., Komari, T. and Kumashiro, T. (1994) Efficient transformation of rice (Oryza sativa L.) mediated by Agrobacterium and sequence analysis of the boundaries of the T-DNA. Plant J. 6, 271–282.

Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M. et al. (2018) Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature, 556, 57–63.

Hua, K., Tao, X., Han, P., Wang, R. and Zhu, J.K. (2019) Genome engineering in rice using Cas9 variants that recognize NGM PAM sequences. Mol. Plant, 12, 1003–1014.

Jiang, H., Lei, R., Ding, S.W. and Zhu, S. (2014) Skewer: a fast and accurate short-read mapper. Nat. Methods, 11, 137–141.

Jin, S., Seo, J., Lee, K.S., Lee, J., Kim, S. et al. (2020) Rational design of APOBEC3B cytosine base editors with improved specificity. Mol. Cell, 79(5), 728–740.e6.

Jones, S.J. (2009) Circos: An information aesthetic for comparative genomics. Genome Res. 19, 1639–1645.

Kolesnikov, A., Knott, G.J., Palumbo, C.M., Lin-Shiao, E., Richter, M.F., Zhao, K.T., Beal, P.A. et al. (2020) DNA capture by a CRISPR-Cas9-guided adenine base editor. Science, 369, 566–571.

Li, J., Xu, R., Qin, R., Liu, X., Kong, F. and Wei, P. (2021) Genome editing mediated by SpCas9 variants with broad non-canonical PAM compatibility in plants. Mol. Plant, 14, 352–360.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M.J., Scelonne, C. et al. (2016) Morphogenic regulators baby boom and Wuschel improve monocot transformation. Plant Cell, 28(9), 1998–2015.

Maher, M.F., Nasti, R.A., Vollbrecht, M., Starker, C.G., Clark, M.D. and Voytas, D.F. (2020) Plant gene editing through de novo induction of meristems. Nat. Biotechnol. 38, 84–89.

Mail, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E. et al. (2013) RNA-guided human genome engineering via Cas9. Science, 339, 823–826.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garinella, K. et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 20, 1297–1303.

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.

Negishi, K., Kaya, H., Abe, K., Hara, N., Saika, H. and Toki, S. (2019) An adenine base editor with expanded targeting scope using SpCas9-N2gV1 in rice. Plant Biotechnol. J., 17, 1476–1478.

Nishimasa, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T. et al. (2018) Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science, 361, 1259–1262.

Nishimura, A., Aichi, I. and Matsuoka, M. (2006) A protocol for Agrobacterium-mediated transformation in rice. Nat. Protoc. 1, 2796–2802.

Qin, R., Li, J., Liu, X., Xu, R., Yang, J. and Wei, P. (2020) SpCas9-SG self-targets the sgRNA sequence in plant genome editing. Nat. Plants, 6, 197–201.

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, J., Meng, X., Hu, F., Liu, Q., Cao, Y., Li, H., Yan, C. et al. (2021a) Expanding the scope of genome editing with Spg and SpPy variants in rice. Science China Life Sci. 64(10), 1784–1787.

Ren, Q., Sretenovic, S., Liu, G., Zhong, Z., Wang, J., Huang, L., Tang, X. et al. (2021b) Improved plant cytosine base editors with high editing activity, purity, and specificity. Plant Biotechnol. J. 19(10), 2052–2068.

Ren, Q., Sretenovic, S., Liu, S., Tang, X., Huang, L., He, Y., Liu, L. et al. (2021c) PAM-less plant genome editing using a CRISPR-SpyX toolkit. Nat. Plants, 7, 25–33.

Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thrushel, B.W., Wilson, C. et al. (2020) Phase-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat. Biotechnol. 38, 883–891.

Rodriguez-Leal, D., Lemmon, Z.H., Man, J., Bartlett, M.E. and Lippman, Z.B. (2017) Engineering quantitative trait variation for crop improvement by genome editing. Cell, 171(2), 470–480.e8.

Stewart, C.N. Jr and Via, L.E. (1993) A rapid CTAB DNA isolation technique useful for RAPD fingerprinting and other PCR applications. Biotechniques, 14, 748–750.

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X. et al. (2018) A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Crpf1 (Cas12a) nucleases in rice. Genome Biol. 19, 84.

Tang, X., Lowder, L.G., Zhang, T., Malzahn, A.A., Zheng, X., Voytas, D.F., Zhong, Z. et al. (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. et al. (2019) Single transcript unit CRISPR 2.0 systems for robust Cas9 and Cas12a mediated plant genome editing. Plant Biotechnol. J. 17, 1431–1445.

Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 14, 178–192.

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.
Wang, B., Zhaohui, Z., Huanhuan, Z., Xia, W., Binglin, L., Lijia, Y., Xiangyan, H. et al. (2019a) Targeted mutagenesis of NAC transcription factor gene, OsNAC041, leading to salt sensitivity in rice. Rice Sci. 26, 98–108.

Wang, M., Wang, Z., Miao, Y., Lu, Y., Yang, R., Tao, X. and Zhu, J.K. (2019b) Optimizing base editors for improved efficiency and expanded editing scope in rice. Plant Biotechnol. J. 17, 1697–1699.

Wang, Z., Liu, X., Xie, X., Deng, L., Zheng, H., Pan, H., Li, D. et al. (2021) ABE8e with polycistronic trRNA-gRNA expression cassette significantly improves adenosine base editing efficiency in Nicotiana benthamiana. Int. J. Mol. Sci. 22, 5663.

Wu, C., Wang, C., Jia, M., Guo, H.X., Luo, P.Y., Wang, M.G., Zhu, J.K. et al. (2021) Efficient generation of homozygous substitutions in rice in one generation utilizing an rABE8e base editor. J. Integr. Plant Biol. 63, 1595–1599.

Wilm, A., Aw, P.P.K., Bertrand, D., Yeo, G.H.T., Ong, S.H., Wong, C.H., Khor, C.C. et al. (2012) LoFreq: a sequence-accuracy aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 40, 11189–11201.

Xing, S., Chen, K., Zhu, H., Zhang, R., Zhang, H., Li, B. and Gao, C. (2020) Fine-tuning sugar content in strawberry. Genome Biol. 21, 230.

Xu, Z., Kuang, Y., Ren, B., Yan, D., Yan, F., Spetz, C., Sun, W. et al. (2021) SpRY greatly expands the genome editing scope in rice with highly flexible PAM recognition. Genome Biol. 22, 6.

You, Q., Zhong, Z., Ren, Q., Hasan, F., Zhang, Y. and Zhang, T. (2018) CRISPRMatch: an automatic calculation and visualization tool for high-throughput CRISPR genome-editing data analysis. Int. J. Biol. Sci. 14, 858–862.

Yu, Y., Leete, T.C., Born, D.A., Young, L., Barrera, L.A., Lee, S.J., Rees, H.A. et al. (2020) Cytochrome b6f base editors with minimized unguided DNA and RNA off-target events and high on-target activity. Nat. Commun. 11, 2052.

Zeng, D., Li, X., Huang, J., Li, Y., Cai, S., Yu, W., Li, Y. et al. (2020a) Engineered Cas9 variant tool expand targeting scope of genome and base editing in rice. Plant Biotechnol. J. 18, 1348–1350.

Zeng, D., Liu, T., Tan, J., Zhang, Y., Zheng, Z., Wang, B., Zhou, D. et al. (2020b) PhCeCBEs: Plant high-efficiency cytidine base editors with expanded target range. Mol. Plant. 13, 1666–1669.

Zhang, C., Wang, Y., Wang, F., Zhao, S., Song, J., Feng, F., Zhao, J. et al. (2021) Expanding base editing scope to near-PAMless with engineered CRISPR/Cas9 variants in plants. Mol. Plant. 14, 191–194.

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. et al. (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. and Zhou, J. et al. (2016) Effective screen of CRISPR/Cas9-induced mutants in rice by single-strand conformation polymorphism. Plant Cell Rep. 35, 1545–1554.

Zhong, Z., Sretenovic, S., Ren, Q., Yang, L., Bao, Y., Qi, C., Yuan, M. et al. (2019) Improving plant genome editing with high-fidelity xCas9 and non-canonical PAM-targeting Cas9-NG. Mol. Plant. 12, 1027–1036.

Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L. et al. (2019a) Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. Nature, 571, 275–278.

Zhou, J., Deng, K., Cheng, Y., Zhong, Z., Tian, L., Tang, X., Tang, A. et al. (2017) CRISPR-Cas9 based genome editing reveals new insights into microRNA function and regulation in rice. Front. Plant Sci. 8, 1598.

Zhou, J., Xin, X., He, Y., Chen, H., Li, Q., Tang, X., Zhong, Z. et al. (2019b) Multiplex QTL editing of grain-related genes improves yield in elite rice varieties. Plant Cell Rep. 38, 475–485.

Zhou, J., Yuan, M., Zhao, Y., Quan, Q., Yu, D., Yang, H., Tang, X. et al. (2021) Efficient deletion of multiple circle RNA loci by CRISPR-Cas9 reveals Os06circ02797 as a putative sponge for OsMIR408 in rice. Plant Biotechnol. J. 19, 1240–1252.

Zhu, J., Zhang, R., Jia, X., Tang, X., Guo, Y., Yang, H., Zhang, X. et al. (2022) CRISPR-Cas9 mediated OsMIR166a knockdown reveals its pleiotropy in rice. Plant Biotechnol. J. 20, 310–322.

Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L. et al. (2019) Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science, 364, 289–292.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Guide RNA-dependent off-target mutagenesis by SpRY

Figure S2 Guide RNA-dependent off-target mutagenesis by nSpRY-ABE8e at OsPDS-gr04-TAA site

Figure S3 Sequencing reads indicative for T-DNA self-editing by SpRY constructs

Figure S4 Sequencing reads indicative for T-DNA self-editing by nSpRY-PmCDAl constructs

Figure S5 Sequencing reads indicative for T-DNA self-editing by nSpRY-ABE8e constructs

Figure S6 INDEL mutations in all sequenced samples and their genome-wide distributions

Figure S7 Comparison of deletion sizes among all mutations induced by different genome editing systems

Figure S8 Genome-wide distribution of mutations (SNVs+INDELs) from all sequenced sample

Figure S9 Allele frequencies of A-to-G SNVs identified in nSpRY-ABE8e (n = 8) and nSpRY-ABE8e_backbone (n = 2) samples

Figure S10 Sequence signature of ABE8e based genome-wide off-target mutations

Figure S11 Venn diagram showing mutations shared between individual plants

Table S1 Vector and sample information.

Table S2 Summary of WGS samples and statistics.

Table S3 Genome-wide analysis of guide RNA-dependent off-target mutations.

Table S4 Oligos used in this study.