



Article

# Exploring C-to-G and A-to-Y Base Editing in Rice by Using New Vector Tools

Dongchang Zeng<sup>1,2,†</sup>, Zhiye Zheng<sup>1,†</sup>, Yuxin Liu<sup>1</sup>, Taoli Liu<sup>1</sup>, Tie Li<sup>1</sup>, Jianhong Liu<sup>1</sup>, Qiyu Luo<sup>1</sup>, Yang Xue<sup>1</sup>, Shengting Li<sup>1</sup>, Nan Chai<sup>1</sup>, Suize Yu<sup>1</sup>, Xianrong Xie<sup>1,2</sup>, Yao-Guang Liu<sup>1,2</sup> and Qinlong Zhu<sup>1,2,\*</sup> 

- <sup>1</sup> State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Life Sciences, South China Agricultural University, Guangzhou 510642, China; dong@scau.edu.cn (D.Z.); xinaxin@stu.scau.edu.cn (Z.Z.); liuyuxin@stu.scau.edu.cn (Y.L.); liutaoli@stu.scau.edu.cn (T.L.); litie@stu.scau.edu.cn (T.L.); liujianhong@stu.scau.edu.cn (J.L.); qiyuluo@scau.edu.cn (Q.L.); xueyang@stu.scau.edu.cn (Y.X.); lishengting@stu.scau.edu.cn (S.L.); chinan@scau.edu.cn (N.C.); ysz@stu.scau.edu.cn (S.Y.); xiexianrong@scau.edu.cn (X.X.); ygliu@scau.edu.cn (Y.-G.L.)
- <sup>2</sup> Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou 510642, China
- \* Correspondence: zhuql@scau.edu.cn
- † These authors contributed equally to this work.

**Abstract:** CRISPR/Cas9-based cytosine base editors (CBEs) and adenine base editors (ABEs) can efficiently mediate C-to-T/G-to-A and A-to-G/T-to-C substitutions, respectively; however, achieving base transversions (C-to-G/C-to-A and A-to-T/A-to-C) is challenging and has been rarely studied in plants. Here, we constructed new plant C-to-G base editors (CGBEs) and new A-to-Y (T/C) base editors and explored their base editing characteristics in rice. First, we fused the highly active cytidine deaminase evoFENRY and the PAM-relaxed Cas9-nickase variant Cas9n-NG with rice and human uracil DNA N-glycosylase (rUNG and hUNG), respectively, to construct CGBE-rUNG and CGBE-hUNG vector tools. The analysis of five NG-PAM target sites showed that these CGBEs achieved C-to-G conversions with monoallelic editing efficiencies of up to 27.3% in T<sub>0</sub> rice, with major byproducts being insertion/deletion mutations. Moreover, for the A-to-Y (C or T) editing test, we fused the highly active adenosine deaminase TadA8e and the Cas9-nickase variant SpGn (with NG-PAM) with *Escherichia coli* endonuclease V (EndoV) and human alkyladenine DNA glycosylase (hAAG), respectively, to generate ABE8e-EndoV and ABE8e-hAAG vectors. An assessment of five NG-PAM target sites showed that these two vectors could efficiently produce A-to-G substitutions in a narrow editing window; however, no A-to-Y editing was detected. Interestingly, the ABE8e-EndoV also generated precise small fragment deletions in the editing window from the 5'-deaminated A base to the SpGn cleavage site, suggesting its potential value in producing predictable small-fragment deletion mutations. Overall, we objectively evaluated the editing performance of CGBEs in rice, explored the possibility of A-to-Y editing, and developed a new ABE8e-EndoV tool, thus providing a valuable reference for improving and enriching base editing tools in plants.

**Keywords:** CGBE; ABE8e-EndoV; CBE; ABE; CRISPR/Cas9; rice



**Citation:** Zeng, D.; Zheng, Z.; Liu, Y.; Liu, T.; Li, T.; Liu, J.; Luo, Q.; Xue, Y.; Li, S.; Chai, N.; et al. Exploring C-to-G and A-to-Y Base Editing in Rice by Using New Vector Tools. *Int. J. Mol. Sci.* **2022**, *23*, 7990. <https://doi.org/10.3390/ijms23147990>

Academic Editor: Zsófia Bánfalvi

Received: 10 June 2022

Accepted: 18 July 2022

Published: 20 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cytosine base editors (CBEs) and adenine base editors (ABEs) based on CRISPR/Cas9 variants have applications in animal and plant functional genomics and genetic engineering. From BE1 to evoBE4max of the different generations of CBEs [1], researchers have mainly paid attention to improving C-to-T substitution efficiency and developing a series of effective CBEs, such as PhieCBEs [2], hA3A-PBE [3], Anc689BE4max-nCas9 [4], and DisSUGs [5]. From TadA7.10 to TadA8e [6], researchers have improved A-to-G substitution efficiency and developed efficient ABEs, such as PhieABEs [7], rBE53~rBE65 [8], SpG-ABE8e and SpRY-ABE8e [9–11], PTG-ABE8e [12], and ABEmax-nCas9 [4]. Efforts have been made to improve the efficiency of ABEs and CBEs for achieving high-efficiency and stable base

editing in plant genomes, and these tools have been widely used in gene function research and crop genetic improvement [13–18].

The currently available base editors can implement C-to-T/G-to-A and A-to-G/T-to-C substitutions; however, achieving base transversion remains challenging. For CBEs, cytosine (C) is deaminated to uracil (U), and U is recognized as thymine (T) at the DNA level, leading to C-to-T substitution after DNA repair and replication [19]. However, via the base excision repair (BER) pathway in eukaryotic cells, the U bases of U:G mismatches can be easily removed by uracil DNA N-glycosylase (UNG), thus resulting in apurinic/aprimidinic (AP) sites. The AP sites are hydrolyzed by AP lyase via DNA repair to achieve C-to-G/C-to-A transversions and insertions/deletions (indels). Based on the principle of BER, the uracil DNA glycosylase inhibitor (UGI) domain in the CBE system is replaced by the UNG domain, and high-efficiency C-to-G transversion is produced in mammalian cells [20,21]. Although studies have evaluated CGBE editing in protoplasts and stable transformations in rice and poplar [22,23], the target sites and data are insufficient to objectively assess the editing performance of CGBE in plants.

For ABEs, adenine (A) in the editing window is deaminated to inosine (I), and I is recognized as guanine (G) at the DNA level, ultimately achieving A-to-G substitution after DNA repair and replication [19]. Because of the lack or low activity of glycosylase on the I base in eukaryotic cells, the ABE system contains few A-to-Y (C or T) or indel byproducts [6,7]. However, a study indicated that cells are subjected to natural hydrolysis and nitrosative stress upon exposure to endogenous and/or exogenous agents, leading to the deamination of A to I and I:T mismatches that stimulate both BER and alternative excision repair (AER) pathways [24]. In the BER pathway, alkyladenine DNA glycosylase (AAG) catalyzes the cleavage of the glycosidic bond of inosine to remove the I base from the DNA strand [24,25]. In the AER pathway, endonuclease V (EndoV) hydrolyzes the second phosphodiester bond located at 3' to inosine in the DNA or RNA strand [24,26–28]. In the *E. coli* model, an EndoV-dependent AER pathway for the removal of deoxyinosine from DNA has been identified. However, the AER pathway has not yet been identified in mammalian cells. According to a biochemical investigation of mammalian EndoV, it possesses considerable inosine 3' endonuclease activity on inosine-containing single-stranded DNA (ssDNA) and poor activity on double-stranded DNA (dsDNA) [24,26–28]. Studies on the BER and AER pathways based on I:T mismatches in the ABE system have been limited. Therefore, we integrated hAAG and EndoV by using ABE tools and explored whether they could achieve A-to-Y or A-to-indel editing.

In this study, based on the U:G mismatch BER pathway, we used rice UNG (rUNG) and human UNG (hUNG) to construct PevoFRNY-Cas9n-NG-rUNG (CGBE-rUNG) and PevoFRNY-Cas9n-NG-hUNG (CGBE-hUNG), respectively, and evaluated their C-to-G editing performance in rice. In addition, based on the I:T mismatch repair pathway, we used human AAG (hAAG) of the BER pathway and *E. coli* EndoV (EndoV) of the AER pathway to construct ABE8e-hAAG and ABE8e-EndoV vectors and explored their editing performance of A-to-Y transversions or indels in rice.

## 2. Results

### 2.1. CGBE-rUNG and CGBE-hUNG Achieved C-to-G Base Editing

The amino acid sequences of UNGs are highly conserved between prokaryotes and eukaryotes; hence, UNGs from humans (1AKZ\_A), rice (XP\_015634353.1), Arabidopsis (NP\_188493.1), and *E. coli* (EFN1850840.1) were screened by NCBI BLAST. We selected hUNG reported in mammalian cells and rUNG in plant cells as the candidates. We first optimized the nucleic acid sequences of hUNG and rUNG according to the plant codon preferences of rice, followed by chemical synthesis of these sequences (Sequence S1). PevoFRNY-Cas9n-NG, with high efficiency and broader genome-targeting protospacer-adjacent motifs (PAMs) (recognizes NG-PAM), was selected as the basic vector for further modification [2]. The 2xUGI domains of PevoFRNY-Cas9n-NG were replaced by hUNG and rUNG, respectively, to construct PevoFRNY-Cas9n-NG-rUNG (CGBE-rUNG) and

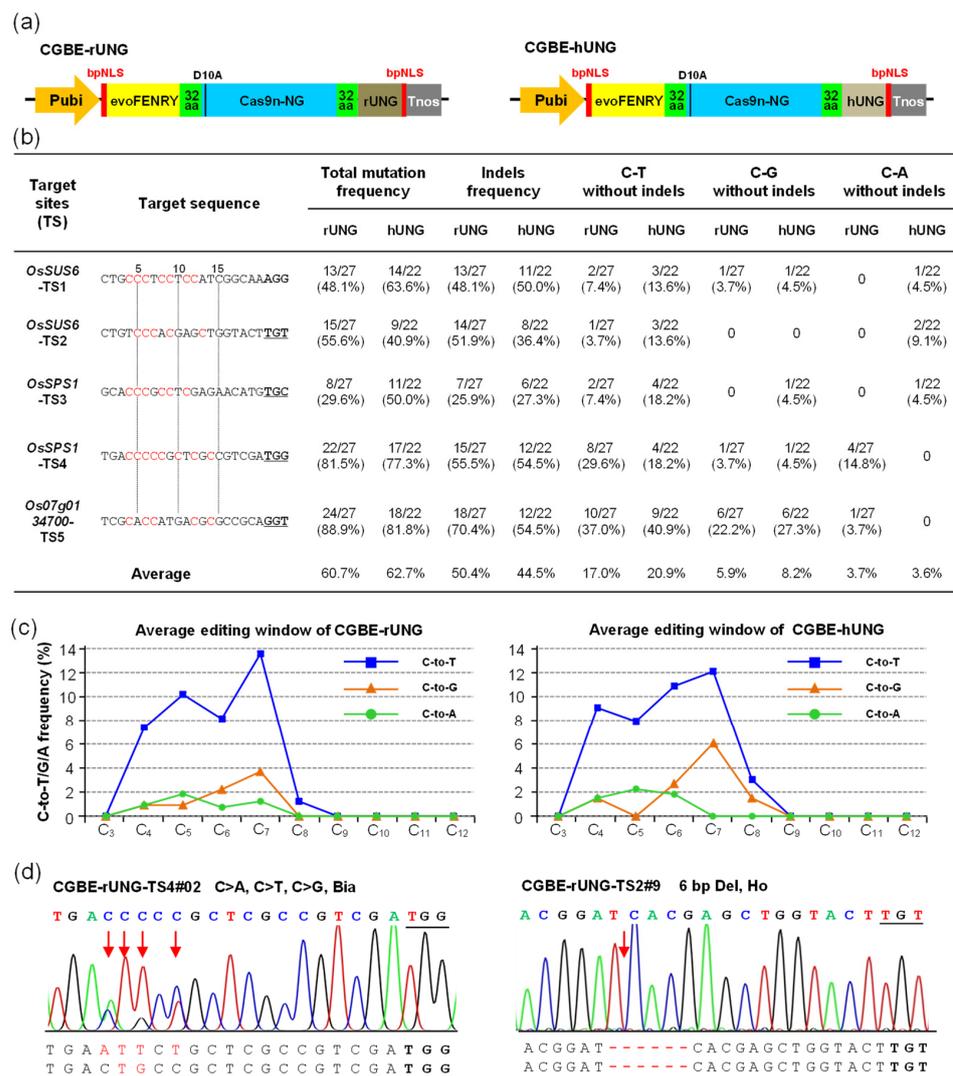
PevoFRNY-Cas9n-NG-hUNG (CGBE-hUNG) (Figure 1a). CGBE-rUNG and CGBE-hUNG have a structure similar to that of the reported CGBE system [20]. Five NG-PAM targets from the rice genome with sgRNA were assembled into CGBE-rUNG and CGBE-hUNG to explore and compare their editing performance in rice. The overall editing efficiencies of CGBE-rUNG (29.6–88.9%) and CGBE-hUNG (40.9–81.8%) were measured in five targets, with the average editing efficiency being 60.7% and 62.7%, respectively (Figure 1b). However, the mutation types were mainly C-to-indels with both CGBE-rUNG (50.4%) and CGBE-hUNG (44.5%) editing (Figure 1b, Tables S1 and S2). We then calculated the C-to-T/G/A editing efficiency without indels, as indels generally cause nonsense mutations. To ensure the inheritance of the base-edited mutants, the alleles containing C-to-T/G/A substitutions with indels were excluded, such as CGBE-rUNG-TS1#06 and CGBE-rUNG-TS5#07 Allele 1 (A1) (Tables S1 and S2). The base editing results showed that the average C-to-T/G/A base editing efficiency of CGBE-hUNG (20.9%/8.2%/3.6%) was slightly higher than that of CGBE-rUNG (17.0%/5.9%/3.7%); however, both CGBE-hUNG and CGBE-rUNG base editing types were mainly C-to-T substitutions, and the C-to-G transversion average efficiency was less than 10% (Figure 1b). These results are inconsistent with those obtained in mammalian cells, in which CGBE systems significantly reduced C-to-T substitution and improved C-to-G transversion without excessive indel byproducts [20]. However, our results are almost consistent with those of studies on rice and poplar [22,23]. The allelic variations in CGBE-rUNG and CGBE-hUNG are mainly biallelic mutations and heterozygous mutations, with homozygous mutations in a small part (Figures 1d, S1 and S2). The C-to-T editing window of CGBE-rUNG and CGBE-hUNG was mainly C<sub>4</sub>-C<sub>7</sub>, which is narrower than that of PevoFRNY-Cas9n-NG (C<sub>3</sub>-C<sub>8</sub>, extended to C<sub>3</sub>-C<sub>10</sub>) [2]. The C-to-G editing window of for CGBE-rUNG (C<sub>4</sub>-C<sub>7</sub>) was slightly narrower than that of CGBE-hUNG (C<sub>4</sub>-C<sub>8</sub>); however, the C-to-A window of CGBE-rUNG (C<sub>4</sub>-C<sub>7</sub>) was slightly wider than that of CGBE-hUNG (C<sub>4</sub>-C<sub>6</sub>) (Figures 1c and S1).

To determine whether replacing the UGI domain with the UNG domain would increase the off-target effects, we selected off-target sequences from five targets of CGBE with a mismatch of  $\leq 3$  bases as candidate off-target sites for further sequencing analysis (Figure S3a) [29–31]. The sequencing results showed that among the 15 candidate off-target sites of CGBE, except for TS4-off2 of CGBE-hUNG (4.5%) with a discovered mutation, no mutations were found in other candidate off-target sites (Figure S3a). The results suggested that the candidate off-target sequence, which had more than two bases' difference from the on-target sequence of the UNG-based CGBE system, also did not increase the off-target effects. Thus, in terms of the C-to-G editing window, editing efficiency, and product purity, the performance of hUNG was slightly superior to that of rUNG. Although the C-to-G efficiency of CGBE-TS5 was as high as 22.2% (rUNG) and 27.3% (hUNG), achieving C-to-G transversion without byproducts at most of the targets is a great challenge (Figure 1b).

## 2.2. A-to-Y Transversions Were Not Detected by ABE8e-EndoV and ABE8e-hAAG

hAAG and EndoV play a role in the I:T mismatch BER and AER pathways; hence, we selected human AAG (hAAG) and *E. coli* EndoV (EndoV) to optimize codons and chemically synthesize their DNA sequences according to the codon preferences of rice (Sequence S2). The ABE8e-SpG developed by our team with high efficiency and broad targets (to recognize NG-PAM) was selected as the basic vector for further modification [7]. hAAG and EndoV were fused to the C-terminus of ABE8e-SpG to construct ABE8e-SpGn-EndoV (ABE8e-EndoV) and ABE8e-SpGn-hAAG (ABE8e-hAAG), respectively (Figure 2a). Five targets from the rice genome were selected to explore their editing performance; ABE8e-EndoV and ABE8e-hAAG exhibited stable editing efficiency in the five targets, and the overall average editing efficiency of ABE8e-EndoV (62.7%) was slightly higher than that of ABE8e-hAAG (55.6%). Moreover, the average A-to-G editing efficiency of ABE8e-EndoV (56.7%) and ABE8e-hAAG (55.6%) without indels was the same as that of ABE8e-SpG (59.7%), indicating that the addition of the hAAG and EndoV domains did not affect the deaminase activity of ABE8e-SpG [7]. In addition to A-to-G substitution, ABE8e-EndoV generated monoallelic

indel mutations with an average efficiency of 17.3%, indicating that EndoV works on ABE system-mediated inosine repair. However, for ABE8e-hAAG, except for one indel mutation (4.3%) at TS4, the other mutants were A-to-G substitutions, which might be because of the low activity of hAAG. A-to-Y transversions were not detected in ABE8e-EndoV and ABE8e-hAAG, probably because I bases are strictly recognized as G at the DNA level or the activity of hAAG and EndoV is not sufficiently high (Figure 2b). The allelic variations in ABE8e-EndoV and ABE8e-hAAG were mainly biallelic and heterozygous mutations, with homozygous mutations in a small part (Figure S3b and Tables S2 and S3). The A-to-G editing window of ABE8e-EndoV and ABE8e-hAAG was concentrated in A<sub>4</sub>–A<sub>7</sub> (extended to A<sub>4</sub>–A<sub>8</sub>). Moreover, ABE8e-EndoV and ABE8e-hAAG had narrower editing windows than hyABE8e-SpGn (A<sub>4</sub>–A<sub>8</sub>, extended to A<sub>4</sub>–A<sub>11</sub>) [7] (Figure 2c).



**Figure 1.** The editing efficiency and mutation types of CGBE-rUNG and CGBE-hUNG in T<sub>0</sub> rice. (a) Structures of the CGBE-rUNG and CGBE-hUNG expression cassettes. evofERNY, a cytosine deaminase with efficient C-T substitution; Cas9n-NG, an SpCas9 variant that recognizes NG-PAM and only has cleavage activity on the target strand; rUNG, uracil N-glycosylase from rice; hUNG, uracil N-glycosylase from humans. (b) The efficiency and mutation type of CGBE-rUNG and CGBE-hUNG in T<sub>0</sub> rice, total mutations, and indel frequency were calculated by edited plants/total transgenic positive plants; for C-to-T, C-to-G, and C-to-A substitutions, base substitutions with indel alleles were eliminated. (c) The average editing window and editing efficiency of C-to-T/G/A substitution; the 20th base at the 5' end of PAM is numbered C<sub>1</sub>. (d) Sanger sequencing chromatogram for some typical mutation types. Bia, biallelic mutation.



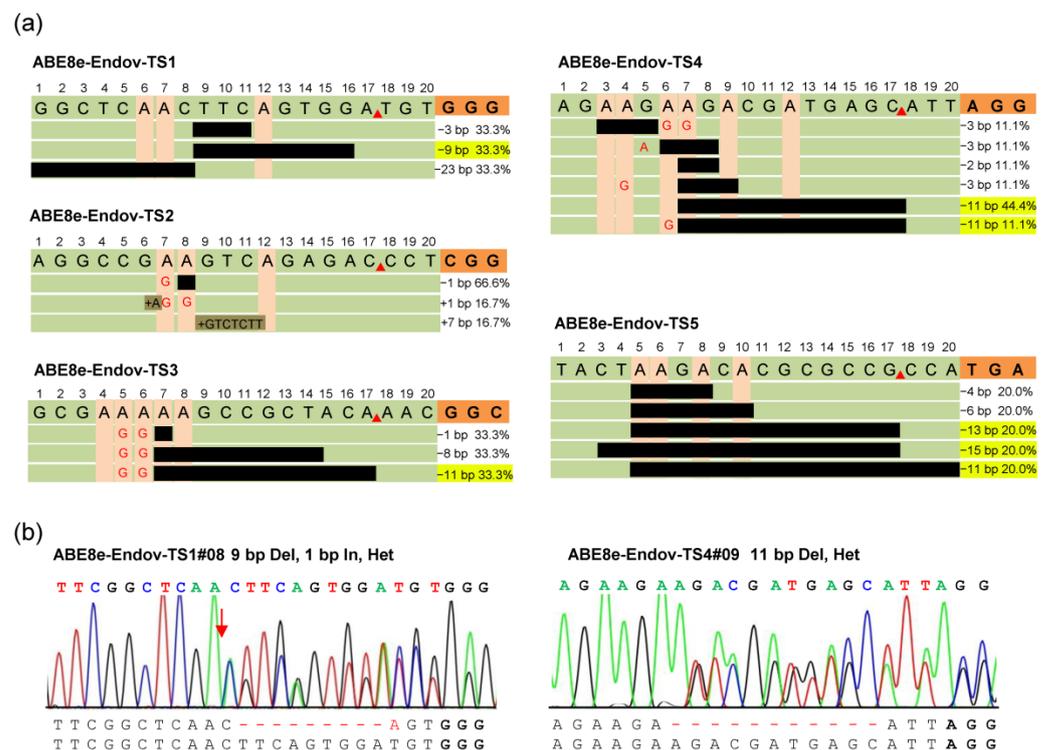
**Figure 2.** The editing efficiency and mutation type of ABE8e-EndoV and ABE8e-hAAG in T<sub>0</sub> rice. (a) Expression cassette structures of ABE8e-EndoV and ABE8e-hAAG. TadA8e, adenine deaminase with efficient A-to-G substitution; SpGn, an SpCas9 variant that recognizes NG-PAM and inactivates the cleavage activity of non-target strands; EndoV, *E. coli* endonuclease V, which works in the inosine alternative excision repair (AER) pathway, hydrolyzes the second phosphodiester bond located at 3' to inosine in the DNA strand; hAAG, human alkyladenine DNA glycosylase, which works in the inosine BER pathway, removes inosine, and generates an apurinic/aprimidinic (AP) site. (b) The efficiency and mutation type of ABE8e-EndoV and ABE8e-hAAG in T<sub>0</sub> rice, total mutations, and indel frequency were calculated by editing plant/total transgenic positive plants; for the A-to-G, A-to-C, and A-to-T substitutions, the base substitutions with indel alleles were eliminated. (c) Average editing windows and average editing efficiency of A-to-G/C/T substitution.

To evaluate whether fusing EndoV and hAAG domains would increase the off-target efficiency, we used the same off-target analysis method as that used for CGBEs and screened 13 candidate off-target sites (Figure S3b). The sequencing results showed that except for one mutant (TS2-off1: 3.3%) of ABE8e-EndoV with off-target mutation, no mutations were found in other candidate off-target sites (Figure S3b). The off-target analysis of ABE8e-EndoV and ABE8e-hAAG also indicated that the fusion of EndoV and hAAG with ABE8e did not result in a high frequency of off-target mutations.

### 2.3. ABE8e-EndoV Produces Predictable Small DNA Fragment Deletions

CRISPR/Cas9 has HNH and RuvC nuclease domains, which cleave the target and non-target strands, respectively, leading to DNA double-strand breaks and random insertion and deletion mutations. In animal and plant genomes, several *cis*-acting elements composed of small-nucleotide fragments play a key role in the spatiotemporal-specific expression of genes [16]. CRISPR/Cas9 requires NGG-PAM recognition and random indel production, which is a challenge for functional studies of specific small nucleotide fragments. In this study, ABE8e-EndoV was detected with an average of 17.3% indel mutations. Further

analysis of these indels indicated that ABE8e-EndoV generated the precise deletion of small fragments from 5'-deaminated A bases to the SpGn nickase cleavage site in TS1 (33.3%), TS3 (33.3%), TS4 (55.5%), and TS5 (60%) (Figure 3a). The deletion window mainly focuses on the A<sub>5</sub>-A<sub>8</sub> to the SpGn cleavage site, leading to the deletion of 9–13 bp nucleotides (Figure 3a,b). These results indicated that ABE8e-EndoV could also accurately predict the deletion of small fragments, in addition to an efficient A-to-G substitution, and the principle of predicting precise deletion is similar to that of the previously reported APOBEC-Cas9 fusion-induced deletion systems [32]. ABE8e-EndoV can be applied to the study of A/T-rich elements. In addition, ABE8e-EndoV was fused with a broadly targeted Cas9 variant, SpGn, which helped achieve precise base editing and small-fragment deletions.

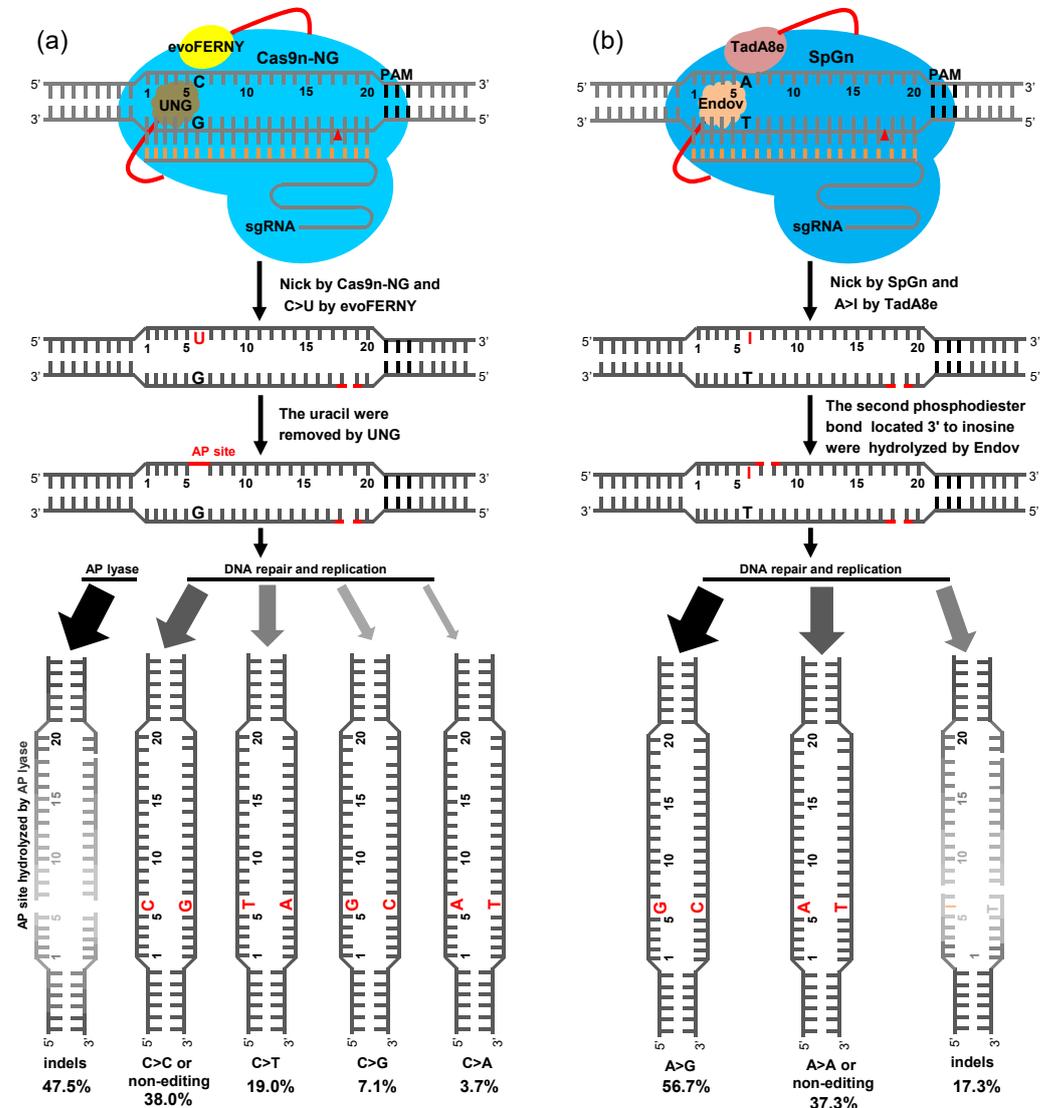


**Figure 3.** ABE8e-EndoV exhibits predictable small DNA fragment deletions. (a) ABE8e-EndoV generates precise deletions from the 5'-deaminated A base to the SpGn nicking site. The proportion of deletions versus total indels is shown below PAM, and the uniform precise deletion ratio is highlighted in yellow; PAMs are highlighted in orange, and black bars represent the deleted nucleotide fragment. (b) Sanger sequencing chromatogram for some typical mutation types of ABE8e-EndoV. Het, heterozygous mutation; Del, deletion; In, insertion.

#### 2.4. BER Pathway of Plant CGBE and ABE8e-EndoV

We objectively evaluated the editing features of CGBE in rice. To more intuitively reveal the repair pathway of CGBE and ABE8e-EndoV, we summarize their base repair processes. In the U:G BER pathway of CGBE, C bases were deaminated by cytosine deaminase evoFERNY to U, U bases were excised by rUNG or hUNG to form an AP site, and then two repair pathways were initiated in the plants. The first pathway involves hydrolysis of the AP site by AP lyase and its subsequent activation by non-homologous end repair, resulting in several indels (47.5%). The position of indel mutations is mainly in the AP site within the editing window, most deletions extend to the 3'-terminus nick site of Cas9n-NG, and a small part extends to the 5'-terminus of the AP site (Figure 4a and Tables S1 and S2). In the second pathway, the AP site generates C-to-T (19.0%), C-to-G (7.1%), and C-to-A (3.7%) substitutions or the original C base after undergoing DNA repair and replication pathways (Figure 4a). Although 2xUGI was replaced by rUNG or hUNG, the base transversion efficiency of C-to-G and C-to-A was slightly improved, and that of

C-to-T was significantly reduced. These results indicate that achieving base transversions without generating excessive indels in the editing window, such as in CGBE-TS1, CGBE-TS2, and CGBE-TS3, is challenging (Figures 1a and S1).



**Figure 4.** Plant CGBE- and ABE8e-EndoV-mediated excision repair pathway. (a) In the U base excision repair (BER) pathway of CGBE-rUNG and CGBE-hUNG, C bases in the editing window are deaminated by cytosine deaminase evoFERNY to U, U bases are excised by UNG to form an apurinic/aprimidinic (AP) site, and then two repair pathways are initiated. In one of these pathways, the AP site is hydrolyzed by AP enzymes and activated by non-homologous end repair, leading to several indels, whereas in the other pathway, the AP site generates C-to-T, C-to-G, and C-to-A substitutions or restores original C after DNA repair and replication pathways. (b) The I base is the alternative BER (AER) pathway of ABE8e-EndoV. Bases are deaminated by adenine deaminase TadA8e to I, and the second phosphodiester bond located at the 3'-terminus of I bases is hydrolyzed by EndoV, resulting in A-to-G substitutions or the restoration of original A after DNA repair and replication. In addition, some indel mutations are generated after non-homologous end-joining repair. The gray area indicates the degree to which the indel frequency is generated.

In the editing of ABE8e-hAAG, only 4.3% deletion mutations were detected at the TS4 site (Figure 2b); hence, determining whether hAAG is less active or not active for the I:T BER pathway was difficult in this study. However, ABE8e-EndoV certainly played a key role in the I:T AER pathway; hence, only the AER pathway of ABE8e-EndoV is shown

here. In the ABE8e-EndoV editing system, A bases were deaminated by adenine deaminase TadA8e to I, the I:T mismatch activated the AER pathway, and the second phosphodiester bond located at the 3'-terminus of I was hydrolyzed by EndoV. Finally, I-to-G (56.7%) or I-to-original A base substitutions were achieved through DNA repair and replication. In addition, indel mutations (17.3%) from DNA double-strand break and non-homologous end-joining repair were achieved. The region of indel mutations was mainly in the I base site within the editing window; a part of the deletion extended to the 3'-terminal nick site of SpG, and another part extended to the 5'-terminus of the I base (Figure 4b and Tables S3 and S4). The cleavage site of EndoV was at the second base of the 3' of I, which did not generate an AP site with the breakage of the DNA strand. This could be the reason for the inefficiency of ABE8e-EndoV in generating the A-to-Y transversion.

### 3. Discussion

High-efficiency ABE and CBE base editing systems are vital in the study of gene function, genetic improvement, de novo domestication, and alternative splicing of mRNA [19,33]. However, base substitution types in ABE and CBE systems are limited and cannot meet the needs of various applications. Although CGBEs in mammalian cells effectively switch C-to-T substitutions to C-to-G transversions without causing more indel byproducts and effectively expand the applied range of base editors in animal genomes, in this study, both CGBE-rUNG and CGBE-hUNG produced indel byproducts, and the average C-to-G monoallelic editing efficiency, excluding indel mutations, was 7.1%; hence, CGBEs in plants should be further improved. Koblan et al. tested different UNG orthologs and found that UNG from *Mycobacterium smegmatis* (UdgX) moderately improved the purity of the C-to-G transversion product and the transversion efficiency, and the authors also reported that DNA polymerase D2 (POLD2) and RNA-binding motif protein X-linked (RBMX) would further improve the C-to-G editing efficiency [34]. Kurt et al. fused rAPOBEC1 (R33A) and *E. coli* eUNG to different positions of Cas9n, producing eUNG-rAPOBEC1-Cas9n (CGBE1) or removing the 2xUGI domain without adding the UNG domain, resulting in a high C-to-G editing efficiency [21]. However, whether these methods are effective for CGBE improvement in plants remains to be determined. Nevertheless, we speculated that the high activity of AP lyase in plants is one of the reasons for the significant indel byproducts. Inhibition of the AP lyase activity may effectively decrease the frequency of indels and increase the frequency of C-to-G transversion. We speculate that another reason for the significant indels is that consecutive C bases ( $\geq 2$ ) within the editing window were synchronously deaminated to U bases in C-rich targets, and more than two consecutive U bases were removed by UNG to form consecutive AP sites. Multiple AP sites simultaneously hydrolyzed by AP lyase were more likely to result in high-frequency indels. However, in this case, we could not change the GC environment of the target sequence.

In eukaryotes, the low activity of BER enzymes for I:T mismatches results in the high purity of the ABE editing product. However, recent studies have reported that I:T mismatches will activate the BER and AER pathways [24–28]. In this study, ABE8e-EndoV, based on the AER pathway, achieved an A-to-indel breakthrough, which provides a basis for generating A-to-Y transversions in the next step. ABE8e-hAAG, based on the BER pathway, only detected 4.3% indels at TS4, suggesting that hAAG might play a role in I excision repair, but hAAG activity is possibly ineffective and requires further optimization. Using a protein-assisted evolution system is a common strategy to enhance the activity of hAAG and obtain superior AAG [1]. Another strategy is the screening of homologous AAGs from other species with high activity. Regarding the inefficiency of ABE8e-EndoV in producing A-to-Y transversion, we speculate that the cleavage site of EndoV at the second base of the 3' of I, which does not generate an AP site with the DNA strand break, is the main reason. Based on the aforementioned analysis, we hypothesize that A-to-Y transversion could be achieved by screening high-activity AAG and constructing highly active ABE8e-AAG based on the BER pathway.

Overall, we systematically explored the editing performance of the CGBE system in rice, although we did not observe prominent C-to-G performance, as observed in previous studies on mammalian cells. We found some differences in the U base excise repair system between animals and plants, and our results provide a valuable reference for the improvement of CGBEs in plants. In the process of exploring the A-to-Y system, we found that ABE8e-EndoV can generate predictable precise small-fragment deletions of 9–13 bps, in addition to generating efficient A-to-G substitutions at the target sites. We conclude that ABE8e-EndoV can facilitate the study of *cis*-element and protein domains in plants.

#### 4. Materials and Methods

##### 4.1. Construction of Plant CGBE-rUNG, CGBE-hUNG, ABE8e-hAAG, and ABE8e-EndoV Vectors

To construct CGBE-rUNG and CGBE-hUNG vectors, we performed an NCBI BLAST search to obtain the amino acid sequences of UNG from humans (Sequence ID: 1AKZ\_A) and rice (Sequence ID: XP\_015634353.1) and then codon-optimized hUNG and rUNG amino acid sequences to nucleotide sequences based on the codon preference of rice, which were chemically synthesized by GeneCreate (Wuhan, China) (Sequence S1). PevoFERNY-Cas9n-NG preserved by our team was used as the basic vector for modification [2], and 2xUGI of evoFERNY-Cas9n-NG was replaced by rUNG and hUNG with overlapping PCR to obtain two fragments, namely, evoFERNY-Cas9n-NG-rUNG and evoFERNY-Cas9n-NG-hUNG. The two fragments were digested by *Pst* I and *Bam*H I. Subsequently, pYLCRISPR/Cas9Pubi-H developed by our team was digested with *Pst* I and *Bam*H I [35], and the 1300 binary vector backbone was extracted through agarose gel electrophoresis. Finally, the evoFERNY-Cas9n-NG-rUNG and evoFERNY-Cas9n-NG-hUNG fragments were ligated to the 1300 backbone by using T4 DNA ligase, and PevoFERNY-Cas9n-NG-rUNG (CGBE-rUNG) and PevoFERNY-Cas9n-NG-hUNG (CGBE-hUNG) vectors were obtained.

To construct ABE8e-hAAG and ABE8e-EndoV vectors, we also searched the protein sequences of hAAG from humans (Sequence ID: 1BNK\_A) and EndoV from *E. coli* (Sequence ID: EGI38768.1) by using NCBI BLAST and then optimized the nucleic acid sequences according to the rice codon preferences and chemically synthesized them (Sequence S2). ABE8e-SpG preserved by our team was used as the basic vector for further modification [7]. hAAG and EndoV were respectively fused to the C-terminus of ABE8e-SpG by overlapping PCR. Finally, the ABE8e-SpGn-hAAG and ABE8e-SpGn-EndoV fragments were assembled into the 1300 binary vector through *Pst* I and *Bam*H I digestion and T4 DNA ligation to construct ABE8e-SpGn-hAAG (ABE8e-hAAG) and ABE8e-SpGn-EndoV (ABE8e-EndoV).

##### 4.2. Target Design and gRNA Expression Cassette Assembly

The five target sequences of CGBE-rUNG and CGBE-hUNG were derived from the coding DNA sequence regions of *LOC\_Os02g58480* (*OsSUS6*), *LOC\_Os01g69030* (*OsSPS1*), and *LOC\_Os07g04230* (*Os07g0134700*), and then the CRISPR-GE web-based online tool [30] was used to analyze and select five C-rich targets (Figure 1b) [30]. The five targets of ABE8e-hAAG and ABE8e-EndoV were obtained from *LOC\_Os01g68870* (*OsMSP1*), *LOC\_Os05g06480* (*Chalk5*), and *LOC\_Os01g10110* (*OsCKX2*) (Figure 2b), and CRISPR-GE was also used to analyze and select five A-rich targets. The sgRNA expression cassette was assembled into binary vectors by referring to the published multi-target assembly methods of CRISPR/Cas9 [35].

##### 4.3. Genetic Transformation

All of the CGBE-rUNG, CGBE-hUNG, ABE8e-hAAG, and ABE8e-EndoV plasmids containing sgRNA expression cassettes were transformed into *Agrobacterium tumefaciens* (strain *EHA105*) through electroporation and then transformed into rice (*Oryza sativa*) (Nipponbare variety calli) by using the *Agrobacterium*-mediated method. T<sub>0</sub> transgenic rice was obtained by callus induction and subculture, callus infection, co-culture, screening, differentiation of resistant callus, and rooting and transplantation of strong seedlings, according to previously reported [36].

#### 4.4. On-Target and Off-Target Efficiency Analysis of T<sub>0</sub> Rice

The genomic DNA of T<sub>0</sub> plants was extracted using the SDS method. Target-specific primers with on-target and off-target efficiency for PCR and sequencing were designed using online CRISPR-GE tools. Our previous study results showed that selecting the on-target sequence difference with candidate off-target  $\geq 3$  bases can greatly reduce the off-target effects [29]. Candidate off-target sites with one to three base variations compared with the on-target sites of CGBE-rUNG, CGBE-hUNG, ABE8e-hAAG, and ABE8e-EndoV were analyzed and selected using the off-target subprogram of CRISPR-GE. To evaluate the on-target and off-target efficiencies, the target regions were amplified by PCR from the T<sub>0</sub> plants by using site-specific primers, and barcode adapter primers and PCR product libraries were constructed. These PCR product libraries were sequenced using Illumina high-throughput deep sequencing (ANOROAD, Yiwu, China). The deep sequencing data were analyzed using Hi-TOM, which is an effective genome editing assessment tool [31].

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23147990/s1>.

**Author Contributions:** Conceptualization: Q.Z.; investigation: D.Z., Z.Z., Y.L., T.L. (Taoli Liu), T.L. (Tie Li), J.L., Q.L., Y.X., S.L., N.C., S.Y. and X.X.; methodology: D.Z., Z.Z., Y.L., T.L. (Taoli Liu), J.L., Y.X., S.L. and N.C.; writing—original draft preparation: Q.Z. and D.Z.; writing—review and editing: Y.-G.L. and Q.Z.; funding acquisition: Y.-G.L., Q.Z. and D.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by grants from the Laboratory of Lingnan Modern Agriculture Project (NT2021002), the National Natural Science Foundation of China (31991222; 31971915), the Major Program of Guangdong Basic and Applied Research (2019B030302006), the Guangdong Basic and Applied Basic Research Foundation (2020A1515110334), and the Guangdong Special Support Program of Young Top-Notch Talent in Science and Technology Innovation (2019TQ05N147).

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Thuronyi, B.W.; Koblan, L.W.; Levy, J.M.; Yeh, W.; Zheng, C.; Newby, G.A.; Wilson, C.; Bhaumik, M.; Shubina-Oleinik, O.; Holt, J.R.; et al. Continuous evolution of base editors with expanded target compatibility and improved activity. *Nat. Biotechnol.* **2019**, *37*, 1070–1079. [[CrossRef](#)]
2. Zeng, D.; Liu, T.; Tan, J.; Zhang, Y.; Zheng, Z.; Wang, B.; Zhou, D.; Xie, X.; Guo, M.; Liu, Y.G.; et al. PhieCBEs: Plant High-Efficiency Cytidine Base Editors with Expanded Target Range. *Mol. Plant* **2020**, *13*, 1666–1669. [[CrossRef](#)]
3. Zong, Y.; Song, Q.; Li, C.; Jin, S.; Zhang, D.; Wang, Y.; Qiu, J.; Gao, C. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat. Biotechnol.* **2018**, *36*, 950–953. [[CrossRef](#)] [[PubMed](#)]
4. Wang, M.; Wang, Z.; Mao, Y.; Lu, Y.; Yang, R.; Tao, X.; Zhu, J.K. Optimizing base editors for improved efficiency and expanded editing scope in rice. *Plant Biotechnol. J.* **2019**, *17*, 1697–1699. [[CrossRef](#)] [[PubMed](#)]
5. Xu, W.; Yang, Y.; Liu, Y.; Kang, G.; Wang, F.; Li, L.; Lv, X.; Zhao, S.; Yuan, S.; Song, J.; et al. Discriminated sgRNAs-Based SurroGate System Greatly Enhances the Screening Efficiency of Plant Base-Edited Cells. *Mol. Plant* **2020**, *13*, 169–180. [[CrossRef](#)] [[PubMed](#)]
6. Richter, M.F.; Zhao, K.T.; Eton, E.; Lapinaite, A.; Newby, G.A.; Thuronyi, B.W.; Wilson, C.; Koblan, L.W.; Zeng, J.; Bauer, D.E.; et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat. Biotechnol.* **2020**, *38*, 883–891. [[CrossRef](#)]
7. Tan, J.; Zeng, D.; Zhao, Y.; Wang, Y.; Liu, T.; Li, S.; Xue, Y.; Luo, Y.; Xie, X.; Chen, L.; et al. PhieABEs: A PAM-less/free high-efficiency adenine base editor toolbox with wide target scope in plants. *Plant Biotechnol. J.* **2022**, *20*, 934–943. [[CrossRef](#)]
8. Yan, D.; Ren, B.; Liu, L.; Yan, F.; Li, S.; Wang, G.; Sun, W.; Zhou, X.; Zhou, H. High-efficiency and multiplex adenine base editing in plants using new TadA variants. *Mol. Plant* **2021**, *14*, 722–731. [[CrossRef](#)]
9. Ren, Q.; Sretenovic, S.; Liu, S.; Tang, X.; Huang, L.; He, Y.; Liu, L.; Guo, Y.; Zhong, Z.; Liu, G.; et al. PAM-less plant genome editing using a CRISPR–SpRY toolbox. *Nat. Plants* **2021**, *7*, 25–33. [[CrossRef](#)]
10. Wei, C.; Wang, C.; Jia, M.; Guo, H.X.; Luo, P.Y.; Wang, M.G.; Zhu, J.K.; Zhang, H. Efficient generation of homozygous substitutions in rice in one generation utilizing an rABE8e base editor. *J. Integr. Plant Biol.* **2021**, *63*, 1595–1599. [[CrossRef](#)]
11. Li, J.; Xu, R.; Qin, R.; Liu, X.; Kong, F.; Wei, P. Genome editing mediated by SpCas9 variants with broad non-canonical PAM compatibility in plants. *Mol. Plant* **2021**, *14*, 352–360. [[CrossRef](#)] [[PubMed](#)]

12. Wang, Z.; Liu, X.; Xie, X.; Deng, L.; Zheng, H.; Pan, H.; Li, D.; Li, L.; Zhong, C. ABE8e with Polycistronic tRNA-gRNA Expression Cassette Significantly Improves Adenine Base Editing Efficiency in *Nicotiana benthamiana*. *Int. J. Mol. Sci.* **2021**, *22*, 5663. [[CrossRef](#)] [[PubMed](#)]
13. Veillet, F.; Perrot, L.; Guyon-Debast, A.; Kermarrec, M.; Chauvin, L.; Chauvin, J.; Gallois, J.; Mazier, M.; Nogu e, F. Expanding the CRISPR Toolbox in *P. patens* Using SpCas9-NG Variant and Application for Gene and Base Editing in Solanaceae Crops. *Int. J. Mol. Sci.* **2020**, *21*, 1024. [[CrossRef](#)] [[PubMed](#)]
14. Veillet, F.; Perrot, L.; Chauvin, L.; Kermarrec, M.; Guyon-Debast, A.; Chauvin, J.; Nogu e, F.; Mazier, M. Transgene-Free Genome Editing in Tomato and Potato Plants Using Agrobacterium-Mediated Delivery of a CRISPR/Cas9 Cytidine Base Editor. *Int. J. Mol. Sci.* **2019**, *20*, 402. [[CrossRef](#)]
15. Liu, X.; Qin, R.; Li, J.; Liao, S.; Shan, T.; Xu, R.; Wu, D.; Wei, P. A CRISPR-Cas9-mediated domain-specific base-editing screen enables functional assessment of ACCase variants in rice. *Plant Biotechnol. J.* **2020**, *18*, 1845–1847. [[CrossRef](#)]
16. Zeng, D.; Liu, T.; Ma, X.; Wang, B.; Zheng, Z.; Zhang, Y.; Xie, X.; Yang, B.; Zhao, Z.; Zhu, Q.; et al. Quantitative regulation of Waxy expression by CRISPR/Cas9-based promoter and 5'UTR-intron editing improves grain quality in rice. *Plant Biotechnol. J.* **2020**, *18*, 2385–2387. [[CrossRef](#)]
17. Khan, Z.A.; Kumar, R.; Dasgupta, I. CRISPR/Cas-Mediated Resistance against Viruses in Plants. *Int. J. Mol. Sci.* **2022**, *23*, 2303. [[CrossRef](#)]
18. Tripathi, L.; Ntui, V.O.; Tripathi, J.N. Control of Bacterial Diseases of Banana Using CRISPR/Cas-Based Gene Editing. *Int. J. Mol. Sci.* **2022**, *23*, 3619. [[CrossRef](#)]
19. Kantor, A.; McClements, M.; MacLaren, R. CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *Int. J. Mol. Sci.* **2020**, *21*, 6240. [[CrossRef](#)]
20. Zhao, D.; Li, J.; Li, S.; Xin, X.; Hu, M.; Price, M.A.; Rosser, S.J.; Bi, C.; Zhang, X. Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* **2021**, *39*, 35–40. [[CrossRef](#)]
21. Kurt, I.C.; Zhou, R.; Iyer, S.; Garcia, S.P.; Miller, B.R.; Langner, L.M.; Gr unewald, J.; Joung, J.K. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* **2021**, *39*, 41–46. [[CrossRef](#)] [[PubMed](#)]
22. Sretenovic, S.; Liu, S.; Li, G.; Cheng, Y.; Fan, T.; Xu, Y.; Zhou, J.; Zheng, X.; Coleman, G.; Zhang, Y.; et al. Exploring C-To-G Base Editing in Rice, Tomato, and Poplar. *Front. Genome Ed.* **2021**, *3*, 756766. [[CrossRef](#)] [[PubMed](#)]
23. Tian, Y.; Shen, R.; Li, Z.; Yao, Q.; Zhang, X.; Zhong, D.; Tan, X.; Song, M.; Han, H.; Zhu, J.K.; et al. Efficient C-to-G editing in rice using an optimized base editor. *Plant Biotechnol. J.* **2022**, *20*, 1238–1240. [[CrossRef](#)] [[PubMed](#)]
24. Kuraoka, I. Diversity of Endonuclease V: From DNA Repair to RNA Editing. *Biomolecules* **2015**, *5*, 2194–2206. [[CrossRef](#)] [[PubMed](#)]
25. Hedglin, M.; O'Brien, P.J. Human Alkyladenine DNA Glycosylase Employs a Processive Search for DNA Damage. *Biochemistry* **2008**, *47*, 11434–11445. [[CrossRef](#)]
26. Wu, J.; Samara, N.L.; Kuraoka, I.; Yang, W. Evolution of Inosine-Specific Endonuclease V from Bacterial DNase to Eukaryotic RNase. *Mol. Cell* **2019**, *76*, 44–56. [[CrossRef](#)]
27. Kiyonari, S.; Egashira, Y.; Ishino, S.; Ishino, Y. Biochemical characterization of endonuclease V from the hyperthermophilic archaeon, *Pyrococcus furiosus*. *J. Biochem.* **2014**, *155*, 325–333. [[CrossRef](#)]
28. Dalhus, B.; Alseth, I.; Bj or as, M. Structural basis for incision at deaminated adenines in DNA and RNA by endonuclease V. *Prog. Biophys. Mol. Biol.* **2015**, *117*, 134–142. [[CrossRef](#)] [[PubMed](#)]
29. Liu, T.; Zeng, D.; Zheng, Z.; Lin, Z.; Xue, Y.; Li, T.; Xie, X.; Ma, G.; Liu, Y.G.; Zhu, Q. The ScCas9++ variant expands the CRISPR toolbox for genome editing in plants. *J. Integr. Plant Biol.* **2021**, *63*, 1611–1619. [[CrossRef](#)]
30. Xie, X.; Ma, X.; Zhu, Q.; Zeng, D.; Li, G.; Liu, Y.G. CRISPR-GE: A Convenient Software Toolkit for CRISPR-Based Genome Editing. *Mol. Plant* **2017**, *10*, 1246–1249. [[CrossRef](#)]
31. Liu, Q.; Wang, C.; Jiao, X.; Zhang, H.; Song, L.; Li, Y.; Gao, C.; Wang, K. Hi-TOM: A platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. *Sci. China Life Sci.* **2019**, *62*, 1–7. [[CrossRef](#)] [[PubMed](#)]
32. Wang, S.; Zong, Y.; Lin, Q.; Zhang, H.; Chai, Z.; Zhang, D.; Chen, K.; Qiu, J.L.; Gao, C. Precise, predictable multi-nucleotide deletions in rice and wheat using APOBEC-Cas9. *Nat. Biotechnol.* **2020**, *38*, 1460–1465. [[CrossRef](#)] [[PubMed](#)]
33. Qin, L.; Li, J.; Wang, Q.; Xu, Z.; Sun, L.; Alariqi, M.; Manghwar, H.; Wang, G.; Li, B.; Ding, X.; et al. High-efficient and precise base editing of C•G to T•A in the allotetraploid cotton (*Gossypium hirsutum*) genome using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* **2019**, *18*, 45–56. [[CrossRef](#)] [[PubMed](#)]
34. Koblan, L.W.; Arbab, M.; Shen, M.W.; Hussmann, J.A.; Anzalone, A.V.; Doman, J.L.; Newby, G.A.; Yang, D.; Mok, B.; Replogle, J.M.; et al. Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. *Nat. Biotechnol.* **2021**, *39*, 1414–1425. [[CrossRef](#)] [[PubMed](#)]
35. Ma, X.; Zhang, Q.; Zhu, Q.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.; Li, H.; Lin, Y.; et al. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol. Plant* **2015**, *8*, 1274–1284. [[CrossRef](#)]
36. Tan, J.; Wang, Y.; Chen, S.; Lin, Z.; Zhao, Y.; Xue, Y.; Luo, Y.; Liu, Y.; Zhu, Q. An Efficient Marker Gene Excision Strategy Based on CRISPR/Cas9-Mediated Homology-Directed Repair in Rice. *Int. J. Mol. Sci.* **2022**, *23*, 1588. [[CrossRef](#)]