Genome Biology

BRIEF REPORT

Open Access

# Efficient glycosylase-mediated base editing with minimal off-target effects in mammalian embryos

Yinghui Wei[1,2*†], Xi Cao[1,2†], Shuhong Huang[1,2†], Yang Yue[1,2†], Muhua Luo[1,2†], Zujiang Liu[1,2†], Bo Liu[1,2], Qijing Zhang[1,2], Yangsheng Wu[3], Liqin Wang[3], Jiapeng Lin[3], Yulin Chen[1,2], Wenxin Zheng[4], Weiwei Wu[5*], Kun Xu[1,2*] and Xiaolong Wang[1,2*]

†Yinghui Wei, Xi Cao, Shuhong Huang, Yang Yue, Muhua Luo and Zujiang Liu contributed equally to this work.

*Correspondence: yinghuiwei@nwafu.edu.cn; wuweiweigp@foxmail.com; xukunas@nwafu.edu.cn; xiaolongwang@nwafu.edu.cn

[1] Hainan Institute of Northwest A&F University, Sanya, Hainan 572025, China
[2] International Joint Agriculture Research Center for Animal Bio-Breeding of Ministry of Agriculture and Rural Affairs, College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi 712100, China
[3] Institute of Biotechnology, Xinjiang Academy of Animal Science, Urumqi, Xinjiang 830011, China
[4] Institute of Animal Husbandry Quality Standards, Xinjiang Academy of Animal Science, Urumqi, Xinjiang 830011, China
[5] Institute of Animal Science, Xinjiang Academy of Animal Science, Urumqi, Xinjiang 830011, China

## Abstract

Developing glycosylase-based base editors (gBEs) to broaden the editing scope is highly desirable for biomedical research and agricultural applications. However, the off-target effects and applicability of gBEs need further investigation. We employ GOTI to detect rare DNA off-target events in mouse embryos injected with N-methylpurine glycosylase-based AYBE and gGBE. Transcriptome-wide RNA analysis reveals that TadA8e-V106W, derived from AYBE, induces low-frequency RNA off-target editing. Both base editors efficiently induce A/G-to-Y editing in mouse and sheep embryos, and in newborn lambs. The robust efficiency and specificity of AYBE and gGBE underscore their potential for clinical applications and genetic improvement in livestock.

## Background

Base editors hold significant promise in basic research, therapeutic applications, and the enhancement of traits in plant and animal breeding. In recent years, cytosine base editors (CBEs) [1] and adenine base editors (ABEs) [2] were developed by fusing cytosine and adenosine deaminases with Cas9 nickase to achieve the C-to-U(T) and A-to-I(G) base transitions. The addition of an uracil glycosylase inhibitor (UGI) to Cas9 nickase is essential for CBEs to effectively suppress base excision repair (BER) and ensure efficient C-to-T base editing [1]. Conversely, replacing the UGI in CBEs with uracil DNA N-glycosylase (UNG or UDG) enhances the excision of U bases, leading to the development of novel C-to-G base editors (CGBEs) [3–6]. Similarly, the fusion of hypoxanthine (I) excision proteins such as N-methylpurine DNA glycosylase (MPG) [7] or mouse alkyladenine DNA glycosylase (mAAG) [8] with ABEs has resulted in the creation of a novel A-to-C/T base editor (AYBE) (Additional file 1: Fig. S1A). A deaminase-free glycosylase-based guanine base editor (gGBE), capable of inducing G-to-C/T base transversions,

was developed by directly fusing the engineered gMPGv6.3 with the Cas9 nickase (Additional file 1: Fig. S1B) [9]. Recently, similar deaminase-free glycosylase-based gTBEs and gCBEs were also developed by fusing Cas9 nickase with engineered variants of uracil-DNA glycosylase (UNG), cytosine-DNA glycosylase (CDG), or thymine-DNA glycosylase (TDG) for targeted excision of T or C bases [10–13]. Overall, the evolution of DNA glycosylases like MPG, UNG, CDG, and TDG allows for the direct excision of G, I, T, and C bases, creating AP sites and initiating BER to efficiently repair the loss of these bases.

AYBE enables A-to-K base editing in rice and tomato [14–16], as well as A-to-Y base editing in mice and maize [8, 17]. Meanwhile, gGBE demonstrates efficient G-to-Y base editing in mice and high-purity G-to-T base editing in rice [9, 18]. However, the feasibility and editing profiles of AYBE and gGBE in large animals are remain to be fully characterized. More importantly, the off-target effects of these two editors have yet to be comprehensively evaluated. The Genome-wide Off-target analysis by Two-cell embryo Injection (GOTI) assay enables highly sensitive and unbiased detection of sgRNA-independent off-target effects by editing one blastomere of two-cell stage mouse embryos. As the embryo develops, rare off-target sites in the edited blastomere are replicated in daughter cells, thereby amplifying the off-target signal in sequencing analysis, as previously reported for BE3 and DdCBE architectures [19, 20]. Here, we employed GOTI and RNA-seq to systematically evaluate the genome-wide DNA and transcriptome-wide RNA off-target effects of AYBE-V106W and gGBE. We also characterized the editing efficiency, editing profiles, and indel frequencies of these two base editors in mouse embryos, sheep embryos, and lambs. These findings provide valuable evidence supporting the safety and feasibility of AYBE- and gGBE-mediated base editing for both clinical research and agricultural breeding.

## Results and discussion

To assess the editing efficiency and indel profiles of the AYBE and gGBE base editors, we performed comprehensive evaluations at multiple genomic loci in mouse embryos. C57BL/6J mouse embryos were microinjected with mRNAs (100 ng/μL) encoding AYBE, AYBE-V106W, or gGBE, along with different sgRNAs (50 ng/μL) targeting the mouse *Tyr* gene. The AYBE and AYBE-V106W showed comparable A-to-C/T editing efficiencies and purities at three target sites (Additional file 1: Fig. S2A, B). Furthermore, our data revealed that A > Y and G > Y (Y = C or T) editing efficiencies, along with the associated indel profiles mediated by the three base editors, varied markedly across genomic loci and among individual mouse embryos (Additional file 1: Fig. S2A-D). Additionally, A > C and G > C edits exhibited higher efficiencies, and although AYBE-V106W induced slightly higher average indel frequencies than gGBE, both remained below 10% (Additional file 1: Fig. S2C, D).

Next, we utilized the GOTI method to assess genome-wide off-target effects of TadA8e-V106W and the engineered MPG in the context of AYBE and gGBE systems. We first in vitro transcribed AYBE-V106W and gGBE mRNAs targeting the mouse *Tyr* gene and injected them, along with the Cre mRNA, into one blastomere of two-cell embryos derived from Ai9 background (CAG-LoxP-Stop-LoxP-tdTomato) mice, while leaving the other blastomere un-injected (Fig. 1A). The embryos injected only with Cre

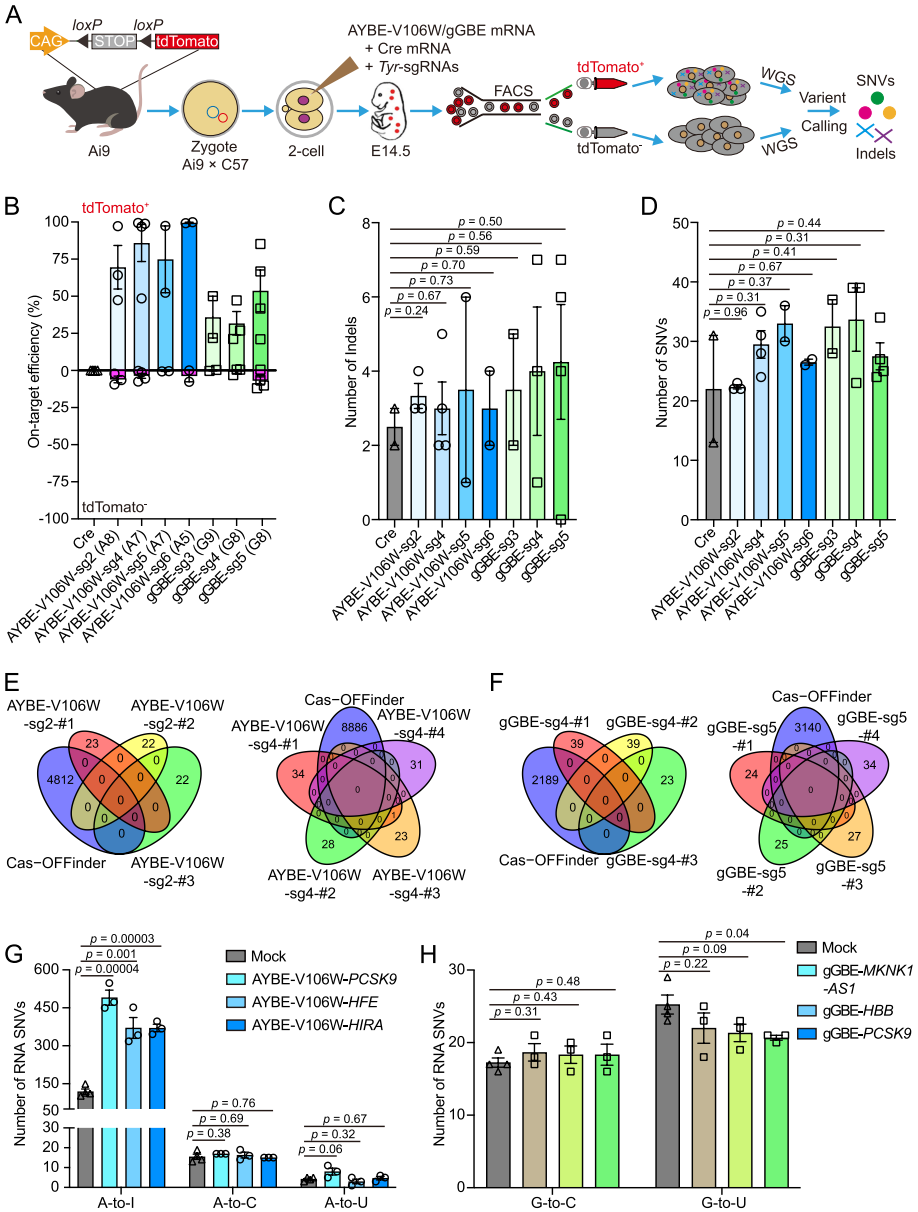


**Fig. 1** Minimal off-target effects induced by AYBE-V106W and gGBE in mouse embryos. **A** GOTI workflow for analyzing the off-target profiles of N-methylpurine glycosylase-based base editors (AYBE-V106W and gGBE). **B** On-target efficiency of tdTomato$^+$ and tdTomato$^-$ cells in Cre-, AYBE-V106W-, and gGBE-treated groups at multiple *Tyr* sites, based on targeted deep sequencing. **C**, **D** Comparison of total detected indels (**C**) and SNVs (**D**) in Cre-, AYBE-V106W-, and gGBE-injected groups at multiple *Tyr* sites, as determined by WGS. **E**, **F** Overlap among SNVs detected by GOTI with predicted off-target sites from Cas-OFFinder in AYBE-V106W-treated samples (**E**) and gGBE-treated samples (**F**). **G**, **H** The bar graph shows the number of A > I (**G**), A > C (**G**), A > U (**G**), G > C (**H**), and G > U (**H**) mutations detected in RNA from mCherry-, AYBE-V106W-, and gGBE-treated HEK293T cells at multiple endogenous sites. Data are presented as means ± s.e.m. *p* values were evaluated using a two-tailed unpaired Student's t-test. All experiments included at least two biological replicates. Two Cre samples were obtained from previous published studies [19, 21]

mRNA served as the control group. The restoration of tdTomato fluorescence signal by Cre-mediated removal of the "-Stop-" component served as an indicator of successful delivery of the mRNAs for these two editors and Cre (Additional file 1: Fig. S3A, S4A).

A subset of injected embryos was retained to monitor developmental progression, and approximately half of the un-transplanted 4-cell or 8-cell embryos exhibited tdTomato expression (Additional file 1: Fig. S3A, S4A). At 14.5 days post-transfer of injected 2-cell embryos into surrogate mothers, E14.5 embryos were collected from both the AYBE-V106W and gGBE groups. Embryos were digested into single cells sorted by fluorescence-activated cell sorting (FACS) into tdTomato$^+$ and tdTomato$^-$ populations for genotyping of base transversion outcomes at the *Tyr* gene, followed by whole-genome sequencing (WGS) analysis (Fig. 1A and Additional file 1: Fig. S3B, S4B). Sanger sequencing and WGS results revealed that A-to-other on-target editing efficiencies in the AYBE-V106W group ranged from 47.37% to 99.90%, while G-to-other editing efficiencies in the gGBE group ranged from 21.00% to 85.29% (Fig. 1B and Additional file 1: Fig. S3C, S4C). Additionally, we observed a slight editing efficiency (< 10%) due to the leakage of injected mRNA components into the un-injected blastomere in both AYBE-V106W and gGBE tdTomato$^-$ cells (Fig. 1B and Additional file 1: Fig. S3C, S4C). Next, WGS datasets were used to assess the off-target editing by examining the number of indels or single-nucleotide variants (SNVs) in the AYBE-V106W and gGBE treated groups through a stringent variant filter pipeline. The number of indels did not differ significantly between the treated groups and the Cre-only group (Fig. 1C). These results indicate that AYBE-V106W and gGBE base editors did not induce any detectable indels above background frequency in mouse embryos.

To rule out the contribution of Cre injection to SNV formation, we examined Cre-only samples lacking AYBE-V106W or gGBE and observed approximately 20 SNVs (Fig. 1D). Moreover, we did not observe any significant differences in SNV frequency in the groups treated with AYBE-V106W or gGBE compared to the Cre-only group (Fig. 1D). Specifically, AYBE-V106W induced $22.33 \pm 0.47$, $29.5 \pm 4.03$, $33.0 \pm 3.0$, and $26.5 \pm 0.5$ SNVs at the four *Tyr* gene sites, while gGBE induced $32.50 \pm 4.50$, $33.67 \pm 7.54$, and $27.5 \pm 3.91$ SNVs at another three sites, with all values comparable to the $22.0 \pm 9.0$ SNVs observed in the Cre-only group (Fig. 1D). Additionally, none of the mutations identified in AYBE-V106W- or gGBE-treated embryos were shared among individuals or overlapped with off-target sites predicted by Cas-OFFinder (mismatch threshold = 6) (Fig. 1E, F and Additional file 1: Fig. S5A, B). Considering the editing profiles of TadA8e-V106W and engineered MPGs, we analyzed all SNVs identified in AYBE-V106W- and gGBE-targeting samples. We found no significant enrichment of A > G/T > C base transitions (Additional file 1: Fig. S5C) or A > C/T > G and A > T/T > A base transversions (Additional file 1: Fig. S5D) in the AYBE-V106W group compared to the Cre-only group, nor of G > C/C > G and G > T/C > A base transversions (Additional file 1: Fig. S5E) in the gGBE group. We also performed an orthogonal R-loop assay to further evaluate the sgRNA-independent off-target effects of the MPG-based AYBE-V106W and gGBE base editors. These findings demonstrated that AYBE-V106W (TadA8e-V106W) induces minimal off-target activity, with observed frequencies below 2% (Additional file 1: Fig. S6A). In contrast, the data showed that TadA8e, derived from ABE8e, exhibited substantial off-target activity as detected by the R-loop assay (Additional file 1: Fig. S6A). Furthermore, no detectable off-target activity was observed for the MPG-based gGBE base editor or the two UNG-based CGBE editors (Additional file 1: Fig. S6B). While R-loop analysis revealed low-level off-target activity of AYBE-V106W, GOTI detected none, likely due

to differences in R-loop formation duration and the delivery methods used (plasmid vs. mRNA). Given that ABE8e-V106W is known to minimize off-target RNA editing [22], we evaluated the RNA off-target effects of AYBE-V106W and gGBE at three sites using RNA-seq. AYBE-V106W caused modestly increased RNA off-target effects in HEK293T cells, mainly A-to-I editing mediated by TadA8e-V106W, while MPG did not significantly induce A-to-C or A-to-U edits (Fig. 1G and Additional file 1: Fig. S7A). Additionally, the MPG-mediated gGBE base editor showed no significant RNA off-target editing activity, including G-to-C and G-to-U conversions (Fig. 1H and Additional file 1: Fig. S7B). Taken together, these findings suggest that AYBE-V106W and gGBE combine high on-target editing efficiency with minimal off-target activity.

We proceeded to investigate the feasibility of using AYBE-V106W and gGBE for generating genetically modified large animals. Four target sites for AYBE-V106W and five for gGBE within the sheep Myostatin (*MSTN*) gene were selected, and corresponding all-in-one expression vectors containing protein and various sgRNA architectures were constructed. Sheep fetal fibroblast cells (sFFCs) were transfected with their respective expression vectors, and transfection-positive cells were collected 72 h later using FACS. These cells underwent targeted deep sequencing to assess the editing efficiency at each target site. The results showed successful A-to-Y and G-to-Y base transversions at all targets in sFFCs by AYBE-V106W and gGBE, respectively (Fig. 2A, B). Although the patterns of A-to-other or G-to-other base editing in sFFCs varied depending on the target and the position of the target bases, AYBE-V106W was observed to predominantly induce A-to-G base transitions, followed by A-to-C base transversions (Fig. 2A), while gGBE generally induced G-to-Y editing (Fig. 2B). We also observed that AYBE-V106W and gGBE induced only low levels of indels (~ 5%) in sFFCs (Fig. 2A, B). In addition, we assessed the editing efficiency and indel patterns of AYBE-V106W and gGBE in sheep embryos, yielding results consistent with those obtained in mice (Fig. 2C, D and Additional file 1: Fig. S8). AYBE-V106W paired with *MSTN*-sg3 ($5.05 \pm 7.16\%$ A > T transversion at A7, leading to premature termination of MSTN protein expression) and gGBE paired with *MSTN*-sg2 ($4.03 \pm 8.67\%$ G > T transversion at G5, also resulting in premature termination of MSTN expression), both of which exhibited relatively high editing efficiencies at the tested sites in sheep embryos, were selected for the generation of gene-edited lambs (Fig. 2C, D and Additional file 1: Fig. S8).

In vitro transcribed AYBE-V106W or gGBE mRNA, along with their respective sgRNAs targeting the *MSTN* gene, was injected into fertilized eggs from Merino sheep. The injected embryos were briefly cultured in vitro to the two-cell stage prior to transplantation into surrogate sheep (Fig. 2E). Following a gestation period of approximately 150 days, six lambs edited with AYBE-V106W and three lambs edited with gGBE were born, all exhibiting normal growth and development. Blood samples were collected from each lamb for DNA extraction and genotyping. Targeted deep sequencing confirmed successful editing using the AYBE-V106W and gGBE systems in all lambs (Additional file 1: Fig. S9A, B). In gene-edited lambs, similar to AYBE-V106W in sFFCs, A-to-C editing (up to 49.48%) and A-to-T (up to 31.38%) editing are also detected at position A4 of *MSTN*-sg3 (Fig. 2F, G). However, AYBE-V106W exclusively induced A-to-C editing (up to 88.67%) at position A7 of *MSTN*-sg3, showing a quite different editing pattern compared to that observed in sFFCs and sheep embryos (Fig. 2F, G). Interestingly, all

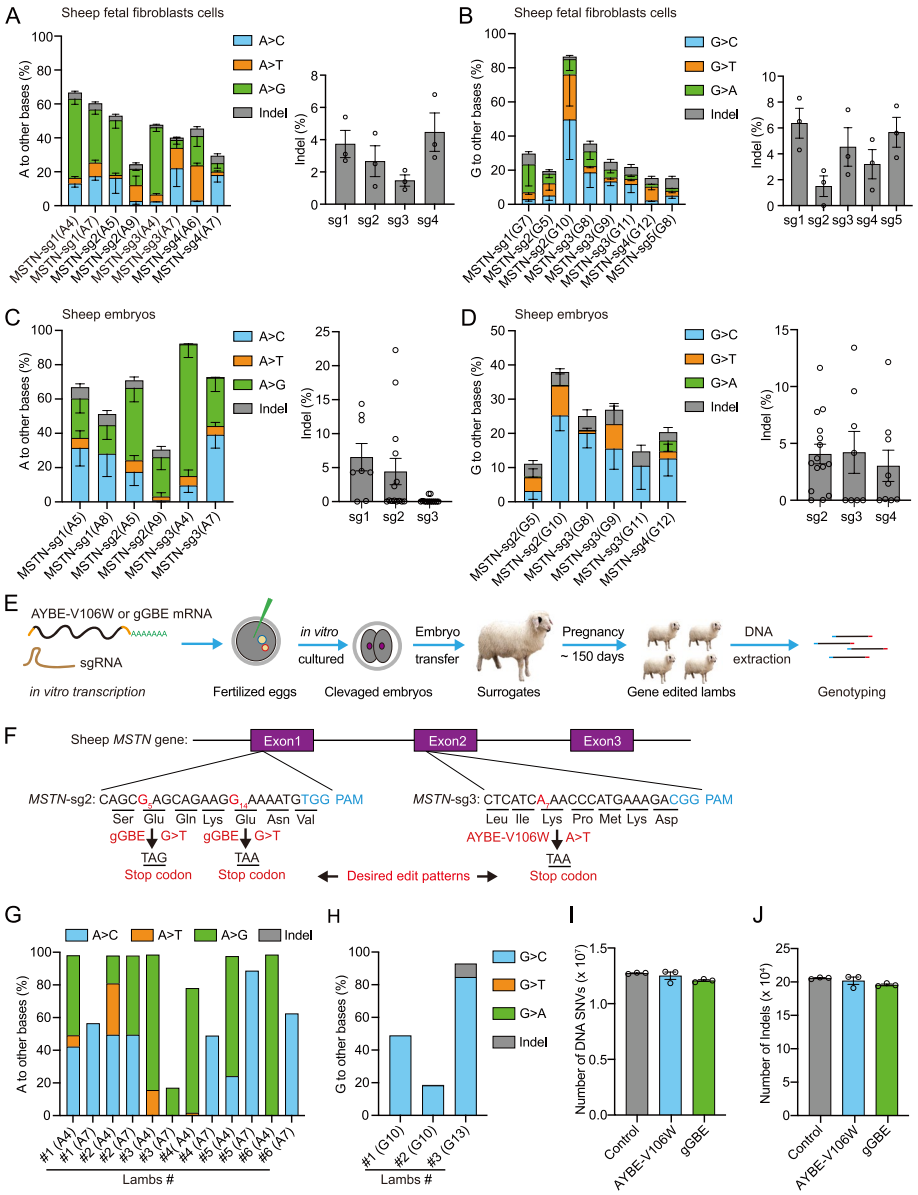


**Fig. 2** Efficient and specific base editing was observed in sheep embryos treated with AYBE-V106W and gGBE. **A**, **B** Editing efficiency and indel profiles of different sgRNAs targeting the *MSTN* gene using AYBE-V106W (**A**) and gGBE (**B**) in sheep fetal fibroblasts (sFFCs). **C**, **D** Editing efficiency and indel profiles of different sgRNAs targeting the *MSTN* gene using AYBE-V106W (**C**) and gGBE (**D**) in sheep embryos. **E** Schematic representation of the generation of *MSTN*-edited sheep using AYBE-V106W and gGBE. **F** Schematic illustration of target sites and intended editing patterns for MSTN disruption in sheep using AYBE-V106W and gGBE systems. **G**, **H** Genotypes of six base-edited lambs generated using AYBE-V106W (**G**) and three using gGBE (**H**). **I**, **J** Comparison of total detected SNVs (**I**) and indels (**J**) in control, AYBE-V106W-, and gGBE-edited lambs, as determined by WGS. All experiments included at least three biological replicates

three gGBE-edited lambs exhibited exclusively the G-to-C editing genotype (ranging from 18.40% to 84.69%) at position G10 of *MSTN*-sg2, demonstrating a distinctly different editing profile compared to that in sFFCs and sheep embryos (Fig. 2H). Additionally, all the gene-edited lambs exhibited a low indel frequency, except for lamb #3 edited with gGBE at *MSTN*-sg2, which showed a slightly higher indel frequency (8.31%)

(Fig. 2H). We performed WGS to assess DNA off-target effects in AYBE-V106W- and gGBE-treated lambs by quantifying indels and SNVs. Our results indicated that treatment with either AYBE-V106W or gGBE did not lead to a significant increase in indel or SNV counts in gene-edited lambs relative to wild-type controls (Fig. 2I, J). Our initial objective was to use the AYBE-V106W and gGBE editors to introduce nonsense mutations (A > T or G > T) in the *MSTN* gene; however, the resulting edits were predominantly missense mutations (A > C or G > C) (Additional file 1: Fig. S10A). Furthermore, we conducted phenotypic analyses to evaluate whether these missense mutations affected MSTN function and resulted in distinct phenotypes. Our findings showed that the missense mutations introduced by these two base editors did not result in significant differences in birth weight or average daily gain (0–90 days) in gene-edited lambs compared to wild-type lambs (Additional file 1: Fig. S10B-D). Collectively, these results demonstrate that AYBE-V106W and gGBE induced A-to-Y and G-to-Y base transversions in both sFFCs and sheep embryos. Notably, a preferential A-to-C editing by AYBE-V106W and G-to-C editing by gGBE was observed in the resulting gene-edited lambs.

Base editors enable highly efficient and programmable nucleotide conversions without the need for donor templates or the risk of inducing double-strand breaks (DSBs). Established base editors, such as CBEs and ABEs, mediate C-to-T and A-to-G conversions by deaminating cytosine and adenine bases, respectively. However, achieving other types of base substitutions via base editing has remained a significant challenge. In recent years, a range of base editors incorporating engineered glycosylases have been developed, including CGBE, AYBE, ACBE, gGBE, gTBE (DAF-TBE), and gCBE (DAF-CBE).

Several studies have demonstrated that CBEs employing APOBEC1 (a single-strand deaminase) or DddA (a double-strand deaminase) can cause substantial off-target DNA editing, as detected by the highly sensitive and unbiased GOTI method [19, 20]. AYBEs were generated by integrating additional glycosylases (hMPG or mAAG) into ABEs, whereas the deaminase-free, glycosylase-based gGBE was developed by directly fusing engineered gMPG to Cas9 nickase. Hence, it is crucial to investigate whether glycosylase utilization leads to potential genome-wide off-target events. In this study, the GOTI assay was employed to assess the off-target effects of two novel base editors, revealing that embryos edited with AYBE-V106W and gGBE exhibited no detectable genome-wide indels or SNVs. SNVs identified by the GOTI assay did not overlap with each other or with the off-target sites predicted by Cas-OFFinder, suggesting that off-target events were both rare and gRNA-independent. Additionally, some tdTomato-negative cell samples exhibited editing activity below 10%. These observations may be attributed to two possible factors: (1) intercellular communication between the two blastomeres, and (2) the occurrence of Cre activation without base editing, or vice versa, in cleavage-stage blastomeres. While R-loop analysis detected minimal off-target activity of AYBE-V106W, the GOTI assay revealed none, likely reflecting differences in R-loop formation duration and delivery methods (plasmid versus mRNA). Furthermore, the team led by David R. Liu has shown that delivering base editors like BE3 and ABE8e as ribonucleoprotein (RNP) complexes markedly decreases off-target editing relative to plasmid-based delivery in HEK293T cells [22, 23]. Regarding UNG-based GBEs, a previous study demonstrated that both eUNG-based eOPTI-CGBE and cUNG-based cOPTI-CGBE exhibited no detectable off-target editing, as assessed by the GOTI method [6]. Given

that ABE8e-V106W is known to minimize off-target RNA editing [22], our findings align with this observation, showing that AYBE-V106W induced only a slight increase in RNA off-target effects—mainly A-to-I edits mediated by TadA8e-V106W—while MPG did not significantly trigger A-to-C or A-to-U editing.

Over the past decade, our team has successfully generated *MSTN*-edited sheep using CRSIPR/Cas9 [24–26]. In this study, we aimed to employ AYBE and gGBE to induce nonsense mutations (A-to-T and G-to-T transversions) in the *MSTN* gene, marking the first use of these base editors in gene-edited livestock. We believe these two base editors hold particular promise for generating additional gene-edited sheep, especially those targeting SNPs linked to traits such as tail length (*TBXT*, c.334G > T) and live body weight (*MC4R*, c.138A > C). However, the gene-edited lambs predominantly exhibited missense mutations arising from A-to-C and G-to-C editing.

These results highlight the ongoing need to improve the purity and editing efficiency of A-to-Y or G-to-Y transversions. Potential strategies for improvement include co-expressing TLS polymerase Polη to enhance the purity of A/G-to-T editing [7], and continuously optimizing TadA-8e or MPG to improve the efficiency of A/G-to-Y editing [8]. Overall, these findings underscore the promising applications of AYBE-V106W and gGBE for both gene disruption and genetic improvement in large animals.

## Conclusions

This study employed the GOTI method to rigorously assess potential off-target editing events in mouse embryos treated with AYBE-V106W and gGBE, revealing no detectable DNA off-target effects beyond the level of spontaneous mutations. Transcriptome-wide RNA analysis demonstrated that AYBE-V106W induced low levels of RNA off-target editing, whereas gGBE exhibited none. We also observed that these two base editors mediate highly efficient A-to-Y and G-to-Y editing in mouse and sheep embryos, as well as in newborn lambs. These findings suggest that AYBE-V106W and gGBE combine high on-target editing efficiency with minimal off-target activity, underscoring their potential applications in fundamental research, therapeutic interventions, and genetic improvement of large animals.

## Methods

### Plasmid construction

The base editor constructs used in this study were cloned into a mammalian expression plasmid backbone under the control of an EF1α promoter using standard molecular cloning techniques. The plasmids targeting the mouse *Tyr* and sheep *MSTN* genes were constructed based on AYBE and gGBE (gifts from Huawei Tong; Addgene plasmids #193967 and #202629). The sgRNA oligos were annealed and ligated into the corresponding BpiI-digested backbone vectors. The spacer sequences are listed in Additional file 1: Table S1.

### Cell culture, transfection and genotyping

Sheep fetal fibroblast cells (sFFCs) were isolated from the trunk of a 40-day-old sheep fetus. HEK293T cells were obtained from the American Type Culture Collection (ATCC). All cell lines were routinely tested for mycoplasma. HEK293T cells and

fibroblasts were cultured in DMEM medium (Gibco) supplemented with 10% FBS (Gibco) and 1% penicillin–streptomycin (Gibco) until they reached 80–90% confluency, at which point they were used for transfection. Transfection was performed using Lipofectamine 3000 Reagent (Invitrogen), following the manufacturer's instructions. Briefly, fibroblasts were transfected with *MSTN* sgRNAs (1.0 μg) along with 1.0 μg of AYBE-V106W or gGBE plasmid using Lipofectamine 3000 in a 24-well culture plate. 72 h after transfection, five thousand sorted cells were harvested for genomic DNA extraction by adding 20 μL of lysis buffer (Vazyme), following the manufacturer's manual. The genomic regions of the target sites were amplified using Phanta Max Super-Fidelity DNA Polymerase (Vazyme) through nested PCR (Additional file 1: Table S2), with the use of primers containing barcodes. Purified PCR products were subjected to targeted deep sequencing and analyzed as described in the "Targeted deep sequencing analysis" section of the methods. For Cas9-independent off-target evaluation, the orthogonal R-loop assay was adapted using a dSaCas9–sgRNA plasmid. Similarly, 300 ng of SpCas9 sgRNA plasmid, 400 ng of base editor plasmid (ABE8e, AYBE-V106W, gGBE, eOPTI-CGBE, cOPTI-CGBE), and 300 ng of dSaCas9–sgRNA plasmid were co-transfected into HEK293T cells using polyethyleneimine (PEI, Polysciences). PCR primers, target protospacers, and amplicons used in orthogonal R-loop assays are indicated in Additional file 1: Table S4. For the RNA sequencing experiment, HEK293T cells were seeded in 10-cm dishes and transfected with 20 μg of either AYBE-V106W-CMV-mCherry or gGBE-CMV-mCherry plasmids using PEI.

#### GOTI assay

A mixture of Cre mRNA, AYBE-V106W mRNA or gGBE mRNA with their corresponding *Tyr*-sgRNAs, was injected into one blastomere of a 2-cell embryo derived from mating Ai9 male mice with wild-type C57BL/6 female mice. The embryos injected with the mRNA mixture were transplanted into surrogate ICR mice within 2 to 4 h after injection. The action of Cre is expected to generate a chimeric embryo where half of the cells are labeled with tdTomato (colored red). At E14.5, $tdTomato^+$ and $tdTomato^-$ cells were isolated from the chimeric embryos using fluorescence-activated cell sorting (FACS) for whole-genome sequencing (WGS). Off-target SNVs and indels were identified by comparing $tdTomato^+$ and $tdTomato^-$ cells using three variant calling algorithms (Mutect2, Lofreq, and Strelka for SNVs; Mutect2, Scalpel, and Strelka for indels). SNVs were represented as colored dots and indels as crosses in Fig. 1A.

#### RNA sequencing experiments

Seventy-two hours after transfection, HEK293T cells were washed with 1 × PBS and harvested using 0.25% trypsin (Gibco). Approximately 500,000 cells with the top 15% mCherry-positive signal were collected using a FACSAria III (BD Biosciences). Total RNA was extracted using a TRIzol-based method, then fragmented and reverse transcribed into cDNA using the HiScript Q RT SuperMix (Vazyme), following the manufacturer's instructions. Total RNA integrity was quantified using an Agilent 2100 Bioanalyzer. RNA-seq libraries were sequenced on the DNBSEQ-T7 platform.

### Animals

All experiments involving mice were approved by the Executive Committee on Laboratory Animal Management and Ethical Review at Northwest A&F University. Super ovulated 4-week-old C57BL/6 females were mated with heterozygous Ai9 males (full name B6.Cg-Gt (ROSA) 26Sortm9 (CAG-td-Tomato) Hze/J; JAX strain 007909). Females from the ICR strain were used as foster mothers. The mice were housed in a controlled barrier facility with a 12-h light/dark cycle, maintained at temperatures ranging from 20 °C to 26 °C and humidity levels between 40 and 60%. Food and water were readily available at all times.

### In vitro transcription of mRNA and sgRNA

The AYBE-V106W and gGBE plasmids were linearized using the FastDigest HindIII restriction enzyme (Thermo Fisher), purified with Gel Extraction Kit (Omega), and employed as the template for in vitro transcription (IVT) using the mMESSAGE mMACHINE T7 Ultra kit (Life Technologies). The T7 promoter was incorporated into the sgRNA template by PCR amplification of pCX330 (a gift from Feng Zhang; Addgene plasmid #42230) using the primer pair provided in Additional file 1: Table S3. The PCR products purified with Gel extraction Kit (Omega) as templates were transcribed using the MEGAshortscript T7 Kit (Life Technologies). The Cre mRNA, AYBE-V106W and gGBE mRNA, and *Tyr*/*MSTN*-sgRNAs were purified using the MEGAclear kit (Life Technologies) and eluted in RNase-free water. In vitro transcribed RNAs were aliquoted and stored at −80 °C until use. Prior to microinjection, the mixture of mRNA was prepared by centrifugation for 10 min at 14,000 rpm at 4 °C and the supernatant was transferred to 0.2 mL RNA-free PCR tubes for injection.

### Embryo injection, embryo culturing and embryo transplantation

Super ovulated 4-week-old C57BL/6 females were mated with heterozygous Ai9 males, and fertilized embryos were collected from oviducts 23 h after hCG injection. For the 2-cell injection, a mixture of AYBE-V106W mRNA (100 ng/μL) or gGBE mRNA (100 ng/μL) along with *Tyr*-sgRNA (50 ng/μL) was co-injected with Cre mRNA (2 ng/μL) into the cytoplasm of one blastomere of a 2-cell embryo 48 h after hCG injection. The injection was performed in a droplet of M2 medium containing 5 μg/mL cytochalasin B (CB) using a FemtoJet microinjector (Eppendorf) with constant flow settings. The injected embryos were cultured in M16 medium with amino acids at 37 °C under 5% $CO_2$ in air for 2 h. Subsequently, they were transferred into the oviducts of pseudo-pregnant ICR foster mothers at 0.5 days post-coitum (d.p.c.).

Healthy ewes (aged 3–5 years) with regular estrus cycles served as donors for zygote collection. Zygote collection and treatment of donors were conducted following previously described procedures [26]. One-cell stage zygotes (approximately 10–16 h after the last insemination) were surgically collected and immediately transferred into TCM199 medium (Gibco). The injection procedure was consistent with that used for mouse embryos. Microinjected embryos were cultured in Quinn's Advantage Cleavage Medium (Sage) for 24 h and then transferred into surrogate mothers, as previously reported [26]. Around 3–4 2-cell stage embryos were transferred into the ampullary-isthmic junction of the oviducts of recipient ewes. After approximately 150 days of pregnancy, nine

newborn lambs were genotyped. Comprehensive care and monitoring were provided to the lambs following delivery.

#### Genotyping of E14.5 mouse embryos and F0 sheep founders

The genotypes and off-target analysis of mutant E14.5 mouse embryos were determined by PCR of genomic DNA extracted from tdTomato$^+$ and tdTomato$^-$ cells. Additionally, peripheral venous blood samples from two-week-old lambs were collected and subjected to genomic DNA extraction. The Phanta Max Super-Fidelity DNA Polymerase (Vazyme) was first activated at 95 °C for 5 min. PCR amplification was conducted for 35 cycles with the following conditions: 95 °C for 30 s, 60 °C for 30 s, and 72 °C for 30 s, followed by a final extension at 72 °C for 5 min. Subsequently, PCR products underwent Sanger or targeted deep sequencing. Peripheral venous blood samples from two-week-old lambs were collected and subjected to genomic DNA extraction. PCR primers for genotype analysis are listed in Additional file 1: Table S2.

#### Fluorescence-activated cell sorting

To isolate tdTomato$^+$ and tdTomato$^-$ cells, whole E14.5 mouse embryos were dissected and enzymatically dissociated in 5 mL of Trypsin–EDTA (0.05%) at 37 °C for 30 min. The enzymatic dissociation was stopped by adding 5 mL of DMEM medium containing 10% fetal bovine serum (FBS). The fetal tissues were homogenized by passing them through a 1 mL pipette tip 30–40 times. The cell suspension was subsequently centrifuged at 800 rpm for 6 min, and the resulting pellet was resuspended in DMEM medium supplemented with 10% FBS. Finally, the cell suspension was filtered through a 40-μm cell strainer, and tdTomato$^+$ and tdTomato$^-$ cells were isolated by fluorescence-activated cell sorting (FACS). Samples for WGS analysis was confirmed to be > 95% pure following a second round of flow cytometry and fluorescence microscopy analysis.

#### Targeted deep sequencing analysis

The target sites were amplified using nested PCR with Phanta Max Super-Fidelity DNA Polymerase (Vazyme), along with primers containing barcodes. The PCR products were pooled and purified using a Gel extraction kit (Omega). The amplicon-seq libraries were prepared using the VAHTS Universal DNA Library Prep Kit (Vazyme), purified, and then sequenced on an Illumina NovaSeq 6000 platform with 150-bp paired-end reads. The sequencing data were initially demultiplexed using Cutadapt (v.2.8) and then processed by CRISPResso2 [27] for quantifying base editing efficiency. Refer to Table S1 for the target site sequence information.

#### Whole genome sequencing (WGS) and data analysis

Briefly, genomic DNA was extracted from sorted tdTomato$^+$ or tdTomato$^-$ cells using the DNeasy blood and tissue kit (Qiagen) following the manufacturer's instructions. WGS was performed at an average coverage of 30 × using the Illumina NovaSeq 6000 platform. We used FastQC (v0.12.1) and Trimmomatic (v0.39) to trim low-quality reads and adapter sequences from the raw FASTQ files. Qualified reads were mapped to the mouse reference genome (GRCm39, GCF_000001635.27) using BWA (v0.7.17) with the mem -M option. Picard-tools (v2.27.5) were then employed to reorder, sort, add read

groups, and mark duplicates in the aligned BAM files. Subsequently, Strelka2 (v2.9.2), Lofreq (v2.1.5), and Mutect2 (v4.3.0) were utilized to identify genome-wide de novo variants with high confidence. Concurrently, Mutect2 (v4.3.0), Scalpel (v0.5.4), and Strelka2 (v2.9.2) were independently employed to detect whole-genome de novo indels. Variants identified by all three algorithms would be regarded as true SNVs or indels. During variant calling, we employed tdTomato$^-$ data as a control to identify mutations observed in tdTomato$^+$ data for each pair of samples. Off-target sites predicted using CRISPR RGEN Tools (Cas-OFFinder; http://www.rgenome.net/cas-offinder/) [28], allowing up to six mismatches based on the *Tyr*-sgRNAs, were subsequently compared with the SNVs identified through GOTI analysis.

Genomic DNA samples from AYBE-V106W– and gGBE-edited lambs, as well as from wild-type lambs, were sequenced using the DNBSEQ-T7 platform following standard experimental procedures. Raw sequencing reads from each individual were quality-trimmed using Fastp (v0.22.0) and aligned to the domestic sheep reference genome (ARS-UI_Ramb_v2.0, GCA_016772045.1) using BWA-MEM (v0.7.17). The subsequent steps of read alignment, variant calling, and variant filtering were similar to those used in transcriptome-wide RNA analysis. To identify high-confidence (consensus) mutations, variants in the edited sheep were first called using both GATK and bcftools. The resulting variant sets were then intersected to derive a consensus set, ensuring that only mutations consistently detected by both tools were retained as high-confidence.

#### Transcriptome-wide RNA analysis

For transcriptome-wide RNA off-target analysis, raw RNA-seq data were first processed using fastp (v0.22.0) to remove adapter sequences and low-quality reads. Cleaned reads were analyzed using the RADAR pipeline. To eliminate ribosomal RNA contamination, reads were first aligned to the human RNA45SN5 reference, and those mapped to rRNA were discarded. The remaining reads were then aligned to the human reference genome using both BWA (v0.7.17-r1188) and HISAT2 (v2.2.1). Variant calling was performed using GATK HaplotypeCaller (v4.3.0.0), and the resulting VCF files were further processed and filtered using bcftools (v1.21). During VCF filtering, variants present in public databases such as dbSNP b151, the 1000 Genomes Project, and the EVS database, or located within Alu repeat regions, were excluded. Only variants on autosomal chromosomes were retained. Additional filtering criteria included: QD < 2.0, QUAL < 30.0, FS > 30.0, MQ < 40.0, MQRankSum < –12.5, ReadPosRankSum < –8.0, SOR > 3.0, and INFO/DP < 10. Sites supported by fewer than five mutated or non-mutated reads were also excluded. From the resulting high-confidence variants, AYBE-mediated (A-to-C/T/G) and gGBE-mediated (G-to-C/T) RNA editing events were extracted and visualized using the ggplot2 package in R.

#### Statistical analysis

Data from two or three biological replicate experiments are presented as the mean ± s.e.m. Statistical analyses were conducted using Graph-Pad Prism 9 (v.9.5.1) with an unpaired, two-tailed Student's t-test. A $p < 0.05$ was considered to be statistically significant.

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03838-6.

Additional file 1: Supplementary figures S1-S10 and tables S1-S4.

Additional file 2. Details of sequencing data.

**Acknowledgements**
We appreciate the support from the Gene Editing Scientific Teaching (NWAFU-GEST), High-Performance Computing and Life Science Research Core Services platforms (M. Zhou, K.R. Huang, X.R. Liu, L. Chen, and L.Q. Li) at Northwest A&F University (NWAFU). We also thank the members of Yang lab, Tong lab, and HuidaGene Therapeutics Co., Ltd. for their contributions in supplying experimental materials and valuable discussions.

**Peer review information**
Veronique van den Berghe was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

**Authors' contributions**
Y.H.W., X.W., and K.X. designed and performed the experiments. Y.H.W. and S.H. performed data analysis. Y.Y., Z.L., and M.L. performed cell transfection and FACS. Y.H.W., W.W., X.C., B.L., L.W., Q.Z., J.L., Y.S.W., and W.Z. performed animal experiments and micro-injection. Y.H.W. and K.X. wrote the manuscripts. Y.H.W., K.X., Y.C., W.W., and X.W. supervised the project. All authors read and approved the final manuscript.

**Funding**
This work was supported by the Major Science and Technology Project of Xinjiang (grant nos. 2024A02004-1 to Y.W. and 2023A02011-2 to X.W.), the National Natural Science Foundation of China (grant nos. 32441080 and 32301251 to Y.W. and 32272848 to X.W.), the Biological Breeding-Major Projects (grant nos. 2023ZD04051 to Y.W., 2023ZD04074 to K.X. and 2022ZD04014 to X.W.), the China Agricultural Research System (CARS-39–03 to X.W.), and local grants (QCYRCXM-2023–104 and 2025NC-YBXM-109 to Y.W., and 2022GD-TSLD-46 to X.W.).

**Data availability**
All data supporting the findings of this study are available in this article and its supplementary information files. No new code was generated in this study, and all analyses were performed using existing software packages. CRISPResso2, an open-source software pipeline for rapid and intuitive interpretation of genome editing experiments, is available on GitHub (https://github.com/pinellolab/CRISPResso2). The BWA software package, used for mapping DNA sequences against a large reference genome, is available on GitHub (https:/github.com/lh3/bwa). Strelka2, Mutect2, and LoFreq are SNV-calling tools based on distinct algorithms, available on GitHub (https:/github.com/Illumina/strelka); (https:/github.com/NCI-GDC/mutect2-cwl); (https://github.com/CSB5/lofreq). Scalpel, a software package for detecting INDELs (insertions and deletions), is available at (https://scalpel.sourceforge.net). Cas-OFFinder, a tool for predicting potential off-target sites of Cas9 RGENs in various species and genomes, was used via its web version at (http://www.rgenome.net/cas-offinder). RNA-seq data were analyzed using the open-source RNA-editing analysis pipeline RADAR (https:/github.com/YangLab/RADAR). HISAT2 (https://github.com/DaehwanKimLab/hisat2) and BWA (https://github.com/lh3/bwa) were used for read alignment to the reference genome. Samtools (https://github.com/samtools/samtools) was applied for data manipulation and file conversion, and GATK (https://github.com/broadinstitute/gatk) was employed for variant calling and downstream analyses. The raw WGS and RNA-seq FASTQ files used for the analysis have been deposited in the National Center for Biotechnology Information (NCBI) under accession numbers PRJNA1132788 [29] and PRJNA1293385 [30]. Detailed information on the raw data is provided in Additional file 2.

## Declarations

**Ethics approval and consent to participate**
Experiments involving mice and sheep were approved by the Executive Committee on Laboratory Animal Management and Ethical Review at Northwest A&F University. All animal studies were conducted in accordance with the relevant ethical guidelines for animal research: NWAFU-DK2023035 and NWAFU-DK2024063.

**Consent for publication**
Not applicable.

**Competing interests**
The authors declare no competing interests.

Received: 24 July 2024 Accepted: 14 October 2025
Published online: 21 October 2025

**References**

1. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533:420–4.
2. Gaudelli NM, Komor AC, Rees HA, Packer MS, Badran AH, Bryson DI, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. Nature. 2017;551:464–71.

3. Kurt IC, Zhou R, Iyer S, Garcia SP, Miller BR, Langner LM, et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. Nat Biotechnol. 2021;39:41–6.
4. Zhao D, Li J, Li S, Xin X, Hu M, Price MA, et al. Glycosylase base editors enable C-to-A and C-to-G base changes. Nat Biotechnol. 2021;39:35–40.
5. Koblan LW, Arbab M, Shen MW, Hussmann JA, Anzalone AV, Doman JL, et al. Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning. Nat Biotechnol. 2021;39:1414–25.
6. Yuan T, Yan N, Fei T, Zheng J, Meng J, Li N, et al. Optimization of C-to-G base editors with sequence context preference predictable by machine learning methods. Nat Commun. 2021;12:4902.
7. Tong H, Wang X, Liu Y, Liu N, Li Y, Luo J, et al. Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase. Nat Biotechnol. 2023;41:1080–4.
8. Chen L, Hong M, Luan C, Gao H, Ru G, Guo X, et al. Adenine transversion editors enable precise, efficient A•T-to-C•G base editing in mammalian cells and embryos. Nat Biotechnol. 2024;42:638–50.
9. Tong H, Liu N, Wei Y, Zhou Y, Li Y, Wu D, et al. Programmable deaminase-free base editors for G-to-Y conversion by engineered glycosylase. Natl Sci Rev. 2023;10:nwad143.
10. Ye L, Zhao D, Li J, Wang Y, Li B, Yang Y, et al. Glycosylase-based base editors for efficient T-to-G and C-to-G editing in mammalian cells. Nat Biotechnol. 2024;42:1538–47.
11. He Y, Zhou X, Chang C, Chen G, Liu W, Li G, et al. Protein language models-assisted optimization of a uracil-N-glycosylase variant enables programmable T-to-G and T-to-C base editing. Mol Cell. 2024;84:1257-70.e1256.
12. Tong H, Wang H, Wang X, Liu N, Li G, Wu D, et al. Development of deaminase-free T-to-S base editor and C-to-G base editor by engineered human uracil DNA glycosylase. Nat Commun. 2024;15:4897.
13. Yi Z, Zhang X, Wei X, Li J, Ren J, Zhang X, et al. Programmable DNA pyrimidine base editing via engineered uracil-DNA glycosylase. Nat Commun. 2024;15:6397.
14. Li X, Xie J, Dong C, Zheng Z, Shen R, Cao X, et al. Efficient and heritable A-to-K base editing in rice and tomato. Hortic Res. 2024;11:uhad250.
15. Li Y, Li S, Li C, Zhang C, Yan L, Li J, et al. Engineering a plant A-to-K base editor with improved performance by fusion with a transactivation module. Plant Commun. 2023;4:100667.
16. Wu X, Ren B, Liu L, Qiu S, Li X, Li P, et al. Adenine base editor incorporating the N-methylpurine DNA glycosylase MPGv3 enables efficient A-to-K base editing in rice. Plant Commun. 2023;4:100668.
17. Zhong D, Pan H, Li K, Zhou Y, Zhao F, Ye L, et al. Targeted A-to-T and A-to-C base replacement in maize using an optimized adenine base editor. Plant Biotechnol J. 2024;22:541–3.
18. Tian Y, Li X, Xie J, Zheng Z, Shen R, Cao X, et al. Targeted G-to-T base editing for generation of novel herbicide-resistance gene alleles in rice. J Integr Plant Biol. 2024;66:1048–51.
19. Zuo E, Sun Y, Wei W, Yuan T, Ying W, Sun H, et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science. 2019;364:289–92.
20. Wei Y, Li Z, Xu K, Feng H, Xie L, Li D, et al. Mitochondrial base editor DdCBE causes substantial DNA off-target editing in nuclear genome of embryos. Cell Discov. 2022;8:27.
21. Xie L, Feng H, Li Z, Li D, Yang X, Yuan T, et al. Undetectable off-target effects induced by FokI catalytic domain in mouse embryos. Genome Biol. 2024;25:51.
22. Richter MF, Zhao KT, Eton E, Lapinaite A, Newby GA, Thuronyi BW, et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat Biotechnol. 2020;38:883–91.
23. Rees HA, Komor AC, Yeh WH, Caetano-Lopes J, Warman M, Edge ASB, et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery. Nat Commun. 2017;8:15790.
24. Wang X, Niu Y, Zhou J, Yu H, Kou Q, Lei A, et al. Multiplex gene editing via CRISPR/Cas9 exhibits desirable muscle hypertrophy without detectable off-target effects in sheep. Sci Rep. 2016;6:32271.
25. Zhou S, Kalds P, Luo Q, Sun K, Zhao X, Gao Y, et al. Optimized Cas9:sgRNA delivery efficiently generates biallelic MSTN knockout sheep without affecting meat quality. BMC Genomics. 2022;23:348.
26. Wang X, Yu H, Lei A, Zhou J, Zeng W, Zhu H, et al. Generation of gene-modified goats targeting MSTN and FGF5 via zygote injection of CRISPR/Cas9 system. Sci Rep. 2015;5:13878.
27. Clement K, Rees H, Canver MC, Gehrke JM, Farouni R, Hsu JY, et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nat Biotechnol. 2019;37:224–6.
28. Bae S, Park J, Kim JS. Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics. 2014;30:1473–5.
29. Wei Y, Cao X, Huang S, Yue Y, Luo M, Liu Z, Liu B, Zhang Q, Wu Y, Wang, L, Lin J, Chen Y, Zheng W, Wu W, Xu K, Wang X. Efficient glycosylase-mediated base editing with minimal off-target effects in mammalian embryos. Mouse GOTI: Datasets. NCBI BioProject PRJNA1132788. 2024. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1132788.
30. Wei Y, Cao X, Huang S, Yue Y, Luo M, Liu Z, Liu B, Zhang Q, Wu Y, Wang, L, Lin J, Chen Y, Zheng W, Wu W, Xu K, Wang X. Efficient glycosylase-mediated base editing with minimal off-target effects in mammalian embryos. Mouse GOTI, HEK293T RNA-seq, and Sheep WGS: Datasets. NCBI BioProject PRJNA1293385. 2025. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1293385.


## Publisher's Note