Optimized gGBE enables robust G-to-Y conversion in mammalian cells and zebrafish embryos

Min Li[1,#], Yaoge Jiao[1,#], Rui Tao[2], Yun Hu[1], Junyi Fei[1], Chunting Wang[1], Shaohua Yao[1,3,*]

[1]Laboratory of Biotherapy, National Key Laboratory of Biotherapy, Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

[2]Institute of Kidney Diseases, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

[3]Tianfu Jincheng Laboratory, Chengdu 610095, Sichuan, China

[#] These authors contributed equally

[*] To whom correspondence should be addressed. Email: shaohuayao@scu.edu.cn

Zoological Research

ABSTRACT

Base editing is a powerful and precise genome-editing tool capable of introducing single-base substitutions with high resolution. Depending on the base-modification domains employed, base editors can mediate both transition and transversion base editing in mammalian cells. While deaminase-derived base editors have enabled efficient transition editing in the zebrafish genome, glycosidase-derived transversion editing remains a significant challenge. In this study, guided by structural insights into the Cas9 complex, we engineered a series of glycosidase guanine base editor(gGBE) constructs by inlaying the N-methylpurine DNA glycosylase(MPG) domain into Cas9 protein. Among these variants, we identified a high-performance gGBE in which MPG was inserted between residues G1246 and S1247 of the PI domain in *Streptococcus pyogenes* Cas9, which we named gGBE-PIGS. This variant improved editing efficiency by up to 3.5-fold in mammalian cells and 13-fold in zebrafish embryos, without compromising editing precision. Furthermore, we developed a gGBE-editing-responsive GFP reporter to monitor editing efficiency and enrich for gGBE-edited cells and embryos. Using a combination of the optimized gGBE-PIGS and the GFP reporter, we successfully generated zebrafish models carrying mutations homologous to human amyotrophic lateral sclerosis(ALS)-associated variants. Together, these findings highlight the robustness and versatility of the gGBE platform for precise base editing and underscore its promising potential for *in vivo* disease modeling and therapeutic genome engineering.

Keywords: CRISPR/Cas9; Base editing; Glycosidase; Zebrafish; Amyotrophic lateral

sclerosis

INTRODUCTION

Point mutations are a major cause of human genetic disorders, as they can disrupt normal gene expression and function, leading to pathological processes(Cao et al., 2020). To gain a deeper understanding of the mechanisms by which these mutations contribute to disease pathogenesis and to design effective therapeutic strategies, the development of efficient and precise genome editing tool is of paramount importance(Knott and Doudna, 2018). Zebrafish serve as a widely used model organism for studying gene function and embryonic development(Howe et al., 2013; Yan et al., 2018). Genomic analysis revealed that approximately 70% of human genes have at least one identifiable zebrafish ortholog (Howe et al., 2013), which makes zebrafish an ideal model system for studying human genetic diseases (Asakawa et al., 2021; Lieschke and Currie, 2007; Oliveira et al., 2023; Perles et al., 2015).

CRISPR/Cas9-mediated base editing technology enable precise single-base conversions without inducing double-strand breaks (DSBs), provides a rapid and efficient tool for generating transgenic zebrafish, enabling the precise simulation of pathogenic mutations associated with human diseases(Kantor et al., 2020). Base editors by fusing Cas9 with cytosine or adenine deaminases can efficiently introduce C-to-T or A-to-G base transitions respectively(Gaudelli et al., 2017; Komor et al., 2016). Based on these systems, cytosine glycosylase base editors(CGBEs) and, adenine glycosylase base editors(AYBEs) were developed by introducing a uracil(U) or inosine(I) specific glycosylase enzyme that recognize deaminated U or I respectively. These enzymes excise the target base to generate an apurinic/apyrimidinic(AP) site, enabling targeted

C-to-U-to-G/A or A-to-I-to-C/T conversions (Kurt et al., 2021; Tong et al., 2023). Recently, engineered uracil-DNA glycosylase(UNG) and N-methylpurine DNA glycosylase(MPG) were developed to recognize normal cytosine(C) or thymine(T) and guanine(G) respectively to directly excise these bases from target site guided by Cas9 to produce transversion base editing(He et al., 2024; Tong et al., 2023; Ye et al., 2024).

Although deaminase derived transition base editors, including both cytosine base editor(CBE) and adenine base editor(ABE), have been proved efficient in zebrafish, it remains unclear whether glycosidase derived transversion base editors can also produce efficient genome editing in zebrafish(Qin et al., 2024; Qin et al., 2025; Qin et al., 2018; Xue et al., 2023; Zhang et al., 2024; Zhang et al., 2017; Zheng et al., 2025; Zhong et al., 2025). This limited base conversion spectrum has hindered the utility of zebrafish in modeling a broader range of human pathogenic mutations. Consequently, the development of efficient transversion base editing strategies with robust efficiency in zebrafish is critically needed to facilitate the generation of more diverse and physiologically relevant disease models *in vivo*.

In this study, we re-engineered glycosidase guanine base editor(gGBE) by embedding the MPG domain at multiple Cas9 docking sites, identifying a high-performance variant gGBE-PIGS with enhanced editing efficiency in mammalian cell lines and zebrafish embryos. We also developed a universal gGBE reporter system to monitor editing efficiency and enrich edited cells and embryos. Combining optimized gGBE-PIGS with this reporter, we successfully generated mutant zebrafish carrying amyotrophic lateral sclerosis(ALS)-associated human homologous mutations. Our

work confirms gGBE functionality in zebrafish embryos and establishes an optimized protocol for efficient disease modeling.

MATERIALS AND METHODS

Zebrafish husbandry and maintenance

Zebrafish were raised and bred at 28.5°C in a circulating system according to standard methods under a cycle of 14h:10h light/dark. The development of embryos was staged by standard morphological criteria. All animal procedures were conducted in accordance with protocols approved by the Institutional Animal Care and Use Committee (IACUC) of West China Hospital, Sichuan University (No. 20250304175)

In vitro transcription of mRNA

mRNAs were synthesized via T7 High Yield RNA Transcription Kit (TR101-02, Vazyme, China). Briefly, the coding fragments of each protein were cloned into mRNA production plasmid. Capping of the *in vitro* transcribed mRNAs were performed co-transcriptionally using the trinucleotide cap1 analog, CleanCap (Cap3111, Synthgene, China). The final transcribed mRNAs were purified using RNA clean Kit (DP412, TIANGEN, China).

Embryo microinjection

C-gGBE or gGBE-PIGS mRNA (100 ng/μL) and sgRNA (100 ng/μL) were mixed and co-injected into zebrafish embryos of one-cell stage. Injected embryos were incubated at 28.5°C. Embryos that developed normally at 48h were collected in one group. Fluorescence images were acquired using a stereofluorescence microscope (MZ10F, Leica Microsystems, Germany) . Genomic DNA was extracted by using 50 μL of freshly prepared lysis buffer (10 mM Tris-HCl pH 7.5, 0.05% SDS, 2 mg/μL Proteinase K (ST535, Beyotime, China)) into each well of the tissue culture plate. The

mixture was incubated at 55°C for 2h and then was heat-inactivated at 95°C for another 30min.

Cell culture and transfection

HEK293T cells were cultured in Dulbecco's modified Eagles's medium (SP03, SPERIKON, China), supplemented with 10% (v/v) fetal bovine serum (FSD500, ExCell, China), 1% penicillin/streptomycin (SP00, SPERIKON, China) at 37°C with 5% CO_2. They were seeded on a 96-well plate (TCP011096, BIOFIL, China) 12h–16h before transfection, and each well was seeded with 2×10^5 cells. Transfection was conducted using 400 ng of plasmids (300 ng of gGBE plasmid DNA, 100 ng of sgRNA) and 0.7 μL of lipomax(32012, SUDGEN, China) at a confluence of ~70%–80%, according to the manufacturer's protocol. 72h post transfection, genomic DNA was extracted for further analysis.

Design and construction of plasmid

C-gGBE plasmid used in this study was derived from gGBE plasmid from Yang Hu(Tong et al., 2023), which was constructed in the pVAX vector backbone. Briefly, C-gGBE plasmid was used as a parent plasmid to generate gGBE-PIGS, 1022-gGBE, 1054-gGBE, 1068-gGBE, 1300-gGBE or 769-gGBE. MPG with linkers was PCR amplified and cloned into the C-gGBE derived fragment with C-terminal MPG deleted. SgRNA expression vectors were constructed by ligating annealed oligonucleotide duplexes into pU6 sgRNA cut with Bbs1. Oligos used to generate sgRNA expression plasmids were listed in Supplementary Table S1. All plasmids were verified by Sanger sequencing (Shanghai Sangon, China).

High-throughput DNA sequencing and data analysis

Genomic DNA regions of interest were amplified with Phanta Max Super-Fidelity polymerase (P505, Vazyme, China) with primers harboring individual barcodes to distinguish different samples. Primers used to amplify flanking region of each on- or off-target site are listed in Supplementary Tables S2-S5. The amplicons were purified with Oriscience Fast Gel Recovery kit (NB402, Oriscience, China) according to the manufacturer's instructions. Amplicon samples were sequenced commercially using Illumina HiSeq platform (Shanghai Sangon, China). Frequencies of base conversions and indels were quantitated with CRISPResso2 (Clement et al., 2019).

GFP-based reporter cells sorting

At 72h post-transfection, cells were harvested by centrifugation at 1 200 r/min for 3 min and subsequently resuspended in flow cytometry sorting buffer (PBS supplemented with 1% FBS). Fluorescence images were acquired using an inverted fluorescence imaging system (Ts2R, Nikon, Japan) . GFP-positive cells were then sorted by flow cytometry (BD FACSAria™ III, USA). The sorted cells were used for genomic DNA extraction.

RESULTS

C-gGBE induces G-to-Y base conversion in zebrafish

To investigate whether gGBE can mediate G-to-Y base editing in the zebrafish genome, we tested its performance with a previously validated GFP target for SpCas9 nuclease(Auer et al., 2014). To begin our study, we employed a previously developed gGBE(gGBE6.3) in mammalian cell base editing, in which the engineered MPG that recognize and cleave natural guanine was fused to the C-terminus of Cas9, hereafter referred to as C-gGBE. We generated C-gGBE mRNA and GFP-sgRNA through *in vitro* transcription (Supplementary Figure S1) , and then injected the mixture into one-cell stage embryos from transgenic zebrafish harboring FLK1-GFP(Choi et al., 2007). After 48h post-fertilization, we extracted genomic DNA from the embryos and performed PCR amplification and targeted amplicon high-throughput sequencing (HTS) analysis (Figure 1A). The HTS results revealed that C-gGBE induced approximately 5% G base transitions at the GFP locus (Figure 1B). These editing events included 60% G-to-C, 25% G-to-T, and 15% G-to-A transitions(Supplementary Figure S2).

Next, we selected 18 endogenous zebrafish loci for further characterization, all of which had previously been shown to be efficiently edited using Cas9 nuclease or its derivate base editors (Ear et al., 2016; Lu et al., 2018; Rosello et al., 2021; Sun et al., 2019; Zhang et al., 2017). We co-injected sgRNAs targeting these loci along with C-gGBE mRNA into wildtype AB zebrafish embryos at one-cell stage and then analyzed the editing outcomes at 48h by HTS. The results showed that C-gGBE-mediated editing occurred with an efficiency of ~0.5% to 28% (Figure 1C-F). On average, among the G

editing events, about 60% were G-to-C transitions, ~35% were G-to-T transitions, and ~5% were G-to-A transitions (Supplementary Figure S2). Together with the observations on the GFP target, the major editing outcomes of gGBE in zebrafish embryos are G-to-C transitions. This is consistent with glycosidase-dependent base editing in human and mouse cells, where G-to-C transitions are preferred by gGBE, AYBE, and gTBE, suggesting that these species share a similar translesion synthesis (TLS) preference against AP site. In summary, these results demonstrated that C-gGBE could produce G to Y base editing in zebrafish embryos, albeit at relatively lower efficiency as compared to CBE and ABE(Lu et al., 2018; Qin et al., 2018).

Domain inlaying gGBE-PIGS improve editing efficiency

The limited efficiency of gGBE in zebrafish embryos prompted us to optimize the system. The workflow of gGBE started with the generation of AP site via the hydrolysis of glycosidic linkage to excise the target guanosine from DNA, which was then repaired through subsequent endogenous DNA repair or replication processes to either cleave the AP site to turn back to G via base excision repair (BER) or to introduce indels in combination with the Cas9 nickase cleavage site, or achieve transversion base editing via error-prone DNA polymerases mediated TLS. We hypothesized that AP sites at a position relative closer to the Cas9 cleavage site might be more prone to TLS mediated AP base editing, potentially improving editing efficiency. To test this hypothesis, we designed an internal inlaid gGBE base editor by inserting MPG in-between the 1246G and 1247S of the spCas9 PI domain, which we previously showed to enlarge and back-ward shift the editing window of CBE(Wang et al., 2019).

Thereafter, the resulting new gGBE was referred to as gGBE-PIGS.

We compared the efficiency of gGBE-PIGS with its original variant, C-gGBE, in HEK293T mammalian cells. The comparison revealed that the gGBE-PIGS system achieved significantly higher G-to-Y conversion rates than the conventional C-gGBE across all targeted genomic loci examined (Figure 2B). On average, the efficiency of gGBE-PIGS was about 1.7-fold higher than that of C-gGBE, with the highest increase reaching 3.5-fold (Figure 2C). However, this enhancement appeared to be unrelated to our hypothesis of a backward-shifting editing window. Because when comparing G base editing efficiency at the same target, we observed little difference in the relative editing ratios between Gs located proximal or distal to the PAM (nick site). We used AlphaFold3 to predict the structures of gGBE-PIGS and C-gGBE. By measuring the distance between the catalytic pocket residue Tyr162 of MPG and the edited base on the non-target strand (NTS), we found that MPG in gGBE-PIGS was localized closer to the NTS (52.9 Å) than in C-gGBE (62.5 Å) (Figure 2A). Therefore, these results indicated that domain inlaying of gGBE-PIGS improved editing efficiency via enhancing the action of MPG rather than shifting the editing window.

Moreover, we compared the indel ratios of all target sites tested and found that C-gGBE and gGBE-PIGS exhibited comparable indel scores (indels VS on-target editing , Figure 2D-F). Furthermore, compared to C-gGBE, gGBE-PIGS exhibited similar base conversion ratios, i.e., the proportions of G-to-C, G-to-A, and G-to-T conversions (Supplementary Figure S3). Off-target analysis revealed that gGBE-PIGS slightly reduced off-target activities in both Cas9-dependent (Supplementary Figure S4) and

independent off-target sites (Supplementary Figure S5).

To assess potential RNA off-target effects of gGBE, we designed sgRNAs harboring G-containing 3'-extended sequence to investigate whether MPG could potentially introduce unintended edits on this extended Gs. This design was motivated by two observations: (1) structural studies have shown that the 3' end of the sgRNA is partially exposed in the Cas9-sgRNA-DNA complex, and (2) in prime editing system, the 3' extension of the pegRNA can hybridize with the nicked non-target DNA strand and is recognized by the C-terminally fused reverse transcriptase. Therefore, if gGBE has RNA off-target editing, the sgRNA extension would be more likely to be edited than untethered RNA molecules. We co-transfected these 3'-extended sgRNAs together with either C-gGBE or gGBE-PIGS and performed deep sequencing of the corresponding cDNAs. To validate the sensitivity and feasibility of this assay, we evaluated its performance using ABE8e, which is known to exhibit substantial RNA off-target activity (Chen et al., 2023; Li et al., 2021; Rees et al., 2019). This analysis revealed that ABE8e induced ~ 15% RNA editing events specifically within the sgRNA extensions, thereby confirming the feasibility of our approach for assessing RNA off-target effects of base editors. We then applied this assay to evaluate the RNA off-target activities of C-gGBE and gGBE-PIGS. Compared with the nCas9 control, neither C-gGBE nor gGBE-PIGS produced detectable RNA-level off-target edits (Supplementary Figure S6).These results indicated that the MPG based glycosylase editors did not produce RNA off-target editing as tethered RNA molecules were not edited, which is consistent with the fact that natural MPG primarily recognizes and excises damaged

bases in DNA rather than RNA (Miao et al., 1998).

Previous studies have demonstrated that single or double strand DNA binding domains can enhance Cas9 mediated and deaminase-based cytosine or adenine base editing efficiency(Yin et al., 2023; Zhang et al., 2020). To further improve the editing efficiency of gGBE-PIGS, we fused it with a commonly used DNA-binding domain derived from *Sulfolobus solfataricus*, sso7d, which has proven effective in both base editor and CRISPRi architectures(Choli et al., 1988; Wang et al., 2024; Wang et al., 2004). Sso7d DNA binding domain was human codon optimized and fused to the C-terminus of gGBE-PIGS. However, compared with gGBE-PIGS, the editing efficiency of sso7d fusion editor showed no obvious improvement (Supplementary Figure S7) .

Characterization of inlaid gGBEs in different docking sites

The improvement in efficiency observed with gGBE-PIGS prompted us to screen additional inlaying sites to further optimize gGBE performance. Guided by the structural information of the SpCas9-RNA-DNA complex (Figure 3A), we designed five additional inlaid gGBEs, with MPG inserted at positions 1022, 1054, 1068, 1300, and 769 of SpCas9 (Figure 3B). These positions were selected because they are located within unstructured loops on the surface of the SpCas9 complex, minimizing the risk of disrupting the Cas9 folding(Liu et al., 2020). Additionally, these sites were positioned on the NTS from different orientations, which diversified the interactions between MPG and the NTS and hopefully led to different editing windows.

These variants were first tested in HEK293T. Compared to the original C-gGBE, the 1022-, 1054-, 1300-, and 769-gGBE variants exhibited higher editing efficiencies

at most target sites, with the maximum improvement reaching up to 4.4-fold (Figure 3C). However, the 1068-gGBE variant showed similar or even lower efficiency compared to C-gGBE. Despite this, all the new variants exhibited lower efficiency than gGBE-PIGS, except for the 1300- and 769-gGBE variants, which demonstrated higher efficiency at certain target sites. Specifically, the 1300-gGBE and 769-gGBE variants outperformed gGBE at *FANCF* and *CHM*, respectively (Figure 3D). This suggests that for specific target loci, different internal inlaid gGBEs should be screened to optimize the editing strategy.

Additionally, product purity analysis revealed that the frequencies of unintended indels induced by different gGBE variants were closely related to G conversion efficiencies in most cases (Supplementary Figure S8). The conversion ratios of G to non-G bases were similar across the different gGBE variants, with G to C being the most frequent (Supplementary Figure S9).

Editing strategy of gGBE-PIGS in zebrafish and modeling pathogenic mutations

Next, we tested the performance of the most efficient editor, gGBE-PIGS, in zebrafish embryos. We first compared the activity of C-gGBE and gGBE-PIGS at the previously characterized GFP target. We found that gGBE-PIGS significantly improved editing at G7 and G10, while reducing editing at G6 (Figure 4A). The improvement with gGBE-PIGS was more pronounced at endogenous target sites, with gGBE-PIGS showing obvious higher efficiencyiy than C-gGBE in 13 out of 18 sites. Especially for the *trmt10c* site, gGBE-PIGS increased the editing efficiency by 10.7-fold (Figure 4A). Together, these results suggest that the insertion of MPG between the

1246G and 1247S of SpCas9 improved gGBE editing in zebrafish.

A comparison between the editing windows of C-gGBE and gGBE-PIGS revealed that both editors showed an active window spanning positions 1–11 (Figure 4B), consistent with the editing window previously reported for C-gGBE in mammalian cells. Notably, we also observed a "skip-editing" pattern at several target sites, where certain editable Gs within the putative window remained unedited despite efficient editing at neighboring positions (Supplementary Figure S10). Interestingly, most of the those skipped Gs were followed by a C or T at their 3' position, suggesting a motif preference of both editors. To test this, we summarized all Gs within the putative editing window across those 19 loci and found that there was a strong NGR motif preference for both editors (N= A, T, G, or C and R=G or A) (Figure 4C; Supplementary Figure S11).

Next, we applied the optimized gGBE to the construction of disease models. ALS is a heterogenous neurodegenerative disease that leads to progressive disability and motor impairments. *Sod1* and *fus* are among the most frequent mutant genes in familial ALS. We designed a gGBE-based editing strategy to introduce G-to-C mutations at the *fus* (R536G) and *sod1* (F21L) loci (Figure 4D, E) to model human disease-associated mutations. Notably, compared to C-gGBE, gGBE-PIGS exhibited enhanced editing efficiency at both target sites (Figure 4F). At the *sod1* locus, the G-to-C editing efficiency increased from 1.26% to 2.66%, and at the *fus* locus, it rose from 2.41% to 3.59%(Supplementary Figure S12).

A GFP-based reporter strategy for enrichment of gGBE-PIGS-edited cells and embryos.

To further improve the efficiency of gGBE-PIGS in disease modeling and gene function studies, we aimed to develop a GFP reporter system responsive to gGBE activity. This system would enable enrichment of edited cells or embryos and facilitate the screening of positive colonies. Like other base editors that generate AP sites, gGBE frequently induces indels during genome editing, likely due to the cleavage of AP sites by AP endonucleases or lyases. We leveraged this property to design an indel-responsive GFP reporter. Under the control of the CMV promoter, a gGBE-PIGS target site was inserted between the coding sequences of mCherry and GFP. This target site was placed out of frame (+1) relative to GFP, thereby suppressing GFP expression. In theory, in-frame indels introduced by gGBE would restore the reading frame, leading to GFP expression (Figure 5A). Therefore, a corresponding reporter system can be established for any target site, except those containing multiple out-of-frame stop codons.

We first validated the reporter targeting siteA in HEK293T cells. Transfection with the reporter plasmid alone did not produce detectable GFP signals under fluorescence microscopy (Supplementary Figure S13A). As expected, GFP signals were observed specifically in cells co-transfection of reporter, gGBE-PIGS, and sgRNA, but not in those transfected with the reporter and gGBE-PIGS (without sgRNA), confirming the specificity of the gGBE reporter (Supplementary Figure S13B). GFP-positive cells were subsequently isolated by fluorescence-activated cell sorting (FACS).

Within the mCherry-positive population, approximately 1.35% showed GFP fluorescence (Supplementary Figure S13C). Importantly, high-throughput quantification of editing efficiency revealed an approximately two-fold increase in the sorted group compared to the unsorted control, indicating substantial enrichment of base-edited cells (Figure 5B).

Next, we replaced the target sequence between mCherry and GFP with the zebrafish fus target site. In HEK293T cells, the *fus* reporter system exhibited robust GFP fluorescence in response to gGBE-PIGS editing (Supplementary Figure S14A). However, when introduced into zebrafish embryos via microinjection, the reporter signal in response to *fus* gGBE-PIGS editing was relatively weak (Supplementary Figure S14B). We hypothesized that the weak reporter signal in zebrafish embryos might be due to the limited efficiency of indel formation and the low transcriptional activity of the CMV promoter in zebrafish. To address this, we replaced the CMV promoter with the zebrafish β-actin promoter. Following microinjection into zebrafish embryos, the modified reporter construct exhibited strong fluorescence in response to *fus* gGBE-PIGS editing (Supplementary Figure S14C).

Interestingly, among the co-injected embryos, we observed two distinct populations: some displayed red fluorescence only, while others exhibited both red and green fluorescence(Figure 5C). HTS analysis of editing outcomes revealed that embryos expressing both red and green fluorescence showed higher editing efficiency at the *fus* locus compared to those with red fluorescence alone(Figure 5D).

The segregation of red and green fluoresce signals might indicate a different distribution of base editor mRNAs and reporter DNAs. To establish an optimized injection strategy, we test the effect of injection time to reporter response. We observed that embryos injected at 15min–25min post-fertilization exhibited a higher proportion of fluorescence segregation compared to those injected at 5min–10min post-fertilization(Figure 5E).

Based on this optimized editing strategy, we injected and raised the mosaic F0 fishes to adulthood, and selected individuals with higher mosaic ratios for germline transmission analysis(Supplementary Figure S15). The offsprings generated by outcrossing these F0 fishes with wild-type partners were raised to 1 month old, and approximately 20 F1 fishes were pooled for initial genotyping by HTS. Pools identified as positive for the intended mutation were further examined by individual genotyping, which led to the identification of three R536G and one R536C F1 fishes out of a total of 80 examined (Figure 5F; Supplementary Figure S15). These heterozygous fish did not exhibit obvious phenotypic changes at 3 months of age and were fertile, which may reflect the fact that the corresponding R521 mutations typically cause disease onset around 40 years of age(Kwiatkowski et al., 2009). Together, these results demonstrate that the intended base edits produced by gGBE can be successfully transmitted through the germline.

DISCUSSION

In summary, we re-engineered the architecture of gGBE by inlaying the MPG domain into different docking sites of Cas9 protein and identified a high-performance gGBE-PIGS that showed the highest editing efficiencies across different targets in both human cells and zebrafish embryos. This variant also exhibited reduced off-target effect as compared to original gGBE version. Moreover, we designed a gGBE-editing-responsive reporter to enrich gGBE-edited cells and embryos, which significantly increased the potions of editing frequencies. Combining our optimized gGBE-PIGS and thegGBE reporter, we introduced mutations homologous to human ALS-associated variants to zebrafish embryos.

Moreover, our analysis of the editing window revealed substantial site-to-site variation in the position of maximal editing. For example, peak editing occurred around G5–10 at the *pkn1a*, *dmd*, and *si:ch73-6k14.2* loci, but shifted toward G1–3 at the *get3*, *mtnr1aa* and *trmt10c* loci. In addition, at some targeting sites such as *ctnnb1*, *rps14* and *alkbh5*, gGBE exhibited a distinct ‘skip-editing’ pattern in which the flanking guanines were edited whereas the intervening guanines remained largely unedited. Notably, most of these skip-editing events can be explained by the NGR motif preference of gGBE: the non-edited internal Gs were predominantly followed by pyrimidines rather than purines. Taken together, these findings indicate that the varied peak positions we observed are governed by at least two factors: (i) differential accessibility of nucleotides within the editing window and (ii) the intrinsic motif preference of gGBE. Both factors are shaped by the local sequence context of each target site.

During gGBE-mediated editing, the MPG domain initiates base editing by excising the target guanine, thereby generating an AP site. This intermediate can be resolved either by BER, restoring the original G or inducing indels, or by TLS, leading to G-to-Y conversions(Mullins et al., 2019; Tong et al., 2023). we initially hypothesized that AP sites located closer to the DNA cleavage site are more likely to undergo TLS, thereby enhancing editing efficiency. Previous studies have demonstrated that insertion of a deaminase at position 1246 of Cas9 shifts the editing window downstream without significantly affecting editing efficiency(Wang et al., 2019). In our study, insertion of MPG at the same site did not shift the editing window but resulted in a marked increase in editing efficiency in mammalian cells. This improvement is unlikely to be solely attributed to spatial proximity because the same or similar docking site did not improve the efficiencies of deaminases derived base editors(Chen et al., 2024; Wang et al., 2019). We proposed that internal insertion at this site may alter the DNA-binding dynamics of Cas9, which altered the disassociation of Cas9 protein from the target after the base excision and TS cleavage, thereby influencing the selection of DNA repair pathways promoting TLS pathway. We also hypothesized that insertion of other domains—such as the UNG domain from CGBE—or even catalytically inactive domains at position 1246 has similar effect by reshaping the local repair environment.

Notably, we observed a marked discrepancy in gGBE editing efficiency between mammalian cells and zebrafish embryos. While Cas9 nucleases and canonical base editors such as CBEs and ABEs generally maintain robust activity across both systems(Qin et al., 2018; Schmidt et al., 2021; Zhang et al., 2017), gGBE exhibit

significantly lower efficiency in zebrafish, contrasting sharply with their performance in mammalian cells. Several factors may contribute to this inconsistency. First, the MPG domain incorporated into gGBEs is sourced from the human genome and may exhibit suboptimal enzymatic activity at the lower incubation temperatures (28℃) typically used for zebrafish development. Second, the DNA damage response and repair festures in zebrafish embryos may differ from that of mammalian cells. For example, embryonic cells in zebrafish have a short cell cycle of approximately 20min, lacking the intervening gap phases, G1 and G2(Kimmel et al., 1995; Verduzco and Amatruda, 2011) . DNA repair pathway choice is tightly regulated by cell cycle as many DNA damage responsive proteins and repair proteins are specifically induced or activated during specific cell cycles(Kostyrko et al., 2015; Langerak and Russell, 2011; Symington and Gautier, 2011). For example, the levels of TLS polymerases, such as pol η and Rev1, are tightly regulated during the cell cycle in eukaryotes. In *S. cerevisiae*, the levels of pol η and Rev1 peak during the G2/M phase compared to the G1 and early S phases. Specifically, the levels of pol η and Rev1 in the G2/M phase are approximately 3-fold and 50-fold higher, respectively, than those in the G1 and early S phases(Pabla et al., 2008; Plachta et al., 2015; Waters and Walker, 2006; Wiltrout and Walker, 2011; Zhao and Washington, 2017).

Our observation of the differences in fluorescent reporter segregation at various injection time suggest that mRNA and DNA exhibit distinct distribution and diffusion patterns upon microinjection. In particular, injections performed within the first 15min post-fertilization show greater asynchrony between mRNA and DNA dispersion,

potentially due to limited cytoplasmic volume or spatial constraints. Based on these observations, we recommend co-injection of mRNA and DNA after 15min post-fertilization to improve consistency in distribution for better enrichment.

DATA AVAILABILITY

All data supporting the findings of this study are available in the article and its supplementary figures and tables. Deep-sequencing data are available under BioProject ID PRJNA1299155.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

S.Y. and M.L. have submitted a patent application (No. 202511026602.9) to the China National Intellectual Property Administration based on the results reported in this study. The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

S.Y. conceived this study; S.Y. and M.L. designed the experiments; M.L. and Y. J. construct expression plasmids and performed cell transfection; M.L., Y. J., and J.F. performed HTS sequencing and data analysis; R.T. and Y.H. performed zebrafish embryos micro-injection; M.L. wrote the manuscript; S.Y. and C.W. edited the manuscript. All authors read and approved the final version of the manuscript.

ACKNOWLEDGMENTS

This research was supported by National Key R&D Program of China (2023YFC3403200), Sichuan Science and Technology Program (2024NSFTD0029 and 2025NSFTD0030), China Postdoctoral Science Foundation(2024M762207) and the Postdoctor Research Fund of West China Hospital, Sichuan University (2024HXBH127).

REFERENCE

Asakawa K, Handa H, Kawakami K. 2021. Illuminating ALS Motor Neurons With Optogenetics in Zebrafish. *Frontiers in Cell and Developmental Biology,* **9**:640414.

Auer TO, Duroure K, De Cian A, et al. 2014. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Research,* **24**(1):142-153.

Cao Y, Li L, Xu M, et al. 2020. The ChinaMAP analytics of deep whole genome sequences in 10,588 individuals. *Cell Research,* **30**(9):717-731.

Chen L, Hong M, Luan C, et al. 2024. Adenine transversion editors enable precise, efficient A*T-to-C*G base editing in mammalian cells and embryos. *Nature Biotechnology,* **42**(4):638-650.

Chen L, Zhang S, Xue N, et al. 2023. Engineering a precise adenine base editor with minimal bystander editing. *Nature Chemical Biology,* **19**(1):101-110.

Choi J, Dong L, Ahn J, et al. 2007. FoxH1 negatively modulates flk1 gene expression and vascular formation in zebrafish. *Developmental Biology,* **304**(2):735-744.

Choli T, Henning P, Wittmann-Liebold B, et al. 1988. Isolation, characterization and microsequence analysis of a small basic methylated DNA-binding protein from the Archaebacterium, Sulfolobus solfataricus. *Biochimica et Biophysica acta,* **950**(2):193-203.

Clement K, Rees H, Canver MC, et al. 2019. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology,* **37**(3):224-226.

Ear J, Hsueh J, Nguyen M, et al. 2016. A Zebrafish Model of 5q-Syndrome Using CRISPR/Cas9 Targeting RPS14 Reveals a p53-Independent and p53-Dependent Mechanism of Erythroid Failure. *Journal of Genetics and Genomics,* **43**(5):307-318.

Gaudelli NM, Komor AC, Rees HA, et al. 2017. Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature,* **551**(7681):464-471.

He Y, Zhou X, Chang C, et al. 2024. Protein language models-assisted optimization of a uracil-N-glycosylase variant enables programmable T-to-G and T-to-C base editing. *Molecular Cell,* **84**(7):1257-1270.

Howe K, Clark MD, Torroja CF, et al. 2013. The zebrafish reference genome sequence and its relationship to the human genome. *Nature,* **496**(7446):498-503.

Kantor A, McClements ME, MacLaren RE. 2020. CRISPR-Cas9 DNA Base-Editing and Prime-Editing. *International Journal of Molecular Sciences,* **21**(17):6240.

Kimmel CB, Ballard WW, Kimmel SR, et al. 1995. Stages of embryonic development of the zebrafish. *Developmental Dynamics,* **203**(3):253-310.

Knott GJ, Doudna JA. 2018. CRISPR-Cas guides the future of genetic engineering. *Science,* **361**(6405):866-869.

Komor AC, Kim YB, Packer MS, et al. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature,* **533**(7603):420-424.

Kostyrko K, Bosshard S, Urban Z, et al. 2015. A role for homologous recombination proteins in cell cycle regulation. *Cell Cycle,* **14**(17):2853-2861.

Kurt IC, Zhou R, Iyer S, et al. 2021. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nature Biotechnology,* **39**(1):41-46.

Kwiatkowski TJJ, Bosco DA, Leclerc AL, et al. 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (New York, N.Y.),* **323**(5918):1205-1208.

Langerak P, Russell P. 2011. Regulatory networks integrating cell cycle control with DNA damage checkpoints and double-strand break repair. *Philosophical Transactions of the Royal Society of London.*

Series B, Biological Sciences, **366**(1584):3562-3571.
Li J, Yu W, Huang S, et al. 2021. Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. *Nature Communications,* **12**(1):2287.
Lieschke GJ, Currie PD. 2007. Animal models of human disease: zebrafish swim into view. *Nature Reviews Genetics,* **8**(5):353-367.
Liu Y, Zhou C, Huang S, et al. 2020. A Cas-embedding strategy for minimizing off-target effects of DNA base editors. *Nature Communications,* **11**(1):6073.
Lu X, Liu Y, Yan G, et al. 2018. Optimized Target-AID system efficiently induces single base changes in zebrafish. *Journal of Genetics and Genomics,* **45**(4):215-217.
Mullins EA, Rodriguez AA, Bradley NP, et al. 2019. Emerging Roles of DNA Glycosylases and the Base Excision Repair Pathway. *Trends in Biochemical Sciences,* **44**(9):765-781.
Oliveira N, Pinho BR, Oliveira J. 2023. Swimming against ALS: How to model disease in zebrafish for pathophysiological and behavioral studies. *Neuroscience and Biobehavioral Reviews,* **148**:105138.
Pabla R, Rozario D, Siede W. 2008. Regulation of Saccharomyces cerevisiae DNA polymerase eta transcript and protein. *Radiation and Environmental Biophysics,* **47**(1):157-168.
Perles Z, Moon S, Ta-Shma A, et al. 2015. A human laterality disorder caused by a homozygous deleterious mutation in MMP21. *Journal of Medical Genetics,* **52**(12):840-847.
Plachta M, Halas A, McIntyre J, et al. 2015. The steady-state level and stability of TLS polymerase eta are cell cycle dependent in the yeast S. cerevisiae. *DNA Repair (Amst),* **29**:147-153.
Qin W, Liang F, Lin SJ, et al. 2024. ABE-ultramax for high-efficiency biallelic adenine base editing in zebrafish. *Nature Communications,* **15**(1):5613.
Qin W, Lin S, Zhang Y, et al. 2025. Rationally Designed TadA-Derived Cytosine Editors Enable Context-Independent Zebrafish Genome Editing. *Advanced science (Weinheim, Baden-Wurttemberg, Germany),* **12**(39):e9800.
Qin W, Lu X, Liu Y, et al. 2018. Precise A*T to G*C base editing in the zebrafish genome. *Bmc Biology,* **16**(1):139.
Rees HA, Wilson C, Doman JL, et al. 2019. Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Science Advances,* **5**(5):x5717.
Rosello M, Vougny J, Czarny F, et al. 2021. Precise base editing for the in vivo study of developmental signaling and human pathologies in zebrafish. *Elife,* **10**:e65552.
Schmidt MJ, Gupta A, Bednarski C, et al. 2021. Improved CRISPR genome editing using small highly active and specific engineered RNA-guided nucleases. *Nature Communications,* **12**(1):4219.
Sun Y, Zhang B, Luo L, et al. 2019. Systematic genome editing of the genes on zebrafish Chromosome 1 by CRISPR/Cas9. *Genome Research,* **30**(1):118-126.
Symington LS, Gautier J. 2011. Double-strand break end resection and repair pathway choice. *Annual Review of Genetics,* **45**:247-271.
Tong H, Liu N, Wei Y, et al. 2023. Programmable deaminase-free base editors for G-to-Y conversion by engineered glycosylase. *National Science Review,* **10**(8):d143.
Tong H, Wang X, Liu Y, et al. 2023. Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase. *Nature Biotechnology,* **41**(8):1080-1084.
Verduzco D, Amatruda JF. 2011. Analysis of cell proliferation, senescence, and cell death in zebrafish embryos. *Methods in Cell Biology,* **101**:19-38.
Wang X, Li L, Guo L, et al. 2024. Robust miniature Cas-based transcriptional modulation by engineering Un1Cas12f1 and tethering Sso7d. *Molecular Therapy.* **32**(4):910-919.

Wang Y, Prosen DE, Mei L, et al. 2004. A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro. *Nucleic Acids Research,* **32**(3):1197-1207.
Wang Y, Zhou L, Liu N, et al. 2019. BE-PIGS: a base-editing tool with deaminases inlaid into Cas9 PI domain significantly expanded the editing scope. *Signal Transduction and Targeted Therapy,* **4**:36.
Waters LS, Walker GC. 2006. The critical mutagenic translesion DNA polymerase Rev1 is highly expressed during G(2)/M phase rather than S phase. *Proceedings of the National Academy of Sciences of the United States of America,* **103**(24):8971-8976.
Wiltrout ME, Walker GC. 2011. Proteasomal regulation of the mutagenic translesion DNA polymerase, Saccharomyces cerevisiae Rev1. *DNA Repair (Amst),* **10**(2):169-175.
Xue N, Liu X, Zhang D, et al. 2023. Improving adenine and dual base editors through introduction of TadA-8e and Rad51DBD. *Nature Communications,* **14**(1):1224.
Yan L, Chen J, Zhu X, et al. 2018. Maternal Huluwa dictates the embryonic body axis through beta-catenin in vertebrates. *Science,* **362**(6417).
Ye L, Zhao D, Li J, et al. 2024. Glycosylase-based base editors for efficient T-to-G and C-to-G editing in mammalian cells. *Nature Biotechnology,* **42**(10):1538-1547.
Yin S, Zhang M, Liu Y, et al. 2023. Engineering of efficiency-enhanced Cas9 and base editors with improved gene therapy efficacies. *Molecular Therapy,* **31**(3):744-759.
Zhang X, Chen L, Zhu B, et al. 2020. Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain. *Nature Cell Biology,* **22**(6):740-750.
Zhang Y, Liu Y, Qin W, et al. 2024. Cytosine base editors with increased PAM and deaminase motif flexibility for gene editing in zebrafish. *Nature Communications,* **15**(1):9526.
Zhang Y, Qin W, Lu X, et al. 2017. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nature Communications,* **8**(1):118.
Zhao L, Washington MT. 2017. Translesion Synthesis: Insights into the Selection and Switching of DNA Polymerases. *Genes (Basel),* **8**(1):24.
Zheng S, Liu Y, Xia X, et al. 2025. Sequence Context-Agnostic TadA-Derived Cytosine Base Editors for Genome-Wide Editing in Zebrafish. *Advanced Science (Weinheim, Baden-Wurttemberg, Germany),* **12**(14):e2411478.
Zhong C, Yu L, Zhao T, et al. 2025. A Plug-in system for reprogramming the editing patterns of base editors. *Nature Communications,* **17**(1):910.

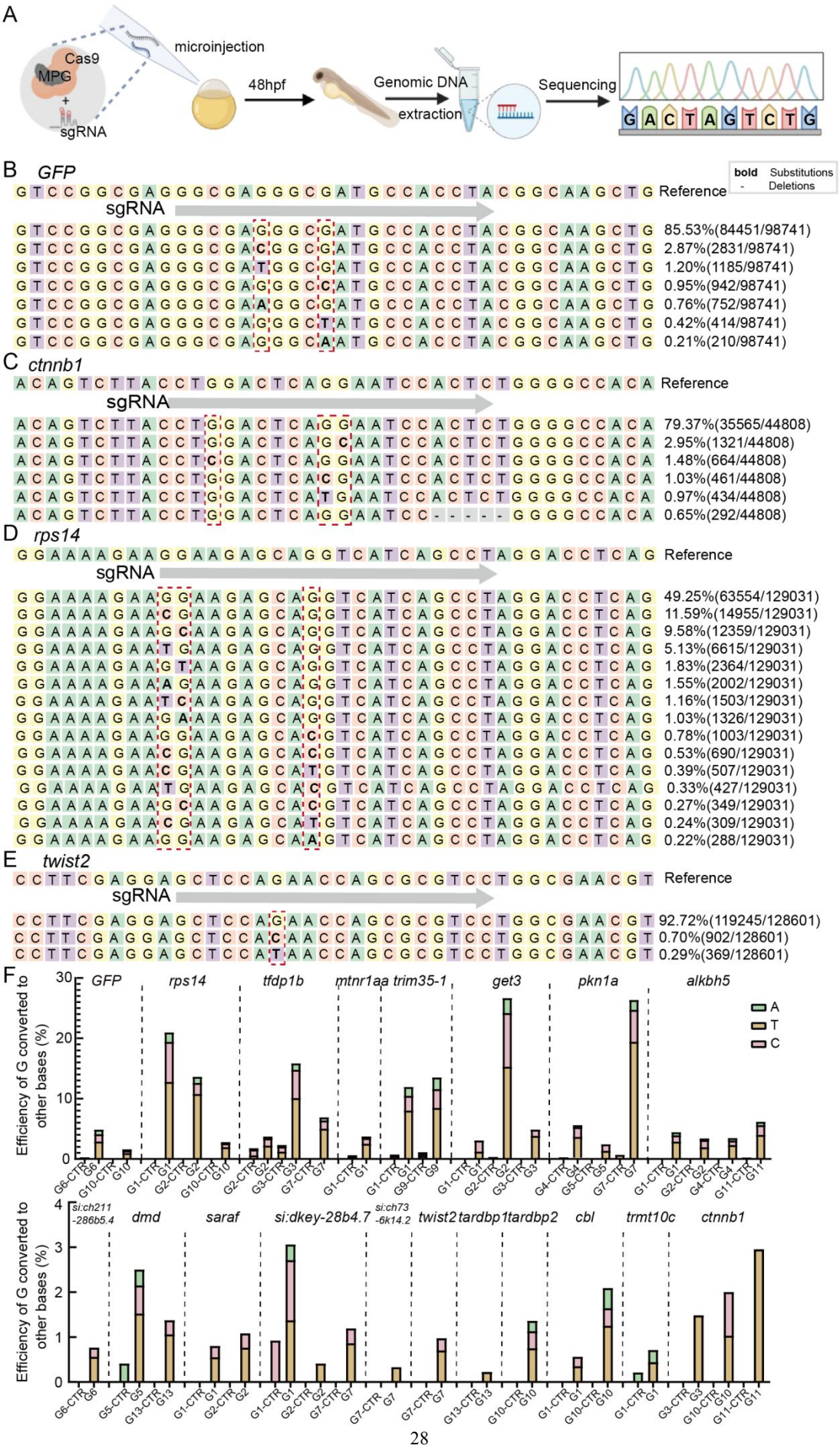


Figure1 C-gGBE base editing in zebrafish

A: Workflow diagram of C-gGBE genome editing in zebrafish. B: HTS analysis of C-gGBE-mediated editing at the GFP locus. C-E: The HTS results of the *ctnnb1*, *rps14*, and *twist2* gene in the C-gGBE system. The edited G is highlighted within the red dashed box. F: Bar plot summarizing HTS results of C-gGBE editing at 19 loci in zebrafish.

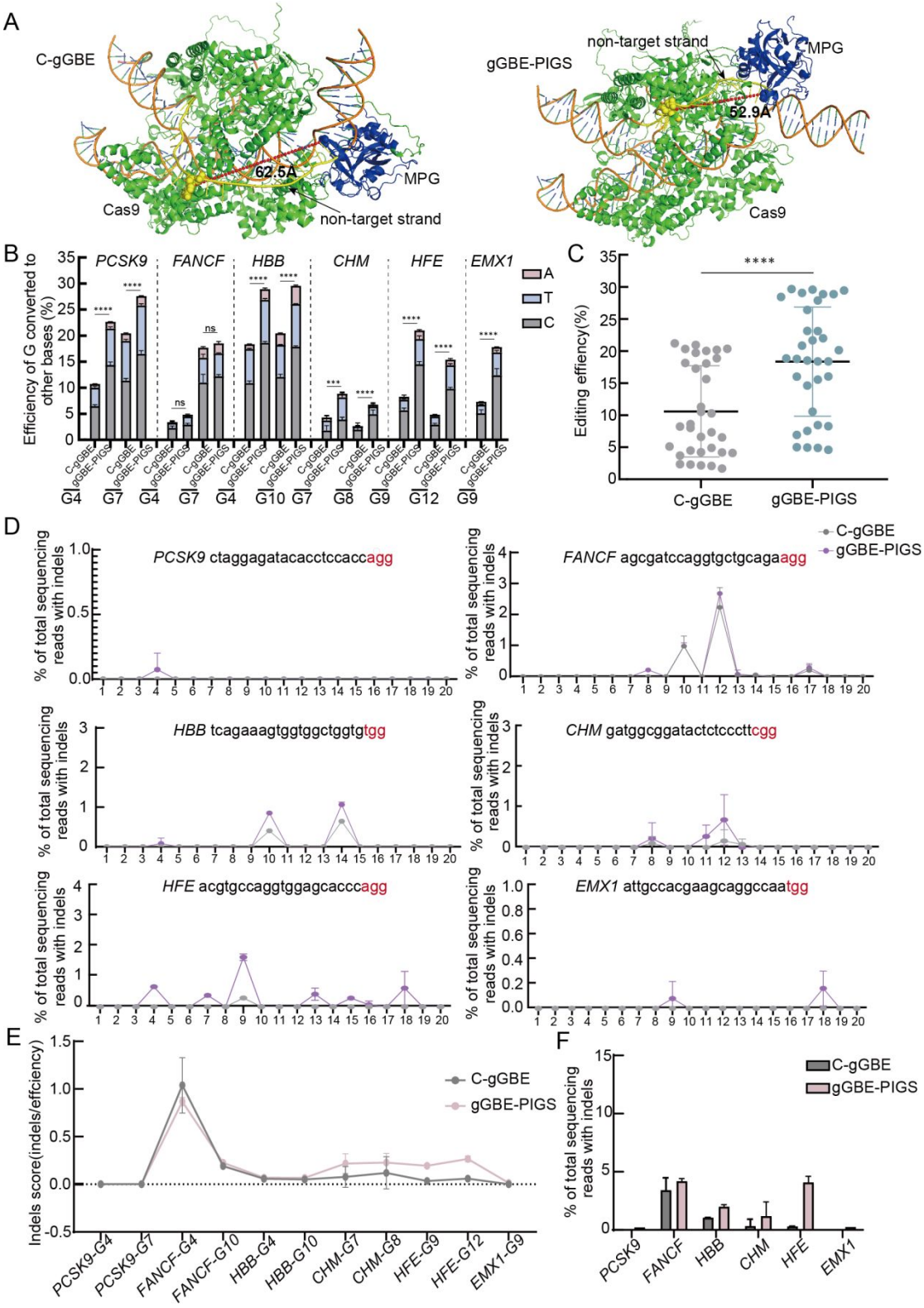


Figure2 Inlaying MPG into internal Cas9 improved the editing efficiency

A: Cartoon predicted structure of C-gGBE (left) and gGBE-PIGS(right) with DNA and sgRNA complex by AlphaFold3 at *HFE*. The NTS is traced in yellow. Tyr162 of MPG

is marked as blue spheres, and the editable guanine (G12) on the NTS is depicted as yellow spheres. B: Comparison of the editing effciency of C-gGBE and gGBE-PIGS at 6 endogenous human genomic loci. C: Quantification of base editing efficiencies of each target G. D: Variation of indels at 6 endogenous human genomic loci. E: Calculated indels score of C-gGBE and gGBE-PIGS across all loci tested in Figure 2B. “Indels score” was calculated as the number of amplicon-seq reads with indels per the number of amplicon-seq reads with pricisely base edit. F: Comparison of indel frequencies generated by C-gGBE and gGBE-PIGS at six endogenous human genomic loci. All data were generated from three independent experiments and represented as mean±SD. ns: Not significant; *: $P<0.05$; **: $P<0.01$; ***: $P<0.001$; ****: $P<0.0001$.

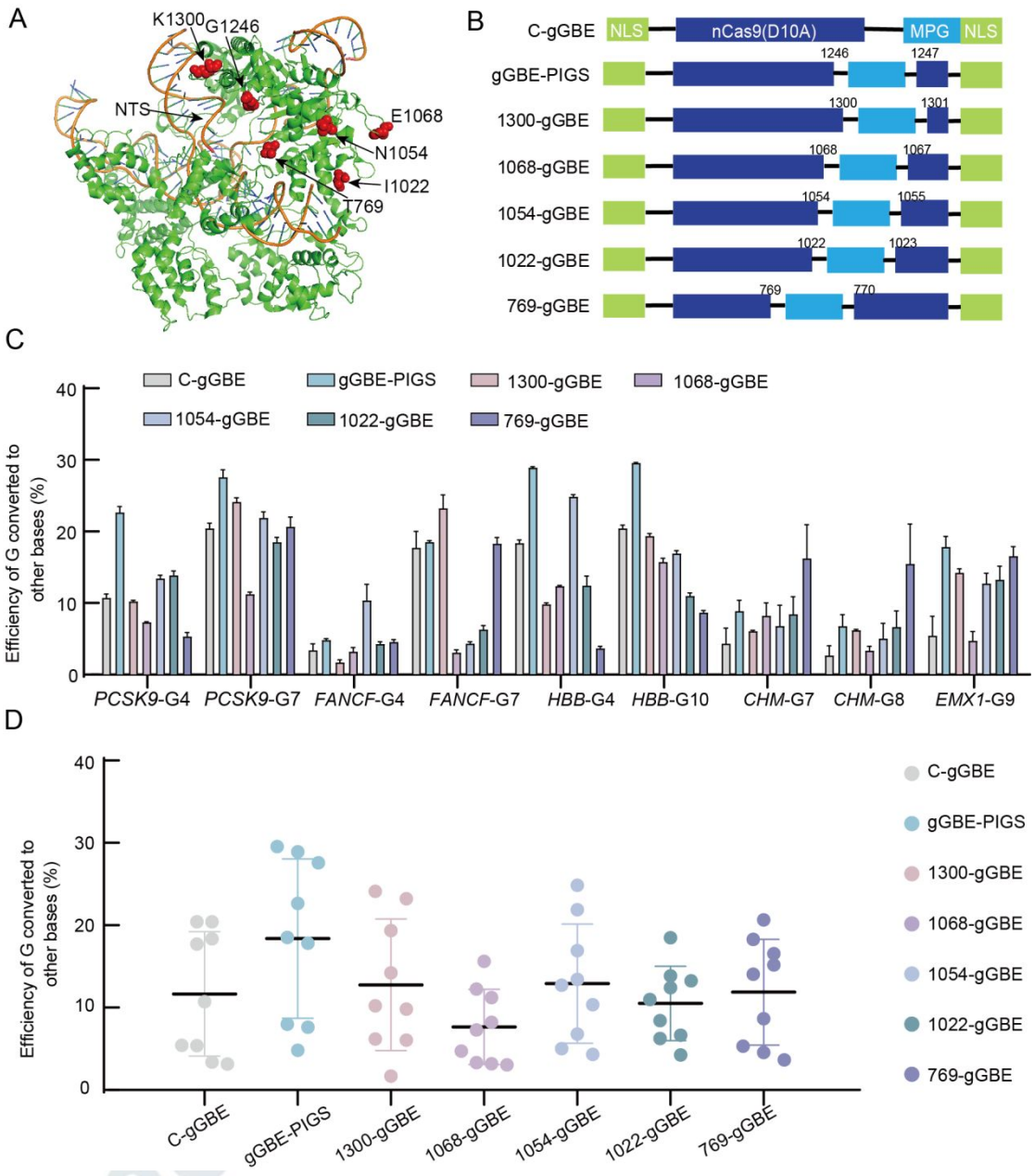


Figure3 Different internal inlaid gGBEs improve the editing efficiency

A: Cartoon representation of the crystal structure of the SpCas9–DNA–sgRNA complex (PDB: 8G1I). Docking sites, including G1246, I1022, N1054, E1068, K1300, and T769, are highlighted in red and shown as spheres. B: Schematic diagram of internally inlaid gGBEs. SpCas9 is shown in dark blue, and the MPG domain is shown in light blue. C: Comparison of editing efficiencies between C-gGBE and internally inlaid gGBEs at five endogenous human genomic loci. D: Quantification of base editing

efficiencies at each target G. Data represent the mean±SD from three independent experiments.

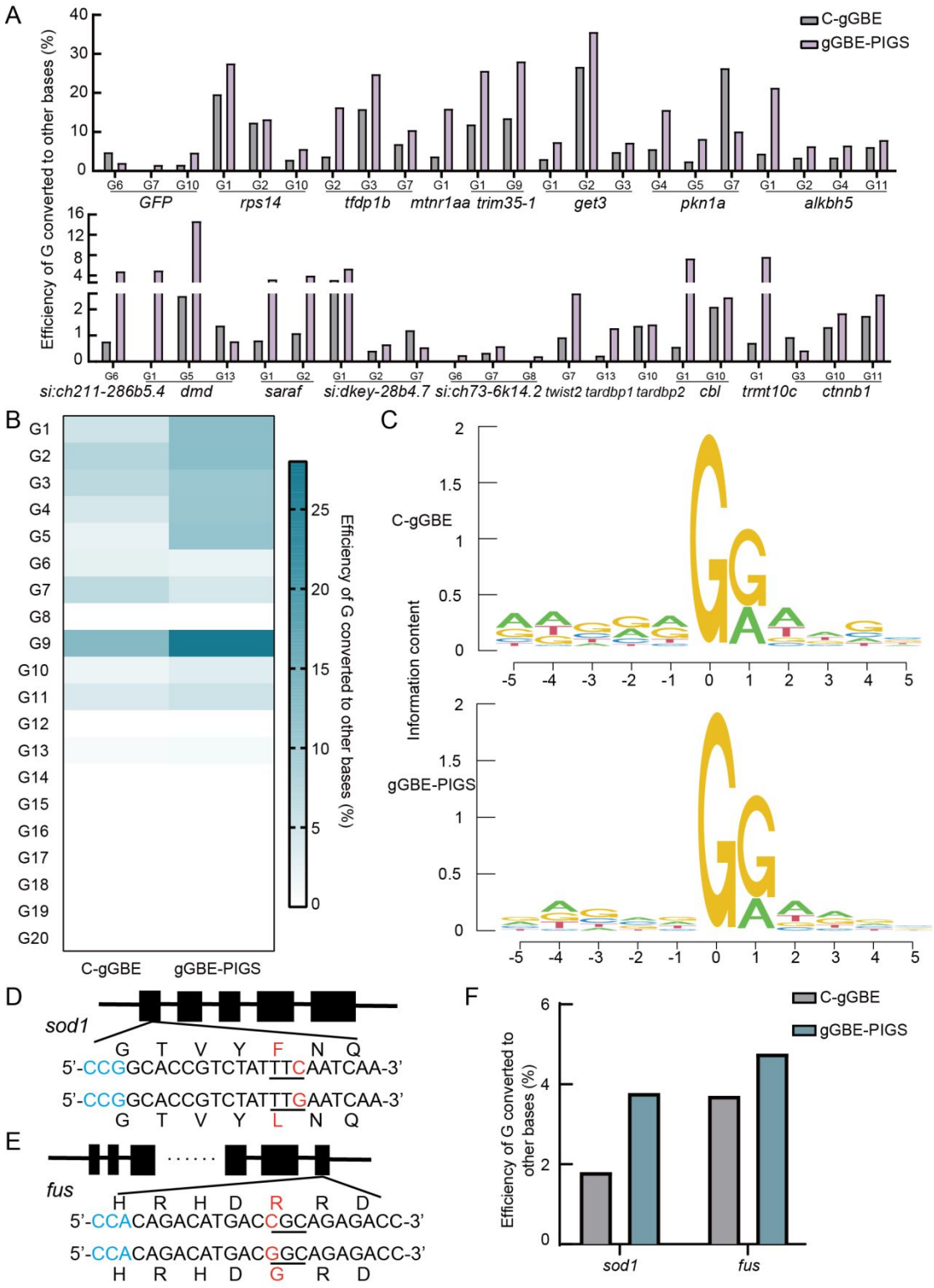


Figure 4 Evaluation of gGBE-PIGS activity in zebrafish embryos

A: Comparison of the editing efficiency produced by C-gGBE and gGBE-PIGS at 19 zebrafish genomic loci. B: Comparison of the editing windows of C-gGBE and gGBE-PIGS. The positional distribution of edited G across the protospacer was quantified to

characterize and compare the editing windows of C-gGBE and gGBE-PIGS. C: Sequence preferences of C-gGBE and gGBE-PIGS. Editing outcomes within the ±5 bp region flanking each target site were quantified by high-throughput sequencing. D: Schematic of the target site in the zebrafish *sod1* gene. The C-to-G conversion on the coding strand, derived from a G-to-C edit on the template strand, is underlined in black; this nucleotide change results in an F21L amino acid substitution. The sgRNA sequence listed in Table 1 targets the template strand. E: Schematic of the target site in the zebrafish fus gene. The C-to-G conversion on the coding strand, derived from a G-to-C edit on the template strand, is underlined in black; this nucleotide change results in an R536G amino acid substitution. The sgRNA sequence listed in Table 1 targets the template strand. F: Comparison of base editing efficiencies between C-gGBE and gGBE-PIGS at a pathogenic site associated with ALS.

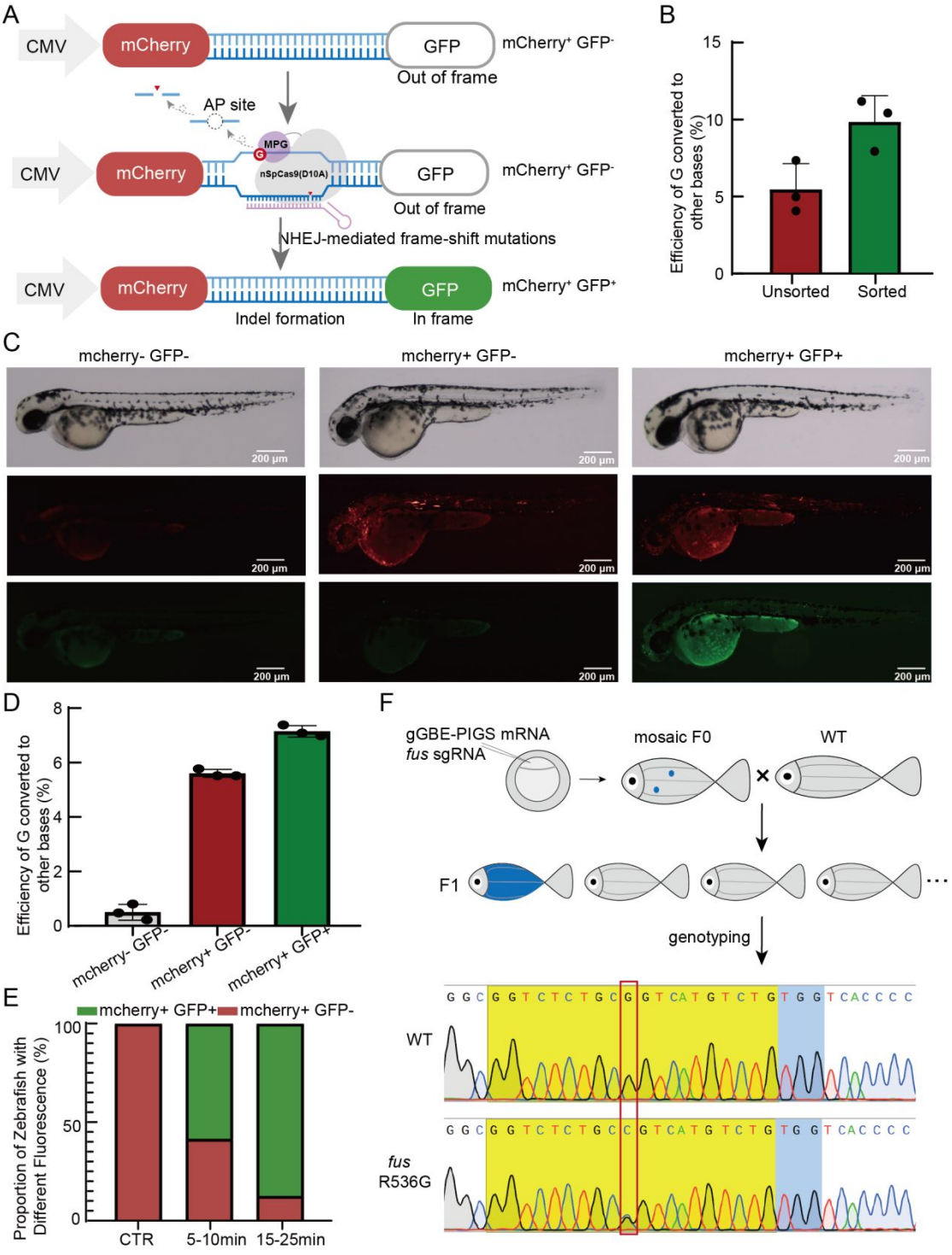


Figure 5 A GFP-based reporter strategy enriches gGBE-edited cells and embryos

A: Schematic Design of Reporting System Based on GFP. B: Comparison of editing efficiencies of the reporter targeting siteA in HEK293T cells after gGBE-PIGS editing, with or without fluorescence-activated cell sorting (FACS) enrichment. C: Assessment

of gGBE-PIGS editing activity in zebrafish embryos using the optimized reporter system. Fluorescence imaging revealed three distinct patterns: no fluorescence, red fluorescence only, and both red and green fluorescence. Green: GFP; Red: mCherry. Scale bar = 200 μm. D: Editing efficiencies were compared among zebrafish embryos showing no fluorescence, red fluorescence only, or both red and green fluorescence. E: Variation in the proportion of fluorescence segregation in zebrafish embryos injected at different time points post-fertilization(n=110). F: Schematic workflow for constructing *fus* R536G mutant zebrafish: Fertilized zebrafish embryos were co-injected with gGBE-PIGS mRNA and *fus*-targeting sgRNA, yielding mosaic F0 fish (with blue dots indicating edited cells). F0 founders were outcrossed with wild-type (WT) zebrafish to produce F1 progeny. Genotyping of F1 individuals via Sanger sequencing confirmed the introduction of the desired G-to-C base substitution (red box) at the target locus, resulting in the R536G mutation.

Supplementary Information for

Optimized gGBE enables robust G-to-Y conversion in mammalian cells and zebrafish embryos

Min Li[1,#], Yaoge Jiao[1,#], Rui Tao[2], Yun Hu[1], Junyi Fei[1], Chunting Wang[1], Shaohua Yao[1,3,*]

[1]Laboratory of Biotherapy, National Key Laboratory of Biotherapy, Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

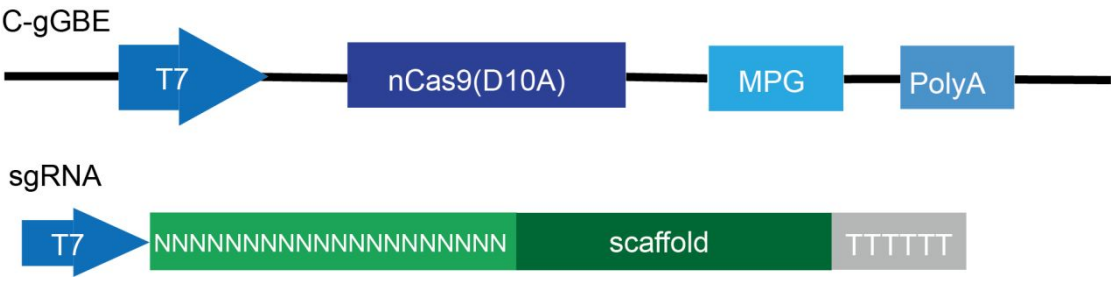
[2]Institute of Kidney Diseases, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

[3]Tianfu Jincheng Laboratory, Chengdu 610095, Sichuan, China

[#] These authors contributed equally

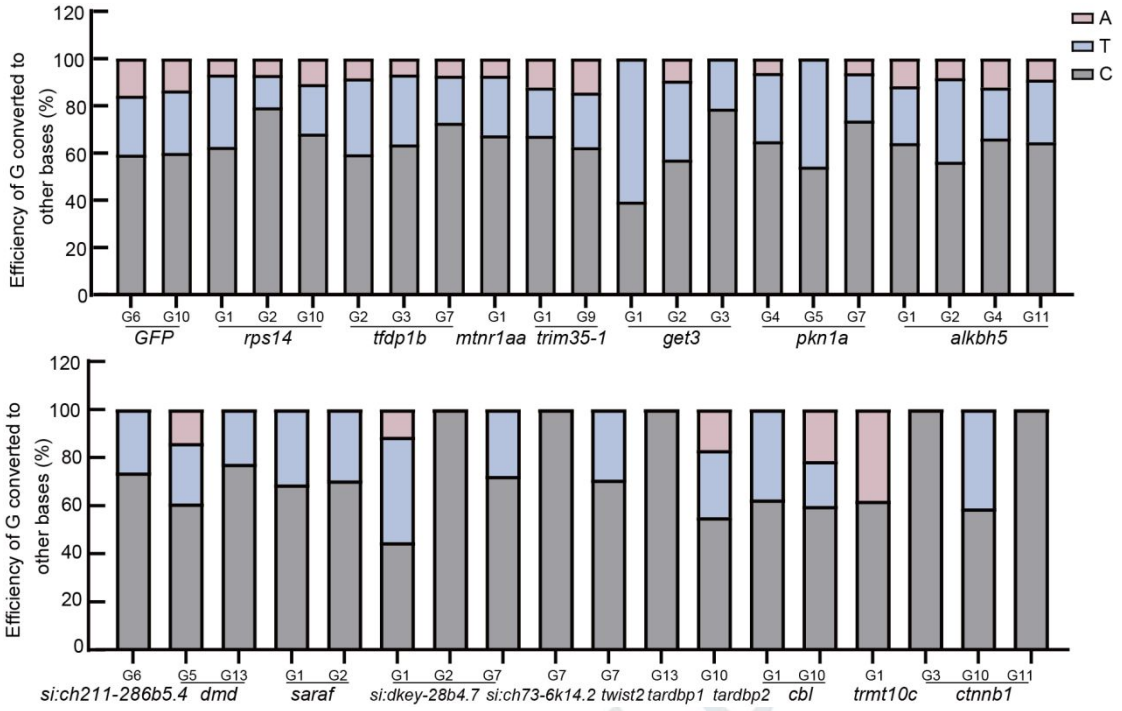
[*] To whom correspondence should be addressed. Email: shaohuayao@scu.edu.cn

Contents



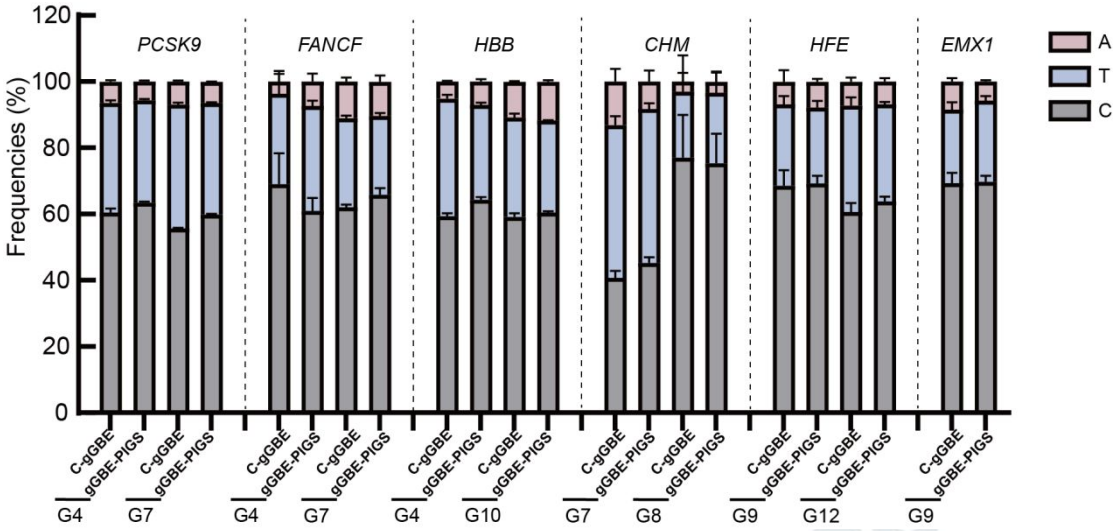
Supplementary Figure S1 Schematic diagram of transcription template

"20 N" represents the spacer sequences of sgRNAs targeting distinct genes in zebrafish.



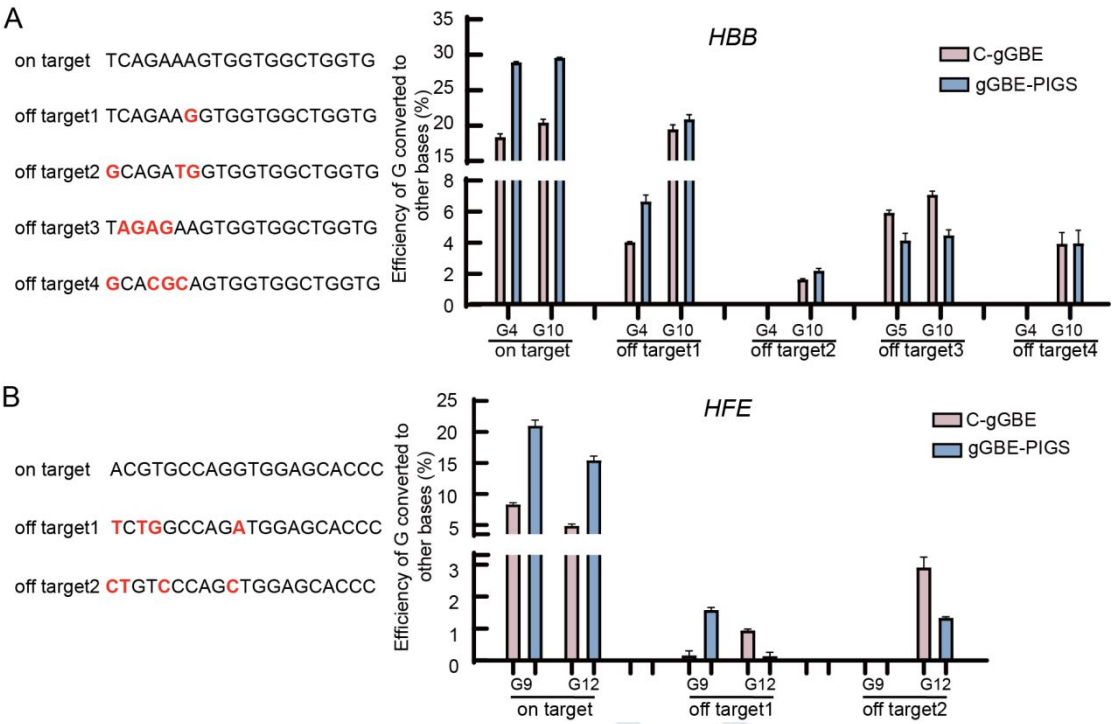
Supplementary Figure S2 Base editing purity mediated by C-gGBE in zebrafish

Bar plots showing G-to-base conversion frequencies at each position in zebrafish, as determined by high-throughput sequencing (HTS).



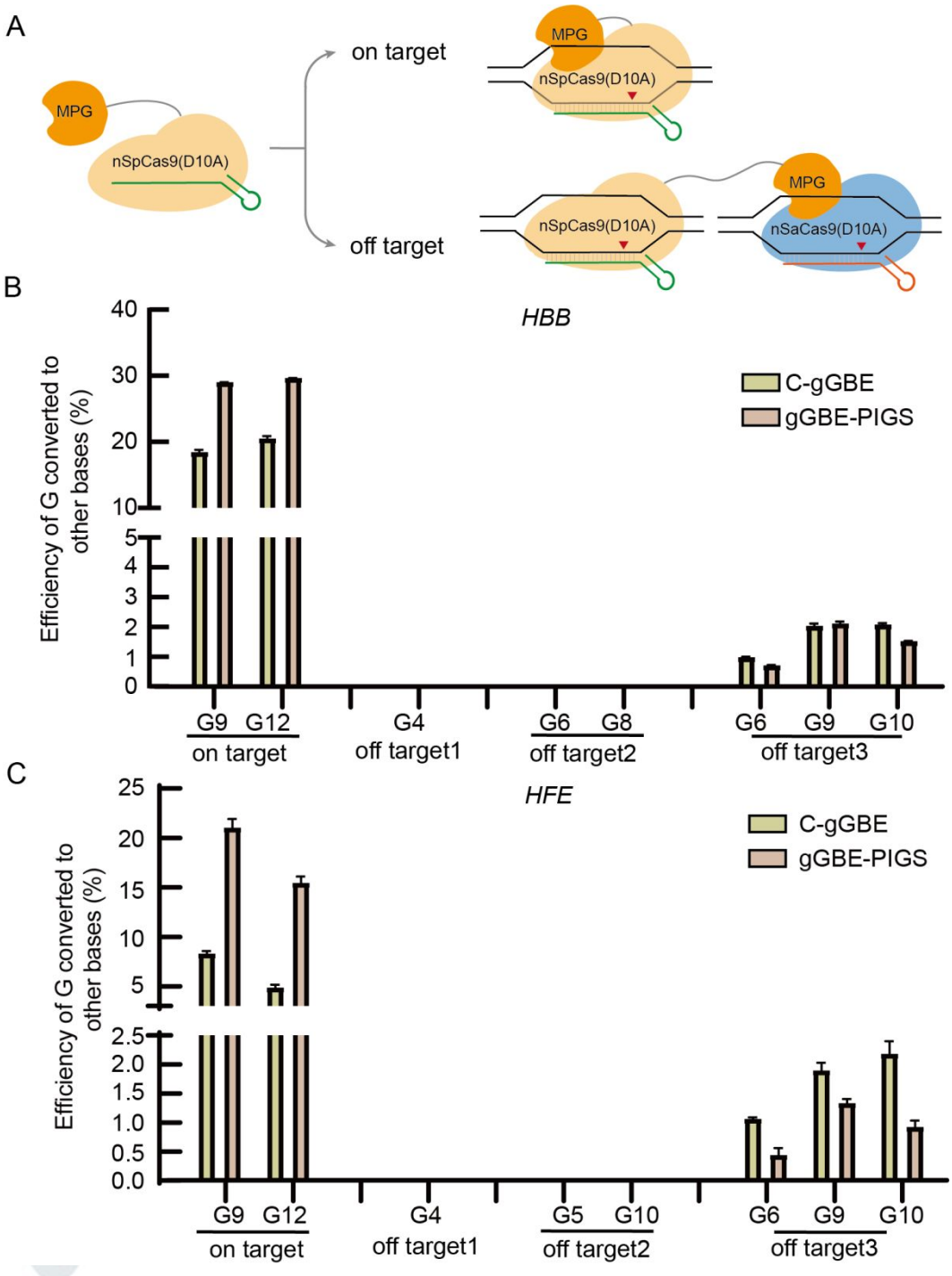
Supplementary Figure S3 Comparison of base editing purity between C-gGBE and gGBE-PIGS

Bar plots showing G-to-base conversion frequencies at each position within the indicated genomic sites in HEK293T cells, as determined by HTS. All data are from three independent experiments and are presented as mean±SD.



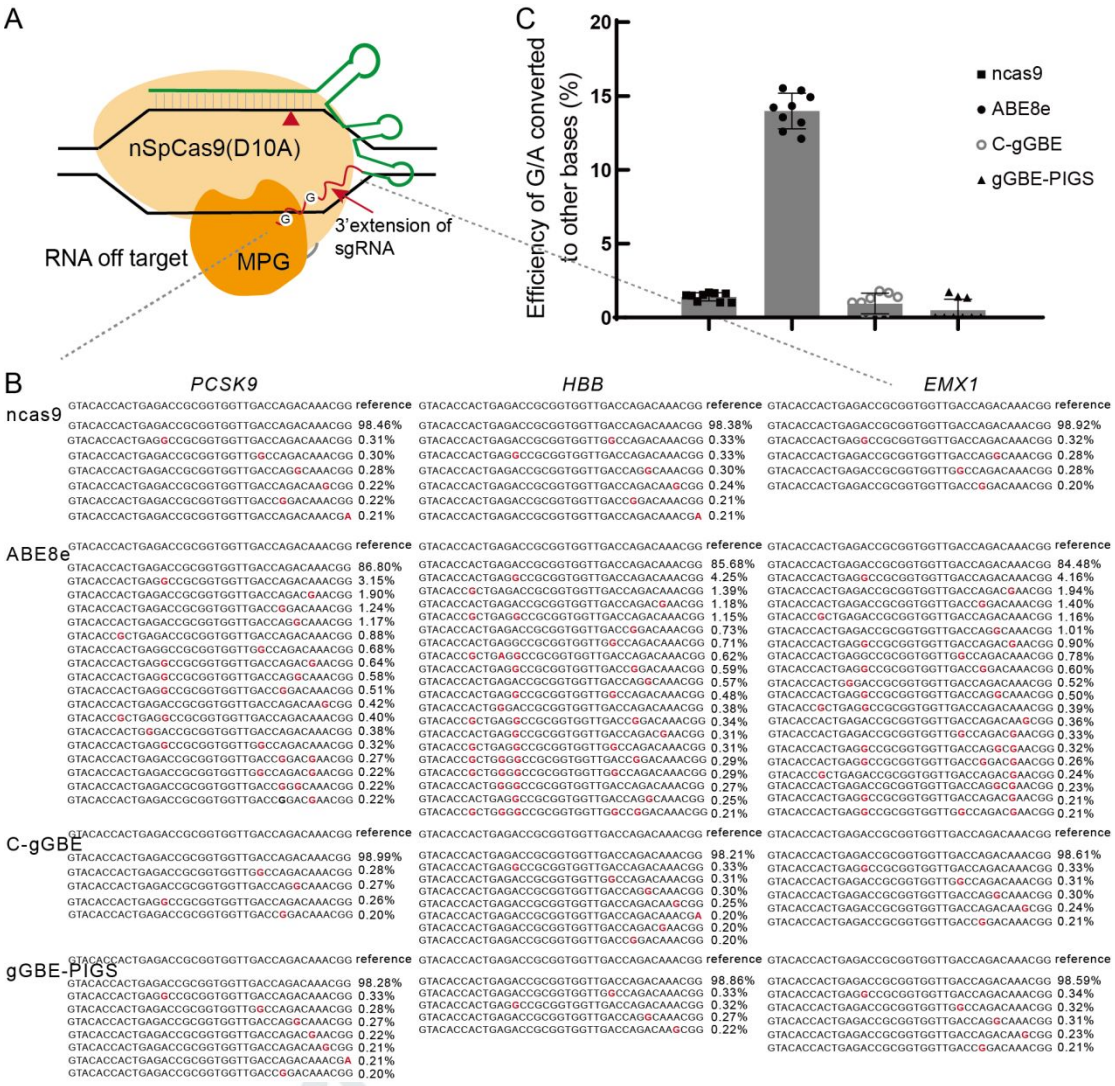
Supplementary Figure S4 Cas9 dependent DNA off-target editing of gGBE-PIGS

Analysis Cas9 dependent DNA off-target editing by gGBE-PIGS at *HBB* (**A**) and *HFE* (**B**). Left panels showed the sequences of on-target sites and predicted off-targets. Mismatched bases were shown in red. On-target and off-target efficiencies were determined by HTS. All data were generated from three independent experiments and represented as mean±SD.



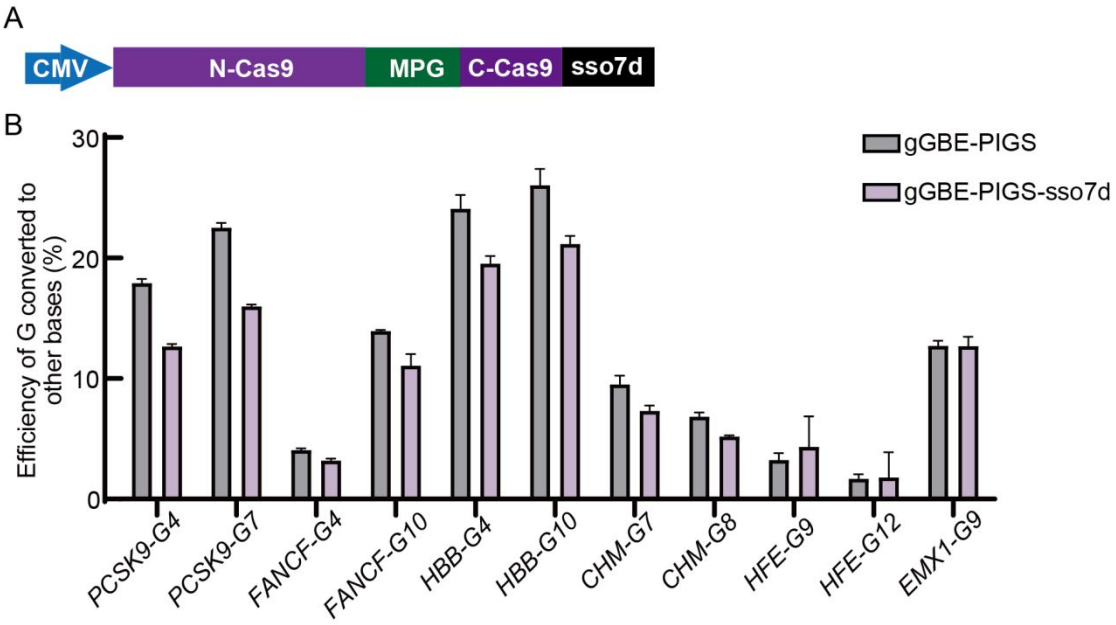
Supplementary Figure S5 Cas9 independent DNA off-target editing of gGBE-PIGS

A: Schematic diagram showing the mechanism of artificial R-loop assays. Off-target R-Loop consisted of nickase SaCas9 and corresponding sgRNAs, while on-target R-loops involve SpCas9-derived gGBEs and their respective sgRNAs. B-C: Comparison of DNA off-target editing by C-gGBE and gGBE-PIGS at two endogenous loci. Plasmids encoding paired R-loops were co-transfected into HEK293T cells, and on-target and off-target editing efficiencies were quantified by HTS. Data represent mean±SD from three independent experiments.



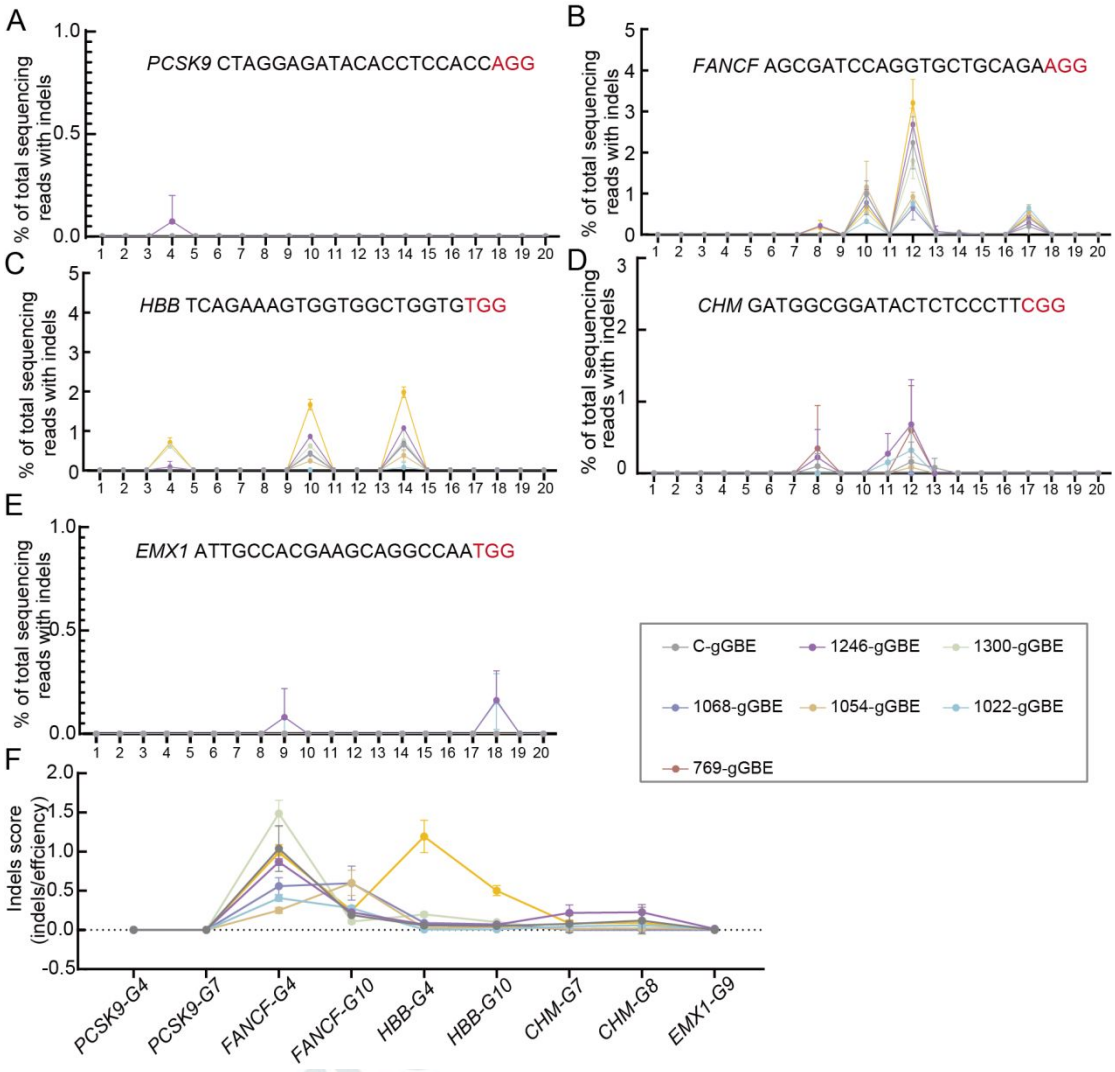
Supplementary Figure S6. Evaluation of RNA off-target editing of C-gGBE and gGBE-PIGS

A: Design of the gGBE RNA off-target evaluation using an sgRNA engineering strategy. A G-containing extension, putatively accessible to MPG, was appended to the 3' end of the sgRNA to assess potential RNA off-target editing. B: HTS analysis of representative RNA off-target editing at sgRNAs containing identical 3' extended tails in cells expressing nCas9, ABE8e, C-gGBE, or gGBE-PIGS. ABE8e induced robust RNA off-target editing specifically within the extended tails (reference sequence shown on the top). Edited nucleotides are highlighted in red. C: Quantification of RNA off-target editing efficiencies derived from the HTS data in (B); RNAs containing any type of nucleotide conversion were counted as off-target editing events. Data are shown as mean±SD from three independent experiments.



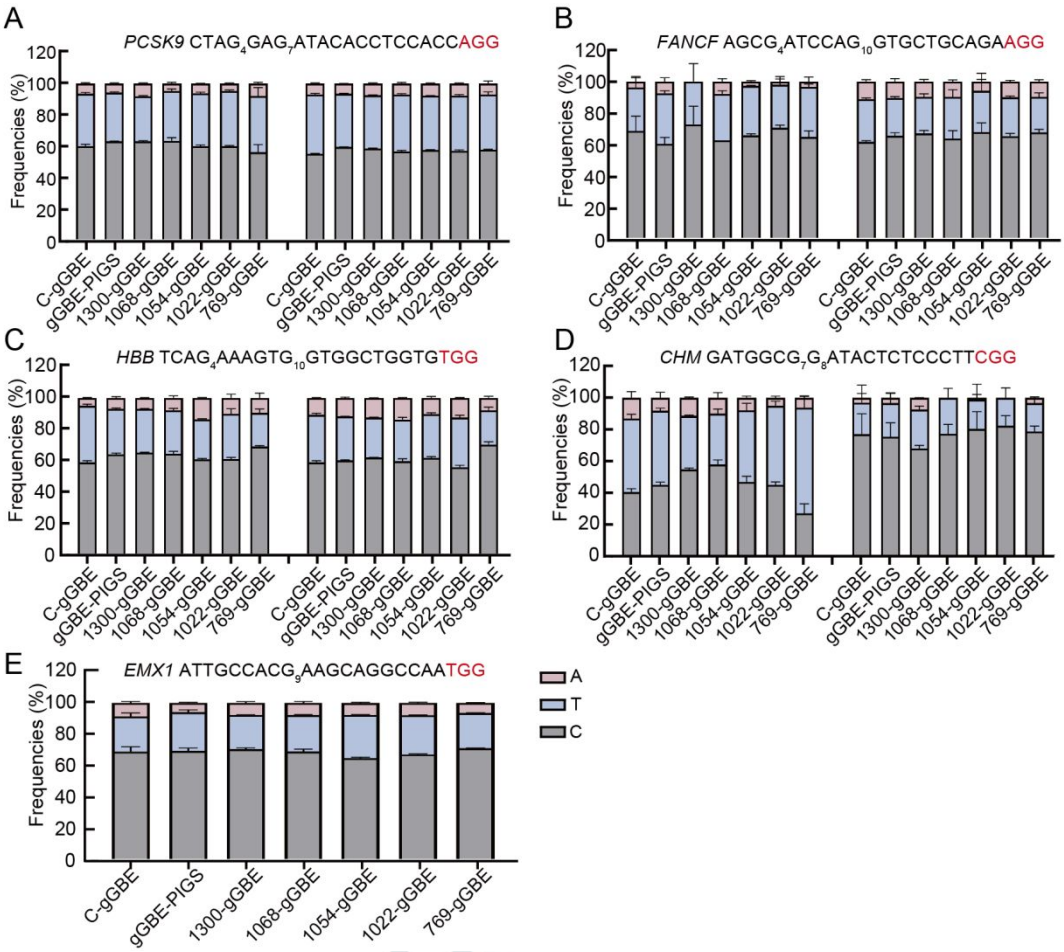
Supplementary Figure S7 Comparison of base editing efficiencies between gGBE-PIGS and gGBE-PIGS-sso7d

A: Schematic diagram of gGBE-PIGS-sso7d. B: Comparison of the editing efficiency produced by gGBE-PIGS and gGBE-PIGS-sso7d at six endogenous loci. Efficiencies were determined by HTS. All Data were generated from three independent experiments and represented as mean±SD.



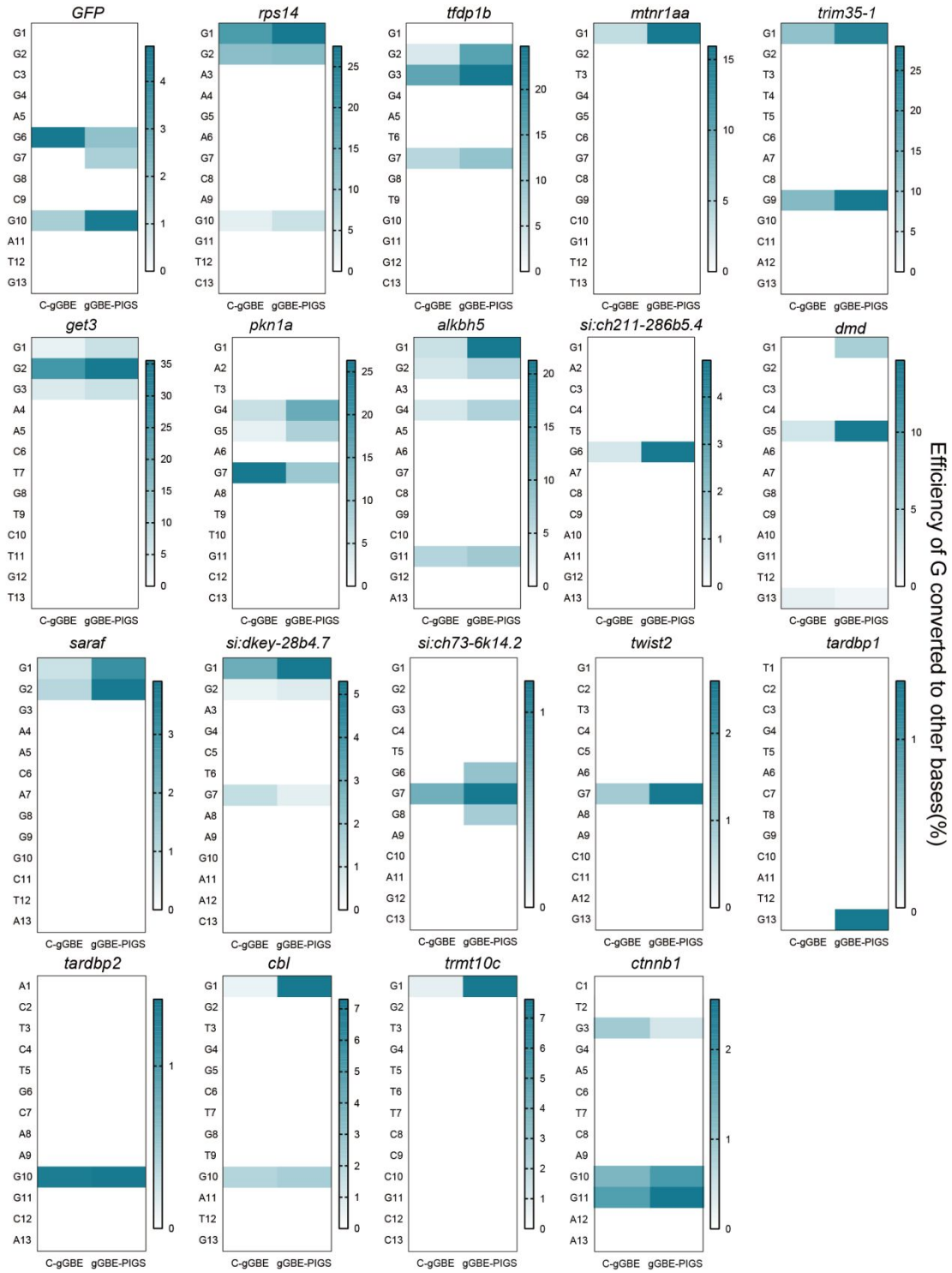
Supplementary Figure S8. Comparison of indels frequencies between C-gGBE and internally inlaid gGBEs

A-E: Comparison of indels frequency produced by C-gGBE and internally inlaid gGBEs at five endogenous genomic loci. F: Calculated indels score of C-gGBE and internally inlaid gGBEs across all loci tested in Supplementary Figure6A. "Indels score" was calculated as the number of amplicon-seq reads with indels per the number of amplicon-seq reads with pricise base edit. All data were generated from three independent experiments and represented as mean±SD.



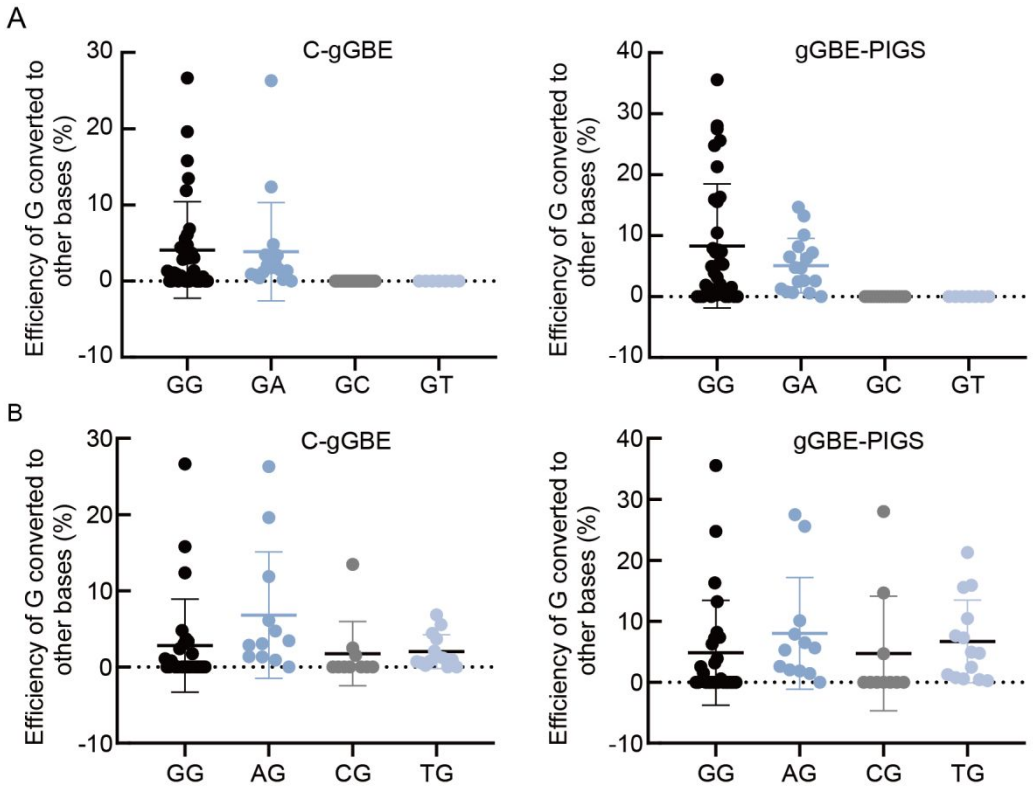
Supplementary Figure S9 Comparison of base editing purity between C-gGBE and internally inlaid gGBEs

A-E: Bar plots showing a comparative analysis of G base editing efficiencies mediated by C-gGBE and its multiple internally inserted variants across various endogenous genomic loci , ac determined by HTS. All data represent mean±SD from three independent experiments.



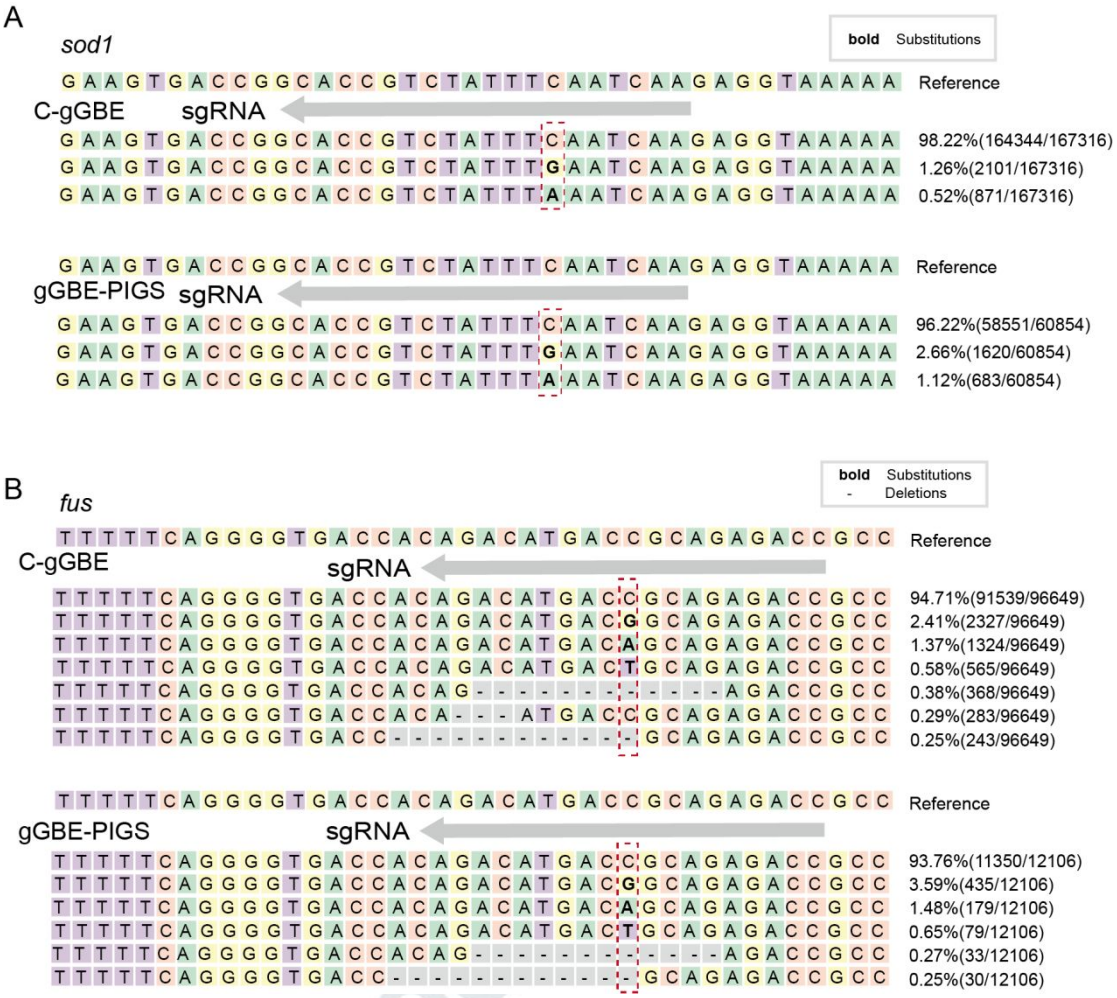
Supplementary Figure S10 Comparison of base editing efficiencies between C-gGBE and gGBE-PIGS across 19 loci in zebrafish embryos

Heatmaps showing the G-to-other (G to C/T/A) conversion efficiencies generated by C-gGBE and gGBE-PIGS at 19 loci in zebrafish embryos. Each panel represents an individual target site. The color scale reflects the percentage of G conversions quantified by HTS.



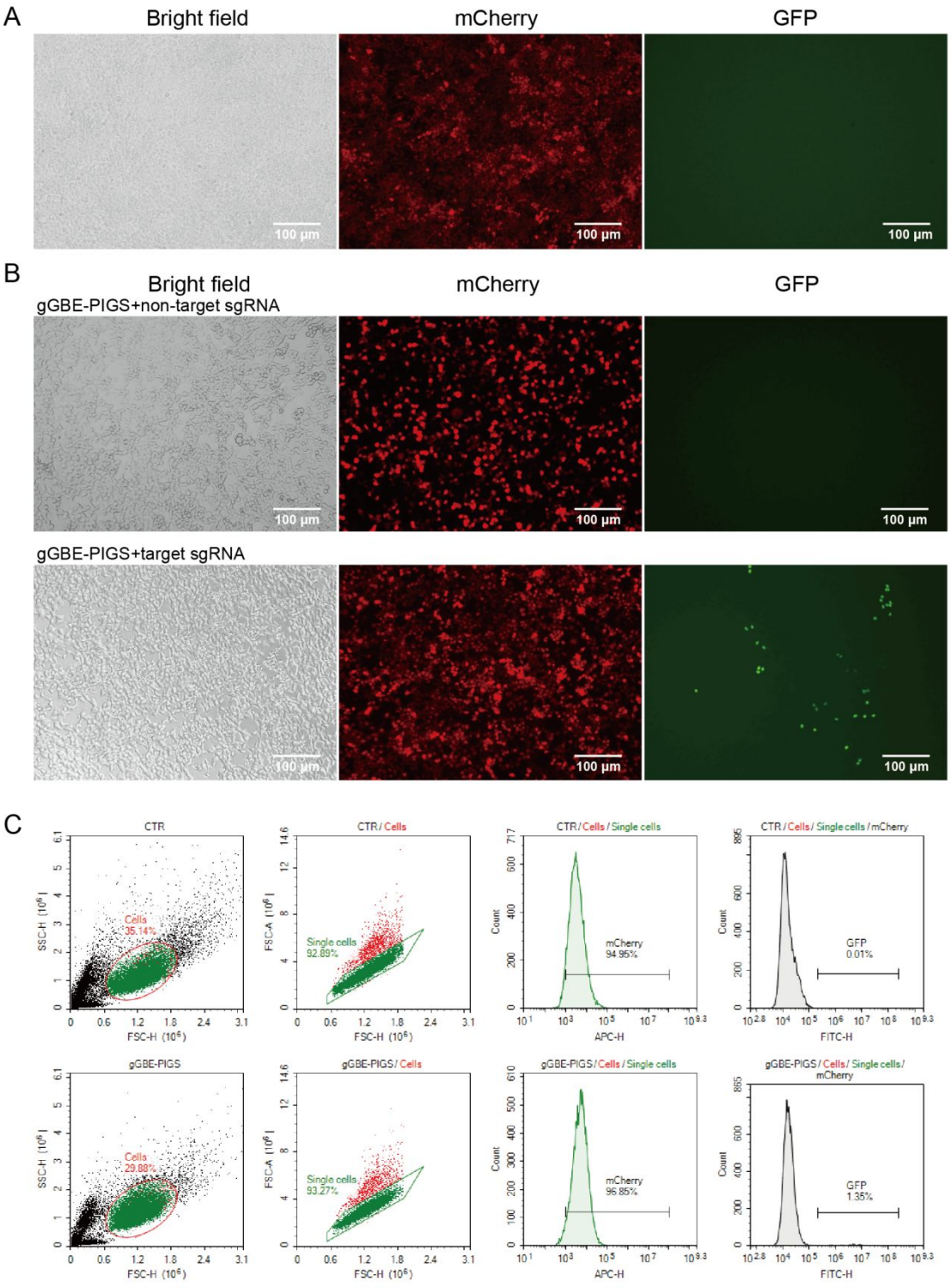
Supplementary Figure S11 Comparison of G-to-other base conversions induced by C-gGBE and gGBE-PIGS across different nucleotide contexts

A: Scatter plots summarizing the efficiencies of G-to-other base conversions observed at endogenous loci within GN sequence contexts. B: Scatter plots summarizing the efficiencies of G-to-other base conversions observed at endogenous loci within NG sequence contexts. Each dot represents an individual target G, and horizontal bars indicate mean±SD.



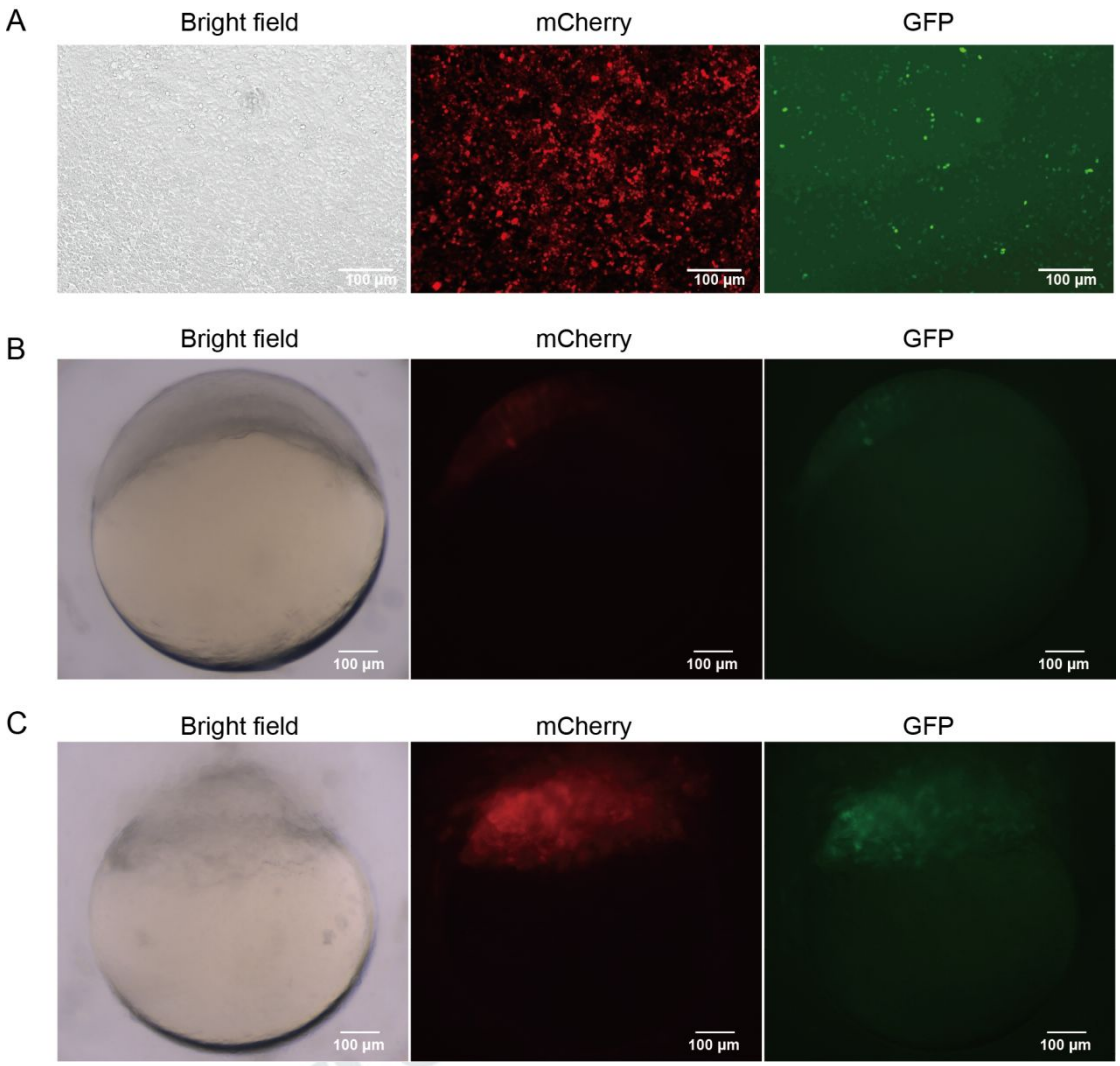
Supplementary Figure S12 gGBE-PIGS induces G base conversion at *sod1* and *fus* loci in zebrafish

A: Allele frequencies at the *fus* genomic locus following treatment with C-gGBE and gGBE-PIGS. The edited G nucleotide is highlighted within the red dashed box. B: Allele frequencies at the *sod1* genomic locus following treatment with C-gGBE and gGBE-PIGS. The edited G nucleotide is highlighted within the red dashed box.



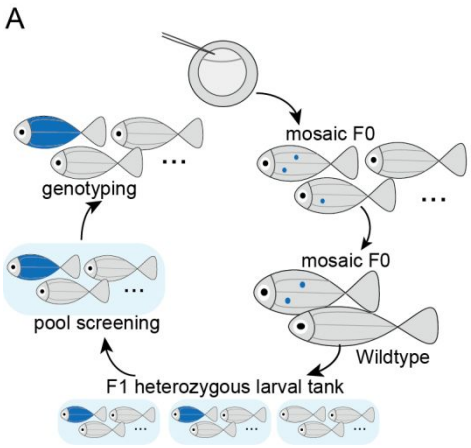
Supplementary Figure S13 Fluorescence analysis and sorting of HEK293T reporter cells transfected with gGBE-PIGS and sgRNAs

A: Fluorescence microscopy images of HEK293T reporter cells. B: Fluorescence microscopy images of HEK293T reporter cells transfected with gGBE-PIGS and either a targeting or non-targeting sgRNA. Scale bar = 100 μm. C: Flow cytometry sorting of the cells shown in (B), isolating GFP-positive cells for downstream editing frequency analysis.



Supplementary Figure S14 gGBE reporter containing a zebrafish *fus* target sequence in HEK293T cells and zebrafish embryos

A: Fluorescence of the CMV promoter-driven reporter containing the *fus* target sequence between mCherry and GFP in HEK293T cells. B: Fluorescence of the CMV promoter-driven reporter containing the *fus* target sequence between mCherry and GFP in zebrafish embryos. C: Fluorescence of the β-actin promoter-driven reporter containing the *fus* target sequence between mCherry and GFP in zebrafish embryos. Scale bar = 100 μm.



B

Sample ID	F0 mosaic editing rate (%)
F0-1	11.35
F0-2	4.59
F0-3	4.22
F0-4	8.53
F0-5	0.44
F0-6	1.96
F0-7	2.56

crossed with wildtype fish

Sample ID	Editing rates of pooled F1 larvae(%)
F1-A	2.71
F1-B	1.51
F1-C	0.00
F1-D	2.36

Genotype	Wildtype	R536G heterozygotes	R536C heterozygotes
Count	76/80	3/80	1/80

Supplementary Figure S15 Generation and identification of heritable *fus* mutations in zebrafish

A: Editing efficiencies of individual mosaic F0 fishes (top left) and of the pooled F1 offspring derived from the highest-mosaic F0 crossed with a wild-type fish (top right). B: Individual genotyping of F1 larvae (bottom) identified four positive larvae, including three carrying the R536G edit and one carrying the R536C edit.

Supplementary Table S1 List of the targets tested in this study.

sgRNA	Target sequence	Oligo-F	Oligo-R	Reference
PCSK9	CTAGGAGATACACCTCCACC**AGG**	CACCGCTAGGAGATACACCTCCACC	AAACGGTGGAGGTGTATCTCCTAGC	(Tong et al., 2023)
FANCF	AGCGATCCAGGTGCTGCAGA**AGG**	CACCGAGCGATCCAGGTGCTGCAGA	AAACTCTGCAGCACCTGGATCGCTC	(Tong et al., 2023)
HBB	TCAGAAAGTGGTGGCTGGTG**TGG**	CACCGTCAGAAAGTGGTGGCTGGTG	AAACCACCAGCCACCACTTTCTGAC	(Tong et al., 2023)
CHM	GATGGCGGATACTCTCCCTT**CGG**	CACCGATGGCGGATACTCTCCCTT	AAACAAGGGAGAGTATCCGCCATC	(Tong et al., 2023)
HFE	ACGTGCCAGGTGGAGCACCC**AGG**	CACCGACGTGCCAGGTGGAGCACCC	AAACGGGTGCTCCACCTGGCACGTC	(Tong et al., 2023)
EMX1	ATTGCCACGAAGCAGGCCAA**TGG**	CACCGATTGCCACGAAGCAGGCCAA	AAACTTGGCCTGCTTCGTGGCAATC	This study
GFP	GGCGAGGGCGATGCCACCTA**CGG**	CACCGGCGAGGGCGATGCCACCTA	AAACTAGGTGGCATCGCCCTCGCC	This study
ctnnb1	CTGGACTCAGGAATCCACTC**TGG**	CACCGCTGGACTCAGGAATCCACTC	AAACGAGTGGATTCCTGAGTCCAGC	(Rosello et al., 2021)
rps14	GGAAGAGCAGGTCATCAGCCT**AGG**	CACCGAAGAGCAGGTCATCAGCCT	AAACAGGCTGATGACCTGCTCTTC	(Qin et al., 2018)
twist2	GCTCCAGAACCAGCGCGTCC**TGG**	CACCGCTCCAGAACCAGCGCGTCC	AAACGGACGCGCTGGTTCTGGAGC	(Qin et al., 2018)

Supplementary Table S1 List of the targets tested in this study.

sgRNA	Target sequence	Oligo-F	Oligo-R	Reference
tfdp1b	GGGGATGGTGGGCCTCTGAG**GGG**	CACCGGGGATGGTGGGCCTCTGAG	AAACCTCAGAGGCCCACCATCCCC	(Sun et al., 2019)
trmt10c	GGTGTTTCCCGCCGGTCAGA**CGG**	CACCGGTGTTTCCCGCCGGTCAGA	AAACTCTGACCGGCGGGAAACACC	(Sun et al., 2019)
mtnr1aa	GGTGGCGCGCGTTCCTCCGG**TGG**	CACCGGTGGCGCGCGTTCCTCCGG	AAACCCGGAGGAACGCGCGCCACC	(Sun et al., 2019)
trim35-1	GGTTTCACGGCAGCGCCCGG**TGG**	CACCGGTTTCACGGCAGCGCCCGG	AAACCCGGGCGCTGCCGTGAAACC	(Sun et al., 2019)
get3	GGGAACTGTCTGTCTTCTCA**AGG**	CACCGGGAACTGTCTGTCTTCTCA	AAACTGAGAAGACAGACAGTTCCC	(Sun et al., 2019)
si:ch211-286b5.4	GACCTGACCAAGAGTTTACA**GG**	CACCGACCTGACCAAGAGTTTAC	AAACGTAAACTCTTGGTCAGGTC	(Sun et al., 2019)
dmd	GGCCGAAGCAGTGAACCCAC**AGG**	CACCGGCCGAAGCAGTGAACCCAC	AAACTGGGTTCACTGCTTCGGCC	(Sun et al., 2019)
saraf	GGGAACAGGGCTAGAGCGGC**CGGG**	CACCGGGAACAGGGCTAGAGCGGC	AAACGCCGCTCTAGCCCTGTTCCC	(Sun et al., 2019)
si:dkey-28b4.7	GGAGCTGAAGAACTTTGATC**TGG**	CACCGGAGCTGAAGAACTTTGATC	AAACGATCAAAGTTCTTCAGCTCC	(Sun et al., 2019)
si:ch73-6k14.2	GGGCTGGGACAGCTCCCGAA	CACGGGCTGGGACAGCTCCCGAA	AAACTTCGGGAGCTGTCCCAGC	(Sun et al., 2019)

Supplementary Table S1 List of the targets tested in this study.

sgRNA	Target sequence	Oligo-F	Oligo-R	Reference
	TGG		CC	
pkn1a	GATGGAGATTGCCGTGTAC**TGG**	CACCGATGGAGATTGCCGTGTAC	AAACGTACACGGCAATCTCCATC	(Sun et al., 2019)
alkbh5	GGAGAAGCGCGGACCCGGGC**AGG**	CACCGGAGAAGCGCGGACCCGGGC	AAACGCCCGGGTCCGCGCTTCTCC	(Sun et al., 2019)
cbl	GGTGGCTGTGATGGGGCTTT**GGG**	CACCGGTGGCTGTGATGGGGCTTT	AAACAAAGCCCCATCACAGCCACC	This study
tardbp1	TCCGTACTGCATGAAGAACTGA**CGG**	CACCGTCCGTACTGCATGAAGAACTGA	AAACTCAGTTCTTCATGCAGTACGGAC	This study
tardbp2	ACTCTGCAAGGCAGCCTGGG**CGG**	CACCGACTCTGCAAGGCAGCCTGGG	AAACCCCAGGCTGCCTTGCAGAGTC	This study
fus	GGTCTCTGCGGTCATGTCTG**TGG**	CACCGGTCTCTGCGGTCATGTCTG	AAACCAGACATGACCGCAGAGACC	This study
sod1	TTGATTGAAATAGACGGTGC**CGG**	CACCGTTGATTGAAATAGACGGTGC	AAACGCACCGTCTATTTCAATCAAC	This study

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
1-GFP-C-gGBE	ACTGTTGatcctggtcgagctggac	ggcggacttgaagaagtcgt	Figure1
1-*ctnnb1*-C-gGBE	TAGCTCTagctggaaatggccatgga	agcaacagaatggccaca	Figure1
1-*rps14*-C-gGBE	AGACATGtctga ctgttgactgtgtg	acttgccagacagatcag	Figure1
1-*twist2*-C-gGBE	AGAGCTAagctccgtcaataaacgtaa	atcttgctgagtttatccga	Figure1
1-*tfdp1b*-C-gGBE	TCAACGCtcatctgcaggtggtcagca	gaagaggagagtgtcagtgc	Figure1
1-*trmt10c*-C-gGBE	CGTAACCgctgatgatgtttctgcgct	acagacttccagatgtcca	Figure1
1-*mtnr1aa*-C-gGBE	GCCTTAGcagtgcagagctgttcagaga	tcgcacaacttacccagacc	Figure1
1-*trim35-1*-C-gGBE	GGCCTAAcagtgcagagctgttcagaga	acacggctgttcatcctcat	Figure1
1-*get3*-C-gGBE	GATCTCAgttcgtaaaagccacataga	ttacctggtatggtggcgg	Figure1
1-*si:ch211-286b5.4*-C-gGBE	GACTACAgttcgtaaaagccacataga	accctatgaaagcacagact	Figure1
1-*dmd*-C-gGBE	ATTGCTCacctgatcgagttttgatga	attcactgctggatctgagg	Figure1
1-*saraf*-C-gGBE	AGAGTGCacctgatcgagttttgatga	cgtcgatgccatctgatcct	Figure1
1-*si:dkey-28b4.7*-C-gGBE	CATACGCaacagacgtttacctgtgca	tccatatcactatggtccta	Figure1
1-*si:ch73-6k14.2*-C-gGBE	CCTTATGaacagacgtttacctgtgca	ttccagcatgaagcccagga	Figure1
1-*pkn1a*-C-gGBE	ATAGGCGcgtagtactagaataaagctt	ccctgtggttctagctctag	Figure1
1-*alkbh5*-C-gGBE	GAACCTAcgtagtactagaataaagctt	aaaccctccgggataacacc	Figure1
1-*cbl*-C-gGBE	ATCGAAGttggagaggcagagagtg	agacgtcctctccccagta	Figure1
1-*tardbp1*-C-gGBE	ACGTCTCtggtcatctgcaaaggtg	agcaaggtatagatgaacca	Figure1
1-*tardbp2*-C-gGBE	AGAAGTCagccgagctatatgtttgg	tagctccagcagcttgggaa	Figure1
1-GFP-CTR	ACTCAAGatcctggtcgagctggac	ggcggacttgaagaagtcgt	Figure1
1-*ctnnb1*-CTR	TACAGAGagctggaaatggccatgga	agcaacagaatggccaca	Figure1
1-*rps14*-CTR	AGTCTACtctgactgttgactgtgtg	acttgccagacagatcag	Figure1
1-*twist2*-CTR	AGCAACTagctccgtcaataaacgtaa	atcttgctgagtttatccga	Figure1
1-*tfdp1b*-CTR	TTGCAGCtcatctgcaggtggtcagca	gaagaggagagtgtcagtgc	Figure1

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
1-*trmt10c*-CTR	CCGGAATgctgatgatgtttctgcgct	acagacttccagatgtcca	Figure1
1-*mtnr1aa*-CTR	GCATAACcagtgcagagctgttcagaga	tcgcacaacttacccagacc	Figure1
1-*trim35-1*-CTR	AATGGCCgtttctgccgtcagtgcatc	acacggctgttcatcctcat	Figure1
1-*get3*-CTR	CTCTGTAgttcgtaaaagccacataga	ttacctggtatggtggcgg	Figure1
1-*si:ch211-286b5.4*-CTR	GTACAGAccagatatgttaatggaagg	accctatgaaagcacagact	Figure1
1-*dmd*-CTR	AGAGCTAacctgatcgagttttgatga	attcactgctggatctgagg	Figure1
1-*saraf*-CTR	TAGAACGcgtgttgcttcgtgatgtcc	cgtcgatgccatctgatcct	Figure1
1-*si:dkey-28b4.7*-CTR	TGCATACaacagacgtttacctgtgca	tccatatcactatggtccta	Figure1
1-*si:ch73-6k14.2*-CTR	GCATAACgctcatgctggggaatctgg	ttccagcatgaagcccagga	Figure1
1-*pkn1a*-CTR	AATGGCCcgtagtactagaataaagctt	ccctgtggttctagctctag	Figure1
1-*alkbh5*-CTR	GATTCTGgcccctgcgaaacaagtac	aaaccctccgggataacacc	Figure1
1-*cbl*-CTR	AGAGTGCttggagaggcagagagtg	agacgtcctctccccagta	Figure1
1-*tardbp1*-CTR	ACGTCTCtggtcatctgcaaaggtg	agcaaggtatagatgaacca	Figure1
1-*tardbp2*-CTR	AGAAGTCagccgagctatatgtttgg	tagctccagcagcttgggaa	Figure1
PCSK9-C-gGBE-rep1	CTCGTATtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
PCSK9-C-gGBE-rep2	CTGCTAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
PCSK9-C-gGBE-rep3	CTGGATGtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
PCSK9-gGBE-PIGS-rep1	CCAATGAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
PCSK9-gGBE-PIGS-rep2	CCACGATtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
PCSK9-gGBE-PIGS-rep3	CCGTGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure2
FANCF-C-gGBE-rep1	AGACATGtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2
FANCF-C-gGBE-rep2	CATCGACtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2
FANCF-C-gGBE-rep3	CATGATCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2
FANCF-gGBE-PIGS-rep1	CATGCAAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
FANCF-gGBE-PIGS-rep2	CAGTTACtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2
FANCF-gGBE-PIGS-rep3	CTATGGTtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure2
HBB-C-gGBE-rep1	TGTACTCcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
HBB-C-gGBE-rep2	TGTTGACcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
HBB-C-gGBE-rep3	TGCAGTTcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
HBB-gGBE-PIGS-rep1	CAATGTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
HBB-gGBE-PIGS-rep2	CAAGTAGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
HBB-gGBE-PIGS-rep3	CATCTTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure2
CHM-C-gGBE-rep1	CATCGACacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
CHM-C-gGBE-rep2	CATGATCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
CHM-C-gGBE-rep3	CGTTCGAacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
CHM-gGBE-PIGS-rep1	ATCGAAGacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
CHM-gGBE-PIGS-rep2	ATCGTCCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
CHM-gGBE-PIGS-rep3	ATGCTTCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure2
HFE-C-gGBE-rep1	CAGAGTCgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
HFE-C-gGBE-rep2	CAGTACTgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
HFE-C-gGBE-rep3	CAGTCTAgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
HFE-gGBE-PIGS-rep1	CTAGACAgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
HFE-gGBE-PIGS-rep2	CTTGAGTgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
HFE-gGBE-PIGS-rep3	CTCACGAgacgtattgcccaatggggatg	aggcactcctctcaaccc	Figure2
EMX1-C-gGBE-rep1	GAGTTCGcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2
EMX1-C-gGBE-rep2	GAGCGATcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2
EMX1-C-gGBE-rep3	GTATGCAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2
EMX1-gGBE-PIGS-rep1	GTACTTGcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
EMX1-gGBE-PIGS-rep2	GTTCAACcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2
EMX1-gGBE-PIGS-rep3	GTTCGTAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure2
PCSK9-1300-gGBE-rep1	CCAATGAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1300-gGBE-rep2	CCACGATtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1300-gGBE-rep3	CCGTGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1068-gGBE-rep1	CTCGTATtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1068-gGBE-rep2	CTGCTAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1068-gGBE-rep3	CTGGATGtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1054-gGBE-rep1	GAAGGTCtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1054-gGBE-rep2	GATTCTGtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1054-gGBE-rep3	GATTGCTtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1022-gGBE-rep1	CGATTCAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1022-gGBE-rep2	CGTAGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-1022-gGBE-rep3	CGTTCGAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-769-gGBE-rep1	CGATTCAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-769-gGBE-rep2	CGTAGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
PCSK9-769-gGBE-rep3	CGTTCGAtcaccagtggctgggatg	ctcgggcacattctcgaag	Figure3
FANCF-1300-gGBE-rep1	TTGATCCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1300-gGBE-rep2	TTGCAGCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1300-gGBE-rep3	AGGACTTtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1068-gGBE-rep1	GTGTCAAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1068-gGBE-rep2	AACTGTCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1068-gGBE-rep3	AACGTGTtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1054-gGBE-rep1	GTCGTGAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
FANCF-1054-gGBE-rep2	GTGAACAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1054-gGBE-rep3	GTGACTCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1022-gGBE-rep1	CTACAAGtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1022-gGBE-rep2	CTACGTCtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-1022-gGBE-rep3	CTAGCACtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-769-gGBE-rep1	GACAGTAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-769-gGBE-rep2	GACTACAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
FANCF-769-gGBE-rep3	TAGTGCAtccaatcagtacgcagagag	aggtagcgcgcccactgca	Figure3
HBB-1300-gGBE-rep1	CGTCCATcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1300-gGBE-rep2	GAACTACcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1300-gGBE-rep3	TGTACTCcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1068-gGBE-rep1	CAAGTAGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1068-gGBE-rep2	CATCTTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1068-gGBE-rep3	CGTCCATcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1054-gGBE-rep1	TGTTGACcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1054-gGBE-rep2	TGCAGTTcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1054-gGBE-rep3	CAATGTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1022-gGBE-rep1	CGTCCATcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1022-gGBE-rep2	GAACTACcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-1022-gGBE-rep3	TGTACTCcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-769-gGBE-rep1	TGTTGACcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-769-gGBE-rep2	TGCAGTTcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
HBB-769-gGBE-rep3	CAATGTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	Figure3
CHM-1300-gGBE-rep1	TGTGTCAacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
CHM-1300-gGBE-rep2	CATGTCTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1300-gGBE-rep3	CACAAGTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1068-gGBE-rep1	ATCGAAGacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1068-gGBE-rep2	ATCGTCCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1068-gGBE-rep3	ATGCTTCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1054-gGBE-rep1	CATCGACacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1054-gGBE-rep2	CATGATCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1054-gGBE-rep3	CGTTCGAacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1022-gGBE-rep1	TGTGTCAacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1022-gGBE-rep2	CATGTCTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-1022-gGBE-rep3	CACAAGTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-769-gGBE-rep1	CATCGACacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-769-gGBE-rep2	CATGATCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	Figure3
CHM-769-gGBE-rep3	CAAGTAGcccagtttagtagttggac	tttgtcccaaaactcgccactg	Figure3
EMX1-1300-gGBE-rep1	GTACTTGcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1300-gGBE-rep2	GTTCAACcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1300-gGBE-rep3	GTTCGTAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1068-gGBE-rep1	GAGTTCGcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1068-gGBE-rep2	GAGCGATcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1068-gGBE-rep3	GTATGCAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1054-gGBE-rep1	AGAGCTAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1054-gGBE-rep2	AGTACAGcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1054-gGBE-rep3	AGTTCCTcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1022-gGBE-rep1	GTTGCCAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
EMX1-1022-gGBE-rep2	GTGGTACcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-1022-gGBE-rep3	GCATTGTcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-769-gGBE-rep1	GTTGCCAcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-769-gGBE-rep2	GTGGTACcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
EMX1-769-gGBE-rep3	GCATTGTcagaagctggaggaggaag	agcagcaagcagcactctg	Figure3
GFP-C-gGBE	ACTGTTGatcctggtcgagctggac	ggcggacttgaagaagtcgt	Figure4
GFP-gGBE-PIGS	AGATCACatcctggtcgagctggac	ggcggacttgaagaagtcgt	Figure4
ctnnb1-C-gGBE	TAGCTCTagctggaaatggccatgga	agcaacagaatggccaca	Figure4
ctnnb1-gGBE-PIGS	TCAAGCTagctggaaatggccatgga	agcaacagaatggccaca	Figure4
rps14-C-gGBE	AGACATGtctgactgttgactgtgtg	acttgccagacagatcag	Figure4
rps14-gGBE-PIGS	AGACATGtctgactgttgactgtgtg	acttgccagacagatcag	Figure4
twist2-C-gGBE	AGAGCTAagctccgtcaataaacgtaa	atcttgctgagtttatccga	Figure4
twist2-gGBE-PIGS	AGAGCTAagctccgtcaataaacgtaa	atcttgctgagtttatccga	Figure4
tfdp1b-C-gGBE	TCAACGCtcatctgcaggtggtcagca	gaagaggagagtgtcagtgc	Figure4
tfdp1b-gGBE-PIGS	TCGCTAGtcatctgcaggtggtcagca	gaagaggagagtgtcagtgc	Figure4
trmt10c-C-gGBE	CGTAACCgctgatgatgtttctgcgct	acagacttccagatgtcca	Figure4
*trmt10*C-gGBE-PIGS	CGCAATAgctgatgatgtttctgcgct	acagacttccagatgtcca	Figure4
mtnr1aa-C-gGBE	GCCTTAGcagtgcagagctgttcagaga	tcgcacaacttacccagacc	Figure4
mtnr1aa-gGBE-PIGS	GGTATCGcagtgcagagctgttcagaga	tcgcacaacttacccagacc	Figure4
trim35-1-C-gGBE	GGCCTAAcagtgcagagctgttcagaga	acacggctgttcatcctcat	Figure4
trim35-1-gGBE-PIGS	AATCCGGcagtgcagagctgttcagaga	acacggctgttcatcctcat	Figure4
get3-C-gGBE	GATCTCAgttcgtaaaagccacataga	ttacctggtatggtggcgg	Figure4
get3-gGBE-PIGS	GACAGTAgttcgtaaaagccacataga	ttacctggtatggtggcgg	Figure4
si:ch211-286b5.4-C-gGBE	GACTACAgttcgtaaaagccacataga	accctatgaaagcacagact	Figure4

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
si:ch211-286b5.4-gGBE-PIGS	GACTCACgttcgtaaaagccacataga	accctatgaaagcacagact	Figure4
dmd-C-gGBE	ATTGCTCacctgatcgagttttgatga	attcactgctggatctgagg	Figure4
dmd-gGBE-PIGS	ATCAAGCacctgatcgagttttgatga	attcactgctggatctgagg	Figure4
saraf-C-gGBE	AGAGTGCacctgatcgagttttgatga	cgtcgatgccatctgatcct	Figure4
saraf-gGBE-PIGS	ATCCGGAacctgatcgagttttgatga	cgtcgatgccatctgatcct	Figure4
si:dkey-28b4.7-C-gGBE	CATACGCaacagacgtttacctgtgca	tccatatcactatggtccta	Figure4
si:dkey-28b4.7-gGBE-PIGS	CCTATAGaacagacgtttacctgtgca	tccatatcactatggtccta	Figure4
si:ch73-6k14.2-C-gGBE	CCTTATGaacagacgtttacctgtgca	ttccagcatgaagcccagga	Figure4
si:ch73-6k14.2-gGBE-PIGS	CCGCATAaacagacgtttacctgtgca	ttccagcatgaagcccagga	Figure4
pkn1a-C-gGBE	ATAGGCGcgtagtactagaataaagctt	ccctgtggttctagctctag	Figure4
pkn1a-gGBE-PIGS	TAACGCCcgtagtactagaataaagctt	ccctgtggttctagctctag	Figure4
alkbh5-C-gGBE	CGTCCATcgtagtactagaataaagctt	aaaccctccgggataacacc	Figure4
alkbh5-gGBE-PIGS	GAACCTAcgtagtactagaataaagctt	aaaccctccgggataacacc	Figure4
cbl-C-gGBE	ATCCGGAttggagaggcagagagtg	agacgtcctctccccagta	Figure4
cbl-gGBE-PIGS	ATCGAAGttggagaggcagagagtg	agacgtcctctccccagta	Figure4
tardbp1-C-gGBE	ACGTCTCtggtcatctgcaaaggtg	agcaaggtatagatgaacca	Figure4
tardbp1-gGBE-PIGS	ACGTCTCtggtcatctgcaaaggtg	agcaaggtatagatgaacca	Figure4
tardbp2-C-gGBE	AGAAGTCagccgagctatatgtttgg	tagctccagcagcttgggaa	Figure4
tardbp2-gGBE-PIGS	AGAAGTCagccgagctatatgtttgg	tagctccagcagcttgggaa	Figure4
fus-C-gGBE	TGCAAGGagatggactcgaggtgaat	tggattgattactggtccca	Figure4
fus-gGBE-PIGS	TGCAAGGagatggactcgaggtgaat	tggattgattactggtccca	Figure4
sod1-C-gGBE	TGATGCGtctacagtcagcatggtga	agacgtcttcaatctcgat	Figure4
sod1-gGBE-PIGS	TGATGCGtctacagtcagcatggtga	agacgtcttcaatctcgat	Figure4
unsorted-reporter-rep1	TATGCGTacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
unsorted-reporter-rep2	ACGTCTCacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5
unsorted-reporter-rep3	ACTACGTacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5
sorted-reporter-rep1	ACGTTCTacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5
sorted-reporter-rep2	TGCATACacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5
sorted-reporter-rep3	AGAAGTCacagtacgaacgcgccga	tcgccagacacgctgaact	Figure5
mCherry-GFP-rep1	AACTGTCagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry-GFP-rep2	AAGAGCTagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry-GFP-rep3	AAGTCAGagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP-rep1	CGTAGAAagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP-rep2	GTTGCCAagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP-rep3	CTACAAGagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP+rep1	CTACGTCagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP+rep2	CTAGCACagatggactcgaggtgaat	tggattgattactggtccca	Figure5
mCherry+GFP+rep3	GAAGGTCagatggactcgaggtgaat	tggattgattactggtccca	Figure5
S-*PCSK9*-gGBE-PIGS-rep1	CCAATGAtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*PCSK9*-gGBE-PIGS-rep2	CCACGATtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*PCSK9*-gGBE-PIGS-rep3	CCGTGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*PCSK9*-gGBE-PIGS-sso7d-rep1	CGATTCAtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*PCSK9*-gGBE-PIGS-sso7d-rep2	CGTAGAAtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*PCSK9*-gGBE-PIGS-sso7d-rep3	GAAGGTCtcaccagtggctgggatg	ctcgggcacattctcgaag	FigureS7
S-*FANCF*-gGBE-PIGS-rep1	CATGCAAtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7
S-*FANCF*-gGBE-PIGS-rep2	CAGTTACtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
S-*FANCF*-gGBE-PIGS-rep3	CTATGGTtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7
S-*FANCF*-gGBE-PIGS-sso7d-rep1	CTACAAGtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7
S-*FANCF*-gGBE-PIGS-sso7d-rep2	CTACGTCtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7
S-*FANCF*-gGBE-PIGS-sso7d-rep3	CTAGCACtccaatcagtacgcagagag	aggtagcgcgcccactgcaa	FigureS7
S-*HBB*-gGBE-PIGS-rep1	TGTTGACcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*HBB*-gGBE-PIGS-rep2	TGCAGTTcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*HBB*-gGBE-PIGS-rep3	CAATGTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*HBB*-gGBE-PIGS-sso7d-rep1	CAAGTAGcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*HBB*-gGBE-PIGS-sso7d-rep2	CATCTTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*HBB*-gGBE-PIGS-sso7d-rep3	CATCTTGcccagtttagtagttggac	gcaacgtgctggtctgtgt	FigureS7
S-*CHM*-gGBE-PIGS-rep1	ATCGAAGacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-*CHM*-gGBE-PIGS-rep2	ATCGTCCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-*CHM*-gGBE-PIGS-rep3	ATGCTTCacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-*CHM*-gGBE-PIGS-sso7d-rep1	TGTGTCAacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-*CHM*-gGBE-PIGS-sso7d-rep2	CATGTCTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-*CHM*-gGBE-PIGS-sso7d-rep3	CACAAGTacggagggaatgcctaacttg	tttgtcccaaaactcgccactg	FigureS7
S-EMX1-gGBE-PIGS-rep1	GTACTTGcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7
S-EMX1-gGBE-PIGS-rep2	GTTCAACcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7
S-EMX1-gGBE-PIGS-rep3	GTTCGTAcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7
S-EMX1-gGBE-PIGS-sso7d-rep1	GTTGCCAcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7
S-EMX1-gGBE-PIGS-sso7d-rep2	GTGGTACcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7

Supplementary Table S2 HTS primers used for genomic DNA amplification of each target sites.

On-target site	Forward primer	Reverse primer	Figure
S-EMX1-gGBE-PIGS-sso7d-rep3	GCATTGTcagaagctggaggaggaag	agcagcaagcagcactctg	FigureS7
F0-1	CCTATAGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-2	CCTTATGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-3	CCGCATAagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-4	CCTTATGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-5	CCGCATAagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-6	TGCAAGGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F0-7	TTGATCCagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F1-A	AAGTCAGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F1-B	ATCAGTGagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F1-C	CAATAGCagatggactcgaggtgaat	tggattgattactggtccca	FigureS15
F1-D	CATACGCagatggactcgaggtgaat	tggattgattactggtccca	FigureS15

Supplementary Table S3 HTS primers used for genomic DNA amplification of Cas9 dependent off-target sites.

On-target site	Forward primer	Reverse primer	Figure
D-*HBB*-C-gGBE-off1-rep1	CTCGTATtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-C-gGBE-off1-rep2	CTGCTAAtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-C-gGBE-off1-rep3	CTGGAtgtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-C-gGBE-off2-rep1	AACTAGGtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-C-gGBE-off2-rep2	AAGATGCtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-C-gGBE-off2-rep3	GAACCTAtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-C-gGBE-off3-rep1	GACGAATtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-C-gGBE-off3-rep2	GACGTTGtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-C-gGBE-off3-rep3	TACAGAGtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-C-gGBE-off4-rep1	GCATTGTagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4
D-*HBB*-C-gGBE-off4-rep2	GCTAAGAagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4
D-*HBB*-C-gGBE-off4-rep3	GCTCATTagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4
D-*HBB*-gGBE-PIGS-off1-rep1	CCAATGAtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-gGBE-PIGS-off1-rep2	CCACGATtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-gGBE-PIGS-off1-rep3	CCGTGAAtgcatgtatctgcctacctc	agggaaataggtcttctta	FigureS4
D-*HBB*-gGBE-PIGS-off2-rep1	GAAGGTCtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-gGBE-PIGS-off2-rep2	GATTCTGtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-gGBE-PIGS-off2-rep3	GATTGCTtgcaaacccagtggtgaatc	tcacttgcatcaggctgctt	FigureS4
D-*HBB*-gGBE-PIGS-off3-rep1	TAGACGAtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-gGBE-PIGS-off3-rep2	TAGCTCTtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-gGBE-PIGS-off3-rep3	TTCAGCAtgttagagtgagactctgca	tgcgaatttggaagattgaa	FigureS4
D-*HBB*-gGBE-PIGS-off4-rep1	GCTGTAAagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4

Supplementary Table S3 HTS primers used for genomic DNA amplification of Cas9 dependent off-target sites.

On-target site	Forward primer	Reverse primer	Figure
D-*HBB*-gGBE-PIGS-off4-rep2	GCCTCTAagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4
D-*HBB*-gGBE-PIGS-off4-rep3	GCGAGTTagagtttcccaaaagaaggg	atggtcccccagagtcact	FigureS4
D-*HFE*-C-gGBE-off1-rep1	CTCTGTAgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-C-gGBE-off1-rep2	CTGATGTgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-C-gGBE-off1-rep3	CTGTAGAgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-C-gGBE-off2-rep1	AAGAGCTgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4
D-*HFE*-C-gGBE-off2-rep2	AAGTCAGgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4
D-*HFE*-C-gGBE-off2-rep3	ATCAGTGgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4
D-*HFE*-gGBE-PIGS-off1-rep1	GAACACTgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-gGBE-PIGS-off1-rep2	GATCTCAgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-gGBE-PIGS-off1-rep3	GACATCTgatgaatgtgctcatggag	ctccacattgctctgggca	FigureS4
D-*HFE*-gGBE-PIGS-off2-rep1	ATCTGCTgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4
D-*HFE*-gGBE-PIGS-off2-rep2	ATGTGACgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4
D-*HFE*-gGBE-PIGS-off2-rep3	GACAGTAgcaatgtggatagtctggtc	catccagtccttctgctgc	FigureS4

Supplementary Table S4 HTS primers used for genomic DNA amplification of Cas9 independent off-target sites.

On-target site	Forward primer	Reverse primer	Figure
HBB-C-gGBE-off1-rep1	TGAGACTtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-C-gGBE-off1-rep2	TGTCAGTtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-C-gGBE-off1-rep3	TGATCGTtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-C-gGBE-off2-rep1	ATTCTCGaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-C-gGBE-off2-rep2	ATTGCTCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-C-gGBE-off2-rep3	ATCAAGCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-C-gGBE-off3-rep1	GCTAAGAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HBB-C-gGBE-off3-rep2	GCTCATTagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HBB-C-gGBE-off3-rep3	GCTGTAAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HBB-gGBE-PIGS-off1-rep1	TGACAACtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-gGBE-PIGS-off1-rep2	TGAGCAGtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-gGBE-PIGS-off1-rep3	CAAGTAGtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HBB-gGBE-PIGS-off2-rep1	ATCCGGAaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-gGBE-PIGS-off2-rep2	ATGGCGAaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-gGBE-PIGS-off2-rep3	ACTTAGCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HBB-gGBE-PIGS-off3-rep1	GCCTCTAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HBB-gGBE-PIGS-off3-rep2	GCGAGTTagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HBB-gGBE-PIGS-off3-rep3	AACTAGGagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-C-gGBE-off1-rep1	ACTCTGAtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-C-gGBE-off1-rep2	ACTGACTtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-C-gGBE-off1-rep3	AGACTCTtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-C-gGBE-off2-rep1	ATTCTCGaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HFE-C-gGBE-off2-rep2	ATTGCTCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5

Supplementary Table S4 HTS primers used for genomic DNA amplification of Cas9 independent off-target sites.

On-target site	Forward primer	Reverse primer	Figure
HFE-C-gGBE-off2-rep3	ATCAAGCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HFE-C-gGBE-off3-rep1	GCTAAGAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-C-gGBE-off3-rep2	GCTCATTagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-C-gGBE-off3-rep3	GCTGTAAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-gGBE-PIGS-off1-rep1	AGTCACAtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-gGBE-PIGS-off1-rep2	TCAGAGAtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-gGBE-PIGS-off1-rep3	TGACTGAtgcatttgtaggcttgatgc	ttccagcccagccaaactt	FigureS5
HFE-gGBE-PIGS-off2-rep1	ATCCGGAaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HFE-gGBE-PIGS-off2-rep2	ATGGCGAaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HFE-gGBE-PIGS-off2-rep3	ACTTAGCaaggacaggcagcatagact	acgggaaatgaaggcctag	FigureS5
HFE-gGBE-PIGS-off3-rep1	GCCTCTAagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-gGBE-PIGS-off3-rep2	GCGAGTTagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5
HFE-gGBE-PIGS-off3-rep3	AACTAGGagaatcctggacaaggttt	atgtgagcccatcaggtatt	FigureS5

Supplementary Table S5 HTS primers used for amplification of RNA off-target sites.

On-target site	Forward primer	Reverse primer	Figure
R-*CHM*-CTR-off-rep1	GTTGCCActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-CTR-off-rep2	GTGGTACctaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-CTR-off-rep3	GCATTGTctaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-ABE8e-off-rep1	GTTGCCActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-ABE8e-off-rep2	GTTGCCActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-ABE8e-off-rep3	GTTGCCActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-C-gGBE-off-rep1	GCTAAGActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-C-gGBE-off-rep2	GCTCATTctaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-C-gGBE-off-rep3	GCTGTAActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-gGBE-PIGS-off-rep1	GCCTCTActaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-gGBE-PIGS-off-rep2	GCGAGTTctaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*CHM*-gGBE-PIGS-off-rep3	AACTAGGctaggagatacacctcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-CTR-off-rep1	AACCTAGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-CTR-off-rep2	AAGATGCattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-CTR-off-rep3	AAGCGTGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-ABE8e-off-rep1	AACCTAGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-ABE8e-off-rep2	AACCTAGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-ABE8e-off-rep3	AACCTAGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-C-gGBE-off-rep1	ATTCTCGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-C-gGBE-off-rep2	CTAGACAattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-C-gGBE-off-rep3	CTTGAGTattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-gGBE-PIGS-off-rep1	CTCACGAattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-gGBE-PIGS-off-rep2	CTCAGACattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6
R-*EMX1*-gGBE-PIGS-off-rep3	CTCTCAGattgccacgaagcaggcca	aaaaaatagttggtgtagagga	FigureS6

Supplementary Table S5 HTS primers used for amplification of RNA off-target sites.

On-target site	Forward primer	Reverse primer	Figure
R-*HBB*-CTR-off-rep1	GTGACTCtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-CTR-off-rep2	GTGTCAAtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-CTR-off-rep3	AACTGTCtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-ABE8e-off-rep1	GTGACTCtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-ABE8e-off-rep2	GTGACTCtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-ABE8e-off-rep3	GTGACTCtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-C-gGBE-off-rep1	AACGTGTtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-C-gGBE-off-rep2	AAGAGCTtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-C-gGBE-off-rep3	AAGTCAGtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-gGBE-PIGS-off-rep1	ATCAGTGtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-gGBE-PIGS-off-rep2	ATCTGCTtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6
R-*HBB*-gGBE-PIGS-off-rep3	ATGTGACtcagaaagtggtggctggt	aaaaaatagttggtgtagagga	FigureS6

REFERENCE

Qin W, Lu X, Liu Y, et al. 2018. Precise A*T to G*C base editing in the zebrafish genome. *Bmc Biology,* **16**(1):139.

Rosello M, Vougny J, Czarny F, et al. 2021. Precise base editing for the in vivo study of developmental signaling and human pathologies in zebrafish. *Elife,* **10**:e65552.

Sun Y, Zhang B, Luo L, et al. 2019. Systematic genome editing of the genes on zebrafish Chromosome 1 by CRISPR/Cas9. *Genome Research,* **30**(1):118-126.

Tong H, Liu N, Wei Y, et al. 2023. Programmable deaminase-free base editors for G-to-Y conversion by engineered glycosylase. *National Science Review,* **10**(8):d143.

优化的 gGBE 在哺乳动物细胞和斑马鱼胚胎中实现高效 G-to-Y 转换

黎敏[1,#]，焦瑶歌[1,#]，陶瑞[2]，胡云[1]，费俊逸[1]，王春婷[1]，姚少华[1,3,*]

[1]Laboratory of Biotherapy, National Key Laboratory of Biotherapy, Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

[2]Institute of Kidney Diseases, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, China

[3]Tianfu Jincheng Laboratory, Chengdu 610095, Sichuan, China

[#]并列第一作者

*通讯作者，邮箱：shaohuayao@scu.edu.cn

摘要

碱基编辑是一种高效且精准的基因组编辑工具，可实现高分辨率单碱基替换。根据所采用的碱基修饰结构域不同，碱基编辑器在哺乳动物细胞中既可介导转换型编辑，也可实现颠换型编辑。尽管基于脱氨酶的碱基编辑器已在斑马鱼基因组中实现高效转换编辑，但基于糖基化酶的颠换编辑仍然面临较大挑战。本研究基于 Cas9 复合体的结构信息，将 N-甲基嘌呤 DNA 糖基化酶结构域（MPG）嵌入 Cas9 蛋白内部，构建了一系列糖基化酶型鸟嘌呤碱基编辑器（gGBE）。我们发现将 MPG 插入到化脓性链球菌 Cas9（SpCas9）PI 结构域的 G1246 与 S1247 残基得到的 gGBE-PIGS 总体编辑效率最高。相比 N-端融合的 gGBE，gGBE-PIGS 在哺乳动物细胞中的编辑效率最高可提升 3.5 倍，在斑马鱼胚胎中最高可提升 13 倍，且编辑精确度未受影响。此外，我们构建了 gGBE 编辑响应型 GFP 报告系统，该系统可用于监测编辑效率，同时富集经 gGBE 编辑的细胞和胚胎。结合优化后的 gGBE-PIGS 与 该 GFP 报告系统，我们成功构建了携带人类肌萎缩侧索硬化症（ALS）同源突变的斑马鱼疾病模型。综上所述，本研究证明了 gGBE 在精准碱基编辑领域的高效性与多样性，并揭示其在体内疾病模型构建及基因治疗领域的潜在应用前景。

关键词： CRISPR/Cas9；碱基编辑；糖基化酶；斑马鱼；肌萎缩侧索硬化症