Base editors (BEs) are genome engineering tools that can generate nucleotide substitutions without introducing double-stranded breaks (DSBs). A variety of strategies have been developed to improve the targeting scope and window of BEs. In a previous study, we found that a bacteriophage-derived peptide, referred to as G8PPD, could improve the specificity of Cas9 nuclease. Herein, we investigate the applicability of G8PPD as molecular modulators of BEs. We show that G8PPD can improve cytidine base editor (CBEs) and adenine base editor (ABE) to more focused targeting windows. Notably, in a cell-based disease model, G8PPD increases the percentage of perfectly edited gene alleles by BEs from less than 4% to more than 38% of the whole population. In addition, G8PPD can improve the targeting scope of BE in mouse embryos. In summary, our study presents the peptidyl modulators that can improve BEs for precision base editing.
C lustered regularly interspaced short palindromic repeats (CRISPR)-CRISPR-associated genes (Cas) is the bacterial adaptive immune system for protecting host organisms from invading pathogens1-3. Owing to the modular feature, type II CRISPR systems, particularly CRISPR-Cas9, have been widely used for genome editing, transcriptional and epigenetic modulation and molecular imaging4. CRISPR-Cas9 can be directed to human genome by single-guide RNA (sgRNA) to create double-stranded breaks (DSBs) at targeted genomic loci5-7. In human cells, DSBs are repaired by two competing DNA repair pathways: error-prone nonhomologous end joining (NHEJ) and homology directed repair (HDR). The latter repair pathway can facilitate CRISPR-Cas9-mediated gene correction of pathogenic mutations8.

Recent studies have highlighted base editors (BEs) as an efficient genome editing tool for precision gene therapy9-11. BEs are fusion proteins comprising a catalytically inactive Cas nuclease and a nucleobase deaminase12,13. Unlike CRISPR-Cas9 genome editing tools, BEs generate precise base substitutions without introducing DSBs, thus avoiding undesired, competing NHEJ events that incorporate nucleotide insertions and deletions (indels)14. Cytidine base editors (CBEs) and adenine base editors (ABEs) have been developed to realize genomic alterations of C-G to T-A and A-T to G-C, respectively. Besides DNA editing, Cas variants with RNA-binding capacities have been adapted for programmable RNA editing15,16.

The widely used third-generation CBEs, referred to as BE3, are fusion proteins composed of rat APOBEC (rAPOBEC), Cas9 D10A nickase and uracil glycosylase inhibitor (UGI). BE3 has a five nucleotide targeting window ranging from position 4 to 8, counting the protospacer adjacent motif (PAM) trinucleotides as positions 21 to 23. An engineered CBE variant A3A, where rAPOBEC is replaced with human APOBEC3A (hAPOBEC3A), has a wider targeting window at positions 2 to 13 and is capable of editing methylated genomic regions17. Engineering endeavors have been made to improve the editing window of both BE3 (rAPOBEC-nCas9-UGI)18 and A3A (hAPOBEC3A-nCas9-UGI) CBEs19. Compared with ABEs, CBEs have higher genome-wide off-target events20,21. Despite of the considerable studies, the origin of the flexible editing windows and genome-wide off-target of CBEs are not completely understood.

It has been well known that excessive dosage of Cas9 and sgRNA can lead to elevated off-target mutations22. This finding led to the widespread use of directly delivered Cas9-guide RNA (gRNA) ribonucleoproteins (RNP)s, which can improve the genome editing specificity by limiting the exposure of genome to gene-editing agents23. Similarly, delivery of BE-gRNA RNP can increase the base editing specificity compared with transfection of BE-coding plasmids24. Besides the use of gene-editing agents with short intracellular half time23, molecular switches provide compelling opportunities to improve the specificity of CRISPR-Cas via the temporal control of its cellular activity. Indeed, a variety of CRISPR-Cas inhibitors have been developed, including naturally occurring anti-CRISPR (Acr) proteins25-27, synthetic oligonucleotides28 and small molecules29. These inhibitors can improve the genome editing specificity of Cas9 RNP30 and modulate the cellular activity of BE329.

In a previous study, we have identified bacteriophage-derived peptides G8P9D as inhibitors to Streptococcus pyogenes Cas9 (SpCas9)31. Herein, we hypothesized that the broad targeting window of CBEs is attributed, at least in part, to the excess intracellular dosage of base editing agents that could be alleviated by timed delivery of G8P9D. Indeed, we found that overexpression of the peptide inhibitors of SpCas9 in human cells could direct BE3 and A3A CBEs to more focused editing windows and facilitate precision gene correction in a cell model of Marfan syndrome32 and base editing in mouse embryos. 

### Results

**G8P9D peptides inhibit CBE targeting of EGFP reporter.** We have reported in a previous study that the periplasmic domain of the major coat proteins (G8P9D) from M13 and f1 bacteriophages (Fig. 1a) can inhibit the in vitro and cellular activities of SpCas931. Because these peptides suppress SpCas9 activity by disrupting SpCas9-sgRNA binding31, we hypothesized that they could also act as inhibitors to cytidine base editors (CBEs) the activity of which also rely on the sgRNA-guided DNA binding of the SpCas9 domain. In order to evaluate the inhibitory activities of M13 and f1 G8P9D on CBEs, we constructed an enhanced green fluorescent protein (EGFP) reporter carrying an inactivating mutation Y66C at the chromophore (Supplementary Fig. 1a). An sgRNA was designed to target the Y66C mutation site for introducing C-to-T conversion by CBE (Supplementary Fig. 1b). This mutation can correct EGFP-Y66C to wild-type genotype, yielding fluorescent cells that can be readily detected by flow cytometry. Inhibition of A3A CBE by G8P9D peptides will lead to reduced EGFP fluorescence (Supplementary Fig. 1a).

Our previous study has shown that G8P9D peptides bind to sgRNA-free SpCas9 (apo-Cas9) and that maximum SpCas9-inhibiting activity of G8P9D can be achieved by pre-incubating cells with overexpressed G8P9D31. Hence, in the present study we transfected G8P9D plasmids at 24 h prior to the transfection of CBE17, sgRNA and EGFP-Y66C plasmids (Supplementary Fig. 1c). An inactive G8P9D mutant (Mut2)31 (Fig. 1a) was included to control for the cell stress induced by serial transfections and non-specific effects of peptide expression on base editing (Fig. 1b). It was found that A3A CBE targeting of EGFP-Y66C reporter plasmid could efficiently recover EGFP fluorescence (Supplementary Fig. 1d). Pre-incubation of cells with M13 or f1 G8P9D overexpression plasmid reduced more than 50% of GFP activation compared with A3A CBE only or Mut2 G8P9D groups (Fig. 1b). These results supported our hypothesis that M13 and f1 G8P9D could also function as inhibitors to BE. Next-generation sequencing (NGS) analyses of the PCR amplification product of A3A CBE-targeted site revealed evident C-to-T mutations in EGFP reporter (Fig. 1c). M13 and f1 G8P9D, but not the inactive G8P9D mutant, could efficiently inhibit CBE-induced C-to-T mutations in EGFP reporter gene (Fig. 1c). It appeared that f1 G8P9D has significantly higher inhibitory activity than M13 G8P9D (Fig. 1b,c). Thus, subsequent studies are performed with f1 G8P9D.

**G8P9D inhibits CBE targeting at endogenous genomic sites in human cells.** To examine the inhibitory effects of G8P9D on CBE at endogenous sites, we designed sgRNA targeted to human EPPKI, GATA2D2A, DNMT3B, and DDX53 genes respectively for A3A CBE (Fig. 1d). Similar to the procedure in EGFP reporter assay, f1 G8P9D-encoding plasmid was transfected into HEK293T cells at 24 h prior to A3A CBE and sgRNA transfection (Fig. 1e). The sgRNA-encoding and f1 G8P9D-encoding plasmids carried GFP and mCherry reporter genes respectively. A G8P9D-free pcDNA3.1 vector was included to control for cell stress induced by serial transfections. In order to enrich cells with transfected plasmids, we used flow cytometry to enrich GFP and mCherry-dual positive cells (Supplementary Fig. 2) for analyses of CBE-induced genomic mutations.

NGS analyses of sorted HEK293T cells revealed efficient base editing by A3A CBE at the four selected genomic sites (Supplementary Fig. 3a). The presence of f1 G8P9D reduced CBE-mediated C-to-T conversion by various degree across different positions and genomic loci (Supplementary Fig. 3a). Statistical analyses of the editing efficiency at all examined positions revealed significant inhibitory activity of G8P9D toward A3A CBE (Fig. 1f). Comparing
**Fig. 1** G8PPD peptides inhibit CBE targeting of EGFP reporter and endogenous genomic sites in human cells. **a** The design of G8PPD plasmids for mammalian expression. **b–c** Analysis of the effects of G8PPD on A3A CBE using EGFP reporter cells. **b** Normalized flow cytometry results. **c** Next-generation sequencing analysis of C-to-T conversion. For **b–c**, the data are shown as mean ± standard error of mean (SEM) (n = 3 biologically independent replicates). **d** Design of sgRNA targeted to DDX53, DNMT3B, GATAD2A and EPPK1 genomic loci. The PAM sequence and cytidines are highlighted in blue and red, respectively. Red arrows denote the positions of selected target sites. **e** Flow chart showing experimental procedures. The cartoons are blue for G8PPD, grey for CBE and yellow for sgRNA, respectively. **f** Analysis of the frequency of C-to-T conversion at all editing positions in the absence and presence of G8PPD. Each sgRNA contains 3 biologically independent replicates. The center line in each box indicates the median. The lower and upper bounds of each box represent the first quartile (25%) and the third quartile (75%), respectively. The bottom and top of whiskers denote the minimum and maximum, respectively. For **c** and **f**, significant difference is determined using two-tailed Student’s t test. n.s., not significant. Source data are provided as a Source Data file.

Different target sites or sgRNA, we found that the averaged inhibition rates of each target site had no significant difference (Supplementary Fig. 3b), suggesting that the CBE-inhibiting activity of G8PPD was not dependent on genomic loci or sgRNA.

To investigate whether the activity of G8PPD was cell-type dependent, we examined the performance of G8PPD in U-2 OS cells. It was found that G8PPD reduced the editing efficiency of A3A CBE by various degree across different genomic sites and editing positions (Supplementary Fig. 4a) in U-2 OS. Similar to the results in HEK293T cells, the inhibitory activity of G8PPD in U-2 OS cells was not dependent on genomic loci or sgRNA (Supplementary Fig. 4b). Collectively, these results suggested that G8PPD could inhibit the editing activity of A3A CBE at different genomic sites across different cell types.

G8PPD differentially inhibits the on-target and out-of-window editing of A3A CBE. We next analyzed the inhibitory activity of G8PPD at different editing positions within each target site. It was found that the inhibitory effects of G8PPD displayed notable variations along the 20-bp targeting site for each sgRNA. This result was consistently observed in HEK293T and U-2 OS cells (Fig. 2a). This observation prompted us to compare the effects of G8PPD on the editing positions 4 to 8, which are conventionally deemed as the on-target editing window of CBEs, and on the out-of-window editing positions 1 to 3 and 9 to 20. It was found that G8PPD significantly inhibited both the on-target and out-of-window editing of A3A CBE in HEK293T cells, with the latter being reduced to minimum level (Fig. 2b).

Interestingly, G8PPD had minor or little inhibition toward the on-target editing of A3A CBE but significantly inhibited the out-of-window editing (Fig. 2b).

The differential effects of G8PPD at the on-target and out-of-window editing positions were more evident when the inhibition rates were plotted along the 20-bp targeting site (Fig. 2c). Notably, G8PPD exhibited lower inhibition of A3A CBE activity at...
positions 4 to 8, the canonical on-target positions of CBEs (Fig. 2b), than at out-of-window positions. The hot spots of G8PPD-mediated inhibition had significant overlap with the canonical out-of-window editing positions of CBEs (Fig. 2d). Comparative analyses of the inhibition of C-to-T conversion showed that G8PPD exhibited 2- and 10-fold selectivity of inhibition toward out-of-window positions over on-target positions in HEK293T and U-2 OS cells respectively (Fig. 2e).

G8PPD differentially inhibits the on-target and out-of-window editing of BE3 CBE and ABE7.10. Because G8PPD acts as a SpCas9 inhibitor, we envisioned that G8PPD could inhibit different CBEs carrying SpCas9 module as the DNA-binding domain. Hence, we sought to examine the effects of G8PPD on BE3 CBE and ABE7.10 that contain deaminase domains different from that in A3A CBE. We found that f1 G8PPD could suppress BE3 CBE-induced C-to-T conversion though the inhibitory effects appeared to be less prominent than that with A3A CBE (Supplementary Fig. 5a). Similar to the results with A3A CBE, the inhibitory effects of f1 G8PPD displayed varied inhibition rates at different editing positions of BE3 CBE (Fig. 3a), with the on-target positions 4 to 8 exhibiting lower inhibition compared to the out-of-window positions (Fig. 3b). It was observed that G8PPD had 50-fold selectivity of inhibition toward the out-of-window positions over on-target positions (Fig. 3c).

We next examined the effects of G8PPD on ABE7.10. We designed four sgRNAs targeting to different genomic sites (Fig. 3d). Similar to the results with A3A and BE3 CBEs, f1 G8PPD inhibited the A-to-G conversion activity of ABE7.10 in a position-dependent manner with the on-target positions 4 to 7 exhibiting lower inhibition rates than the out-of-window positions (Fig. 3e). f1 G8PPD showed 20-fold selectivity of inhibition toward the out-of-window positions over on-target positions of ABE7.10 (Fig. 3f). These results collectively demonstrated that G8PPD could preferentially inhibit the editing activities of CBE and ABE at out-of-window positions, suggesting the potential application of G8PPD as an agent to improve the targeting scope of BEs. In addition, we examined the effects of G8PPD on A3A CBE editing at the ABE7.10 targeting site VISTA and consistent inhibitory activity of G8PPD was observed at both the on-target and out-of-window targeting sites of A3A CBE.

As previous studies suggested that timed delivery of Acrs could improve the genome-editing specificity of SpCas930, we next sought to investigate the effects of AcrIIA4 on modulating CBE and ABE activities. HEK293T cells were co-transfected with AcrIIA4 plasmid and sgRNA- and BE-coding plasmids (Supplementary Fig. 6a) and the inhibitory activity of AcrIIA4 was...
examined. It was found that AcrIIA4 suppressed the on-target editing. These results together suggested that peptide inhibitors of SpCas9 such as G8PPD are compatible with and may even improve high-fidelity base editing.

Co-injection of G8PPD mRNA mediates precision base editing in mouse embryos. In order to avoid serial transfection of G8PPD for practical applications, we investigated the effects of co-transfected G8PPD-encoding plasmid on base editing. It was found that pretransfection, but not co-transfection, of G8PPD plasmid could reduce the out-of-window editing of A3A CBE (Supplementary Fig. 10a-b). This result was consistent with the previous finding that timed delivery of G8PPD was important for achieving different effects on SpCas9-based genome editing too31.

We next examined the effects of co-injected G8PPD-encoding mRNA on A3A CBE-mediated base editing in mouse embryos. Expression of G8PPD from mRNA could bypass the transcription

Fig. 3 The inhibitory activities of f1 G8PPD toward BE3 CBE and ABE7.10 at the endogenous genomic sites in human cell lines. a The inhibitory activity of f1 G8PPD toward BE3 CBE in HEK293T cells. b Heat map showing position-dependent inhibition of BE3 CBE by G8PPD. c Comparison of the inhibition rates of G8PPD at the on-target and out-of-window editing positions of BE3 CBE. For a–c, DDX53, DNMT3B and EPPI1 groups contain 2 biologically independent replicates and NANOGP8 groups contain 3 biologically replicates. d Design of sgRNA for ABE7.10 targeted genomic loci. The PAM sequence and adenine positions are highlighted in blue and red, respectively. e Heat map showing the position-dependent inhibition of ABE7.10 by G8PPD. f Comparison of the inhibition rates of G8PPD at the on-target and out-of-window editing positions of ABE7.10. For c and f, the center line in each box indicates the median. The lower and upper bounds of each box represent the first quartile (25%) and the third quartile (75%), respectively. The bottom and top of whiskers denote the minimum and maximum, respectively. Significant difference is determined using two-tailed Student’s t test. n.s., not significant. g Analysis of the effects of G8PPD on A3A CBE targeting at the ABE7.10 targeting site VISTA. The data are shown as mean ± SEM. For e–g, each sgRNA contains 3 biologically independent replicates. Source data are provided as a Source Data file.
process and resemble pre-transfection of G8PPD. A previously described sgRNA targeting to mouse Tyr gene \(^{24}\) was constructed for introduction of a stop codon (Fig. 5a). G8PPD-encoding mRNA was in vitro transcribed and then co-injected into one-cell stage mouse embryos with sgRNA and A3A CBE-coding plasmids (Fig. 5b). Compared with the mock group containing pcDNA, G8PPD mRNA did not compromise blastocyst development or genomic amplification efficiency (Fig. 5c). NGS analysis of successfully amplified gene alleles showed that the use of G8PPD could facilitate the generation of blastocysts carrying perfectly edited gene alleles despite of the slightly lower overall editing efficiency compared to mock group (Fig. 5d and Supplementary Fig. 11).

To expand the generality and applicability of our finding, we performed microinjection experiments using sgRNA and CBE-coding mRNAs. In these experiments, we used chemically synthesized G8PPD peptide to ensure the presence of sufficient G8PPD for modulating the activity of transcribed A3A and BE3 CBE. In consistency with the above observations, G8PPD peptide improved the efficiency of generating perfectly edited blastocysts using A3A or BE3 CBE-coding mRNA (Fig. 5d and Supplementary Fig. 12-13). Thus, G8PPD may function as an efficient tool for generation of genetically modified animals.

**Discussion**

It has well documented that CRISPR-Cas systems are associated with off-target cleavage\(^{35}\), chromosomal rearrangement\(^{36}\) and genotoxicity\(^{37}\). These adverse effects may result from uncontrolled expression of Cas proteins\(^{22,38,39}\). Unfortunately, most therapeutically relevant applications of CRISPR-Cas involve constitutively active Cas nucleases that raises major safety concern\(^{40}\). With the increasing number of therapeutic applications of CRISPR-Cas, the ability to manipulate the activity of CRISPR-Cas in human cells has become urgently needed. One feasible approach is to develop CRISPR-Cas off-switches to control the intracellular activity of Cas proteins. Thus far, synthetic oligonucleotides\(^{28}\), small-molecule inhibitors\(^{29}\) and bacteriophage-derived anti-CRISPR proteins (Acrs)\(^{31}\) have been developed to enable the temporal control of CRISPR-Cas activity. It has been shown that Acrs along with Cas9-sgRNA ribonucleoproteins (RNPs)\(^{30}\).
In our previous study, we discovered that the in vitro and cellular activity of SpCas9 can be inhibited by the periplasmic domain of bacteriophage major coat proteins (G8PPD)\textsuperscript{31}. These G8PPD peptides disrupt Cas9-sgRNA assembly by binding to the PAM interacting (PI) domain of SpCas9\textsuperscript{31}. It was found that time-controlled delivery of overexpressed G8PPD could improve the specificity of SpCas9 in human cells\textsuperscript{31}. In the present study, we hypothesized that the out-of-window base editing of CBE may be attributed to excess CBE fusion proteins. Hence, we sought to explore whether timed delivery of SpCas9-inactivating G8PPD could inhibit and improve the activities of CBE and ABE, in a rationale similar to that with the genome editing of SpCas9\textsuperscript{31}. Our results have demonstrated that G8PPD can preferentially inhibit the out-of-window editing of A3A and BE3 CBEs and ABE7.10, leading to more focused targeting window. To the best of our knowledge, our study represents the first peptide inhibitors that can improve the targeting scope of BEs. Unlike G8PPD that preferentially inhibited out-of-window over on-target editing, AcrIIA4 abolished the activities of CBE and ABE at both on-target and out-of-window sites without any selectivity. These results support the notion that weak inhibitors, rather than strong inhibitors, are optimal agents for modulating the specificity of genome-editing tools. Nevertheless, it is possible that the editing activity of BEs may be improved by limiting the exposure time of SpCas9 to Acrs where Acrs are delivered at later time points during transfection.

It has been reported that the targeting specificity of CBEs and ABEs can be improved by incorporating genetic mutations into deaminase domain\textsuperscript{20,21}. These BE variants are extremely useful when precision base editing is needed for therapeutic applications. In the present study, we have shown that G8PPD has minor yet significant inhibition at the out-of-window editing position of hA3A-eBE-Y130F CBE. Most importantly, G8PPD dose not affect the on-target editing activities of hA3A-eBE-Y130F and BE3-R33A CBEs. These results have established the feasibility of G8PPD not only as an alternative approach to improving the targeting specificity of BEs but also as an additive agent to combine with BE variants. In further studies, it would be
interesting to evaluate the inhibitory activities of G8PPD toward dual ademe and cytidine base editors (A&C-BEs) and prime editors. These studies can advance our understanding of the capability of G8PPD serving as general modulators of CRISPR-Cas genome engineering tools.

As a concept of proof, most of the described applications of G8PPD in the present study was carried out through serial transfection. Although co-transfection of G8PPD plasmid with sgRNA and BE-coding plasmids did not improve the targeting specificity of A3A CBE, co-injection of G8PPD mRNA with sgRNA and BE-coding plasmids into mouse embryos could facilitate precision base editing for generation of blastocysts carrying perfectly edited genes. Similarly, G8PPD peptide could facilitate precision base editing in mouse embryos using A3A or BE3 CBE-coding mRNA. It must be noted, however, that further studies with larger-scale analysis of the effects of G8PPD on embryonic base editing would be important to reveal the full potential of G8PPD. In addition, it may be also feasible to genetically fuse G8PPD-coding sequence with BEs to facilitate the applications of G8PPD. However, due to the specific mechanism of action of G8PPD, which allosterically inhibits SpCas9, the structural organization of G8PPD-BE fusion protein may need to be carefully investigated.

Although only examined f1 and M13 G8PPD for inhibition of CBE activity in the current study, there might be other G8PPD from inoviridea phages, which can function as CBE off-switches. Comparative analyses can be performed on these different G8PPD peptides to determine the sequence-activity relationship. This information would facilitate the design and development of next-generation peptide off-switches of CBEs. Moreover, it can be also important to characterize the mechanism of action of G8PPD on CBE inhibition. Unlike most Acrs that inhibit Cas activities by disrupting DNA or RNA binding, G8PPD functions as an allosteric inhibitor of SpCas9. This unique mechanism may be critical for G8PPD to selectively modulate BE activities at on-target and out-of-window positions. The in-depth mechanism studies could facilitate the development of next-generation CRISPR modulators for precision gene correction.

Methods

Animal ethics statement. The use and care of animals were complied with the guidelines of the Institutional Animal Care and Use Committee (IACUC) of GemPharmatech Co., Ltd, Nanjing, Jiangsu, China. Mice were maintained in a SPF (specific pathogen-free) facility under a 12 h dark-light cycle.

Plasmid construction. sgRNA was cloned into pGL3-sgRNA expression vector carrying a U6 promoter and an EGFP reporter gene (Addgene, #107721). CBE-targeted genomic sites are indicated in Supplementary Table 1. The sequences of sgRNA-encoding oligonucleotides were listed in Supplementary Table 2. Human codon-optimized DNA sequences encoding M13, f1 G8PPD and Mut2 G8PPD were cloned into the BamHI/XbaI sites of pCDNA3.1 (+) by plasmid recombination kit Clone Express (Vazyme Biotech, Nanjing, China). These G8PPD peptides carry a C-terminal SV40 nuclear localization signal (NLS) for co-localization with Cas9 proteins. G8PPD peptides were cloned into pLV-EFluc-mCherry plasmid harboring mCherry fluorescent protein marker. For construction of EGFP-Y66C reporter plasmid, Y66C mutation was introduced by quickchange PCR. Mut2 G8PPD carrying a C-terminal SV40 nuclear localization signal (NLS) was cloned into A3A CBE expression vector.

Cell culture and transfection. HEK293T and U-2 OS cells were obtained from ATCC and maintained in Dulbecco’s Modified Eagle Medium (DMEM) (Cat. No. SH30243, Hyclone, Logan, USA) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin at 37 °C with 5% CO2. Transfection was performed using Lipofectamine 2020 Reagent (Cat. No. 11668019, ThermoFisher Scientific, Waltham, USA) according to manufacturer’s instructions. HEK293T and U-2 OS cells were seeded on to poly-D-lysine (Cat. No. A-003-E, Sigma, St. Louis, USA) pre-coated 24-well plates at 24 h prior to transfection.

For EGFP reporter assay, 1 μg G8PPD-encoding plasmid was transfected into HEK293T cells. After 24 h, 0.25 μg sgRNA plasmid, 0.5 μg A3A CBE plasmid and 0.05 μg the EGFP-Y66C reporter plasmid were co-transfected into G8PPD-expressing cells. At 48 h after the second transfection, cells were harvested and analyzed by flow cytometry. The efficiency of GFP activation was calculated based on sorted cells. To calculate the C-to-T conversion in EGFP gene, the cells were first sorted by EGFP flow cytometry and then subjected to base editing at endogenous genes, unless noted otherwise, cells of approximately 70% confluency were transfected with 1 μg G8PPD plasmid that carries an mCherry selection marker at 12 h after seeding. At 24 h after G8PPD transfection, 0.25 μg sgRNA plasmid that carries a GFP selection marker and 0.5 μg CBE plasmid were co-transfected into G8PPD-expressing cells. At 48 h after CBE and sgRNA plasmids transfection, cells were harvested and analyzed by Beckman Coulter CyteFLEX (Beckman Coulter, Brea, USA) or sorted by BD FACSaria III flow cytometer (BD Biosciences, New York, USA). Flowjo VX or Flowjo 10 was used to analyze flow cytometry data. At least 2,000 mCherry and GFP dual-positive cells were collected for subsequent analyses. Analyses of the Sanger sequencing results of the PCR product of edited EGFP reporter were performed using EdiTaq and analyzed by NGS.

For AcrI4A co-transfection experiments, HEK293T cells were transfected with 0.5 μg of CBE plasmid and 0.25 μg sgRNA-expressing and 1 μg of AcrI4A expressing plasmids that carry GFP and mCherry reporters, respectively. At 72 h after transfection, mCherry and GFP dual positive cells were collected for subsequent analyses.

In vitro transcription. G8PPD mRNA was transcribed from f1 G8PPD-coding PCR products with a 5′ T7 promoter sequence using HiScibe T7 ARCA mRNA kit (NEB). The transcription was performed at 37 °C overnight and then purified by phenol: chloroform extraction, followed by ethanol precipitation. To generate BE3 and A3A CBE mRNA, BE3 and A3A-coding plasmids were linearized by AgeI (NEB, R3552L) and then used as templates for in vitro transcription using T7 ULTRA (Ambion, AM1345). BE3 and A3A mRNAs were purified using RNeasy Mini Kit (Qiagen, 74104). The RNA of T7-targeting sgRNA was amplified and transcribed from a sgRNA-coding plasmid along with a T7 promoter using MEGA short transcript T7 KIT (Ambion, AM1345). The sgrRNA were purified using MEGA clear Kit (Ambion, AM1908) and recovered by alcohol precipitation.

Microinjection of mouse one-cell embryos. Female C57BL/6 mice of 3.5–4 weeks old and male C57BL/6 mice of 3–6 months old were used in our experiments. All mice were housed in the specific pathogen-free (SPF) animal facility at GemPharmatech Co., Ltd in accordance with institutional guidelines under the following conditions: 23 °C ambient temperature, 40–70% humidity, 12 h dark/light cycle and free access to water and rodent chow. Superovalutized female C57BL/6 mice were mated to C57BL/6 j males, and zygotes were collected from the oviducts at E0.5. For DNA injection, 100 ng G8PPD mRNA, 50 ng A3A mRNA and 25 ng sgRNA were mixed and injected into the cytoplasm of zygotes with well recognized pronuclei. For RNA injection, 0.5 ng G8PPD peptide, 50 ng A3A mRNA and 25 ng sgRNA were mixed and injected into the cytoplasm of zygotes with well recognized pronuclei. Injected zygotes were cultured to blastocysts and then harvested for genomic analysis.

Extraction and PCR amplification of genomic DNA. Genomic DNA was extracted using QuickExtract® DNA Extraction Solution (Lucigen, USA). Genomic PCRs were performed using 100 ng genomic DNA, corresponding primers (Supplementary Table 4), Phanta Max Super-fidelity DNA Polymerase (Cat. No. P505-d1, Vazyme) or KOD plus (Cat. No. F0934K, Takara, Kyoto, Japan) with a reaction volume of 25 μl. PCR products were purified using the VAHTSTM Universal DNA Library Prep Kit or TruSeq NanoDNA LT Library Prep Kit (Illumina) for end repairing, 5′ phosphorylation and dA-tailing in one step. PCR products were then ligated into an appropriate vector and used as template for in vitro transcription using T7 polymerase. Shaded PCR products were purified using MEGAClear Kit (Ambion, AM1908) and recovered by alcohol precipitation.

NGS analyses of PCR amplicons. PCR was performed using the NGS primers (Supplementary Table 4). PCR products were purified using Gel Extraction Kit (Cat. No. D2500-02, OMEGA) before construction of NGS libraries. HiSeq3000 SBS6Cluster high-throughput NGS library preparation kit (Cat. No. FC-410-1002, Illumina, San Diego, CA, USA), (VAHTS Universal DNA Library Prep Kit (ND608-01, Vazyme) or TruSeq NanoDNA LT Library Prep Kit (Illumina) were used to prepare single-indexed amplified libraries. Briefer, more than 50 ng purified PCR fragment was used for direct library preparation. The fragments were treated with End Prep Enzyme Mix (Cat. No. WE0229, Illumina) for end repairaging, 5′ phosphorylation and A-tailing in one step reaction, followed by a T-A ligation to add adaptors to both ends. Size selection of adaptor-ligated DNA was then performed using VAHTST SM DNA Clean beads (Cat. No. N411-03, Vazyme) or Beckman AMPPure XP beads (Cat. No. A63882, Illumina). Each sample was then amplified by PCR for 8 cycles using P5 and P7 primers (Supplementary Table 4). Both P5 and P7 primers carry the sequences that can anneal with flow cells for bridge PCR. In addition, P7 primer carries a six-base complement of P5 primer. The PCR products were cleaned using VAHTST SM DNA Clean beads (Cat. No. N411-03, Vazyme) or Beckman AMPPure XP beads (Cat. No. A63882, Illumina), validated using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and quantified by Qubit2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) or Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific). Two or three μl of PCR products were mixed with 7 μl of Buffer P505-d1, Vazyme or Personalbio (Shanghai, China) or Quant-iT PicoGreen dsDNA Assay Kit (ThermoFisher Scientific) and then used for PCR amplification. The PCR products were then purified using P505-d1, Vazyme or TruSeq NanoDNA LT Library Prep Kit (Illumina) and used for subsequent NGS analyses.
Performed using GraphPad Prism 8 with two-tailed Student's t-test. 3. Marrafin, E. et al. CRISPR-Cas9-mediated editing in methylated regions with a human engineered Cas9 ribonucleoprotein. *Nat. Biotechnol.* **31**, 827–832 (2013).

**Statistical analyses.** For cell-based assay, three biological replicates are generally performed and the results are shown as mean ± standard error of the mean (SEM). For NGS analyses, two or three biological replicates were included for each experimental condition. In practice, three biological replicates were performed at first place and samples sent for NGS analysis. If two replicates could be successfully obtained then the data will be collected. Otherwise, the whole experiments will be repeated to keep consistency between parallel groups. The exact number of samples are reflected in each figure by the number of data points. Statistical analyses were performed using GraphPad Prism 8 with two-tailed Student’s t-test unless otherwise noted.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** NGS data have been deposited in the NCBI Sequence Read Archive database under the accession code SRP312256 (cell lines; https://www.ncbi.nlm.nih.gov/era?term=SRP312256) and PRJNA798574 (embryos; https://www.ncbi.nlm.nih.gov/era?term=PRJNA798574). Source data are provided with this paper.

**Received:** 5 February 2020; **Accepted:** 11 March 2022; **Published online:** 29 March 2022

**References.**

1. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

2. Brouns, S. J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).

3. Marrafin, E. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).

4. Pickard-Oliver, A. & Gersbach, C. A. The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).

5. Jinek, M. et al. RNA-programmed genome editing in human cells. *Elife* **2**, e0471 (2013).

6. Mali, P. et al. RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).

7. Cong, L. et al. Multiplex genome engineering using CRISPR-Cas systems. *Science* **339**, 819–823 (2013).

8. Cho, S. W., Kim, S., Kim, J. M. & Kim, S. J. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **31**, 230–232 (2013).

9. Liang, P. et al. Correction of beta-thalassemia mutant by base editor in human cells and heterozygous embryos. *Cell* **177**, 1067–1079 (2019).

10. Shin, J. et al. Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* **3**, e1701620 (2017).

11. Zuo, E. et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* **364**, 289–292 (2019).

12. Kluesner, M. G. et al. EditR: A method to quantify base editing from Sanger-derived peptides. *Genome Biol.* **21**, 51 (2020).

13. Zeng, Y. et al. Correction of the Marfan syndrome pathogenic FBN1 mutation by base editing in human cells and heterozygous embryos. *Mol. Ther.* **26**, 2631–2637 (2018).

14. Kollar, L. et al. Improving cytidine and adenosine base editors by expression optimization and ancestral reconstruction. *Nat. Biotechnol.* **36**, 843–848 (2018).

15. Kim, K. et al. Highly efficient RNA-guided base editing in mouse embryos. *Nat. Biotechnol.* **35**, 435–437 (2017).

16. Koo, T., Lee, J. & Kim, J. S. Measuring and reducing off-target activities of programmable nucleases including CRISPR-Cas9. *Mol. Cells* **38**, 475–481 (2019).

17. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).

18. Li, C. et al. HDAD5/35(→+) adenosine vector expressing anti-CRISPR peptides decreases CRISPR/Cas9 toxicity in human hematopoietic stem cells. *Mol. Ther. Methods Clin. Dev.* **9**, 390–401 (2018).

19. Patterson, Y. et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).

20. Fu, Y., Sanders, J. D., Reyon, D., Cасio, V. M. & Joung, J. K. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* **32**, 279–284 (2014).

21. Fallmann, C., Gowen, B. G., Lin, P. C., Doudna, J. A. & Corn, J. E. Cornerstones of CRISPR-Cas in drug discovery and therapy. *Nat. Rev. Drug Discov.* **16**, 89–100 (2017).

22. Zhang, X. H. et al. Dual base editor catalyzes both cytosine and adenine base conversions in human cells. *Nat. Biotechnol.* **38**, 856–U10 (2020).

23. Alzanalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–149 (2019).

24. Kluersner, M. G. et al. EditR: A method to quantify base editing from Sanger sequencing. *CRISPR J.* **1**, 239–250 (2018).

**Acknowledgements.** We thank the High-Throughput Screening Platform and Biomedical Big Data Platform at Shanghai Institute for Advanced Immunochemical Studies (SIAIS) at ShanghaiTech University for the support of flow cytometry experiments and analyses of NGS data. This work is supported by the National Natural Science Foundation of China (31600686 to J.L.) and ShanghaiTech University Startup Fund (20190301-000-01 to J.L.)

**Author contributions.** J.L. conceptualized the study. K.J. and Y.-R.C. designed the experiments, S.H. analyzed next-generation sequencing data. K.J. and Y.-R.C. performed in vivo CBE-inhibiting experiments. Z.L. and P.M. provided critical resources. J.L. and K.J., and Y.-R.C. wrote the manuscript. All authors discussed the results and approved the manuscript.
