Molecular Therapy
Nucleic Acids
Original Article

Engineering of near-PAMless adenine base editor with enhanced editing activity and reduced off-target

Xiaofang Cao,[1,2,3,7] Junfan Guo,[4,5,7] Shisheng Huang,[4,5] Wenxia Yu,[4,5] Guanglei Li,[4] Lisha An,[1,3] Xiangyang Li,[4,5] Wanyu Tao,[4,5] Qing Liu,[6] Xingxu Huang,[4] Xiaohua Jin,[1,2,3] and Xu Ma[1,2,3]

[1]National Research Institute for Family Planning, Beijing 100081, China; [2]Graduate School, Peking Union Medical College, Chinese Academy of Medical Sciences, Beijing 100005, China; [3]National Human Genetic Resources Center, Beijing 102206, China; [4]School of Life Science and Technology, ShanghaiTech University, Shanghai 201210, China; [5]University of Chinese Academy of Sciences, Beijing 100049, China; [6]Department of Medical Genetics, School of Basic Medical Sciences, Southern Medical University, Guangzhou 510515, China

About 47% of pathogenic point mutations could be corrected by ABE-induced A·T-to-G·C conversions. However, the applications of ABEs are still hindered by undesired editing efficiency, limited editing scopes, and off-targeting effects. Here, we develop a new adenine base editor, by embedding TadA-8e monomer into SpRY-nCas9, named as CE-8e-SpRY, which exhibits higher activity at NRN than NYN PAMs favored by SpRY nuclease. CE-8e-SpRY could target nearly all genomic sites in principle and induces the highest targeting efficiency among tested SpRY-based ABEs. In addition, CE-8e-SpRY also shows reduced RNA and DNA off-targeting activities. With optimized sgRNAs, CE-8e-SpRY induces efficient or desired target editing at some disease-relevant loci where conventional ABEs were unable to induce precise and satisfied editing. Taken together, our CE-8e-SpRY could broaden the applicability of ABEs in correcting or introducing pathogenic point mutations.

INTRODUCTION

Spontaneous cytosine deamination in living organisms, if not corrected, will result in C·G to T·A conversion, which contributes to about 47% of pathogenic point mutations for human genetic diseases in the ClinVar database.[1–4] Adenine base editors (ABEs), containing the adenosine deaminase fused to N terminus of D10A nickase Cas9 (nCas9), mediate single base transition of A-to-G or T-to-C at a guide RNA-programmed target site.[1] ABEs perform base editing efficiently without breaking double-strand DNA, and with high product purity, demonstrating their great potential in installing or correcting pathogenic mutations.[1,5,6] However, the target sites edited by ABEs, like by other CRISPR-Cas-derived editors, need a properly positioned PAM. For example, NGG PAM, which is recognized by the widely used adenine base editor, ABEmax, occurs only in 1 of 16 genome sequences.[7] To access more target sequences, more ABEs targeting diverse PAMs are developed either by using other Cas nucleases, like SaCas9, or by engineering Cas9 to target non-canonical PAMs, like NG PAM.[7–9] SpRY-Cas9, developed by structure-guided mutagenesis, is capable of targeting NRN (where R is A or G) and NYN (where Y is C or T) PAMs, and NRN is recognized more efficiently than NYN.[10] However, activity assessments of ABEmax-SpRY were performed at 120 h after transfection, much longer than 72 h for ABEmax-NG and CBE-SpRY, suggesting possible undesirable editing efficiency.[8,10] Moreover, off-target effects, including transcriptome-wide single guide RNA (sgRNA)-independent off-target RNA editing, and DNA off-target induced by SpRY-Cas9 nuclease, should be noticed for ABEmax-SpRY.[10,11]

Previous studies have demonstrated ABEs engineering with other Cas homologs exhibit substantially lower or even no virtual activities, compared with the corresponding cytosine base editors (CBEs), while TadA-8e, a new TadA monomer generated by phage-assisted evolution, offers large improvements in base editing even when coupled to non-SpCas9 Cas effectors.[12–14] However, ABEs containing TadA-8e also display higher off-target editing levels.[12] To develop the more efficient ABE that could break through the limitation of PAM, we developed the 8e-SpRY editor by fusing the TadA-8e monomer (hereafter referred to as TadA-8e) with SpRY-nCas9. Furthermore, to reduce the off-target induced by 8e-SpRY, we embedded TadA-8e into the tolerant region of SpRY-Cas9 to develop CE-8e-SpRY, a new SpRY-derived ABE with enhanced on-target editing activity and reduced off-target effects. We also used CE-8e-SpRY to achieve efficient editing at the *PAH* G728A site, a hotspot mutation in Chinese phenylketonuria (PKU) patients. CE-8e-SpRY could broaden the applicability of ABEs in correcting pathogenic point mutations.

Received 8 October 2021; accepted 29 April 2022;
https://doi.org/10.1016/j.omtn.2022.04.032.

[7]These authors contributed equally

Correspondence: Xiaohua Jin, National Research Institute for Family Planning, Beijing 100081, China.
E-mail: jinxiaohua@nrifp.org.cn
Correspondence: Xu Ma, National Research Institute for Family Planning, Beijing 100081, China.
E-mail: maxu_fg@nrifp.org.cn

RESULTS

8e-SpRY shows higher editing efficiency

8e-SpRY was constructed by substituting the heterodimeric deaminase moiety of ABEmax-SpRY with TadA-8e (Figure 1A). Editing efficiency was compared between ABEmax-SpRY and 8e-SpRY across 47 sites (12 sites harboring NAN, 11 sites harboring NGN, 12 sites harboring NCN and 12 sites harboring NTN) in HEK293T cells. Results showed that 8e-SpRY markedly increased editing efficiency at all tested sites (Figures 1B, 1C and S1A–S1D). Specifically, 8e-SpRY offered nearly four times higher average editing rate than ABEmax-SpRY (Figure 1D, 41.57% versus 10.34%, $p < 0.0001$). Then we analyzed the editable positions, and found the editing window of 8e-SpRY spanned from position 3 to 10, wider than that of ABEmax-SpRY (counting PAM distal position in target site as position 1, Figures 1E and S3A–S3B). These findings demonstrated 8e-SpRY was an efficient adenine base editing tool targeting almost all PAMs.

Engineering 8e-SpRY variants mediate robust base conversion

ABE8e was reported to induce increased DNA and RNA off-target editing compared with ABEmax;[12] SpRY-Cas9 induced increased off-target events due to expanded PAM recognition compared with NGG-Cas9.[10] To overcome the off-target issues of base editors, two strategies, including mutating deaminase domains and embedding the deaminase into nCas9,[12,15–18] were usually employed. Specifically, V106W of TadA-8e domain, a mutation reported to reduce RNA and DNA off-target for ABE8e, was introduced into 8e-SpRY to generate V106W-SpRY.[12] Our previous research showed that insertion of deaminase into the 1048T-1063I region of nCas9 resulted in dramatically decreased off-target and had a minimal impact on on-target editing,[19] so we developed CE-8e-SpRY by embedding the TadA-8e into this tolerant site of SpRY-nCas9. Moreover, four substitutions of SpCas9 high-fidelity variant number 1 (SpCas9-HF1) could eliminate DNA off-target events generated by relaxed PAM tolerances of SpRY-Cas9.[10] Thus, 8e-SpRY-HF and V106W-SpRY-HF were engineered by introducing HF1 substitutions (N497A, R661A, Q695A, and Q926A) to 8e-SpRY and V106W-SpRY, respectively (Figure 2A). Totally, four new 8e-SpRY variants were developed.

The editing activities of these variants at endogenous sites were evaluated. V106W-SpRY and CE-8e-SpRY achieved comparable editing rates to 8e-SpRY, while the other two 8e-SpRY variants, 8e-SpRY-HF and V106W-SpRY-HF, exhibited decreased editing activities at almost all tested sites (Figures 2B, 2C, and S2). Specifically, there were significantly decreased editing rates for 8e-SpRY-HF and V106W-SpRY-HF, compared with 8e-SpRY ($p = 0.0015$ for 8e-SpRY-HF, $p < 0.0001$ for V106W-SpRY-HF, Figure 2D), while both CE-8e-SpRY and V106W-SpRY had comparable efficiency ($p = 0.3056$ for CE-8e-SpRY, $p = 0.3899$ for V106W-SpRY, Figure 2D). PAM preferences were further analyzed. Consistent with previous reports, all three SpRY-based ABEs (8e-SpRY, CE-8e-SpRY, V106W-SpRY) were more effective at NRN target sites than NYN (Figure 2E). CE-8e-SpRY and V106W-SpRY remained the same editing window as 8e-SpRY, but CE-8e-SpRY induced higher editing activities at positions 8–10 (Figures 2F and S3B–S3D). Our results suggested all of the engineering variants showed robust base conversion.

Besides, ABE8 constructs (ABE8s), generated by evolving TadA enzyme contained in ABE7.10, offer increased editing activities than ABE7.10.[20] Given that Cas-embedding strategy could maintain comparable on-target efficiency, we also developed CE-8.17m-SpRY and CE-8.20m-SpRY by inserting the deaminases of ABE8.17m and ABE8.20m, two constructs exhibiting high editing efficiencies when paired with diverse Cas9 nucleases, into SpRY-nCas9.[20] Editing activities of CE-8e-SpRY, CE-8.17m-SpRY and CE-8.20m-SpRY were compared across 14 endogenous sites covering NAN (three sites), NTN (four sites), NCN (four sites), and NGN (three sites) PAMs. The results showed CE-8.17m-SpRY and CE-8.20m-SpRY exhibited significantly decreased editing activities, in comparison with CE-8e-SpRY ($p = 0.0002$ for CE-8.17m-SpRY, $p = 0.0038$ for CE-8.20m-SpRY, Figures S4A and S4B). A-to-non-G edits and indels generated by CE-8e-SpRY and CE-8s-SpRYs were also analyzed. The results showed that CE-8s-SpRYs induced less A-to-non-G edits and indels than CE-8e-SpRY, but without significant differences (Figures S4C and S4D). Therefore, 8e-SpRY variants were selected for subsequent off-target assessments.

CE-8e-SpRY shows reduced RNA off-targets

Next, the RNA off-targets for SpRY-based ABEs were evaluated. DNA on-target results showed all tested editors generated efficiently editing at sgRNA targeting site, while ABEmax-SpRY exhibited relatively weak editing activity (Figures 3A and S5A). Previous studies revealed ABEmax induced tens of thousands of A-to-I (inosine) RNA edits in human cells.[11,15] Similar result was seen for ABEmax-SpRY (Figure 3B). Interestingly, compared with ABEmax-SpRY, higher numbers of RNA total mutations while lower numbers of highly edited mutations (>40%), were observed for 8e-SpRY (Figures 3B–3D). However, CE-8e-SpRY and V106W-SpRY, respectively, generated 2,097 and 9,184 RNA edits, much lower than 22,617 mutations of 8e-SpRY (Figures 3B and 3C). Specifically, RNA edits induced by CE-8e-SpRY were 10.8 times lower than 8e-SpRY, 8.7 times lower than ABEmax-SpRY, and 4.4 times lower than V106W-SpRY (Figure 3C). To further analyze the sequence preferences of RNA edits, WebLogo 3.6.0 software was used. As expected, TadA/TadA* heterodimer for ABEmax-SpRY highly edited tRNA-like CTACGAA sequence at RNA levels, consistent with reported study (Figure S5B),[15] while 8e-SpRY preferred a short TAC motif regardless of RNA editing rates, which was also observed for CE-8e-SpRY and V106W-SpRY (Figure 3E). This phenomenon may suggest that these two variants might reduce RNA off-target by reducing the affinity for RNA instead of altering the preferred sequences.

CE-8e-SpRY also decreased DNA off-targets

For sgRNA-dependent off-target DNA editing, the potential off-target sites were predicted by Cas-OFFinder based on base-pair mismatch.[21] Four sgRNAs with the highest on-target activities at NA/T/C/GN PAMs were chosen as on-target sequences. For each

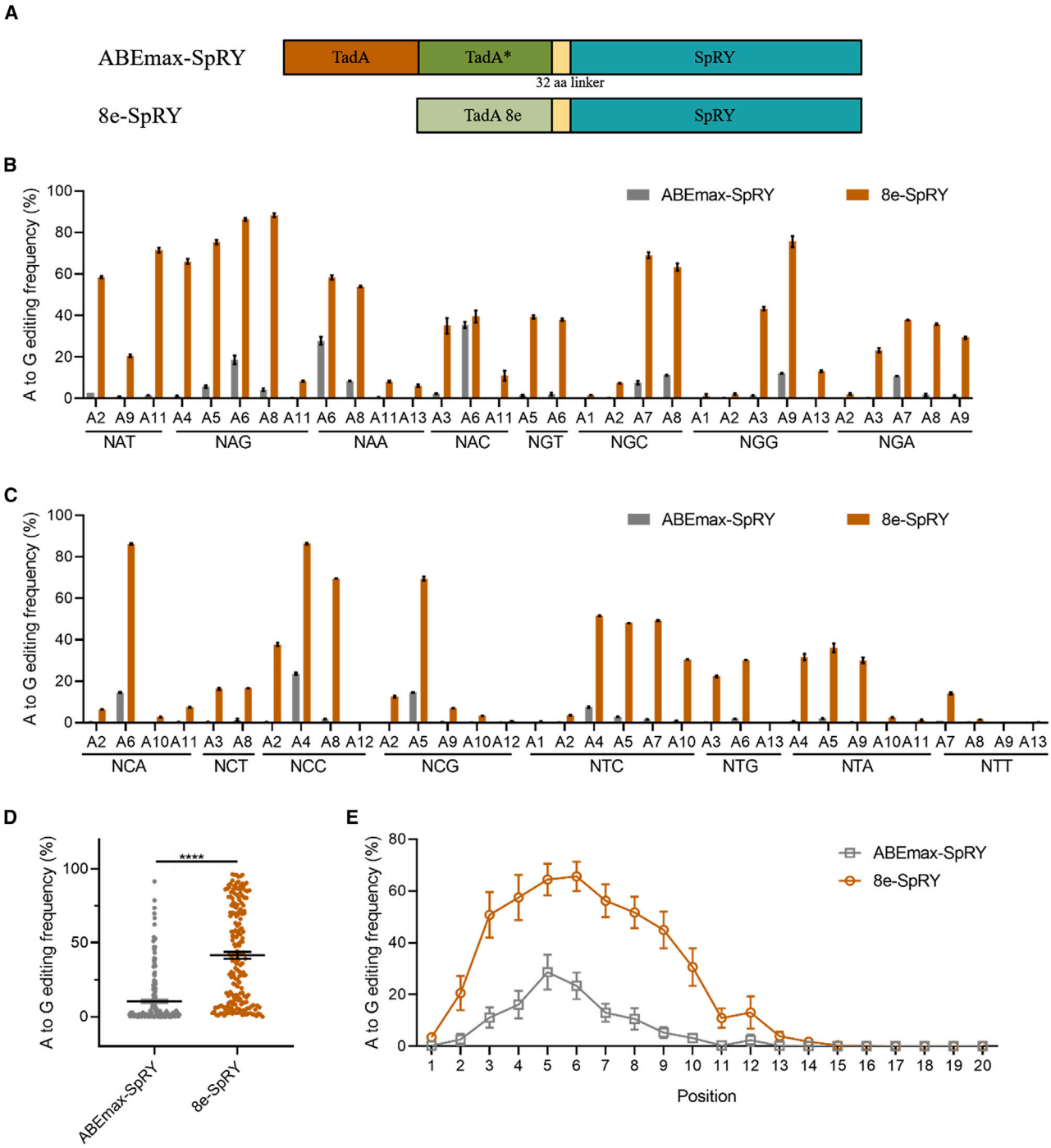
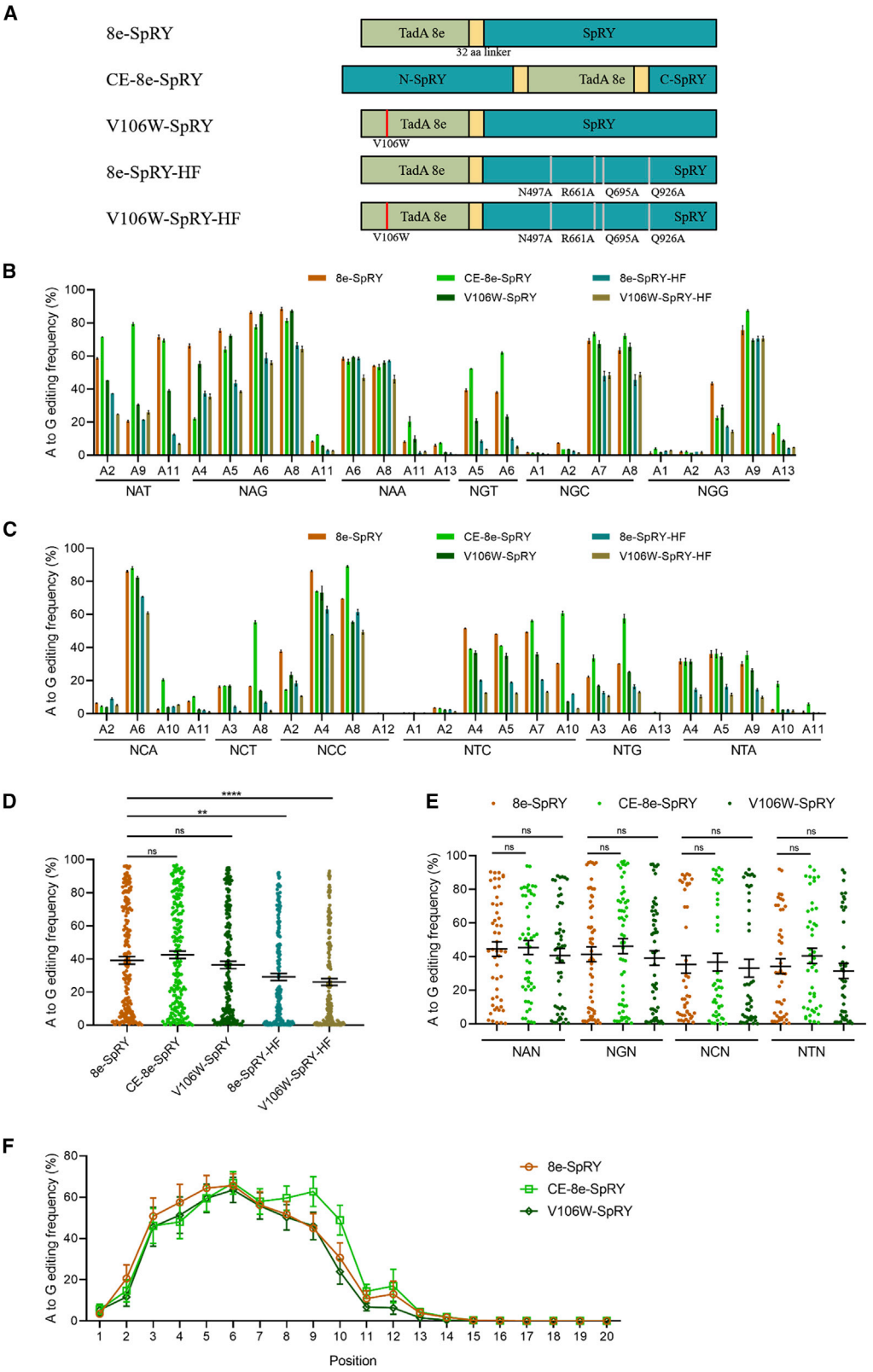


Figure 1. 8e-SpRY can induce A-to-G conversions with higher efficiency

(A) A schematic representation of ABEmax-SpRY and 8e-SpRY. (B) A-to-G base editing of endogenous sites bearing NRN with ABEmax-SpRY and 8e-SpRY in HEK293T cells. (C) A-to-G base editing of endogenous sites bearing NYN with ABEmax-SpRY and 8e-SpRY in HEK293T cells. For (B) and (C), editing efficiency was shown as mean ± SD, n = 3 independent experiments. (D) Editing efficiencies across target sites with NR/YN PAMs in HEK293T cells, summarized from (B) and (C), and Figure S1; $p < 0.0001$ for comparison of ABEmax-SpRY and 8e-SpRY. Editing efficiencies are shown as mean ± SEM. (E) The comparison of editing windows between ABEmax-SpRY and 8e-SpRY. Edited adenines were counted by setting the base distal to PAM as position 1; each data point is shown as mean ± SEM.



(legend on next page)

target site, 10 predicted off-target sites in the genome with the least mismatches and with NA/GN PAMs were selected preferentially. The sequencing was performed at an average depth of 736858 × in our study. Assessment results showed that all tested 8e-SpRY variants except ABEmax-SpRY induced robust on-target editing (Figures 4A–4D). There were no detectable editing activities at all OT sites except 1-OT3, 2-OT4, and 3-OT7 (Figures 4A–4D and S6). Compared with 8e-SpRY, CE-8e-SpRY substantially decreased OT activities at 2-OT4 and 3-OT7 sites except at the protospacer position 10 of 3-OT7 and 1-OT3 (Figures 4A–4D). Besides, introducing V106W and HF1 mutations to 8e-SpRY also improved DNA off-target editing (Figure S7). Moreover, in comparison with ABEmax-SpRY, 8e-SpRY variants induced slightly increased A-to-non G editing and higher indels (Figure S8).

Meanwhile, sgRNA-independent off-target DNA editing was assessed by orthogonal R-loop assay, which detected ABE-induced mutations at single-stranded DNA (ssDNA) regions created by nSaCas9/sgRNA complex. Deep sequencing results indicated that, at three nSaCas9 R-loop sites, all three tested SpRY-ABEs displayed detectable off-target activities at varying editing frequencies (Figure S9B). Furthermore, both CE-8e-SpRY and V106W-SpRY induced decreased off-target activities relative to 8e-SpRY, while the lowest off-target editing frequencies were observed for CE-8e-SpRY (Figure S9B). Taken together, CE-8e-SpRY was selected as the best SpRY-derived ABE with enhanced on-target editing activity and reduced off-target effects.

Application of CE-8e-SpRY in correcting disease-relevant loci

Next, we further explored the potential of CE-8e-SpRY to correct disease-associated mutations in human cells. We first established three mutant cell lines harboring three different pathogenic mutations. G728A substitution in *PAH*, the most prevalent mutation among Chinese patients with phenylketonuria (PKU),[21] was introduced to HEK293T cells (HEK293T^{G728A}); C494T mutation in *MYO7A*, resulting in rare genetic deafness, was introduced to U2OS cells (U2OS^{C494T}); and C1331T in *RAG1*, a pathogenic mutation of severe combined immunodeficiency (SCID), was introduced to HeLa cells (HeLa^{C1331T}). Correction sgRNAs with different PAMs, locating respectively the target base at protospacer position 3–10 were designed.

In HEK293T^{G728A}, using CE-8e-SpRY, three of eight correction sgRNAs successfully generated detectable editing for the pathogenic mutation, among which, sg1 mediated the highest A-to-G conversion with the efficiency of 51.60%, other two sgRNAs (sg2 and sg3) generated average 7.02% and 8.82% editing, respectively (Figures 5B and S10). Notably, the target mutation for sg1 was at protospacer position 10, where was usually inaccessible for canonical ABEs. Besides, base substitution at the undesired site (c.729A) was also observed at a frequency of 19.06%. Fortunately, the undesired base editing resulted in synonymous mutation, producing intact PAH protein. We also assessed the ability of ABEmax-NG, x-ABEmax, and ABEmax-SpRY in correcting G728A mutation in HEK293T^{G728A}, no corrections were observed for the three ABEs (Figure S10).

In U2OS^{C494T}, six of eight designed sgRNAs gave detectable correction using CE-8e-SpRY, at efficiency of 11% to 63% (Figures 5C and S11), compared with ABEmax-NG at efficiency of 33% (Figures 5C and S11). In HeLa^{C1331T}, CE-8e-SpRY induced >40% editing efficiency with six sgRNAs at the target base, of which, sg6 generated precise correction without bystander editing (Figures 5D and S12). sg8 corrected the mutation at a higher efficiency than sg6, accompanied with a synonymous substitution. Meanwhile, ABEmax-NG with sg3, and ABEmax with sg4 also exhibited efficient target base editing (56% and 41%, respectively, Figures 5D and S12).

Application of CE-8e-SpRY in introducing disease-relevant loci

We also applied CE-8e-SpRY to introduce mutations associated with Marfan syndrome (T3202C in *FBN1*) and glycogen storage disease (T1214C in *GAA*) in U2OS cells, long QT syndrome (A944G in *KCNQ1*), and Charcot-Marie-Tooth disease (A494G in *MFN2*) in HeLa cells, and hereditary persistence of fetal hemoglobin (−198T > C at *HBG* promoter) in HEK293T cells. In each case, eight sgRNAs locating the intended edit in protospacer position 3–10 were tested for CE-8e-SpRY. At site T3202C in *FBN1*, where ABEmax-NG failed to generate detectable editing, but CE-8e-SpRY achieved the highest editing efficiency of 44% with sg5 (Figures 6 and S13). At site T1214C in *GAA*, deleterious bystander editing could be avoided by selecting the optimal sgRNAs with comparable efficiency to ABEmax-NG (Figures 6 and S14). At sites A944G in *KCNQ1* and A494G in *MFN2*, the best sgRNA performed simultaneous target editing without collateral edits at much higher efficiency than canonical ABEs (Figures 6, S16 and S17). For the site −198T > C at *HBG* promoter, CE-8e-SpRY mediated higher intended editing, but also higher bystander editing, in comparison with canonical ABEs (Figures 6 and S15). Taken together, these results suggest CE-8e-SpRY outperformed conventional ABEs in generating efficient or desired target editing through sgRNA optimization at disease-relevant loci.

Figure 2. On-target editing characterizations of 8e-SpRYs in HEK293T cells

(A) A schematic representation of 8e-SpRY variants. (B) A-to-G base editing of endogenous sites bearing NRN with 8e-SpRYs in HEK293T cells. (C) A-to-G base editing of endogenous sites bearing NYN with 8e-SpRYs in HEK293T cells. For (B) and (C), editing efficiency is shown as mean ± SD, n = 3 independent experiments. (D) Editing efficiencies across target sites with NR/YN PAMs in HEK293T cells. p = 0.3056 for comparison of 8e-SpRY with CE-8e-SpRY; p = 0.3899 for comparison of 8e-SpRY with V106W-SpRY; p = 0.0015 for comparison of 8e-SpRY with 8e-SpRY-HF; p < 0.0001 for comparison of 8e-SpRY with V106W-SpRY-HF. (E) Editing efficiencies across target sites with NAN, NGN, NCN, and NTN, respectively. For (D) and (E), data are summarized from (B) and (C), and Figure S2. Each data point represents the mean of the three replicates; editing efficiencies are shown as mean ± SEM. (F) Editing windows of 8e-SpRY variants. Edited adenines were counted with setting the base distal to PAM as position 1; each data point is shown as mean ± SEM. Editing data of 8e-SpRY are from Figures 1B, 1C, and S1.

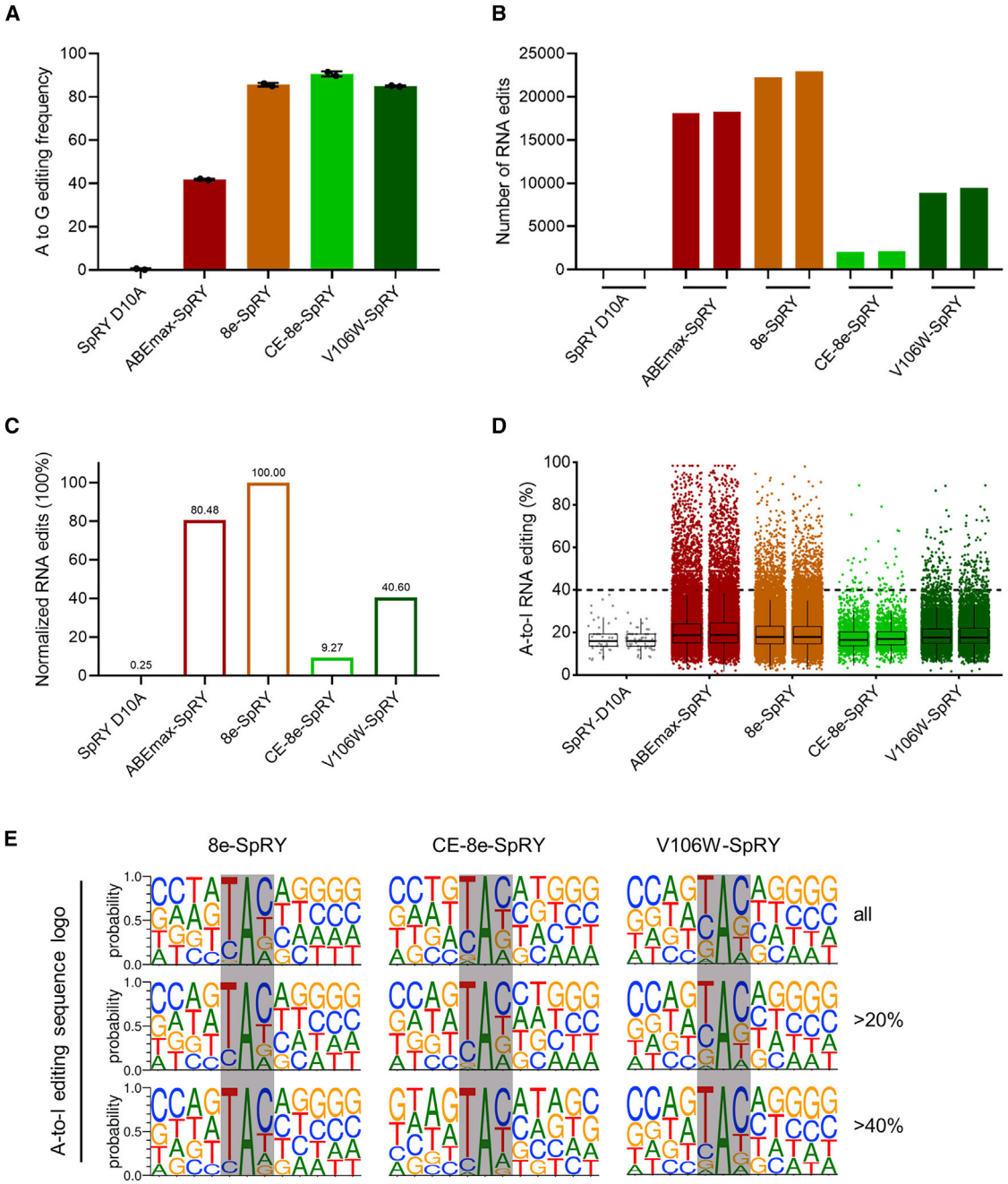


Figure 3. RNA off-target assessments of SpRY-derived ABEs

(A) On-target DNA base editing efficiencies of SpRY-derived ABEs. Editing efficiencies at position 8 of 3-NGC site were analyzed in HEK293T cells used to detect RNA off-target effects. (B) The numbers of RNA A-to-I edits induced by SpRY-derived ABEs. (C) Normalized RNA edit numbers of SpRY-derived ABEs. The number of RNA edits generated by 8e-SpRY was set as 100%. Data are summarized from (B). (D) The efficiency distributions of RNA A-to-I edits induced by SpRY-derived ABEs. Each dot represents an edited adenine in RNA. For (B) and (D), two independent replicates are shown. (E) Sequence logos centered by the edited adenine with indicated editing efficiencies (>40%, >20%, and all) detected by RNA sequencing. Preferred sequences are highlighted in gray. Sequence logos of 8e-SpRY are from Figure S5B.

DISCUSSION

CRISPR-derived base editors perform efficient genetic modifications at specified DNA sites without generating DSBs and excess undesired byproducts. In general, three major classes of base editors are described: CBEs, which convert C·G base pairs to T·A base pairs;[22] ABEs, which catalyze A·T-to-G·C conversions;[1] and CGBEs

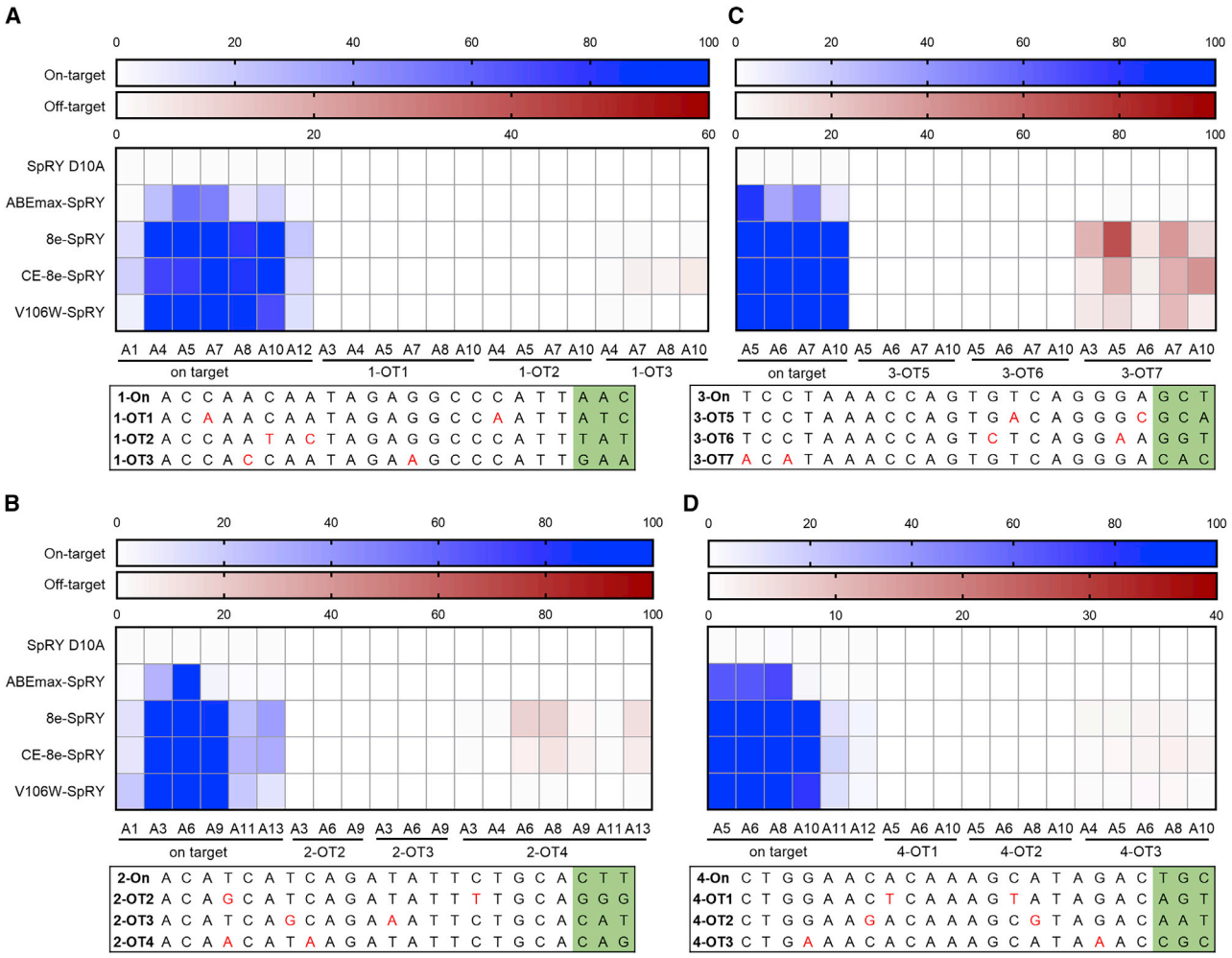


Figure 4. DNA off-target detection of 8e-SpRY variants at predicted sites

The heat maps showed DNA on-target (blue) and off-target (red) editing efficiencies (mean ± SD, n = 3 independent experiments) at 3-NAC (A), 3-NTT (B), 3-NCT (C), and 3-NGC (D), respectively. Off-target sites were predicted by Cas-OFFinder. The table below each figure gives the sequence information of three predicted off-target sites and their corresponding on-target site. Mismatched bases are shown in red, and PAMs are highlighted in green.

(C·G-to-G·C base editors), which mediate C·G-to-G·C base editing.[23] Besides, PEs (prime editors) enable introduction of all 12 types of point mutations in addition to small insertions and deletions.[24] To date, engineered editing tools have made many types of pathogenic mutations editable. Half of known disease-associated gene variants are point mutations, and nearly half of pathogenic point mutations in principle could be corrected by ABEs or PEs induced A·T-to-G·C conversions,[3,4] making these two tools especially useful in correcting the pathogenic genetic variants. However, prime editing applications are still limited by undesired editing efficiency, which is highly dependent on the design of prime editing guide RNAs, in contrast, ABEs could induce more efficient editing when the target nucleotide is present within the base editing window.[24] In this study, we generated CE-8e-SpRY, with expanded targeting scope and significant improvement in editing efficiency. ABEmax-SpRY only induced limited editing at protospacer position 5 to 6, while CE-8e-SpRY generated robust editing activity spanning position 3 to 10, making a wider range of nucleotides editable. Moreover, CE-8e-SpRY could provide many options for spacer sequences regardless of PAM restriction, allowing editing efficiency, DNA specificity, or other considerations to be optimized to satisfy the given application.

We also compared the off-target effects of the new editors. Introduction of rationally engineered point mutations was the commonly adopted method to minimize off-target activities.[25] Our study demonstrated, inserting TadA8e into the tolerate site of SpRY-Cas9 maintained comparable editing activities, making it a more recommended way to generate the ABE targeting NNN PAMs. Besides, Cheng and co-workers engineered ABE variants with diversified editing scopes and reduced RNA off-target activities by inserting

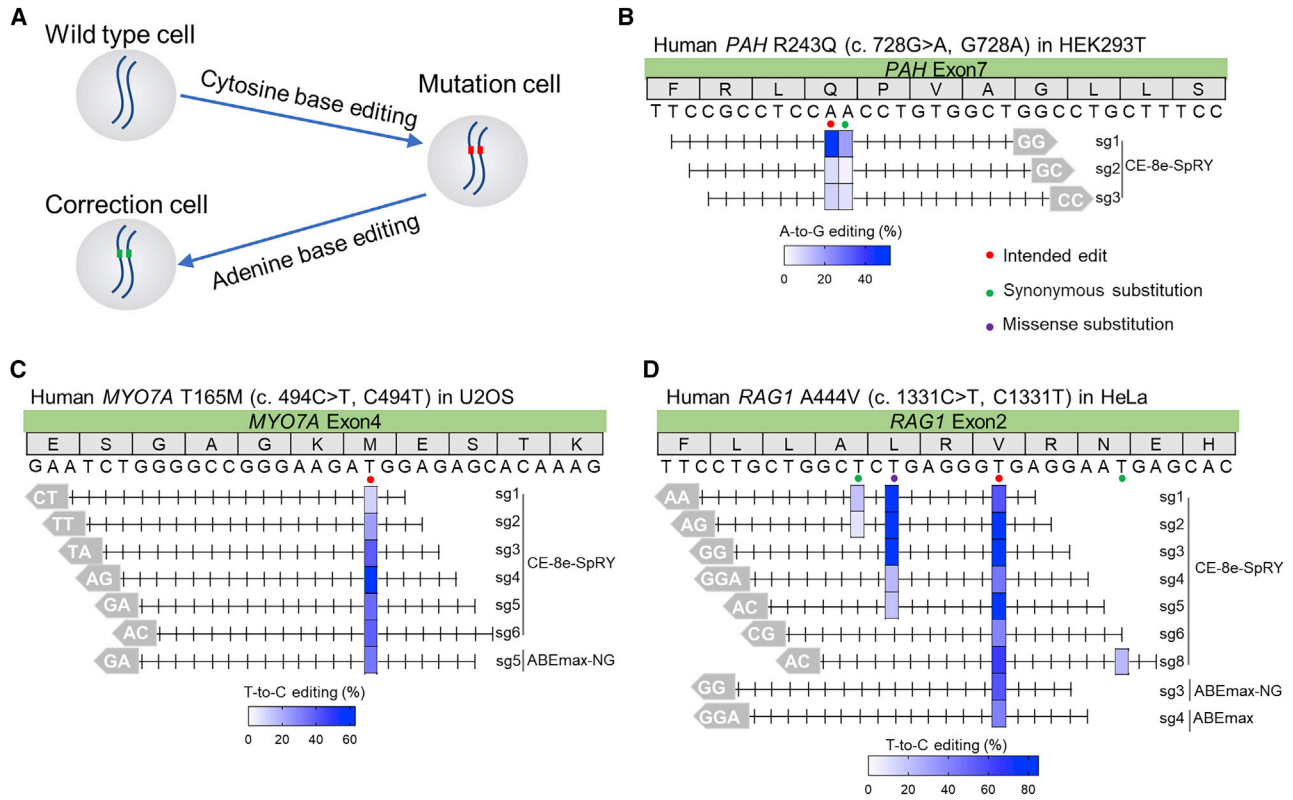


Figure 5. Correction of pathogenic mutation using CE-8e-SpRY in different human cells

(A) Schematic diagram of simulating and correcting the pathogenic mutation in different human cell lines. The disease cell model was established by a cytosine base editing system under the guide of mut-sgRNA, then was corrected by CE-8e-SpRY. (B–D) Schematic diagrams illustrate the design of sgRNA for correction. sgRNAs with detectable editing in Sanger chromatogram are displayed. Heatmaps show the editing efficiencies of bases (mean ± SD, n = 3 independent experiments). For (B), correction efficiency was detected by deep-seq; for (C) and (D), correction efficiency was detected by Sanger sequencing. Red circles indicate pathogenic mutations; green circles indicate bystander bases, which resulted in synonymous substitutions; purple circles indicate missense substitutions.

into different sites of SpCas9,[26] sharing similar editing traits to CE-8e-SpRY. But CE-8e-SpRY itself displayed shifted editing scopes with efficient editing activities (>40%) by recognizing NNN PAMs, and could make every genomic site targetable and editable in principle.

We also found CE-8e-SpRY outperformed other variants in reducing RNA off-target effects. Meanwhile, inserting TadA-8e into SpRY-nCas9 generated comparable unintended editing to introducing HF1 mutations, but higher off-target editing activities than introducing V106W mutation. Thus, combining CE-8e-SpRY with V106W mutation could be expected to achieve further improvements in RNA and DNA off-target effects. Alternatively, using CE-8e-SpRY, a sgRNA that is unique in the genome or displays the fewest DNA unwanted mutations could be selected to perform target editing.

It is notifying, 8e-SpRY variants, including CE-8e-SpRY, induced slightly increased A-to-non-G editing and higher indels than ABEmax-SpRY; unsignificant increases in unwanted edits were also observed for CE-8e-SpRY, compared with CE-8s-SpRYs. Given that the frequencies of unwanted editing varied among target sites, sgRNA optimization might be useful in reducing the formation of byproducts. Alternately, CE-8s-SpRYs could offer another choice for target editing though at a lower frequency than CE-8e-SpRY when indels and non-G edits are special concerns. For therapeutic applications, multiple edits could be avoided or tolerated by optimal sgRNA with CE-8e-SpRY, or selecting ABEs with a narrower activity window to appropriately position the target bases.

Moreover, though ABEs were reported to induce minimal sgRNA-independent DNA off-targets,[27,28] low but detectable off-target editing could be observed by employing orthogonal R-loop assay which used nickase SaCas9 (nSaCas9) to generate ssDNA regions.[16] A recent study also revealed R-loop assay with nSaCas9 detected higher off-target editing than with dSaCas9.[29] Therefore, R-loop assay using nSaCas9 may be more suitable for detecting sgRNA-independent DNA off-targets.

In summary, we developed a SpRY-derived ABE with improved efficiency and specificity by embedding TadA-8e into SpRY-nCas9, and

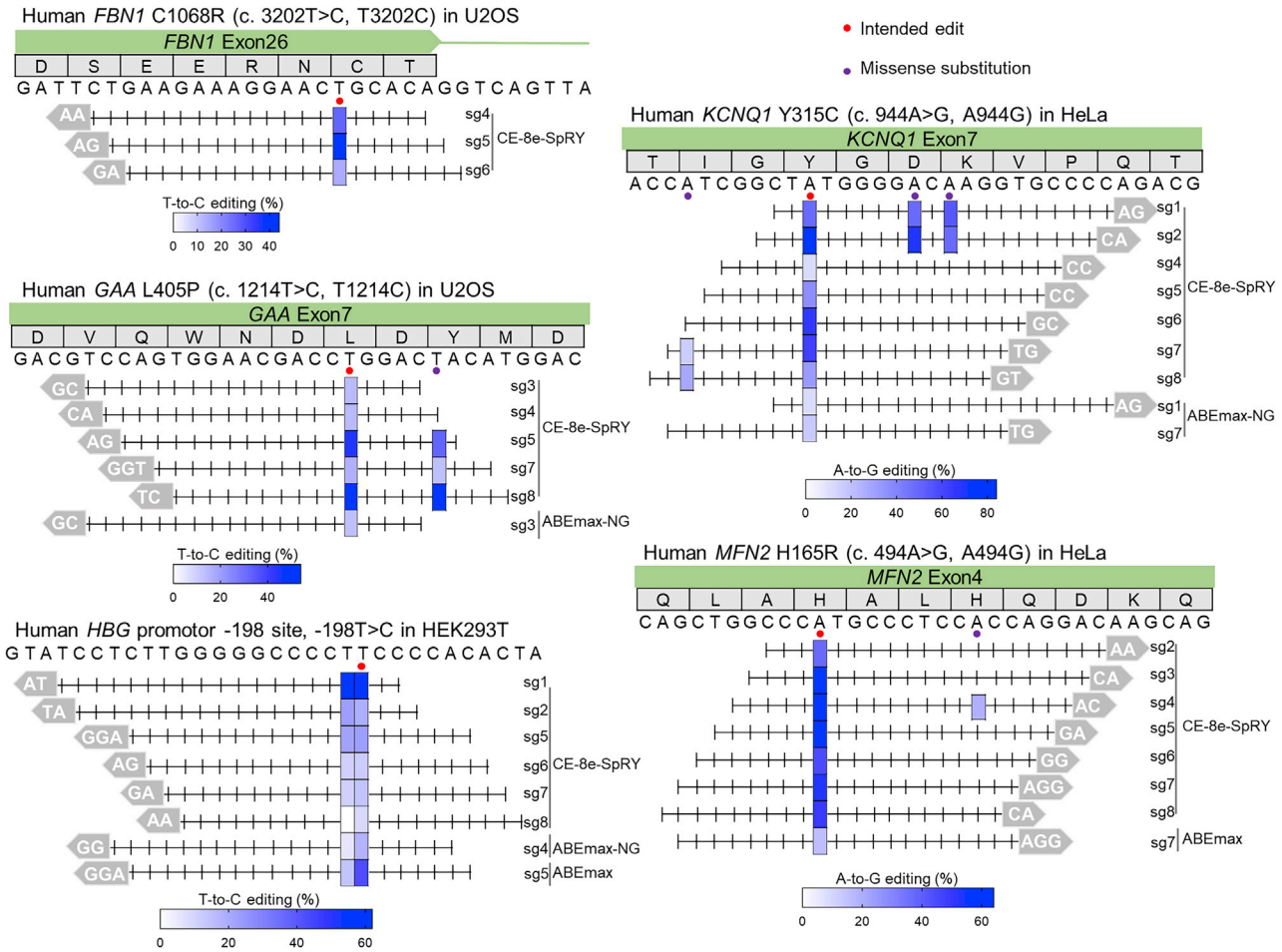


Figure 6. Installation of disease-relevant mutations using CE-8e-SpRY in different human cells

Schematic diagrams illustrate the design of sgRNAs for installing. sgRNAs with detectable editing in Sanger chromatogram are displayed. Editing efficiency was detected by Sanger sequencing and shown by Heatmaps (mean ± SD, n = 3 independent experiments). Red circles indicate pathogenic mutations, and purple circles indicate missense substitutions.

successfully applied it in correcting or introducing multiple pathogenic mutations.

MATERIALS AND METHODS

Plasmid construction

ABEmax-SpRY and 8e-SpRY variants were derived from ABEmax (Addgene 112095). Base editors involved in this study were constructed by introducing mutations or inserting amplified DNA product to ABEmax plasmid using ClonExpress II One Step Cloning Kit (Vazyme, C112-01). DNA sequences of new constructs are provided in Table S1. DNA oligos for constructing sgRNA expression vectors were synthesized, annealed, and cloned into BsaI-digested pGL3-U6-sgRNA-EGFP or pGL3-U6-sgRNA-mCherry. Detailed sgRNA sequences for detecting on-target editing activities and disease-associated sites editing are displayed in Tables S2 and S6, respectively.

Cell culture and transfection

HEK293T, U2OS, and HeLa cells were cultured in DMEM (Gibco) supplemented with 10% fetal bovine serum (v/v) (Gibco) and 1% Penicillin Streptomycin (Gibco). Cells were passaged once every 2 to 3 days and incubated at 37°C with 5% CO_2. For plasmid transfection in this study, unless otherwise specified, HEK293T cells were seeded on 24-well plates (JET-BIOFIL) 1 day before transfection and transfected at approximately 70% to 80% confluence using EZ *trans* (Shanghai Life iLab Biotech Co., Ltd) according to the manufacturer's protocols, then 2 days later, 15% GFP-positive cells were harvested by fluorescence activating cell sorter (FACS). To assess editing activities of ABEs, HEK293T cells were transfected with 600-ng editors and 300-ng corresponding sgRNAs. To evaluate RNA off-target effects, HEK293T cells were seeded on a 6-cm dish (JET-BIOFIL) and transfected with 4 μg editors and 2 μg sgRNAs; 10% GFP-positive cells were isolated. To detect sgRNA-dependent off-target DNA editing, 600 ng editors were co-transfected with 300 ng corresponding

sgRNAs and cells with top 15% GFP signal were collected 3 days later. Orthogonal R-loop assays, used to analyze sgRNA-independent off-target DNA editing, were performed as previously described.[30] In brief, 300 ng tested editors, 200 ng SpRY sgRNAs, 300 ng nSaCas9, and 200 ng SaCas9 sgRNAs were co-transfected into HEK293T cells. Fifteen percent GFP and mCherry double-positive cells were isolated 3 days later.

Detection of DNA on- and off-target editing

Sorted cells by FACS were lysed in the lysis buffer (50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris pH 8.0, 0.5% Nonidet P-40, 0.5% Tween 20, 100 μg/mL protease K) at 65°C for 30 min, then 98°C for 3 min. Fragments containing the target sites were amplified using Phanta Max Super-Fidelity DNA polymerase (Vazyme, P505-d1); cell lysates were used as the template. PCR products were analyzed by Sanger sequencing and deep sequencing on Illumina Nova Seq 6000 platform (2 × 150 PE). Sanger sequencing results were further processed by EditR (https://moriaritylab.shinyapps.io/editr_v10/).[31] Deep sequencing data were processed as previously described.[19] In brief, AdapterRemoval version 2.2.2 was used to remove the adapter pair of the paired-end reads; paired-end read alignments of 11 base pairs or more bases were combined into a single consensus read; then the BWA-MEM algorithm (BWA v0.7.16) was used to map the processed data to target sequences. Mutation rate was calculated using bam-readcount with parameters *-q* 20 *-b* 30. Indels were considered to occur in reads containing at least one inserted or deleted nucleotide in protospacer. Primer used for detecting DNA on-target editing are listed in Table S3. DNA off-target sites are displayed in Table S4. Primer used for detecting DNA off-target editing are shown in Table S5. Primers used for detecting disease-associated sites editing are provided in Table S7.

Detection of RNA off-target editing

For RNA off-target editing assessment, isolated HEK293T cells by FACS were immediately treated with TRIzol reagent (Vazyme, R401-01), and total RNA was extracted according to the manufacturer's instructions. RNA sequencing was performed on Illumina Nova Seq 6000 platform (2 × 150 PE) at a depth of ∼20 million reads per sample. RNA sequencing data were processed as previously described.[19] Briefly, STAR software (Version 2.5.1) was employed to map sequence reads to human reference genome (hg38); GENCODE version V30 was used to make annotation; after removing duplication, variants were called using GATK HaplotypeCaller (version 4.1.2) and filtered with QD (quality by depth) < 2; variants were verified and quantified by bam-readcount with parameters *-q* 20 *-b* 30. The given edits should have at least 10 × depth and at least 99% of reads supporting the reference allele in the wild-type samples. Finally, only A-to-G edits in transcribed strand were used for further analysis.

Modeling and correcting pathogenic mutations

For modeling pathogenic mutation in HEK293T, U2OS, and HeLa cells, 600 ng CBEs and 300 ng corresponding mut-sgRNAs were co-transfected into cells. Two days later, single GFP-positive cells were sorted into a 96-well plate (JET-BIOFIL) by FACS. After culturing 10 to 14 days, part cells of single colonies were used to identify genotype, and the other cells were continuously cultured. Genotype identification was performed by Sanger sequencing and the cell colonies with homozygote mutations were identified as disease cell lines. Then disease cell lines were seeded on 24-well plates and transfected as described above. Correction efficiency was also detected as described above. To introduce disease-relevant mutations in different cells, HEK293T, U2OS, and HeLa cells were seeded, transfected, and detected as described above.

Statistics

Unless otherwise noted, results from two or three independent experiments were presented as the mean ± SD. GraphPad Prism 8.0 was adopted to perform statistical analysis and graphing. Unpaired Student's *t* test (two-tailed) was used for comparisons and $p < 0.05$ was considered to be statistically significant.

DATA AVAILABILITY

High-throughput sequencing data in this study can be accessed in the NCBI Sequence Read Archive database (https://www.ncbi.nlm.nih.gov/sra/PRJNA762984) with accession code PRJNA762984.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2022.04.032.

ACKNOWLEDGMENTS

We thank members of the Huang lab for helpful discussions. We thank Prof. Yunbo Qiao from Guangzhou University for constructive suggestions. We thank the Discovery Technology Platform at Shanghai Institute for Advanced Immunochemical Studies (SIAIS) and the Molecular and Cell Biology Core Facility (MCBCF) at School of Life Science and Technology, ShanghaiTech University, for providing technical support. This work was supported by the National Key Research and Development Program of China (2016YFC1000307) and the Leading Talents of Guangdong Province Program (608285568031).

AUTHOR CONTRIBUTIONS

X.H., X.J., and X.M. conceived and supervised the project. X.C. and J.G. performed most experiments and wrote the paper with the assistance of other authors. G.L., X.H., X.J., and X.M. revised the manuscript. S.H. analyzed the high-throughput sequencing data. W.Y., L.A., X.L., W.T., and Q.L. helped to construct plasmids and sort positive cells by FACS.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature *551*, 464–471. https://doi.org/10.1038/nature24644.

2. Rees, H.A., and Liu, D.R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. Nat. Rev. Genet. *19*, 770–788. https://doi.org/10.1038/s41576-018-0059-1.

3. Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., et al. (2016). ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res. *44*, D862–D868. https://doi.org/10.1093/nar/gkv1222.

4. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. *42*, D980–D985. https://doi.org/10.1093/nar/gkt1113.

5. Ryu, S.M., Koo, T., Kim, K., Lim, K., Baek, G., Kim, S.T., Kim, H.S., Kim, D.E., Lee, H., Chung, E., and Kim, J.S. (2018). Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. Nat. Biotechnol. *36*, 536–539. https://doi.org/10.1038/nbt.4148.

6. Liu, Z., Chen, M., Chen, S., Deng, J., Song, Y., Lai, L., and Li, Z. (2018). Highly efficient RNA-guided base editing in rabbit. Nat. Commun. *9*, 2717. https://doi.org/10.1038/s41467-018-05232-2.

7. Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., and Liu, D.R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. Nature *556*, 57–63. https://doi.org/10.1038/nature26155.

8. Huang, S., Liao, Z., Li, X., Liu, Z., Li, G., Li, J., Lu, Z., Zhang, Y., Li, X., Ma, X., et al. (2019). Developing ABEmax-NG with precise targeting and expanded editing scope to model pathogenic splice site mutations in vivo. iScience *15*, 640–648. https://doi.org/10.1016/j.isci.2019.05.008.

9. Hua, K., Tao, X., and Zhu, J.K. (2019). Expanding the base editing scope in rice by using Cas9 variants. Plant Biotechnol. J. *17*, 499–504. https://doi.org/10.1111/pbi.12993.

10. Walton, R.T., Christie, K.A., Whittaker, M.N., and Kleinstiver, B.P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. Science *368*, 290–296. https://doi.org/10.1126/science.aba8853.

11. Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L., Wei, Y., Hu, X., Zeng, R., et al. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. Nature *571*, 275–278. https://doi.org/10.1038/s41586-019-1314-0.

12. Richter, M.F., Zhao, K.T., Eton, E., Lapinaite, A., Newby, G.A., Thuronyi, B.W., Wilson, C., Koblan, L.W., Zeng, J., Bauer, D.E., et al. (2020). Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. Nat. Biotechnol. *38*, 883–891. https://doi.org/10.1038/s41587-020-0453-z.

13. Li, X., Wang, Y., Liu, Y., Yang, B., Wang, X., Wei, J., Lu, Z., Zhang, Y., Wu, J., Huang, X., et al. (2018). Base editing with a Cpf1-cytidine deaminase fusion. Nat. Biotechnol. *36*, 324–327. https://doi.org/10.1038/nbt.4102.

14. Huang, T.P., Zhao, K.T., Miller, S.M., Gaudelli, N.M., Oakes, B.L., Fellmann, C., Savage, D.F., and Liu, D.R. (2019). Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. Nat. Biotechnol. *37*, 626–631. https://doi.org/10.1038/s41587-019-0134-y.

15. Li, J., Yu, W., Huang, S., Wu, S., Li, L., Zhou, J., Cao, Y., Huang, X., and Qiao, Y. (2021). Structure-guided engineering of adenine base editor with minimized RNA off-targeting activity. Nat. Commun. *12*, 2287. https://doi.org/10.1038/s41467-021-22519-z.

16. Chu, S.H., Packer, M., Rees, H., Lam, D., Yu, Y., Marshall, J., Cheng, L.I., Lam, D., Olins, J., Ran, F.A., et al. (2021). Rationally designed base editors for precise editing of the sickle cell disease mutation. CRISPR J. *4*, 169–177. https://doi.org/10.1089/crispr.2020.0144.

17. Wang, Y., Zhou, L., Liu, N., and Yao, S. (2019). BE-PIGS: a base-editing tool with deaminases inlaid into Cas9 PI domain significantly expanded the editing scope. Signal Transduct. Target. Ther. *4*, 36. https://doi.org/10.1038/s41392-019-0072-7.

18. Nguyen Tran, M.T., Mohd Khalid, M.K.N., Wang, Q., Walker, J.K.R., Lidgerwood, G.E., Dilworth, K.L., Lisowski, L., Pebay, A., and Hewitt, A.W. (2020). Engineering domain-inlaid SaCas9 adenine base editors with reduced RNA off-targets and increased on-target DNA editing. Nat. Commun. *11*, 4871. https://doi.org/10.1038/s41467-020-18715-y.

19. Liu, Y., Zhou, C., Huang, S., Dang, L., Wei, Y., He, J., Zhou, Y., Mao, S., Tao, W., Zhang, Y., et al. (2020). A Cas-embedding strategy for minimizing off-target effects of DNA base editors. Nat. Commun. *11*, 6073. https://doi.org/10.1038/s41467-020-19690-0.

20. Gaudelli, N.M., Lam, D.K., Rees, H.A., Sola-Esteves, N.M., Barrera, L.A., Born, D.A., Edwards, A., Gehrke, J.M., Lee, S.J., Liquori, A.J., et al. (2020). Directed evolution of adenine base editors with increased activity and therapeutic application. Nat. Biotechnol. *38*, 892–900. https://doi.org/10.1038/s41587-020-0491-6.

21. Bae, S., Park, J., and Kim, J.S. (2014). Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. Bioinformatics *30*, 1473–1475. https://doi.org/10.1093/bioinformatics/btu048.

22. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature *533*, 420–424. https://doi.org/10.1038/nature17946.

23. Koblan, L.W., Arbab, M., Shen, M.W., Hussmann, J.A., Anzalone, A.V., Doman, J.L., Newby, G.A., Yang, D., Mok, B., Replogle, J.M., et al. (2021). Efficient C*G-to-G*C base editors developed using CRISPRi screens, target-library analysis, and machine learning. Nat. Biotechnol. *39*, 1414–1425. https://doi.org/10.1038/s41587-021-00938-z.

24. Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature *576*, 149–157. https://doi.org/10.1038/s41586-019-1711-4.

25. Zhang, W., Yin, J., Zhang-Ding, Z., Xin, C., Liu, M., Wang, Y., Ai, C., and Hu, J. (2021). In-depth assessment of the PAM compatibility and editing activities of Cas9 variants. Nucleic Acids Res. *49*, 8785–8795. https://doi.org/10.1093/nar/gkab507.

26. Li, S., Yuan, B., Cao, J., Chen, J., Chen, J., Qiu, J., Zhao, X.M., Wang, X., Qiu, Z., and Cheng, T.L. (2020). Docking sites inside Cas9 for adenine base editing diversification and RNA off-target elimination. Nat. Commun. *11*, 5827. https://doi.org/10.1038/s41467-020-19730-9.

27. Jin, S., Zong, Y., Gao, Q., Zhu, Z., Wang, Y., Qin, P., Liang, C., Wang, D., Qiu, J.L., Zhang, F., and Gao, C. (2019). Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. Science *364*, 292–295. https://doi.org/10.1126/science.aaw7166.

28. Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L.M., Li, Y., and Yang, H. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. Science *364*, 289–292. https://doi.org/10.1126/science.aav9973.

29. Wang, L., Xue, W., Zhang, H., Gao, R., Qiu, H., Wei, J., Zhou, L., Lei, Y.N., Wu, X., Li, X., et al. (2021). Eliminating base-editor-induced genome-wide and transcriptome-wide off-target mutations. Nat. Cell Biol. *23*, 552–563. https://doi.org/10.1038/s41556-021-00671-4.

30. Doman, J.L., Raguram, A., Newby, G.A., and Liu, D.R. (2020). Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. Nat. Biotechnol. *38*, 620–628. https://doi.org/10.1038/s41587-020-0414-6.

31. Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., and Moriarity, B.S. (2018). EditR: a method to quantify base editing from sanger sequencing. CRISPR J. *1*, 239–250. https://doi.org/10.1089/crispr.2018.0014.