A motif preferred adenine base editor with minimal bystander and off-targets editing

Mengyu Shang [1,3], Yinuo Li[1,3], Qiuyu Cao[1,3], Jingxuan Ren[1,3], Yuqiang Zeng[1], Jinxin Wang[1], Xiyu Ge[2], Xiaohui Zhang[1,*]

[1]State Key Laboratory of Common Mechanisms Research for Major Diseases,Suzhou Institute of Systems Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Suzhou 215123, China.

[2]Division of Endocrinology, Stanford University School of Medicine, Stanford, CA, USA.

[3]These authors contributed equally to this work.

*To whom correspondence should be addressed:

Xiaohui Zhang, zhangxiaohui3040@126.com.

Abstract

48% of hereditary disease are caused by single C-to-T base conversion, which makes efficient A-to-G base editing tools (ABEs) have great potential in the treatment of these diseases. However, the existing efficient ABE, while catalyzing A-to-G conversion, will bring more A and C bystander editing and off-target events, which poses safety concerns for their clinical application. Here, we developed ABE10 (ABE8e with TadA-8e A48E) for efficient and accurate editing of As in YA motifs with YAY>YAR (Y=T or C, R=A or G) hierarchy through structure-oriented rational design. Compared with ABE3.1, currently the only motif (YAC) preference ABE version, ABE10 exhibited A-to-G editing efficiency improvement with an average up to 3.1-fold in indicated YA motif while maintaining reduced bystander Cs editing and minimized DNA or RNA off-targets. Also, we showed ABE10 corrected pathogenic mutation with high efficiency and precision in human cells. Moreover, by ABE10, we efficiently and precisely generated hypocholes-terolemia and tail-loss mouse models mimicking human associated disease and mouse *PCSK9* base editing in vivo for hypercholesterolemia gene therapy, indicating their great potential in broad applications for and disease modeling and gene therapy.

35 **Introduction**

36 About 60% of hereditary disease are caused by single base mutations, 48% of these are
37 caused by single C-to-T base conversion[1]. The treatment of such hereditary diseases
38 needs safe and effective genome editing tools. Traditional CRISPR/Cas9-mediated
39 homologous recombination as tools to repair these pathogenic point mutations were
40 very inefficient[1], which leads to the new generation of gene editing tools- base editors,
41 especially the A-to-G base editing tool (ABEs), providing an attractive platform for the
42 treatment of single base mutation diseases[1, 2].

43 Current popular base editors are cytidine base editors (CBEs) and adenine base
44 editors (ABEs), which can convert C•G-to-T•A and A•T-to-G•C, respectively[2, 3]. They
45 can be used to correct disease-causing point mutations that account for 48% and 13% of
46 genetic diseases, respectively[1]. The application of gene editing in clinical treatment
47 often requires accurate and efficient base editing tools. In recent years, scientists around
48 the world have made significant progresses on the accuracy and efficiency of base
49 editing. For CBEs, a series of accurate and effective CBEs have been developed, such
50 as structure-oriented design of cytosine deaminase to construct motif-specific CBEs
51 (TC, CC and WRC)[4-7]; Engineering cytosine base editors with only 1-2nt editing
52 windows[8]; Re-evolving TadA-8e transformed ABEs into more accurate and less off-
53 target CBEs[9-11] etc. For ABEs, some progresses have also been made, such as rational
54 design of accurate ABE9 for only editing A5 or A6 positions[13]. However, limitations of
55 ABEs are still present for current ABEs, takingABE3.1 as an example, the only motif
56 (YAC) preference ABE version, while has low efficiency and induced more bystander
57 Cs editing within editing window and off-targets[2]. Therefore, ABEs that could
58 efficiently catalyze A-to-G conversion in specific motifs without introducing more
59 bystanders within editing window and off-targets still in urgent need.

60 Here, we obtained ABE10 with YA motif sequence preference through structure-
61 oriented rational design of ABE8e mutant. Compared to ABE8e, ABE10 can precisely
62 edit YA sequences with minimized bystanders. More importantly, compared with
63 ABE3.1, the only motif (YAC) preference ABE version, ABE10 exhibited A-to-G
64 editing efficiency improvement with average up to 3.1-fold in indicated motif while
65 maintaining reduced bystander Cs editing and minimized DNA or RNA off-target. We
66 also demonstrated that ABE10 can effectively correct disease-causing point mutations
67 without introducing detrimental mutations in human cells. Moreover, by injecting
68 ABE10 mRNA and sgRNA targeting splice donor of mouse *PCSK9* and *TBXT* into

zygotes, the loss-of-function of PCSK9 mouse models for hypocholes-terolemia, and short-tailed mouse models for simulating mouse evolution were efficiently produced, respectively. Finally, we delivered lipid nanoparticle–packaged ABE10 mRNA and a sgRNA targeting the GT splice donor site of PCSK9 in mice, high base editing efficiency for PCSK9 and substantial reduction the expression of Pcsk9 and LDL-C in plasma were observed.

Results

Structure-based evolution of TadA-8e for YAC motif sequences preference

The earlier reported ABEs with YAC motif preference was ABE3.1[2], but its A-to-G editing efficiency was low (1.2-31.1%) and induced more C bystander editing (Supplementary Fig. 1). Efficiency-wise, ABE8e is currently considered the more efficient version of ABE. However, ABE8e not only has higher editing activity and a wider editing window making it tend to induce more bystander edits and off-targets, but also lose its sequence preference[12]. To overcome this limitation, we designed a series of mutants of ABE8e at the binding and catalytic sites of TadA-8e as well as DNA substrates based on the crystal structure of ABE8e (Fig. 1a), and screened for ABEs to restore its YAC motif sequence preference while maintaining high editing activity. A human endogenous target site (HEK2) with $A_3C\underline{A}_5C\underline{A}_7A_8$ sequence was selected for testing. Among these mutations, ABEs that could edit A_5 while had reducing or no editing efficiency of A_3, A_7 and A_8 were selected as candidates (Fig. 1b). After co-transfected ABE8e mutants and HEK2 sgRNA for 72h in HEK293T cells, we collected genomic DNA of cells for amplicon high-throughput sequencing (HTS). HTS data showed that almost all ABE8e mutants can edit both A_5 and A_7, only ABE8e A48D exhibited reduced A_7 base editing activity without sacrificing on-target A_5 activity (Fig. 1c), indicating A48 maybe the key amino acid of TadA-8e that influences its motif sequence preference. In order to further identify the TadA-8e mutant that can minimize bystander editing, we saturated the mutant A48 amino acid and further tested it with the same target in HEK293T cells. HTS data showed that ABE8e A48E maintains comparable editing activity at A_5 while minimizing bystander editing activity at A_7 (a

3.0-fold decrease) compared to ABE8e A48D (Fig. 1d). We named ABE8e A48E as ABE10.

Characterization of ABE10

To unbiasedly profile the characteristics of ABE10, 23 endogenous targets were tested in HEK293T cells with ABE8e, ABE9 and ABE3.1 as control. Among these targets, they almost cover NAN (N=A, T, C and G) and some targets contained NCN (N=A, T, C and G) within its editing window. After analyzing the A-to-G editing efficiency of the most highly edited A in each target across all 23 target sites, we found that the efficiencies of ABE8e were 43.8-90.9% , which were significant higher than all other ABEs treated group (Fig. 2a,b). The efficiencies of ABE10 were 1.4-90%, whichwere significant higher than that of ABE9 (0.7-78.7%) and ABE3.1 (0-75.3%) (Fig. 2a,b). Through analyzing the A-to-G editing efficiency of each protospacer of all tested targets, we found ABE8e can efficiently edit almost all As without motif preference at major editing window (A_3-A_8) and even edit As outside of the major window (A_2 or A_{10}-A_{13}) with considerable efficiency (Fig. 2a,c). ABE9, ABE3.1 and ABE10 have a similar editing window that was narrower than that of ABE8e, suggesting they are all accurate versions of ABE, suitable for precision gene therapy (Fig. 2a,c). We further analyzed the motif preference of these base editors and no motif preference for ABE8e and ABE9 was observed (Fig. 2d). Notably, ABE3.1 has YAY motif and ABE10 has YA motif preference with YAY>YAR hierarchy (Y=T or C, R=A or G) (Fig. 2a,d). Moreover, the average A-to-G efficiencies of ABE10 at A_4, A_5, A_6 and A_7 were 34.4%, 58.4%, 35.7% and 16.6%, respectively, far higher than the corresponding efficiencies of ABE3.1(Fig. 2a,c). Since ABE8e has been reported to have cytosine deaminase activity[14], we further analyzed the C editing of ABE10 at 12 targets. The results showed that ABE10 induced lower or near background C-to-D editing activity compared to ABE8e. At *FANCF*-sg5, ABE10 decreased the editing efficiency of C induced by ABE8e by up to 10.4 times, which was comparable to ABE9 (Fig. 2e). ABE10 also induced no more indels than that of ABE8e and ABE3.1 (Fig. 2f). Together, ABE10 is highly efficient at YA motif preference with significantly reduced adenine and cytosine bystander editing effects.

Off-target evaluation of ABE10

Next, we performed off-target assessment of ABE10 in three ways: sgRNA-dependent

132 DNA off-target, sgRNA-independent DNA off-target and whole-transcriptomic RNA
133 off-target. 17 off-targets from in silicon prediction or Cas9 known off-targets were used
134 to evaluate the sgRNA-dependent DNA off-target. The results showed that 4/17 off-
135 target were observed for ABE8e and no sgRNA-dependent DNA off-target for both
136 ABE10 and ABE9 (Fig. 3a). Modified R-loop assay was used to evaluate the sgRNA-
137 independent DNA off-target[15]. HTS data showed that ABE10, similar to ABE9, induced
138 no sgRNA-independent DNA off-target, which were far fewer than that ABE8e (Fig. 3b
139 and Supplementary Fig. 3). RNA-Seq was used to evaluate whole-transcriptomic RNA
140 off-target[16, 17]. The results showed that ABE10 induced 5 times lower the off-target rate
141 than that of ABE8e, though slightly higher than that ABE9 (Fig. 3c). In summary,
142 ABE10 is a highly efficient base editing tool with high specificity.
143

144 **Introduction of therapeutic mutation with high precision using ABE10 in human**
145 **cells**

146 To further validate ABE10's potential for gene therapy with high precision and
147 efficiency, we generated stable HEK293T cell lines with 2 pathogenic mutations (*GLAT*
148 c.413C>T, Transferase Deficiency Galactosemia; *OTC* c.533C>T, Ornithine
149 transcarbamylase deficiency) adjacent to bystander mutations that are deleterious in
150 Clin Var database[18]. The results showed that ABE10 exhibited higher desired base
151 editing efficiency than ABE9 at A_6 in *GALT* (45.6%) and A_6 in *OTC* (71.35%), though
152 slightly lower than ABE8e at corresponding sites (Fig. 4a). However, ABE8e also
153 induced deleterious no-desired base mutation with high efficiency, such as A_3 in *GALT*
154 and A_3 in *OTC* which bring the possibility of additional disease. In contrast, ABE10
155 have far lower or no editing efficiency at these sites (Fig. 4a,b).

156 We further counted these pathogenic mutations with the presence of additional As or
157 Ts (the disease-causing As or Ts were on the complementary strand of its DNA) that
158 may present a risk of disease in the ClinVar database. The results showed that there
159 were 12751 pathogenic mutations surrounded by As or Ts with potential security risk in
160 ABE8e`s editing window (Fig. 4c and Supplementary Table 4). Among those mutations,
161 8783 and 1841 pathogenic mutations could be precisely corrected by ABE10 and ABE9,
162 respectively (Fig. 4c and Supplementary Table 4, 5). When PAM was further expanded
163 from NGG to NG, the number of precise disease treatment events was further expanded
164 to 21622 for ABE10 and 6033 for ABE9, respectively (Supplementary Table 4, 5).
165 These data suggested that ABE10 have great potential in precisely targeting specific

base for future clinical gene therapy.

Generation of disease associated mouse models with high precision using ABE10

Base editors have been applied for making disease animal models that is essential for basic medical research. When the base editor is used to make the animal model of disease, it induces not only the target base conversion but also the bystander base editing which potentially perturb the analysis of the relationship between the mouse disease phenotype and base mutations[19]. Therefore, we used ABE10 to generate mouse model to minimize the occurrence of bystander events. We first targeted proprotein convertase subtilisin/kexin type 9 (PCSK9) to generate hypocholesterolemia mouse models by the zygotes microinjection of ABE10 mRNA and sgRNA targeting the splice site of PCSK9 (Fig. 5a and Supplementary Fig. 4)[20]. The results showed ABE 10 and ABE8e induced efficiently base editing in pups, with ABE 10(17/18) and ABE8e (9/11), respectively (Fig. 5b,c and Supplementary Fig. 4). However, only ABE10 induced precisely A-to-G conversion of the opposite DNA strand of GT splice site with minimized bystander editing in pups. On the contrary, not only the target base A_6, but also bystander A_4 was edited in ABE8e treated group (Fig. 5c and Supplementary Fig. 4). The phenotype of the pups carried precise target base editing by ABE10 were further confirmed by the detection of PCSK9 and LDL-C (Fig. 5d,e). These data suggest that ABE, especially the highly efficient version of ABE, enabled efficiently base editing in embryos to produce homozygous mouse with disease phenotype.

However, for some other disease-causing genes that are lethal with homozygous mutations, the highly efficient version of ABE is not suitable. For example, *TBXT*, a gene that influences the evolution of the mouse tail and single exon6-skipped transcript is sufficient to induce a tail-loss phenotype (Fig. 5f)[21]. We designed a sgRNA targeted the opposite of GT splice site of the *TBXT* that its disruption may possible cause exon6 skipping. By co-microinjection ABE8e or ABE10 with sgRNA into mouse zygotes, 8 (8/8) pups were edited in ABE10 treated group (Fig. 5g and Supplementary Fig. 5). Heterozygous mice (#BT03, #BT04) displayed almost no tail phenotype and mice with less editing efficiency (#BT02) showed short-tailed phenotype (Fig. 5h). However, we found no pups were born in ABE8e group (Data not shown) and the probable cause is that homozygous editing by ABE8e in mice leads to complete exon6-skipped and thus death in utero. These data suggest that ABE10 hold great potential in generation of disease model. More importantly, ABE10 has great advantages in animal models of

diseases targeting homozygous lethal genes compared with other high efficiency version of ABEs.

In vivo adenine base editing *PCSK9* in mouse using ABE10

To further validate ABE10's potential for gene therapy in vivo, we targeted proprotein convertase subtilisin/kexin type 9 (*PCSK9*), a therapeutically relevant gene involved in cholesterol homeostasis[20]. We usedlipid nanoparticle–packaged delivery of mRNA encoding an ABE10 and a sgRNA targeting the GT splice donor site of *PCSK9* in mice (Fig. 6a). Three weeks after tail vein injection, higher base editing efficiency (60%) exclusively in the liver was observed (Fig. 6b), and no DNA off-target effects were detected by in silicon-predicated methods (Supplementary Fig.7a). During this time, we also monitored the expression levels of PCSK9 and LDL-C in plasma. The results also showed substantial reduction in the expression levels of Pcsk9 and LDL-C in ABE10 treated group (Fig. 6c,d). Meanwhile, LNPs-packaged ABE10 delivery did not induce liver injury by the detection of serum alanine aminotransferase (ALT) and aspartate transaminase (AST) (Supplementary Fig.7b,c). These data suggested that ABE10 has great potential in clinical gene therapy in vivo.

Discussion

In this study, we have developed a YA motif preferred adenine base editor by structure-oriented rational design of TadA-8e and named it with ABE10. Compared to the efficient ABE version - ABE8e, ABE10 has YA motif preference and a narrower editing window, while reducing bystander A and C editing and DNA/RNA off-target. ABE3.1, currently the only sequence-specific ABE version, has unpredicatable C editing and RNA off-target. For ABE10, it not only effectively alleviates the limitations of ABE, but also has the advantages of high efficiency and small size. Compared with the current accurate ABE-ABE9, ABE10 not only has the same editing efficiency as ABE9 at A_5 and A_6, but also has higher editing efficiency at A_4 and A_7 sites, where ABE9 cannot edit. Therefore, ABE10 is an elegant adenine base editor with high efficiency, precision and small size. However, although ABE10 has high accuracy, it also sacrifices the editing efficiency of on-target. Therefore, it is necessary to further engineer deaminase in the future to improve its efficiency while maintaining its accuracy. For cytosine base editors, TC, CC, WRC motif preference cytosine base editors have been reported, while similar ABE tools are still lacking, and future further exploration is needed to enrich ABE's

toolbox.

Furthermore, we have also demonstrated that ABE10 can be effectively used to correct disease-causing point mutations in human cells, especially in editing windows where there is A mutation with safety risk. By further statistics in the ClinVar database, 7706 disease treatment events can be treated using ABE10 rather than other ABEs by precisely targeting YA motif. When PAM was further expanded from NGG to NG, the number of precise disease treatment events was further expanded to 17722, suggesting that ABE10 have great potential in precisely targeting specific base for future clinical gene therapy. Finally, ABE10 can be used to precisely target YA motif to generate animal models with disease phenotypes and to treat hypercholesterolemia without off-target events.

In summary, ABE10 is a highly efficient and specific base editing tool, which have great potential in animal model making and disease gene therapy.

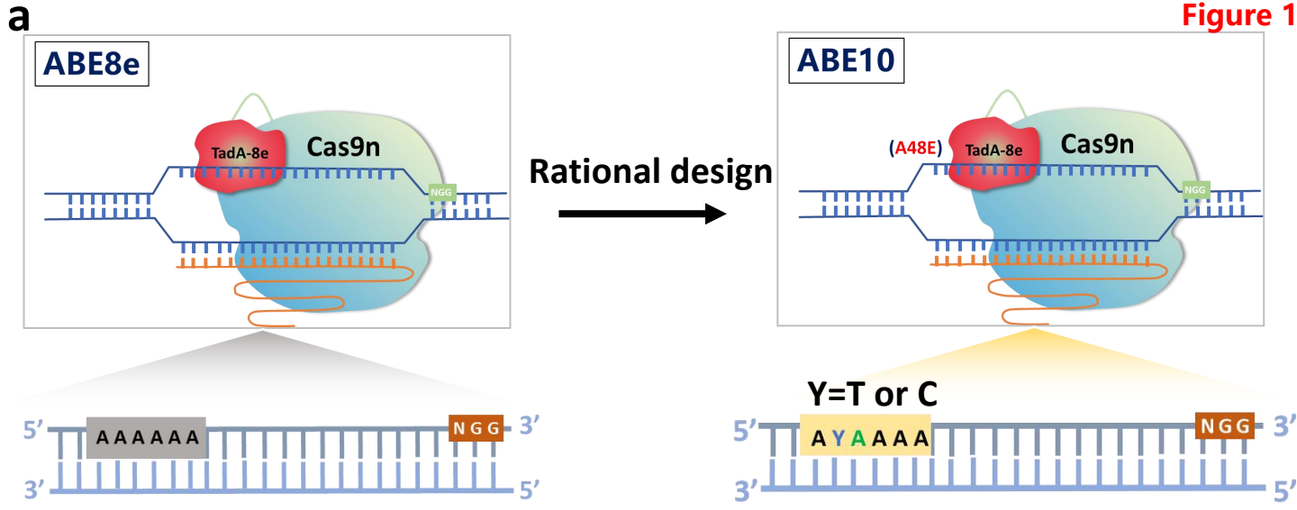
References:

1. Rees, H.A. & Liu, D.R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nature reviews genetics* **19**, 770-788 (2018).
2. Komor, A.C. et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Science advances* **3**, eaao4774 (2017).
3. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A. & Liu, D.R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420-424 (2016).
4. Gehrke, J.M. et al. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nature Biotechnology* **36**, 977-982 (2018).
5. Lee, S. et al. Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects. *Science Advances* **6**, eaba1773 (2020).
6. Liu, L.D. et al. Intrinsic Nucleotide Preference of Diversifying Base Editors Guides Antibody Ex Vivo Affinity Maturation. *Cell Reports* **25**, 884-892.e883 (2018).
7. Liu, Z. et al. Improved base editor for efficient editing in GC contexts in rabbits with an optimized AID-Cas9 fusion. *The FASEB Journal* **33**, 9210-9219 (2019).
8. Kim, Y.B. et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nature Biotechnology* **35**, 371-376 (2017).
9. Chen, L. et al. Re-engineering the adenine deaminase TadA-8e for efficient and specific CRISPR-based cytosine base editing. *Nature Biotechnology* **41**, 663-672 (2022).
10. Lam, D.K. et al. Improved cytosine base editors generated from TadA variants. *Nature Biotechnology* **41**, 686-697 (2023).

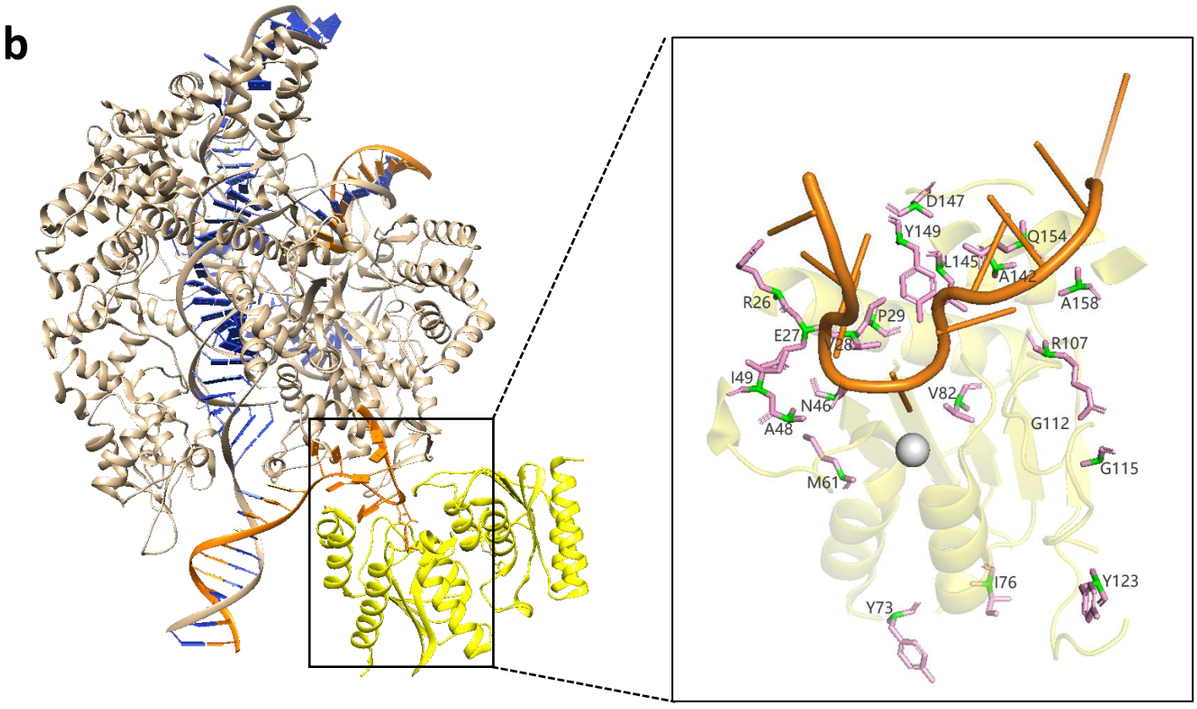
11. Neugebauer, M.E. et al. Evolution of an adenine base editor into a small, efficient cytosine base editor with low off-target activity. *Nature Biotechnology* **41**, 673-685 (2022).
12. Richter, M.F. et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nature Biotechnology* **38**, 883-891 (2020).
13. Chen, L. et al. Engineering a precise adenine base editor with minimal bystander editing. *Nature Chemical Biology* **19**, 101-110 (2022).
14. Jeong, Y.K. et al. Adenine base editor engineering reduces editing of bystander cytosines. *Nature Biotechnology* **39**, 1426-1433 (2021).
15. Doman, J.L., Raguram, A., Newby, G.A. & Liu, D.R.J.N.b. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors. **38**, 620-628 (2020).
16. Grünewald, J. et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. **569**, 433-437 (2019).
17. Zhou, C. et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. **571**, 275-278 (2019).
18. Landrum, M.J. et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic acids research* **44**, D862-D868 (2016).
19. Zhang, X. et al. Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain. *Nature Cell Biology* **22**, 740-750 (2020).
20. Fitzgerald, K. et al. Effect of an RNA interference drug on the synthesis of proprotein convertase subtilisin/kexin type 9 (PCSK9) and the concentration of serum LDL cholesterol in healthy volunteers: a randomised, single-blind, placebo-controlled, phase 1 trial. *The Lancet* **383**, 60-68 (2014).
21. Xia, B. et al. On the genetic basis of tail-loss evolution in humans and apes. *Nature* **626**, 1042-1048 (2024).

Figure 1

a



b



c

HEK site2

	A_2	A_3	C_4	A_5	C_6	A_7	A_8	A_9	A_{12}	A_{14}
ABE8e•R26A	0.60	21.00	0.17	64.67	0.87	62.00	9.63	3.63	4.93	2.03
ABE8e•R26G	0.53	20.23	0.13	63.73	1.33	63.07	10.07	3.87	5.63	3.30
ABE8e•E27A	0.10	1.20	0.83	59.43	8.70	46.57	0.97	0.23	0.70	0.13
ABE8e•E27R	0.10	0.80	0.77	63.27	13.17	36.03	1.10	0.40	0.43	0.13
ABE8e•V28G	0.13	1.03	1.07	62.63	13.10	22.80	0.73	0.33	0.50	0.20
ABE8e•V28A	0.17	2.30	0.50	62.40	2.07	51.37	1.60	0.83	0.80	0.30
ABE8e•P29A	0.23	4.83	0.17	61.57	0.43	55.93	5.87	4.20	0.73	0.23
ABE8e•N46A	0.07	0.80	0.03	5.57	0.23	3.23	0.67	0.27	0.40	0.20
ABE8e•N46L	0.07	1.20	0.63	11.90	2.60	3.53	0.70	0.27	0.37	0.10
ABE8e•I49A	0.27	8.50	0.53	58.57	6.80	57.43	6.23	2.57	4.97	2.33
ABE8e•I49K	0.37	12.60	0.60	59.30	4.23	57.77	15.33	4.37	8.03	1.23
ABE8e•M61I	0.60	19.03	0.20	66.23	1.20	65.07	15.40	7.00	8.70	1.60
ABE8e•Y73A	0.20	5.50	0.33	61.07	3.80	59.07	5.17	2.27	6.73	0.97
ABE8e•Y73H	0.50	11.13	0.17	61.67	1.40	59.93	12.27	4.17	5.60	1.83
ABE8e•Y73S	0.27	9.03	0.30	59.27	2.20	57.67	7.30	3.53	2.97	1.10
ABE8e•I76A	0.57	12.77	0.10	59.30	0.87	57.90	12.03	4.83	6.20	2.23
ABE8e•I76F	1.53	35.27	0.30	56.37	1.37	54.73	18.07	7.13	10.07	1.93
ABE8e•V82T	1.37	27.93	0.90	52.20	7.70	51.70	27.30	12.30	17.70	4.23
ABE8e•R107A	0.53	14.47	0.23	59.00	2.17	57.00	14.23	4.53	4.67	2.07
ABE8e•R107C	0.60	15.67	0.20	59.93	2.03	58.63	19.43	5.33	7.67	2.70
ABE8e•Y123A	0.47	10.57	0.17	60.63	0.73	59.70	7.93	3.70	5.17	1.87
ABE8e•G115M	0.10	1.20	0.10	60.43	0.47	44.07	0.70	0.20	0.63	0.23
ABE8e•A142E	0.40	11.97	0.10	61.67	1.40	59.87	16.97	6.60	1.60	0.87
ABE8e•L145T	0.10	1.90	0.10	54.27	0.57	37.43	1.90	0.77	0.40	0.17
ABE8e•D147R	0.77	27.20	0.17	60.63	1.53	59.33	17.23	6.33	3.30	1.70
ABE8e•Y149A	0.23	3.07	0	61.37	0.33	47.30	2.30	1.43	0.47	0.03
ABE8e•Q154H	0.57	16.50	0.30	57.73	1.20	56.60	14.27	5.07	9.23	2.90
ABE8e•Q154R	1.43	22.27	0.20	58.80	1.23	58.23	17.33	7.07	9.33	3.70
ABE8e•A158S	0.77	18.57	0.23	64.50	1.37	63.83	15.47	5.93	10.80	2.17
ABE8e•A48D	0.10	0.80	0.03	66.77	0.10	19.60	0.33	0.10	0.20	0.13
ABE8e	0.40	9.40	0.07	62.47	0.63	61.63	8.17	2.77	5.17	1.00
ABE9	0.10	0.43	0	52.03	0.03	0.73	0.17	0.13	0.10	0.10
ABE3.1	0.10	0.23	0	55.77	0	1.60	0.10	0.03	0.20	0.10
untreated	0.10	0.20	0	1.17	0.07	0.77	0.13	0.10	0.13	0.13

0 — 66.8

A-to-G/C-to-D(A/G/T)(%)

d

HEK site2

	A_2	A_3	C_4	A_5	C_6	A_7	A_8	A_9	A_{12}
ABE8e A48F	0.47	7.13	0.57	59.93	0.57	59.33	14.20	4.03	5.13
ABE8e A48L	0.30	8.53	0.43	60.83	0.43	60.30	8.47	2.33	5.10
ABE8e A48I	0.33	6.73	0.43	62.13	0.43	61.30	6.37	2.17	9.30
ABE8e A48M	0.87	20.90	0.23	61.03	0.23	60.63	18.00	5.90	10.83
ABE8e A48V	0.47	10.47	0.87	61.10	0.87	60.50	8.67	3.63	8.60
ABE8e A48S	0.40	14.27	1.60	59.87	1.60	59.10	12.80	4.10	6.10
ABE8e A48P	0.13	8.37	0.20	51.57	0.20	50.77	5.93	1.90	2.83
ABE8e A48T	0.33	9.23	1.50	61.63	1.50	60.93	8.47	3.23	4.00
ABE8e A48Y	0.20	3.27	0.17	44.60	0.17	43.27	2.87	0.83	1.83
ABE8e A48H	0.23	7.13	0.83	61.33	0.83	58.43	7.43	3.70	4.57
ABE8e A48Q	0.17	4.70	0.33	60.63	0.33	55.10	3.63	1.60	2.70
ABE8e A48N	0.17	5.27	0.30	35.43	0.30	32.33	4.47	1.73	2.03
ABE8e A48K	0.20	6.30	1.03	61.03	1.03	57.20	6.53	2.87	3.37
ABE8e A48D	0.10	0.80	0.10	66.77	0.10	19.60	0.33	0.10	0.20
ABE8e A48E	0.10	0.47	0.10	61.50	0.10	6.50	0.13	0.17	0.17
ABE8e A48C	0.60	16.53	1.93	55.93	1.93	55.40	14.83	5.70	7.57
ABE8e A48W	0.17	1.73	0.07	46.97	0.07	35.27	1.17	0.40	0.77
ABE8e A48R	0.23	3.10	0.63	58.37	0.63	57.10	3.53	1.33	1.83
ABE8e A48G	0.30	8.10	5.87	59.80	5.87	57.73	7.07	2.23	3.57
ABE8e	0.40	9.40	0.63	62.47	0.63	61.63	8.17	2.77	5.17
ABE9	0.10	0.43	0.03	52.03	0.03	0.73	0.17	0.13	0.10
ABE3.1	0.10	0.23	0	55.77	0	1.60	0.10	0.03	0.20
Untreated	0.10	0.20	0.07	1.17	0.07	0.77	0.13	0.10	0.13

0 — 66.8

A-to-G/C-to-D(A/G/T)(%)

299 **Fig. 1 | Rational design of motif preferred ABE. a**, Chematic illustration of the conceptual design to
300 design an efficient ABE into YA motif preferred ABE. **b**, Overview of the interaction of TadA-8e (yellow)
301 with the single-stranded DNA substrate (orange) (Protein Data Bank (PDB): 6VPC). Cas9n is in gray,
302 sgRNA is in cyan, complementary strand DNA is in orange and noncomplementary strand DNA is in
303 green. Amino acids spatially contacting or adjacent to the substrate DNA are labeled on the enlarged
304 image. **c**, The A-to-G base editing efficiency of ABE8e or ABE8e variants at an endogenous genomic loci
305 containing YAC motif and multiple As with non-YAC motif (HEK2) in HEK293T cells. Data are means
306 (n = 3 independent experiments). **d**, The A-to-G base editing efficiency of ABE8e or ABE8e variants of
307 A48 saturation mutagenesis at HEK2 in HEK293T cells. Data are means (n = 3 independent experiments).

Figure 2

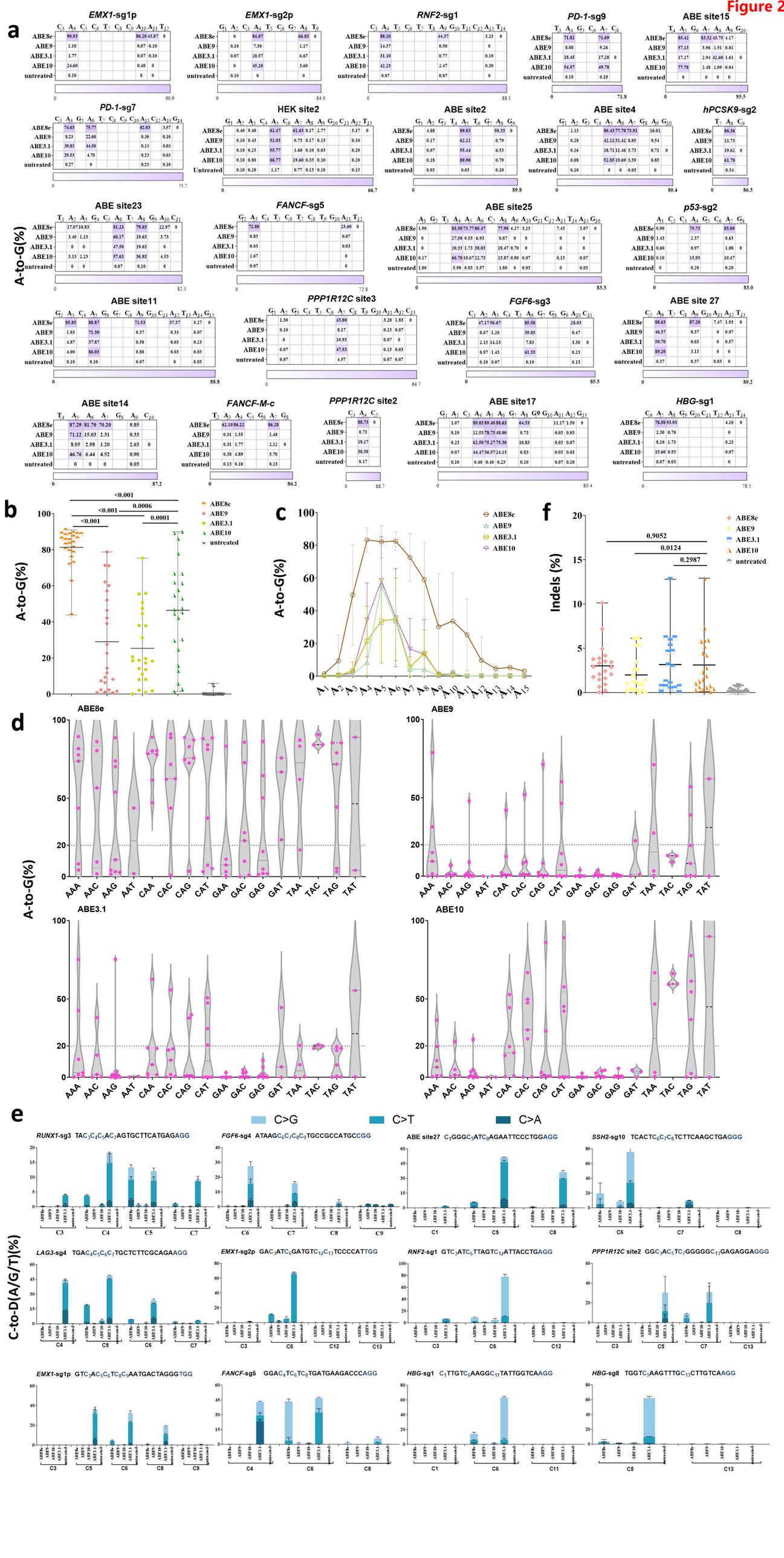
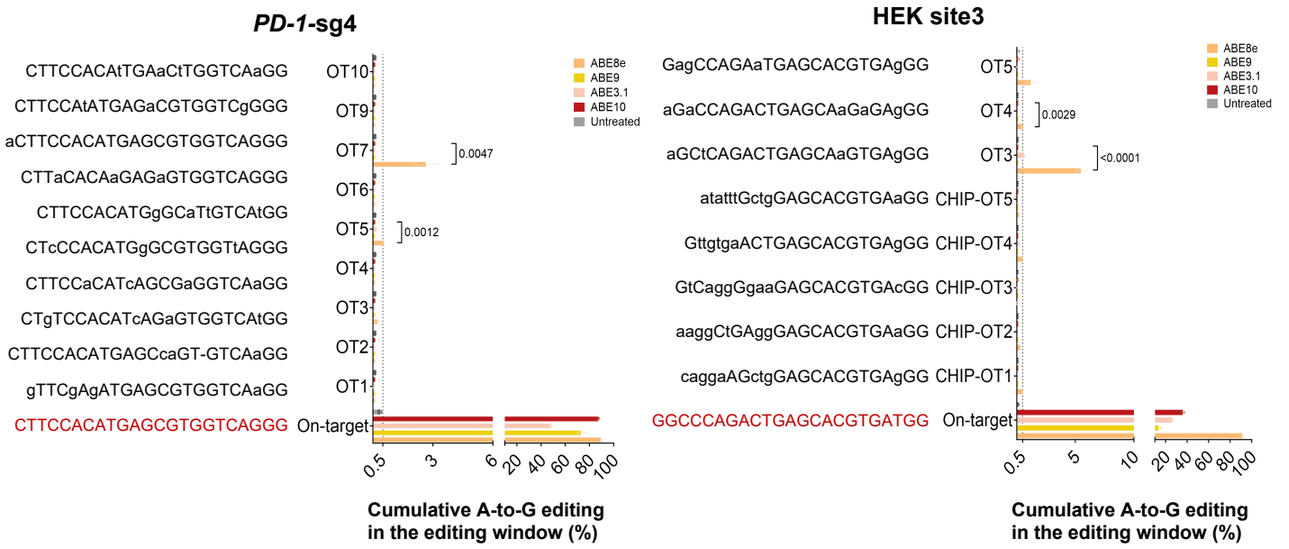

a
EMX1-sg1p
EMX1-sg2p
RNF2-sg1
PD-1-sg9
ABE site15
PD-1-sg7
HEK site2
ABE site2
ABE site4
hPCSK9-sg2
ABE site23
FANCF-sg5
ABE site25
p53-sg2
ABE site11
PPP1R12C site3
FGF6-sg3
ABE site 27
ABE site14
FANCF-M-c
PPP1R12C site2
ABE site17
HBG-sg1
ABE8e
ABE9
ABE3.1
ABE10
untreated
A-to-G(%)
b
A-to-G(%)
<0.001
<0.001
<0.001
0.0006
0.0001
c
A-to-G(%)
f
Indels (%)
0.9052
0.0124
0.2987
d
ABE8e
ABE9
ABE3.1
ABE10
A-to-G(%)
AAA AAC AAG AAT CAA CAC CAG CAT GAA GAC GAG GAT TAA TAC TAG TAT
e
C>G
C>T
C>A
RUNX1-sg3 TAC3C4C5AC7AGTGCTTCATGAGAGG
FGF6-sg4 ATAAGC6C7C8C9TGCCGCCATGCCGG
ABE site27 C1GGGC5ATC8AGAATTCCCTGGAGG
SSH2-sg10 TCACTC6C7C8TCTTCAAGCTGAGGG
C-to-D(A/G/T)(%)
LAG3-sg4 TGAC4C5C6C7TGCTCTTCGCAGAAGG
EMX1-sg2p GAC3ATC6GATGTC12C13TCCCCATTGG
RNF2-sg1 GTC3ATC6TTAGTC12ATTACCTGAGG
PPP1R12C site2 GGC3AC5TC7GGGGGC13GAGAGGAGGG
EMX1-sg1p GTC3AC5C6TC8C9AATGACTAGGGTGG
FANCF-sg5 GGAC4TC6TC8TGATGAAGACCCAGG
HBG-sg1 C1TTGTC6AAGGC11TATTGGTCAAGG
HBG-sg8 TGGTC5AAGTTTGC13CTTGTCAAGG

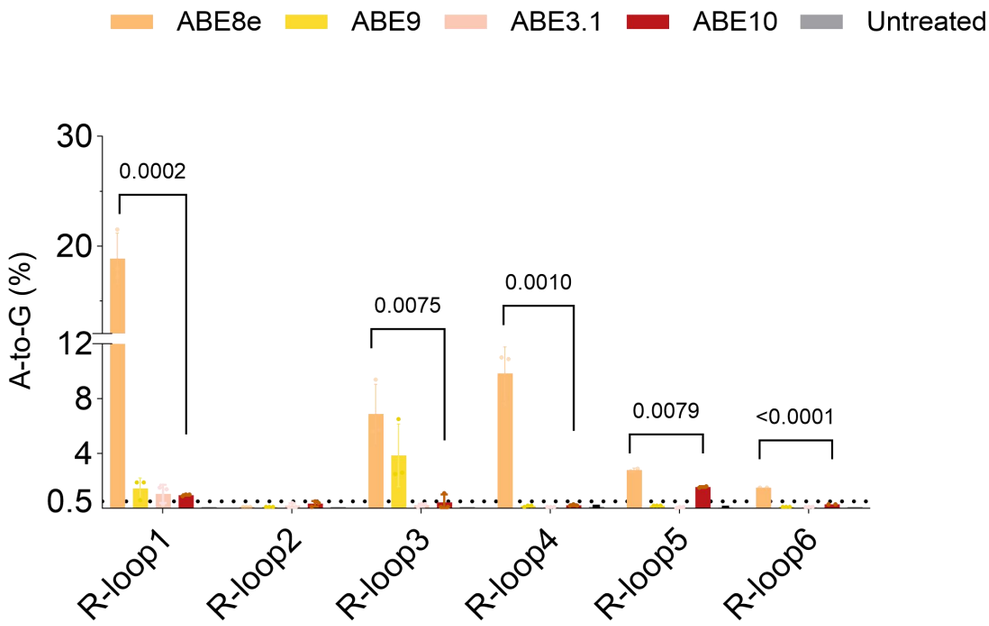
Fig. 2 | Characterization of ABE10. a, The A-to-G editing efficiency of ABE8e, ABE9, ABE3.1 and ABE10 were tested at 23 endogenous target sites in HEK293T cells. Data are means (n = 3 independent experiments). **b**, Summary of the A-to-G editing efficiencies for only the most highly edited adenine induced by ABE8e, ABE9, ABE3.1 and ABE10 were tested at 23 endogenous target sites in HEK293T cells. Each data point represents average A-to-G editing efficiency at all As within the activity window of each target site, calculated from 3 independent experiments. **c**, Average A-to-G editing efficiency of ABE8e, ABE9, ABE3.1 and ABE10 were tested at 23 endogenous target sites in HEK293T cells.Data are means ± SD (n = 3 independent experiments). **d**, The motif preferred analysis of ABE8e, ABE9, ABE3.1 and ABE10 at 23 endogenous target sites in HEK293T cells. Each data point represents average A-to-G editing efficiency at all motifs within the activity window of each target site, calculated from 3 independent experiments. **e**, Average C-to-D(A/C/G) editing efficiency of ABE8e, ABE9, ABE3.1 and ABE10 were tested at 23 endogenous target sites in HEK293T cells. Data are means ± SD (n = 3 independent experiments). **f**, The indels induced by ABE8e, ABE9, ABE3.1 and ABE10 at 23 endogenous target sites in HEK293T cells. Each data point represents average indels frequency at each target site calculated from 3 independent experiments. Significance was tested with two-sided paired Wilcoxon rank-sum test (**b**, **f**).

Figure 3

a


PD-1-sg4
CTTCCACAtTGAaCtTGGTCAaGG OT10
CTTCCAtATGAGaCGTGGTCgGGG OT9
aCTTCCACATGAGCGTGGTCAGGG OT7
CTTaCACAaGAGaGTGGTCAGGG OT6
CTTCCACATGgGCaTtGTCAtGG OT5
CTcCCACATGgGCGTGGTtAGGG OT4
CTTCCaCATcAGCGaGGTCAaGG OT3
CTgTCCACATcAGaGTGGTCAtGG OT2
CTTCCACATGAGCcaGT-GTCAaGG OT1
gTTCgAgATGAGCGTGGTCAaGG
CTTCCACATGAGCGTGGTCAGGG On-target
ABE8e
ABE9
ABE3.1
ABE10
Untreated
0.0047
0.0012
0.5 3 6 20 40 60 80 100
Cumulative A-to-G editing in the editing window (%)
HEK site3
GagCCAGAaTGAGCACGTGAgGG OT5
aGaCCAGACTGAGCAaGaGAgGG OT4
aGCtCAGACTGAGCAaGTGAgGG OT3
atatttGctgGAGCACGTGAaGG CHIP-OT5
GttgtgaACTGAGCACGTGAgGG CHIP-OT4
GtCaggGgaaGAGCACGTGAcGG CHIP-OT3
aaggCtGAggGAGCACGTGAaGG CHIP-OT2
caggaAGctgGAGCACGTGAgGG CHIP-OT1
GGCCCAGACTGAGCACGTGATGG On-target
0.0029
<0.0001
0.5 5 10 20 40 60 80 100
Cumulative A-to-G editing in the editing window (%)

b


ABE8e
ABE9
ABE3.1
ABE10
Untreated
A-to-G (%)
30
20
12
8
4
0.5
0.0002
0.0075
0.0010
0.0079
<0.0001
R-loop1
R-loop2
R-loop3
R-loop4
R-loop5
R-loop6

c

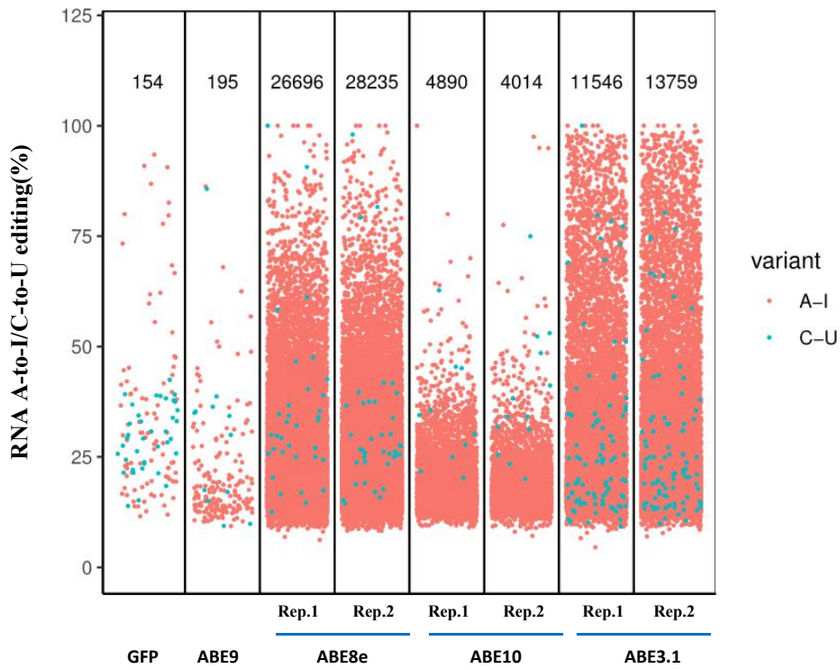

RNA A-to-I/C-to-U editing(%)
125
100
75
50
25
0
154
195
26696
28235
4890
4014
11546
13759
variant
A–I
C–U
Rep.1
Rep.2
Rep.1
Rep.2
Rep.1
Rep.2
GFP
ABE9
ABE8e
ABE10
ABE3.1

Fig. 3 | Off-target assessments of ABE10. a, Cas9-dependent DNA on and off-target analysis of the indicated targets (PD-1-sg4 and HEK3) by ABE8e, ABE9, ABE3.1 and ABE10 in HEK293T cells. Lowercase protospacer sequences represent mismatched bases compared to their corresponding on-target sequences. Data are means ± SD (n = 3 independent experiments). **b**, Cas9-independent DNA off-target analysis of ABE8e, ABE9, ABE3.1 and ABE10 using the modified orthogonal R-loop assay at each R-loop site with nSaCas9-sgRNA. Data are means ± SD (n = 3 independent experiments). **c**, RNA off-target editing activity by ABE8e, ABE9, ABE3.1 and ABE10 using RNA-seq. Each biological replicate is listed on the bottom. Significance was tested by two-tailed Student's t test (**a**, **b**).

Figure 4

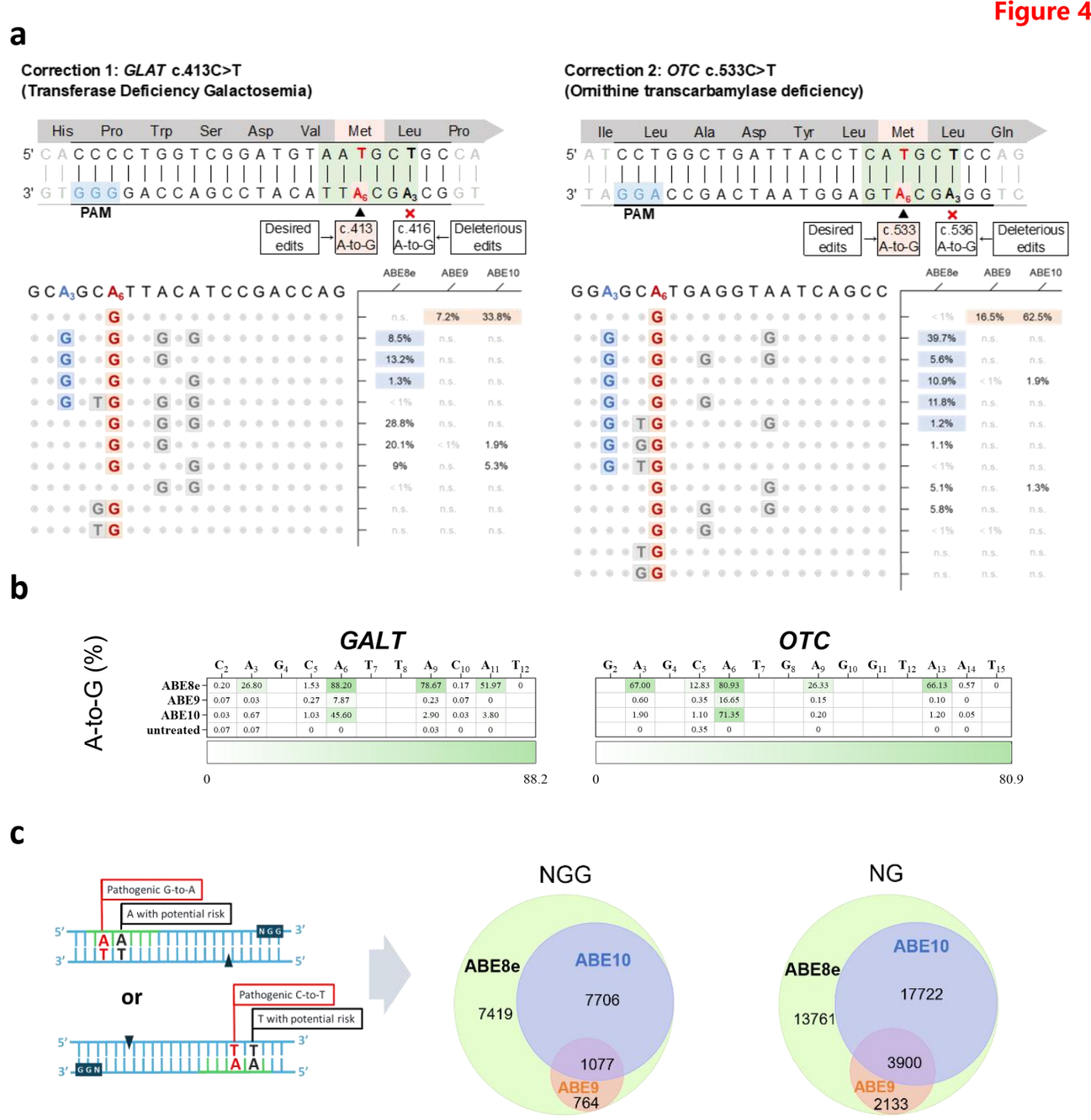
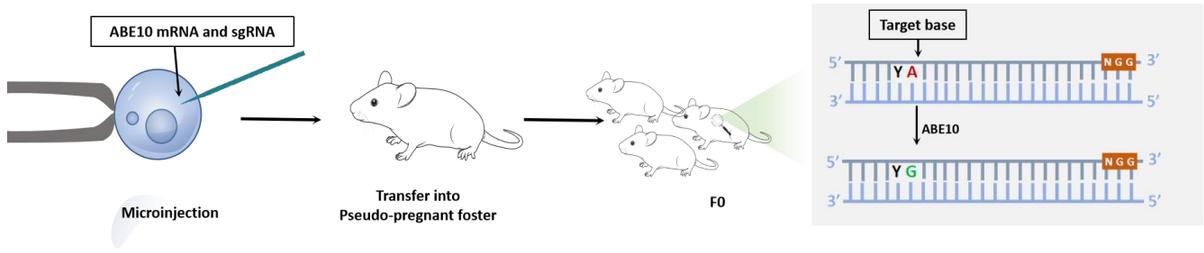

a
Correction 1: GLAT c.413C>T
(Transferase Deficiency Galactosemia)
His Pro Trp Ser Asp Val Met Leu Pro
PAM
Desired edits
c.413 A-to-G
c.416 A-to-G
Deleterious edits
ABE8e ABE9 ABE10
Correction 2: OTC c.533C>T
(Ornithine transcarbamylase deficiency)
Ile Leu Ala Asp Tyr Leu Met Leu Gln
PAM
Desired edits
c.533 A-to-G
c.536 A-to-G
Deleterious edits
ABE8e ABE9 ABE10
b
A-to-G (%)
GALT
OTC
ABE8e
ABE9
ABE10
untreated
0
88.2
0
80.9
c
Pathogenic G-to-A
A with potential risk
or
Pathogenic C-to-T
T with potential risk
NGG
NG
ABE8e
ABE10
ABE9
7419
7706
1077
764
13761
17722
3900
2133

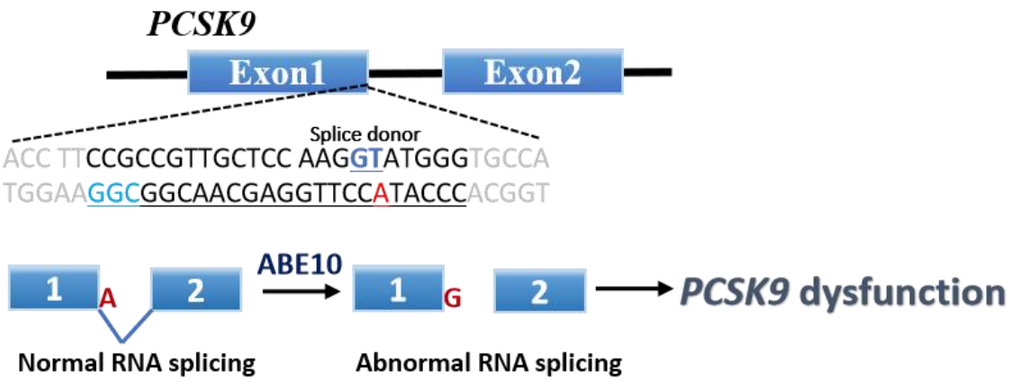
Fig. 4 | Pathogenic mutations corrected with high precision by ABE10. a, The allele base editing by ABE8e, ABE9 and ABE10 at HEK293T stable cells carrying 2 pathogenic mutations (*GLAT* c.413C>T, Transferase Deficiency Galactosemia; *OTC* c.533C>T, Ornithine transcarbamylase deficiency) adjacent to bystander mutations that are deleterious in Clin Var database. **b**, The A-to-G base editing of by ABE8e, ABE9 and ABE10 at the HEK293T stable cells. Data are means (n = 3 independent experiments). **c**, The statistics of pathogenic mutations corrected by ABE8e, ABE9 and ABE10 in Clin Var database (accessed July 16, 2024). The image on the left represents the schematic diagram of the precise correction of pathogenic mutations by ABEs, and the image on the right represents the total number of pathogenic mutations corrected by ABEs as PAM expands from NGG to NGN.

Figure 5

a



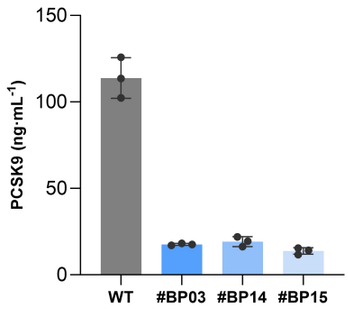
b



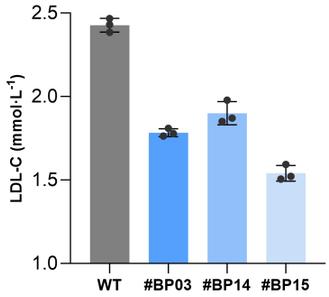
c

WT	GAAGGCGGCAACGAGGTTCCATACCCGGA	Frequency(%)
	ABE8e	
#AP02 ♂	GAAGGCGGCAACGAGGTTCCGTGCCCGGA	54.9
	GAAGGCGGCAACGAGGTTCCGTACCCGGA	20.7
#AP04 ♀	GAAGGCGGCAACGAGGTTCCGTGCCCGGA	69.1
	GAAGGCGGCAACGAGGTTCCGTACCCGGA	4.5
	ABE10	
#BP03 ♂	GAAGGCGGCAACGAGGTTCCGTACCCGGA	89.38
#BP14 ♀	GAAGGCGGCAACGAGGTTCCGTACCCGGA	86.73
#BP15 ♂	GAAGGCGGCAACGAGGTTCCGTACCCGGA	90.09

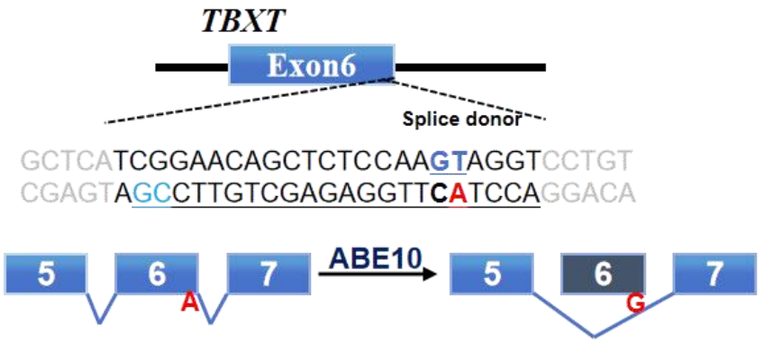
d



e



f



g

WT	AGGACCTACTTGGAGAGCTGTTCCGA	Frequency(%)
#BT01 ♀	AGGACCTGCTTGGAGAGCTGTTCCGA	34.96
#BT02 ♂	AGGACCTGCTTGGAGAGCTGTTCCGA	28.57
#BT03 ♂	AGGACCTGCTTGGAGAGCTGTTCCGA	43.24
#BT04 ♂	AGGACCTGCTTGGAGAGCTGTTCCGA	33.16
#BT05 ♀	AGGACCTGCTTGGAGAGCTGTTCCGA	44.52

h

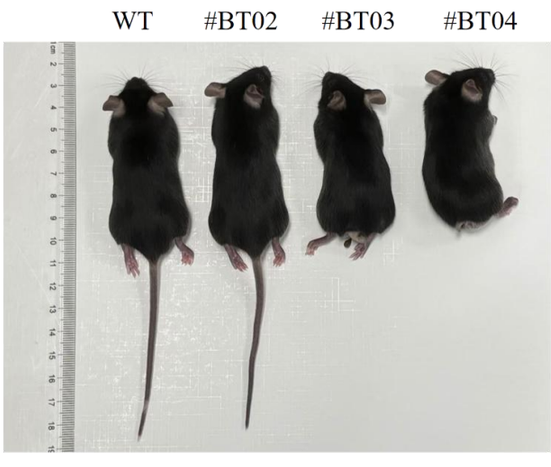


Fig. 5 | The generation of mouse models with high precision by ABE10. a, The workflow for generating mouse models by ABEs. **b**, The target sequence of ABE10 or ABE8e in the proprotein convertase subtilisin/kexin type 9 (*PCSK9*) exon 1 locus. The PAM sequence and sgRNA target sequence are shown in green and bold, respectively. The desired A-to-G conversion to disrupt the opposite DNA strand of the GT splice site is shown in red. **c**, Genotyping of representative F0 pups of *PCSK9* base editing by ABE8e and ABE10. The frequencies of wild-type (WT) and mutant alleles were determined by analyzing HTS using BE-Analyzer. The percentage values on the right represent the frequencies of the indicated mutant alleles. The frequency of the wild-type allele was omitted. **d**, **e**, The expression of PCSK9 and LDL-C in plasma of F0 pups from 8-week-old wild-type and founder (#BP03,#BP14 and #BP15) mice. Data are means ± SD (n = 3 independent experiments) **f**, The target sequence of ABE10 or ABE8e *TBXT* exon 6 locus. The PAM sequence and sgRNA target sequence are shown in green and bold, respectively. The desired A-to-G conversion to disrupt the opposite DNA strand of the GT splice site is shown in red. **g**, Genotyping of representative F0 pups of *TBXT* base editing by ABE10. The frequencies of wild-type (WT) and mutant alleles were determined by analyzing HTS using BE-Analyzer. The percentage values on the right represent the frequencies of the indicated mutant alleles. The frequency of the wild-type allele was omitted. **h**, The tail-loss phenotype of *TBXT* base editing F0 pups from 8-week-old wild-type and founder (#BT02,#BTO3 and #BT04) mice.

Figure 6

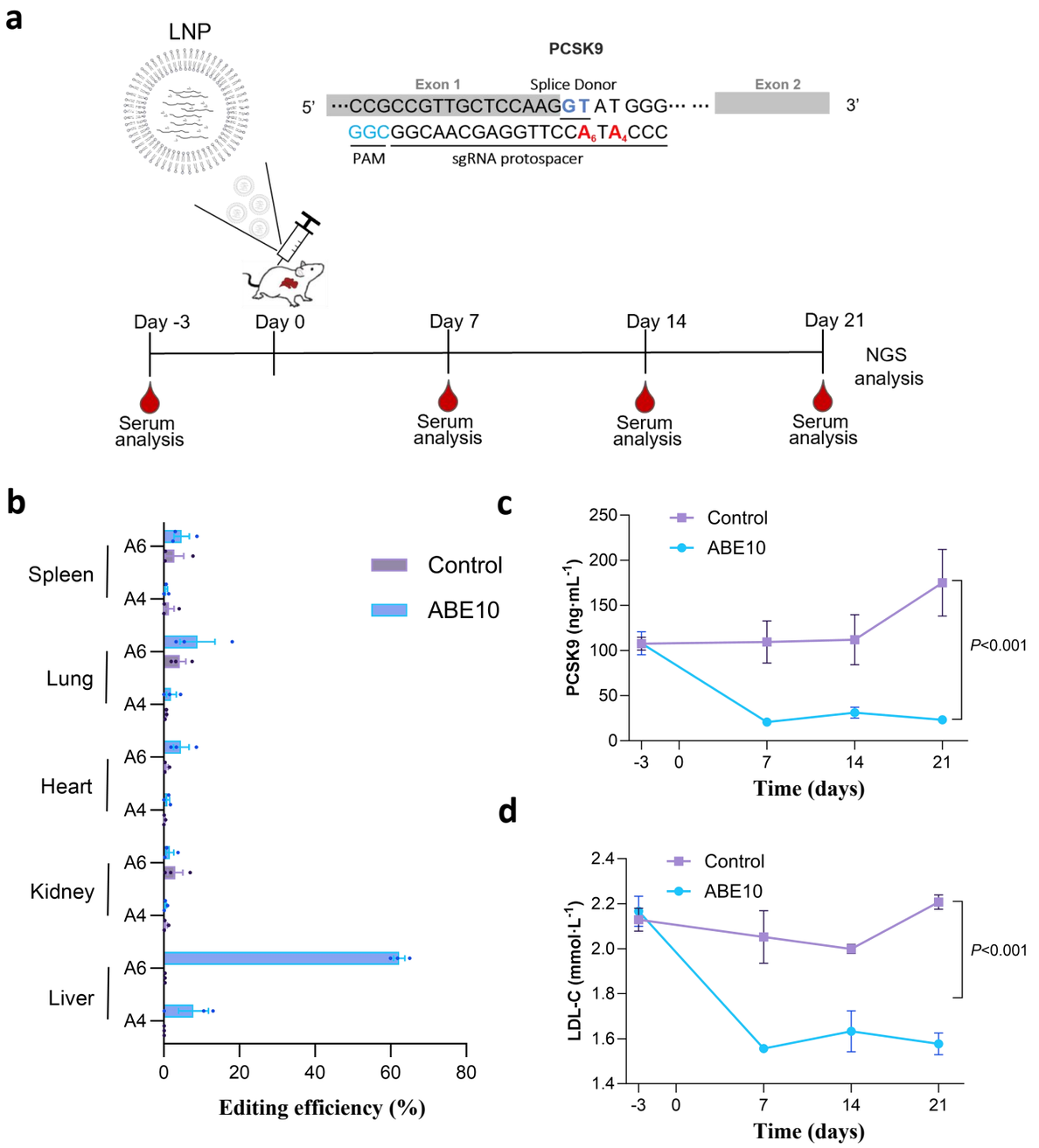

a
LNP
PCSK9
Exon 1
Splice Donor
Exon 2
5'
···CCGCCGTTGCTCCAAGGT AT GGG··· ···
3'
GGCGGCAACGAGGTTCCA6 TA4CCC
PAM
sgRNA protospacer
Day -3
Day 0
Day 7
Day 14
Day 21
NGS analysis
Serum analysis
Serum analysis
Serum analysis
Serum analysis
b
Spleen
Lung
Heart
Kidney
Liver
A6
A4
Control
ABE10
0
20
40
60
80
Editing efficiency (%)
c
Control
ABE10
PCSK9 (ng·mL-1)
250
200
150
100
50
0
-3
0
7
14
21
Time (days)
P<0.001
d
Control
ABE10
LDL-C (mmol·L-1)
2.4
2.2
2.0
1.8
1.6
1.4
-3
0
7
14
21
Time (days)
P<0.001

Fig. 6 | The mouse PCSK9 base editing in vivo using ABE10. a, The workflow for the mouse PCSK9 base editing in vivo by the delivery of LNPs packaged ABE10. The target sequence of ABE10 in the *PCSK9* exon 1 locus. The PAM sequence and sgRNA target sequence are shown in pink and blue, respectively. The desired A-to-G conversion to disrupt the opposite DNA strand of the GT splice site is shown in blod. **b**, The A-to-G editing efficiency of PCSK9 in adult mice 3-weeks after **t**he delivery of LNPs packaged ABE10 and sgRNA. Data are means ± SD (n = 3 different mice) **c**,**d**, The expression of PCSK9 and LDL-C in plasma of adult mice before and after **t**he delivery of LNPs packaged ABE10 and sgRNA. Data are means ± SD (n = 3 for the control group, n = 5 for the ABE10 group).

370 **Methods**

371 **Ethical statement**

372 The research was conducted in accordance with all relevant ethical regulations. All
373 procedures involving mice were meticulously designed to align with established
374 guidelines and have received approval from the Institutional Animal Care and Use
375 Committees (IACUCs) at the Suzhou Institute of Systems Medicine, Chinese Academy
376 of Medical Sciences & Peking Union Medical College, Suzhou, China.

377 **Plasmid construction**

378 The DNA sequences used in this research can be found in Supplementary Sequences.
379 The ABE8e (#138489) and lentiCRISPR v2 (#52961) plasmids were obtained from
380 Addgene. Polymerase chain reaction (PCR) was conducted using KOD-Plus-Neo DNA
381 Polymerase (TOYOBO, Code: KOD-401). The plasmids generated in this article,
382 including those based on ABE8e or lentiCRISPR v2 backbones, were constructed using
383 the ClonExpress MultiS One Step Cloning Kit (Vazyme) (Supplementary Sequence).
384 sgRNA expression plasmids were constructed as described previously[19]. Briefly,
385 oligonucleotides listed in Supplementary Table 4 were denatured at 95°C for 5 min
386 followed by slow cooling to room temperature. Annealed oligonucleotides were ligated
387 into BbsI-linearized U6-sgRNA(sp)-EF1α-GFP for sgRNA expression (Thermo Fisher
388 Scientific).

389 **Human Cell culture**

390 HEK293T (ATCC CRL-3216) cell lines were cultivated in Dulbecco's Modified Eagle's
391 medium (DMEM, Gibco) with 10% (vol/vol) fetal bovine serum (FBS, Gibco) and 1%
392 Penicillin-Streptomycin (Gibco) antibiotic mix. HUDEP-2 cells were maintained and
393 expanded in serum-free expansion medium (Stem Cell Technologies) supplemented
394 with human Stem Cell Factor (SCF, 50 ng ml^{-1}, PeproTech), erythropoietin (EPO, 3 IU
395 ml^{-1}, PeproTech), dexamethasone (DEX, 1 μM, Sigma), doxycycline (DOX, 1 μg ml^{-1},
396 Takara Bio) and 2% penicillin–streptomycin (Gibco). All cell lines used were cultivated
397 at 37 °C, 5% CO_2 in the incubator.

398 **Stable cell line generation**

399 The HEK293T stable cell line was constructed by cloning a 200-bp fragment containing
400 disease-associated mutation flanked by ~100 bp into lentiviral vector with puromycin-
401 resistant gene expression maker. In brief, We used a combination of 12 μg of the
402 specific lentiviral transfer plasmids (Lenti-GALT-sg1 and Lenti-OTC-sg1), along with 6

403 μg of pMD2.G and 9 μg of psPAX2, for co-transfection into HEK293T cells that were
404 nearly 85% confluent in a 10-cm dish. The virus-containing supernatant was harvested
405 at 72 hours post-transfection. The supernatant was subjected to centrifugation at 4,000
406 rpm for 10 minutes at 4 °C, with the objective of precipitating cell debris. Following
407 this, the supernatant was filtered through a 0.45 μm low protein binding membrane
408 (Millipore). And then, it was serially diluted and added to a 24-well plate, each well
409 containing 5×10^4 HEK293T cells. After 24 hours, the lentivirus-transduced cells were
410 spread into new plate wells supplemented with puromycin (2 μg ml^{-1}). Following a
411 week of puromycin selection, the pooled cells were spread into 96-well plate and single
412 clone cells were harvested and expand for cell transfection.

413 **Cell transfection and fluorescence-activated cell sorting (FACS)**

414 In DNA base editing experiments, HEK293T cells were plated in 24-well plates and
415 transfected when they reached about 80% confluency. Next, a mixture of 3 μl
416 polyethyleneimine (PEI, Polysciences), 1 μg plasmid DNA (750 ng ABEs or A&C-BEs
417 expression plasmid and 250 ng sgRNA expression plasmid) and serum-free medium
418 were applied to the cells. After three days, transfected cells were digested with 0.25%
419 trypsin (Gibco) for fluorescence-activated cell sorting (FACS). Genomic DNA was
420 isolated using the QuickExtract™ DNA Extraction Solution (QE09050, Epicentre)
421 according to the manufacturer's instructions. For RNA off-target analysis, HEK293T
422 cells were seeded into 10-cm dishes and transfected with 30 μg of Cas9n-P2A-GFP,
423 ABE8e-P2A-GFP, ABE9-P2A-GFP, ABE10-P2A-GFP and ABE3.1-P2A-GFP using
424 PEI at approximately 80% confluency. After three days, transfected cells were washed
425 with phosphate-buffered saline and digested with 0.25% trypsin (Gibco) for
426 fluorescence-activated cell sorting (FACS). FACS was carried out on a FACSAriaIII
427 (BD Biosciences) using FACSDiva version 8.0.2 (BD Biosciences). Cells were gated on
428 their population via forward/sideward scatter after doublet exclusion (Supplementary
429 Note). Then, cells (300,000 to 400,000 cells) with top 15% GFP signal were collected,
430 and total mRNA was extracted using RNAiso Plus (Takara).

431 **Enhanced orthogonal R-loop assay**

432 In this study, a modified orthogonal R-loop assay was employed for Cas9-independent
433 DNA off-target analysis, which replaced dSaCas9-sgRNA plasmid with nSaCas9-
434 sgRNA plasmid at each R-loop site. For the transfection process, a mixture of 3 μl
435 polyethyleneimine (PEI, Polysciences), 850 ng plasmid DNA (250 ng SpCas9 sgRNA

436 plasmid, 300 ng base editor plasmid (ABE8e, ABE9, ABE10, ABE3.1) and 300 ng
437 nSaCas9 containing sgRNA plasmid were added to the cells. After three days,
438 transfected cells were treated with 0.25% trypsin (Gibco) for sorting and then genomic
439 DNA was extracted using the QuickExtract™ DNA Extraction Solution (QE09050,
440 Epicentre) , adhering to the manufacturer's instructions.

441 **RNA sequencing (RNA-Seq) experiments**

442 RNA sequencing was performed as previously described[19]. A total of 3 μg RNA per
443 sample was used as input material for sample preparations. Sequencing libraries were
444 prepared using a NEBNext Ultra RNA Library Prep Kit for Illumina (NEB) following
445 the manufacturer's recommendations. After assessing library quality with the Agilent
446 Bioanalyzer 2100, index codes were added to assign sequences to individual samples.
447 The index-coded samples were clustered on a cBot Cluster Generation System using the
448 TruSeq PE Cluster Kit v3-cBot-HS (Illumina) according to the manufacturer's
449 guidelines. Following cluster generation, the libraries were sequenced on an Illumina
450 HiSeq platform, generating 125-bp or 150-bp paired-end reads.

451 **RNA sequence variant calling and quality control**

452 RNA sequence variant calling and quality control were performed as previously
453 described as follows[17]. Raw fastq format data were initially processed through in-house
454 Perl Scripts. Adapter sequences were removed, and low-quality bases were trimmed to
455 produce clean data using Trimmomatic. Concurrently, the Q20, Q30, and GC content of
456 the clean data were calculated. All further analyses were conducted using this high-
457 quality clean data. The reference genome index was built using HISAT2 version 2.0.5
458 and paired-end clean reads were aligned to the reference genome (Ensemble GRCh38)
459 with the same HISAT2 version. Single-nucleotide polymorphism calling was carried out
460 using GATK version 4.0. Variant loci in the base editor overexpression groups were
461 filtered to eliminate sites without high-confidence reference genotype calls in the
462 control group.

463 **High-throughput DNA sequencing and data analysis**

464 On- and off-target genomic regions of interest were amplified via PCR using flanking
465 high-throughput sequencing primer pairs listed in the Supplementary Table 2. PCR
466 amplification was carried out with KOD-Plus-Neo DNA Polymerase according to the
467 manufacturer's instructions with 100-150 ng of genomic DNA as a template. For the
468 preparation of high-throughput sequencing (HTS) libraries, site-specific primers

469 containing an adaptor sequence (forward 5′-GGAGTGAGTACGGTGTGC-3′;
470 backward 5′-GAGTTGGATGCTGGATGG-3′) at the 5′ end were employed in the PCR.
471 Subsequently, the products were then subjected to a second-round PCR using primers
472 with different barcode sequences. Then, PCR products with different tags were pooled
473 together for deep sequencing using the Illumina HiSeq platform. For the batch analysis
474 of the resulting FASTQ files, the reference sequences were set to full-length and
475 analyzed as previously described[23.] The base editing or indels efficiencies were
476 quantified using using BE-Analyzer[23] or CRISPResso[24].

477 **Preparation of mRNAs and sgRNAs and Microinjection in mice**

478 sgRNAs with the modification of phosphorothioation and methoxy group were
479 synthesized by GenScript (Nanjing, China) (Supplementary Sequence). mRNA
480 preparation was performed as previously described[19]. In brief, the T7 promoter was
481 introduced into ABE8e or ABE10 template by PCR using the primers T7-ABEs-mRNA-
482 F/R (Supplementary Table 2). The base editor mRNA were transcribed in vitro using
483 mMESSAGE mMACHINE T7 Kit (Invitrogen) and purified using a MEGAclear Kit
484 (Invitrogen)[20].

485 C57BL/6J and ICR mice, housed in a specific pathogen-free environment with a 12-
486 hour light/dark cycle and ad libitum diet, were used as embryo donors and foster
487 mothers, respectively. A 2 nl mixture of ABE8e or ABE10 mRNA (200 ng μl^{-1}) and
488 sgRNA (100 ng μl^{-1}) was co-injected into one-cell stage wild-type embryos. Injected
489 zygotes were transferred into pseudopregnant female mice immediately using an
490 Eppendorf TransferMan NK2 micromanipulator. About 20 days After embryo transplant,
491 genomic DNA from born pups was isolated using the QuickExtract™ DNA Extraction
492 Solution (QE09050, Epicenter) according to the manufacturer's instructions.

493 **Animal Injection and Processing**

494 All mice cohorts were maintained at specific pathogen-free (SPF) facilities in Suzhou
495 Institute of Systems Medicine and approved by Institutional Animal Care and Use
496 Committee. A total 200 μl mix of LNP packaged ABE10 mRNA and a sgRNA targeting
497 the PCSK9 (40 μg: 20 μg) (Starna Therapeutics Co., Ltd., Suzhou) was delivered to 6-8
498 weeks old C57/BL6 mice intravenously via tail vein injection. A control group received
499 an equivalent volume of normal saline. To track the serum levels of PCSK9 and Low-
500 density lipoprotein cholesterol (LDL-C), mice were fasted overnight for 12 hours prior
501 to blood collection via tail tip sampling. Blood samples were allowed to clot at room
502 temperature, after which serum was separated by centrifugation and stored at -20°C for

503 subsequent analysis. For terminal procedures, mice were euthanized using carbon
504 dioxide inhalation. The median lobe of the liver was excised for DNA extraction to
505 evaluate editing efficiency of on-target and off-target sites.

506 **Serum analysis**

507 To minimize batch effects, serum samples from all time points were collected and
508 analyzed concurrently. PCSK9 levels were measured using an ELISA kit (Proteintech,
509 #KE10050), while LDL-C, ALT, and AST levels were assessed using assay kits from
510 Solarbio (#BC5335, #BC1555, #BC1565, respectively). All procedures were conducted
511 in strict adherence to the manufacturers' instructions.

512 **Statistics and reproducibility**

513 Three biologically independent replicates, conducted on different days, were used to
514 calculate means and SD unless stated otherwise. All bar plots and figures were
515 generated using Prism 9.3 (GraphPad). The statistical significance of differences
516 between two groups was assessed using an unpaired two-tailed Student's t-test within
517 Prism 9.3 (GraphPad). The specific P values can be found in the captions accompanying
518 the figures. A P value of less than 0.05 was deemed to indicate significance. RNA-seq
519 data was analyzed using Trim Galore (version 0.6.6), STAR (version 2.7.1a), SAMtools
520 (version 1.14), Picard MarkDuplicates module (version 2.23.9) software. For the
521 analysis of FACS data, FlowJo v.10 was employed.

522 **Data availability**

523 The raw high-throughput sequencing data generated in this study have been submitted
524 to the NCBI sequence Raed Archive database under SUB14831381, SUB14822433,
525 SUB14822417. Additionally, the RNA-seq data are accessible in the NCBI sequence
526 Read Archive database under accession code PRJNA818975. All plasmids described in
527 this work will be deposited to Addgene. Source data are provided with this paper.

528 **Acknowledgments:**

529 We thank Professor Dali Li from East China Normal University for providing many
530 base editors plasmids and some necessary experimental materials to accelerate the
531 process of this project when I set up my research group. We thank Yuanqing He and
532 Naiyun Ma from the Laboratory Animal Science Center of Suzhou Institute of Systems
533 Medicine for their help in generating base-edited mice and the fund of NCTIB Cell and
534 Gene Therapy R&D Platform. This work was partially supported by grants from
535 National Key R&D Program of China (2022YFC3400203 to X.Z. and

2022YFA1103401 to X.Z.), the National Natural Science Foundation of China (No.32201223 to X.Z.), the Non-profit Central Research Institute Fund of Chinese Academy of Medical Sciences (2022-RC180-08 to X.Z.), the CAMS Innovation Fund for Medical Sciences (2022-I2M-1-024 to X.Z., and 2022-I2M-2-004 to Z.G.) and the Suzhou Municipal Key Laboratory (SZS2022005) .

Author contributions:

M. S., Y.L., Q.C., J.R., J.W. and X.G. performed most of the experiments and analyzed the data. X.Z. designed and supervised the project. X.Z. wrote the manuscript with the input from all the authors.

Competing interests:

X.Z., H.L., J.R., Y.L. and Y.H. have submitted patent applications (application no. 202311786372.7, under review) based on the results reported in this study. This patent mainly relates to ABE10 in this paper. The remaining authors declare no competing interests.

References:

22. Yang, L. et al. Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell* **9**, 814–819 (2018).
23. Hwang, G.-H. et al. Web-based design and analysis tools for CRISPR base editing. *BMC Bioinform.* **19**, 542 (2018).
24. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat. Biotechnol.* **37**, 224–226 (2019).